

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Ingénieries Microbienne et Enzymatique

Présentée et soutenue par :

Mme CLEA LACHAUX

le jeudi 20 septembre 2018

Titre :

Remaniement du métabolisme carboné central microbien pour une conversion des sucres lignocellulosiques en acide glycolique avec une conservation optimale du carbone

---

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP)

Directeur de Thèse :

M. JEAN-MARIE FRANCOIS

Rapporteurs :

M. JOHANNES GEISELMANN, UNIVERSITE GRENOBLE ALPES
M. MATTHIEU JULES, AGROPARISTECH

Membres du jury :

MME MURIEL GONDRY, CEA SACLAY, Présidente
M. FABIEN LETISSE, INSA TOULOUSE, Membre
M. JEAN-MARIE FRANCOIS, INSA TOULOUSE, Membre
MME MADELEINE BOUZON, CEA GENOSCOPE, Membre

Résumé

L'acide glycolique est synthétisé à l'échelle industrielle à partir de formaldéhyde toxique dont toute trace est interdite dans les produits à destination de l'industrie cosmétique et médicale. Sa production par fermentation microbienne à partir de carbone renouvelable présente donc un grand intérêt industriel. La lignocellulose est la source de carbone renouvelable la plus abondante sur terre, son hydrolyse libère principalement des hexoses (i.e glucose) et des pentoses (i.e xylose, arabinose). Les procédés microbiens existant de valorisation de la lignocellulose, pour la production d'acide glycolique, requièrent la combinaison de voies métaboliques optimisées pour l'assimilation des pentoses ou des hexoses. Conformément aux principes d'ingénierie métabolique les rendements de ces voies ont été maximisés en augmentant le flux vers l'acide glycolique et en réduisant la formation de sous-produits. Étonnement, aucune de ses études ne considère la production de CO2 comme une perte de carbone majeure pour la production. Nous proposons de développer une nouvelle voie métabolique cyclique qui se démarque des procédés existants par sa capacité à convertir l'ensemble des sucres lignocellulosiques, avec la volonté de réduire l'émission de CO2. Cette thèse présente la conception et l'évaluation de deux approches répondant à cet objectif.

La première approche repose sur l'utilisation de la voie synthétique 'xylulose-1P' (X1P) de conversion du D-xylose en acide glycolique (Alkim *et al.*, 2015, 2016; Cam *et al.*, 2016). Les pentoses (C5) et hexoses (C5) sont naturellement convertis dans *E. coli* en D-xylulose-5P par la voie des pentoses phosphate (PP). L'expression d'une mutase ou phosphatase active sur D-xylulose-5P rendrait la voie X1P accessible au D-glucose et L-arabinose. 7 phosphatases et 3 mutases bactériennes et levuriennes candidates ont été sélectionnées à cet effet.

La deuxième approche présentée est entièrement originale. Le D-ribulose-5P est un métabolite commun au catabolisme des C5 et C6 dans *E.coli* , au même titre que le D-xylulose-5P. Nous avons identifié trois enzymes clefs ayant le potentiel de former une nouvelle voie de production d'acide glycolique à partir de D-ribulose-5P et ainsi valoriser les C5 et C6. Le métabolisme d'*E.coli* a été remanié pour s'affranchir de l'étape de décarboxylation du pyruvate et limiter la perte de carbone sous forme de CO2. Le rendement théorique maximal est de 3 mol d'acide glycolique par mol d'hexoses et 2,5 mol d'acide glycolique par mol de pentoses, il s'agit des plus hauts rendements rapportés à ce jour.

L'évaluation de la faisabilité de ces deux approches a été réalisée sur la base de tests enzymatiques *in vitro* et de screening fonctionnels *in vivo* par test de croissance. La mesure de la croissance permet dans ce cas d'apprécier la capacité des enzymes étudiées à rediriger le flux de carbone vers l'acide glycolique. La deuxième approche a conduit à la conception d'une souche productrice qui a été caractérisée, nous avons démontré la production d'acide glycolique à partir de D-glucose (0,34 g/l) et D-xylose (0,4 g/L). . Les suivis de croissance, l'analyse de l'endométabolome et de l'exométabolome et la réalisation d'un bilan carbone ont permis d'identifier les éléments limitants la production. Ces connaissances sont essentielles à la réalisation de la prochaine étape de développement qui consistera à poursuivre l'ingénierie de l'organisme hôte et développer un procédé de fermentation où la production d'acide glycolique et de biomasse seront découplés. Ce procédé biotechnologique a le potentiel d'atteindre des rendements en acide glycolique supérieurs aux procédés existants, valorisant de façon optimale l'ensemble des sucres lignocellulosiques.

Mots clés

Escherichia coli, conservation du carbone, acide glycolique, ingénierie métabolique

Abstract

On an industrial scale, glycolic acid is produced from formaldehyde of which, due to its high toxicity, no traces are allowed in the end products destined for medical or cosmetic applications. Because of this reason, amongst others, its production by microbial fermentation of renewable carbon sources is of great industrial interest. Lignocellose is the most abundant, renewable carbon source on this planet and its hydrolysis yields mainly hexoses (i.e. glucose) and pentoses (i.e. xylose, arabinose). Hence, the existing microbial processes for the valorization of lignocellulose by the production of glycolic acid require the combination of metabolic pathways, optimised for the assimilation of pentoses and hexoses. Following the principles of metabolic engineering, production yields have been maximized by the combining an increased flux towards the production of glycolic acid with a reduction of by-product formation. Surprisingly, none of these studies consider the significant loss of carbon via the production of CO_2. The here-proposed novel metabolic pathway distinguishes itself from the existing processes by its capacity to convert all lignocellulosic sugars whilst reducing CO_2 emissions. The development and evaluation of two such approaches is presented in this PhD.

The first approach uses the synthetic 'xylulose-1-P' (X1P) pathway for the conversion of D-glucose to glycolic acid (Alkim *et al.*, 2015, 2016; Cam *et al.*, 2016). The pentoses (C5) and hexoses (C5) are naturally converted by *E. coli* into D-xylulose-5P through the pentose phosphate pathway (PP). By expressing a mutase or phosphatase active on D-xylulose-5, the X1P pathway is opened up to D-glucose and L-arabinose. To this end, 7 phosphatases and 3 bacterial/yeast mutases have been selected as candidates.

The second approach is completely novel and starts from D-ribulose-5P which is, much like D-xylulose-5P, a common metabolite in the C5 and C6 catabolism of *E.coli*. Three key enzymes have been identified to constitute a new metabolic pathway for the production of glycolic acid from D-ribulose-5P and thereby valorize the C5 and C6 sugars. *E.coli*'s metabolism has been modified to avoid puyruvate decarboxylation and loss of carbon as CO2. The maximal theoretical yield is the highest reported so far with 3 and 2.5 mol of glycolic acid produced per mol of hexoses and pentoses respectively.

The feasibility of both approached has been tested by *in vitro* enzymatic tests and through *in vivo* functional screening during growth essays. The enzymes' capacity to redirect the carbon flux towards glycolic acid production could be evaluated by measuring growth. The second approach led to the development and characterization of a strain producing 0.34g/L and 0.4g/L of glycolic acid from D-glucose and D-xylose respectively. Follow-up of the growth, the endo- and exometabolome analysis and carbon mass balancing allowed to identify some bottlenecks which should be adressed by advanced engineering of the host metabolism. to obtain a fermentation process in which glycolic acid production and growth are decoupled. Such biotechnological process has the potential to achieve yields superior to those reported for other existing processes and thereby assure an optimal valorization of all lignocellulosic sugars.

Key words

Escherichia coli, carbon conservation, glycolic acid, metabolic engineering

Abréviations

Micro-organismes

E.coli	*Escherichia coli*
C.glutamicum	*Corynebacterium glutamicum*
S.cerevisiae	*Saccharomyces cerevisiae*

Métabolites

A5P	Arabinose-5-phosphate
ADP	Adénosine diphosphate
AG	Acide glycolique
ARN	Acide ribonucléique
ATP	Adenosine triphosphate
C5	Pentose
C6	Hexose
CO2	Dioxyde de carbone
DHA	Dihydroxyacétone
DHAP	Dihydroxyacétone phosphate
E4P	Erythrose-4- phosphate
EG	Ethylène glycol
F6P	Fructose-6-phosphate
FPP	Farnésyl diphosphate
G6P	Glucose-6-phosphaye
GAP	Glycéraldéhyde-3-phosphate
GO	Glycolaldéhyde
IPTG	Isopropyl-beta-D-l-triolgalactopyranoside
KFA	KdsD-Fsa-AldA
KO8P	3-Deoxy-D-manno-octulosonate-8-phosphate
LPS	Lipopolysaccharides
NADH	Nicotinamide adenine dinucleotide
P	Phosphate
PEP	Phosphoenolpyruvate
pNPP	Para-nitrophényl-phosphate

PP	Pentoses Phosphate
Ru5P	Ribulose-5-phosphate
Ru5P	Ribulose-5-phosphate
S7P	Sedoheptulose-7- phosphate
X1P	Xylulose-1P

Enzymes

AldA	Aldéhyde déshydrogénase
AldA	glycolate oxydase/Aldéhyde deshydrogénase
AraA	Arabinose isomérase
AraB	Ribulose kinase
AraD	L-ribulose-5-phosphate-4-épimérase
DTE	D-tagatose-3-epimerase
FLP	Flippase
Fsa	Fructose-6-phosphate aldolase
FucA	Fuculose phosphate aldolase
FucK	Fuculose kinase
GapA	Glycéraldéhyde-3P déshydrogénase
Glk	Glucokinase
HAD	Haloacide déshalogénase
IDH	Isocitrate déshydrogénase
KdsD	Arabinose-5P isomérase
KhkC	Xylulose kinase
Pgi	Phosphoglucose isomérase
PGM	Phosphoglucomutase
PTS	Phosphotransférase
Rha	L-rhamnose
Rpe	Ribulose-5-phosphate-3-épimérase
Tkt	Trancétolase
XylA	Xylose épimérase
XylB	Xylulose kinase

AUTRES

AMM	Autorisation de mise sur le marché
CE	Commission Européenne
DOE	Department of Energy
HPLC	High Performance Liquid Chromatography
METEX	METabolic Explorer
MIT	Massachusetts Institute of Technology
NOG	Non Oxydative Glycolysis
OCDE	Organisation de Coopération et de Développement Economique
ORT	Oxygen transfer rate
PGA	Polyglycolate
RBS	Ribosome Binding Site
TOC	Total Organic Carbon

TABLE DES MATIERES

AVANT PROPOS 1

CHAPITRE I. INTRODUCTION GENERALE 5

1. Contexte bioeconomique 7

1.1 Une economie basée sur l'exploitation de carbone fossile 7

1.2 Une economie nouvelle basée sur l'utilisation de carbone renouvelable 9

2. Choix du substrat d'étude : les sucres lignocellulosiques 11

2.1 La lignocellulose, source de carbone renouvelable 11

2.2 Le bioraffinage 13

3. Choix du microorganisme d'étude pour la valorisation des sucres lignocellulosiques : *Escherichia coli* 13

3.1 Principaux microorganismes producteurs et leurs caractéristiques 13

3.2 Escherichia coli 16

4. Approches d'ingénierie métabolique utilisées pour dépasser le potentiel naturel des microorganismes 19

4.1 Les cinq niveaux d'ingénierie métabolique 19

4.2 Optimisation du flux carboné bactérien 22

4.3 Applications biotechnologiques chez *Escherichia coli* 26

5. Choix de la molécule d'étude : l'acide glycolique 26

5.1 Molécules candidates 26

5.2 L'acide glycolique 27

6. Etat de l'art de la production d'acide glycolique par voie microbienne 28

6.1 Production naturelle par bioconversion de l'éthylène glycol 29

6.2 Niveau 1 d'ingénierie métabolique : production d'AG à partir de glucose par ingénierie du cycle glyocylate naturel 29

6.3 Niveau 2 d'ingénierie métabolique : production d'AG via l'expression de voies hétérologues 38

6.4 Niveau 3 d'ingénierie métabolique : production d'AG à partir de xylose via l'expression de voies non naturelles 40

6.5 Combinaison d'approches d'ingénierie métabolique pour une valorisation des pentoses et hexoses en AG 41

7. Exemples de voies microbiennes de conservation du carbone 43

8. Objectifs principaux de la thèse 45

CHAPITRE II. MATERIELS ET METHODES 51

1. Construction des plasmides 53

1.1 Plasmides utilisés pour les études in vitro, dérivés de pET28 54

1.2 Plasmides utilisés pour les études in vivo dans E.coli 54

2. Construction des souches d'*Escherichia coli* 58

2.1 Délétions des gènes d'E.coli 60

2.2 Préparation de bactéries compétentes et transformation 61

3. Milieux et conditions de cultures 62

3.1 Composition des milieux.....62

3.2 Conditions de culture 63

4. Extraction et quantification des métabolites 64

4.1 Métabolites extracellulaires 64

4.2 Métabolites intracellulaires 64

5. Expression et purification des protéines étiquetées polyhistidine.....65

6. Essais enzymatiques 66

6.1 Activité phosphatasique 66

6.2 Activité mutasique 67

6.3 Mesure d'activité des enzymes clefs de la voie KdsD-Fsa-AldA 67

6.4 Conversion du D-xylose, L-arabinose et D-glucose en acide glycolique *in vitro*.....69

7. Etude de la faisabilité thermodynamique 70

CHAPITRE III. EVALUATION DE L'ACCESSIBILITE DE LA VOIE XYLULOSE-1P AUX HEXOSES ET PENTOSES PAR EXPRESSION D'UNE PHOSPHATASE OU MUTASE POUR LA PRODUCTION D'ACIDE GLYCOLIQUE 71

1. Introduction 73

1.1 La voie synthétique de production d'acide glycolique 'X1P' 73

1.2 L'assimilation des sucres lignocellulosiques chez Escherichia coli 74

1.3 L'intégration d'une D-xylulose-5P mutase ou phosphatase la voie X1P 75

1.4 La superfamille des Haloacid Déshalogénases 76

1.5 Sélection de D-xylulose-5P candidates 83

2. Résultats 88

2.1 Expression et purification des phosphatases 88

2.2 Activité enzymatique des phosphatases sur pNPP *in vitro* 89

2.3 Profils d'activité des phosphatases et mutases sur substrats biologiques *in vitro* 90

2.4 Evaluation de l'activité des phosphatases et mutases sur D-xylulose-5P *in vivo* 97

3. Conclusion 101

CHAPITRE IV. CONSTRUCTION D'UNE VOIE DE CONVERSION DU GLUCOSE, XYLOSE, ARABINOSE DANS *ESCHERICHIA COLI* 103

1. Introduction 105

1.1 Synthèse d'acide glycolique par voie microbienne 105

1.2 Conversion des sucres lignocellulosiques en acide glycolique avec une conservation optimale du carbone 105

2. Résultats 106

2.1 Conception d'une nouvelle voie de production d'acide glycolique 106

2.2 Preuve de concept *in vitro* 111

2.3 Preuve de concept *in vivo* 121

3. Conclusion 147

CHAPITRE V. CONCLUSION ET PERSPECTIVES 149

REFERENCES BIBLIOGRAPHIQUES 159

AVANT-PROPOS

La biomasse végétale est une source très prometteuse de carbone, abondante, peu onéreuse. Elle peut être décomposée en monosaccharides - principalement en glucose et xylose. Un des enjeux majeurs actuels en biologie de synthèse est de construire des microorganismes capables de convertir ces sucres en produits biosourcés. C'est dans ce défi que s'inscrit ma thèse, encadrée par le Prof. Jean Marie François et dédiée à la production d'acide glycolique avec une conservation optimale du carbone. L'acide glycolique est un acide α-hydroxylé largement utilisé dans l'industrie cosmétique, textile et médicale. Son marché représentait un volume de 40 000 tonnes en 2011 et son mode de fabrication à l'échelle industrielle nécessite l'emploi de dérivés du pétrole.

Nous avons conçu dans le cadre de mon projet de thèse une voie synthétique capable de convertir les principaux monosaccharides contenus dans la biomasse végétale en acide glycolique dans *Escherichia coli*. Les microorganismes bactériens possèdent un réseau métabolique d'assimilation et d'utilisation des sucres qui a été façonné par l'évolution afin de permettre une croissance cellulaire optimale en fonction de leur environnement. Au vu des attentes de la bio-économie, on peut se demander si les voies métaboliques identifiées et caractérisées sont optimales. En effet le métabolisme carboné central naturel d'*Escherichia coli*, via la glycolyse, génère une perte de carbone considérable. C'est pourquoi notre objectif n'est pas uniquement de produire de l'acide glycolique mais aussi de développer une souche dont le métabolisme est optimisé pour autoriser une perte de carbone minimale. Le reste du carbone disponible est alors dirigé vers la production d'acide glycolique, permettant une réduction de l'émission de CO_2 et un rendement carbone optimal.

Ce manuscrit se divise en 5 chapitres. Tout d'abord une synthèse bibliographique (**Chapitre I**) présente les informations nécessaires à la contextualisation et à la compréhension de ces travaux. Ainsi après une brève introduction sur la bioéconomie, les principaux paramètres définis lors de la conception du projet de thèse qui porte sur le développement d'un procédé microbien sont présentés : choix du substrat (hexoses et pentoses lignocellulosiques), choix du microorganisme hôte (*Escherichia coli*), choix de la molécule cible (acide glycolique) Une deuxième partie présente les différents outils d'études mis en œuvre au cours de ces travaux de thèse (**Chapitre II**). Les résultats obtenus font l'objet de deux chapitres. Pour répondre aux objectifs posés deux approches ont été conçues et évaluées, chacune fait l'objet d'un chapitre (**Chapitre III, IV**). Ce manuscrit se termine par une conclusion générale synthétisant les principaux résultats obtenus et les perspectives émergentes de ces travaux (**Chapitre V**).

Bonne lecture.

CHAPITRE I

INTRODUCTION GENERALE

1. CONTEXTE BIOECONOMIQUE

1.1 Une économie actuelle basée sur l'exploitation de carbone fossile

Les activités humaines exercent une influence sur le climat, libérant des gaz à effet de serre dans l'atmosphère (utilisation de combustibles fossiles) ou diminuant la captation de CO_2 (exploitation de forêts) (**Figure 1**). Un des principaux défis du XXIème siècle est de répondre aux besoins énergétiques et alimentaires d'une population mondiale en constante augmentation, tout en maitrisant notre impact sur notre environnement. Le processus d'industrialisation commencé au début du siècle dernier dans les pays développés a été corrélé à l'explosion de l'utilisation des ressources fossiles (pétrole, gaz, charbon) qui a fait augmenter la concentration en CO_2 atmosphérique. Le rejet mondial de CO_2 provenant de l'utilisation de combustibles fossiles est maintenant d'environ 7Gt de carbone/an et la concentration atmosphérique est de 400 ppm (Pacala et Socolow, 2004). Si la tendance actuelle se poursuit le rejet de CO_2 devrait doubler d'ici 2015 ce qui conduirait à un réchauffement planétaire global de plus de 4° (Pacala et Socolow, 2004).

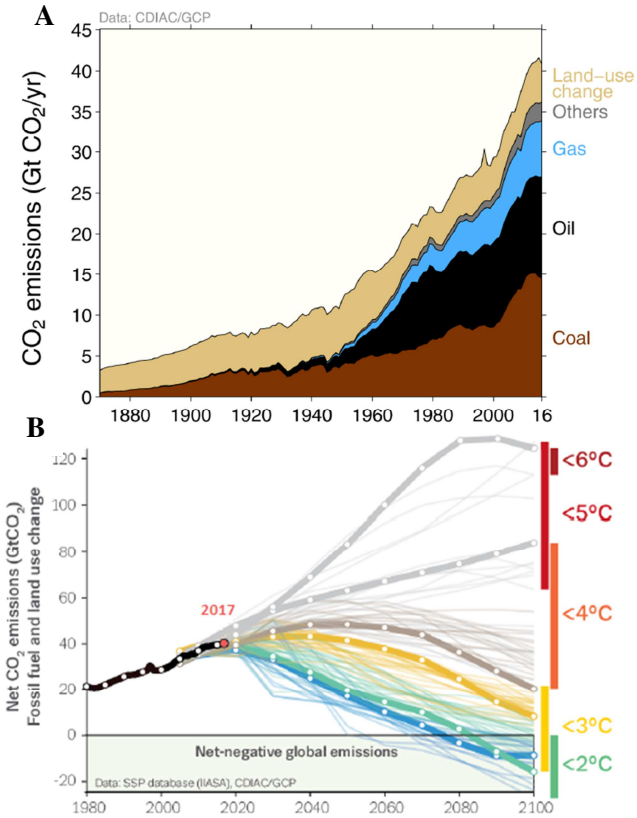


Figure 1.1 : Émission globale de CO_2 en Gt/an. (Pacala et Socolow, 2004).

A. *De 1880 à 2016 : augmentation des émissions dues à l'utilisation de ressources fossiles (charbon, pétrole, gaz) à partir de 1950 qui marque le début de l'industrialisation (International Energy Agency, 2017)*

B. *De 1980 à 2100 : prévisions de l'évolution des émissions de CO_2, et impact sur la température globale (Pacala et Socolow, 2004)*

Si nous voulons éviter d'atteindre le seuil de 4 °C, il faudrait réduire efficacement les émissions de carbone dues à l'utilisation de combustibles fossiles. D'autant plus que la vitesse de formation des ressources fossiles est largement inférieure à la vitesse de leur utilisation, c'est pourquoi elles sont qualifiées de « non renouvelables». L'Agence Internationale de l'Énergie (IEA) prévoit une pénurie dès 2020, d'autres acteurs prévoient des innovations dans le domaine du forage permettant d'atteindre des ressources jusqu'alors inaccessibles mais le prix d'un tel déploiement technologique se répercutera sur les produits dérivés ou réduira la marge de bénéfice pour l'industrie pétrolière (International Energy Agency, 2017). L'exploitation du pétrole est en grande partie liée à la production de carburant (essence, gazole, kérosène, gaz de pétrole liquéfiés), de ce fait des alternatives ont rapidement été trouvées dans le domaine énergétique (voitures électriques, biocarburant) mais le pétrole est à la base de nombreux produits du quotidien.

Le pétrole brut se compose à 90% de carbone et à 10% d'hydrogènes. La distillation du pétrole permet de séparer les hydrocarbures selon leur nombre de d'atomes de carbone, allant de 25 atomes dans le cas du diesel à 1 pour le méthane en passant par le kérosène (13 atomes de carbone), l'essence (7 à 11 carbones), le naphta utilisé pour la fabrication de plastiques (5 à 6 atomes de carbone) et enfin les gaz à 4 atomes de carbone (**Figure 1.2**). Les résidus sont traités pour produire les huiles de lubrification industrielles, des paraffines, du fioul et du bitume. Les hydrocarbures les plus lourds représentent la portion la plus importante, ils sont cassés en molécules plus légères à destination de la chimie. A la sortie de la raffinerie 90% des hydrocarbures sont utilisés pour le transport, le chauffage, la pétrochimie récupère 10% des hydrocarbures légers pour les transformer en bien de la vie courante (Alfke, Irion et Neuwirth, 2007).

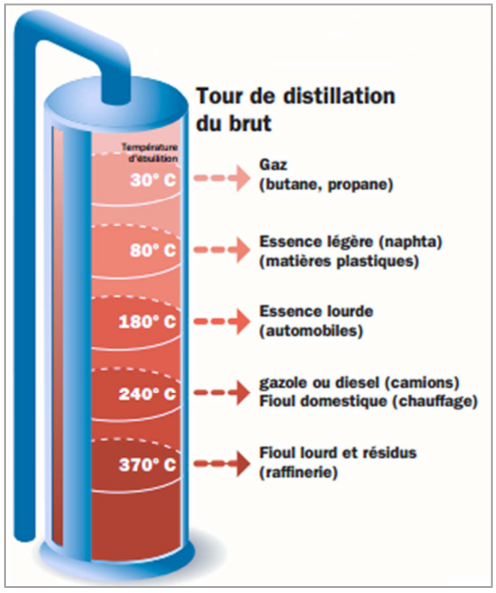


Figure 1.2 : Raffinage de pétrole brut

source : mareesnoires.com

La liste des applications est longue : plastiques (polyéthylène, polypropylène, polystyrène), fibres textiles dites synthétiques, adhésifs, détergents, emballages, cosmétiques, médicaments. Les secteurs de l'économie ne dépendant pas du pétrole sont rares voire inexistants. Le carbone est à la base de notre économie mais les sources de carbone non renouvelables (pétrole, charbon, gaz) sont-elles les seules disponibles et exploitables ? La bioéconomie nait de cette question.

1.2 Une économie nouvelle basée sur l'utilisation de carbone renouvelable

A) La sécurité alimentaire : terrain de réflexion pour la construction d'une bioéconomie

L'utilisation de la biomasse végétale comme source de carbone a éveillé l'intérêt public suite aux crises pétrolières des années 70 avec le développement des biocarburants (biodiesel et bioethanol). Les végétaux riches en saccharose (betterave, canne à sucre,..) ou en amidon (blé, maïs) permettent l'obtention du bioéthanol par fermentation tandis les huiles végétales (colza, soja) via un procédé de transesthérification fournissent du biodiesel. D'un point de vue environnemental c'est une grande avancée : l'emploi de biocarburant permet une diminution des gaz à effet de serre de 40 % pour le bioéthanol par rapport à l'essence et de 70 % pour le biodiesel par rapport au gazole, lorsqu'ils sont utilisés purs (Ministère de la Transition écologique et solidaire, 2017). C'est pourquoi dans le cadre de la lutte contre le réchauffement climatique les gouvernements ont subventionné au début du XXième siècle les cultures à destination des biocarburants. La crise alimentaire mondiale de 2006-2008 souligne qu'une avancée technologique s'intègre dans un environnement complexe (démographique, climatique, économique,..) dont l'équilibre est fragile.

A partir de 2006 le cours de la plupart des matières premières agricoles a brusquement augmenté sur les marchés mondiaux à un rythme qui s'est accéléré en 2007 et 2008 (Nations Unies, 2008) . Cette crise mondiale a été la conséquence d'un déséquilibre entre l'offre et la demande sur le marché à termes et les incitations gouvernementales pour les productions céréalières à destination des biocarburants n'y sont pas étrangères. La demande a fortement augmenté en raison de l'augmentation démographique mondiale et de l'émergence de nouvelles classes moyennes dans les pays émergents comme la Chine. Parallèlement de mauvaises récoltes causées par des évènements climatiques dans les principaux pays producteurs (sécheresses, inondations,..) ont réduit l'offre tandis qu'aux Etats-unis grâce à 6 milliards de subvention 138 millions de tonnes de maïs étaient drainées hors du marché alimentaire (Le Monde.fr | 14.04.2008). La fabrication de biocarburants a donc exacerbé la crise. Ce déséquilibre s'est directement répercuté sur les prix des denrées alimentaires de base, entrainant des famines dans les régions les plus pauvres du globe. Après cet événement l'Union Européenne et l'OCDE (Organisation de Coopération et de Développement Economiques) - qui comprend la majorité des pays développés et démocratiques au monde - ont travaillé au concept de bioéconomie qui ne constituerait pas une menace pour la sécurité alimentaire (CE, 2012a)

B) Définitions de la bioéconomie

La bioéconomie (en anglais « bio-economy » ou « bio-based economy ») est une économie qui s'appuie sur l'utilisation de ressources carbonées durables. Ce terme est utilisé pour la première fois en 1975 par l'économiste roumain Nicholas Georgescu-Roegen, à ce stade il s'agit plus d'une philosophie que d'une économie réelle (Georgescu-Roegen, 1975). En 2005 une première définition est donnée par la commission européenne (CE), la bioéconomie est la transformation durable et efficace des ressources biologiques renouvelables en produits alimentaires, énergétiques et autres produits industriels (Schmid, Padel et Levidow, 2012; Staffas, Gustavsson et McCormick, 2013). C'est à partir de cette même année que le terme commence à être utilisé de façon croissante dans des publications scientifiques (**Figure 1.3**).

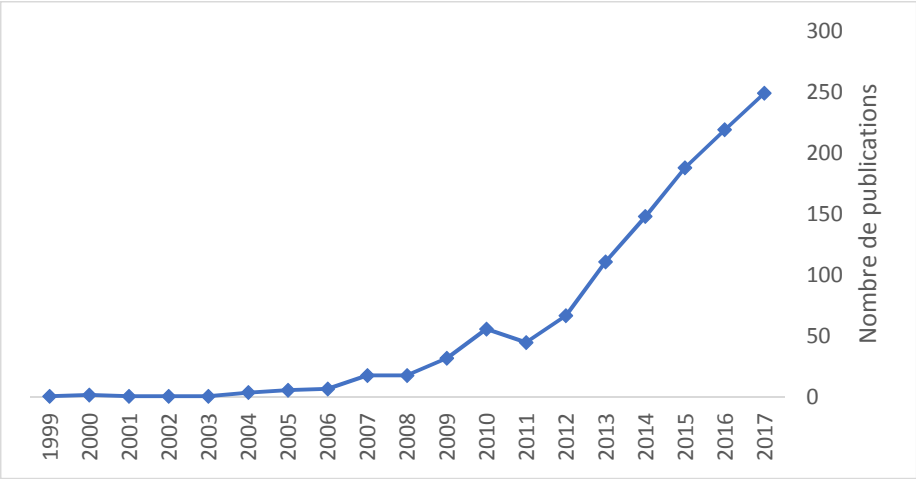


Figure 1.3 : Nombre de citations contenant "bio-based economy", 'biobased economy", "bio-economy" ou "bioeconomy" dans le titre, le résumé ou les mots clefs entre 1999 et 2017 – Scorpus.

A l'échelle de l'OCDE - En 2009, l'OCDE publie une étude qui discute de la nécessité d'un nouveau modèle économique. Le concept bioéconomie est défini comme « l'application des biotechnologies à la production primaire, à la santé et à l'industrie, où la biotechnologie contribue à une part importante de la production économique » (OCDE, 2009). L'OCDE prévoie que la bioéconomie se développera dans les pays OCDE et non-OCDE et propose d'accélérer la recherche dans le secteur des biotechnologies agronomiques et industrielles, en augmentant le budget public alloué à la recherche, réduisant les contraintes administratives et en encourageant les partenariats public-privé.

A l'échelle de l'Union Européenne - De nombreux rapports ont été publiés concernant la bioéconomie en Europe jusqu'à ce qu'une stratégie durable soit présentée en 2012 comprenant 3 piliers : (i) recherche et l'innovation, (ii) économie circulaire (iii) coordination politique (EC, 2012a et 2012b). Les premières définitions mettaient l'accent sur la biomasse et la biotechnologie (en particulier sur la base de modifications génétiques) en omettant les secteurs industriels et économiques qui produisent, gèrent et exploitent les ressources biologiques (agriculture, foresterie, aquaculture,..). Le schéma s'est depuis élargi pour inclure l'ensemble des acteurs, pour former une économie circulaire du carbone où le bioraffinage et le développement de nouveaux produits sont essentiels (McCormick et Kautto, 2013)

Les incitations gouvernementales européennes, et étrangères soutiennent les produits dits de 2eme génération qui sont produits à partir de lignocellulose (bois, feuilles, paille, ..) afin d'éviter la compétition avec l'alimentaire et valoriser les déchets agricoles ou industriels (filière papier,..).

La bioéconomie n'est plus une philosophie mais un modèle de développement économique qui fait appel à l'innovation technologique dans un contexte de croissance économique et de compétitivité. L'investissement dans les secteurs de la bioéconomie pourrait générer 130 000 emplois et 45 billions de plus-value d'ici 2025 (EC, 2012b).

2. CHOIX DU SUBSTRAT D'ETUDE : LES SUCRES LIGNOCELLULOSIQUES

Le bioraffinage est le processus durable de transformation de la biomasse en produits biosourcés (alimentation, produits chimiques et matériaux) et en bioénergie (biocarburants, électricité et chaleur). Les bioraffineries de première génération utilisent de la biomasse alimentaire riche en sucre ou en acides gras (sucre de canne, huile végétale) comme matière première alors que les bioraffineries de deuxième génération emploient de la biomasse non alimentaire dite lignocellulosique (résidus agricoles, forestiers, sous-produits de filières industrielles) (Laurent et al., 2011). Le projet de thèse présenté a été développé dans le cadre de la recherche de nouveaux de procédés de production de 2ème génération.

2.1 La lignocellulose, source de carbone renouvelable

Le terme 'lignocellulose' est employé pour désigner la paroi végétale qui entoure chacune des cellules des plantes (**Figure 1.4**). Cette barrière essentielle - composée majoritairement de polysaccharides (75%), de composés phénoniques et dans une moindre mesure de protéines - constitue une armure pour les végétaux qui leur confère la capacité de résister à des contraintes physiques et biochimiques. La lignocellulose se divise en trois composants majeurs : la cellulose (40-60%), l'hémicellulose (20-40%) et la lignine (10-25%) dans des proportions variables selon le végétal considéré (Wertz, Richel et Gérin, 2015; Zhang, Chen et Xu, 2017).

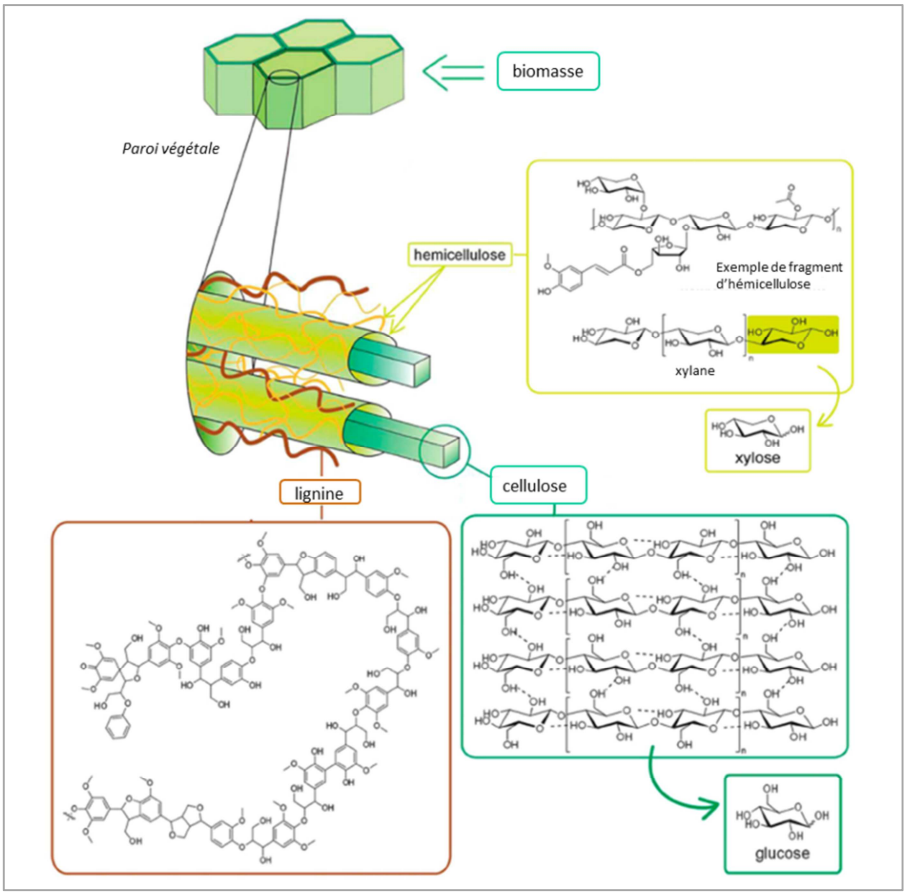


Figure 1.4 : Illustration schématique de la structure et composition de la lignocellulose (Möller et al., 2013, Zhang et al ., 2017)

- La **cellulose** est le constituant majeur de la paroi cellulaire, elle est constituée de chaines linéaires de β-1,4-D-glucopyranose condensées par des liaisons hydrogène en structures cristallines appelées microfibrilles. Chaque microfibrille contient jusqu'à 250 unités de **glucose**. (Kolpak et Blackwell, 1976).

- Les **hémicelluloses** sont un groupe d'hétéro-polysaccharides présents dans toutes les plantes vertes associé à la cellulose et à la lignine (quand elle est présente, les algues vertes n'en contiennent pas par exemple). Leur structure est plus complexe que la cellulose. Elles sont définies structurellement comme des polysaccharides dont le squelette est composé de résidus B-(1,4)-D-pyranose auquel de courtes chaines latérales viennent se fixer pour créer des ramifications. L'hydrolyse de l'hémicellulose ibère majoritairement du D-**xylose** et du L-**arabinose**, avec une quantité variable des monomères de D-glucose, D-galactose, D-mannose, D-glucuronate et 4-O-methyl-D-glucuronate (Scheller et Ulvskov, 2010).

- La **lignine** est un terme regroupant un vaste groupe de polymères aromatiques. (Vanholme *et al.*, 2010; Zhang, Chen et Xu, 2017).

2.2 Le bioraffinage

Les étapes de bioraffinage comprennent le prétraitement de la biomasse, l'hydrolyse des polysaccharides et la fermentation des sucres obtenus de l'hydrolyse (Jørgensen, Kristensen et Felby, 2007). Les monosaccarides obtenus à l'issue de l'étape d'hydrolyse de la cellulose et de l'hémicellulose (glucose, xylose, arabinose) sont séparés de la lignine par fractionnement aux solvants organiques (organosolv) (Sathitsuksanoh, Zhu et Zhang, 2012). Les monosaccharides - principalement glucose, xylose et arabinose - sont ensuite fermentés par des microorganismes naturels ou recombinants afin de produire de l'éthanol et des produits bio-sourcés à haute valeur ajouteé à destination de l'industrie chimique et pharmaceutique.

3. CHOIX DU MICROORGANISME D'ETUDE POUR LA VALORISATION DES SUCRES LIGNOCELLULOSIQUES : *ESCHERICHIA COLI*

3.1 Principaux microorganismes producteurs et leurs caractéristiques

L'Homme utilise les microorganismes depuis des milliers d'années pour conserver le lait, les légumes, les fruits sous forme de produits fermentés comme le fromage, les cornichons, le vin ou la bière. Au début du XXème siècle l'utilisation des microorganismes s'est étendue à la production de solvants, d'acides organiques, de vitamines, d'enzymes et autres produits non-alimentaires. Les microorganismes sont utilisés pour plusieurs raisons :

1) Un rapport surface : volume élevé qui permet une très bonne assimilation des nutriments et captation d'oxygène, supportant un métabolisme efficace de biosynthèse

2) Une grande variété de réactions catalysées

3) Une capacité d'adaptation à des environnements variés (des organismes isolés dans la nature peuvent être cultivés sur des sources de carbone et d'azote peu couteuses en laboratoire ou en usine)

4) La possibilité de manipulations génétiques (transformations, délétions, ...)

5) Une grande diversité des microorganismes connus : pour une même activité plusieurs organismes peuvent être employés ce qui autorise une certaine flexibilité lors la mise en place des conditions de fermentations (pH, température, oxygénation, source carbonée et azotée) (Demain et Adrio, 2008).

Le besoin de nouveaux procédés à destination des bioraffineries est une source d'opportunités pour la microbiologie industrielle et de défis face à de nouveaux substrats à métaboliser et à une grande variété de produits à synthétiser. Il est difficile de trouver un organisme possédant naturellement les qualités nécessaires pour un bioprocédé performant (assimilation du substrat, tolérance des concentrations élevées du produit, pH et température de croissance adaptés,..). Trouver ou concevoir un système microbien optimal pour une production bio-sourcée est l'objectif principal des travaux d'ingénierie industrielle. Le microorganisme idéal en bioraffinerie serait capable de dégrader les composants lignocellulosiques, fermenter efficacement l'ensemble des sucres résultants (hexoses et pentoses), tolérer des concentrations élevées du produit à des températures élevées et à des pH extremes afin d'éviter les étapes de refroidissement et les coûts de stérilisation (Fischer, Klein-Marcuschamer et Stephanopoulos, 2008). Les scientifiques utilisent des outils génétiques et les connaissances à leur disposition pour implémenter les caractéristiques désirées aux microoorganismes déjà connus. Des souches utilisées couramment dans l'industrie comme *Escherichia coli, Saccharomyces cerevisiae* et *Corynebacterium glutamicum* ont été développées pour la production de molécules à haute valeur ajoutée, de polymères et de biocarburants par fermentation de carbone renouvelable (Hong et Nielsen, 2012; Buschke *et al.*, 2013; Paes et Almeida, 2014; Becker et Wittmann, 2016; Baritugo *et al.*, 2018). Les caractéristiques naturelles d'*Escherichia coli, Saccharomyces cerevisiae* et *Corynebacterium glutamicum* seront présentées dans le **Tableau 1.1**, et leurs avantages/inconvénients dans le **Tableau 1.2**

Tableau 1.1 : Caractéristiques de Saccharomyces cerevisiae, Escherichia coli, Corynebacterium glutamicum *En vert les principaux sucres lignocellulosiques.*

Organisme	***Saccharomyces cerevisiae***	***Escherichia coli***	***Corynebacterium glutimicum***
Division	Ascomycète	Protéobacterie	Actinobactérie
Métabolisme respiratoire	Aérobie anaérobie facultatif	Aérobie anaérobie facultatif	Aérobie anaérobie facultatif
Température optimale	30°C	37°C	30°C
pH	2,4< 4,6 <8,6	5,5 < 7 < 8,5	5,5 < 7 < 9,5
Sucres assimilés	**glucose**, fructose, mannose, galactose, ribose, sucrose	**glucose**, fructose, mannose, galactose, ribose, **xylose, arabinose**	**glucose**, fructose, mannose, ribose
Bioproduction	Ethanol	-	L-glutamate
Outils génétiques	Organisme modèle eurocaryote, outils d'ingénierie disponibles pour toute application d'ingénierie métabolique, organisme capable de recombinaison	Organisme modèle procaryote, outils d'ingénierie disponibles pour toute application d'ingénierie métabolique, organisme non doté de recombinase	Relativement peu d'outils comparés aux organismes modèles. Outils adaptés à partir de ceux d'*E.coli* La faible perméabilité de *C.glutamicum* à l'IPTG réduit considérablement le niveau d'expression des systèmes inductibles comparé à *E.coli.* Efficacité de transformation inférieure à *E.coli*
Références	Verduyn *et al.*, 1984 ; Leveau et Bouix, 1996 ; Goffeau *et al.*, 1996 ; Buschke *et al.*, 2013	Blattner, 1997 ; Buschke *et al.*, 2013 ; Pontrelli *et al.*, 2018	Buschke *et al.*, 2013, Baritugo *et al.*, 2018

Tableau 1.2 : Avantages et inconvénients des microorganismes Saccharomyces cerevisiae, Escherichia coli, Corynebacterium glutamicum pour le développement de procédés de production microbiens- *le microorganisme utilisé dans ces travaux de thèse en encadré*

Organisme	***Saccharomyces cerevisiae***	***Escherichia coli***	***Corynebacterium glutamicum***
Avantages	- connaissance détaillée du métabolisme - outils génétiques - tolérance à des pH acides - tolérance aux inhibiteurs de croissance présents dans les hydrolysats lignocellulosiques	- connaissance détaillée du métabolisme - outils génétiques - croissance rapide sur milieux simples - capacité d'assimilation des hexoses et pentoses - capacité d'excrétion de métabolites	- maintien de l'activité métabolique lors de l'arrêt de croissance - tolérance à des concentrations élevées en susbtrat et produit dans le milieu - faible répression catabolique en présence de glucose et autres sources de carbone
Inconvénients	-incapacité à assimiler les pentoses -compartimentation du cytoplasme -capacité excrétoires limités	-sensibilité aux pH acides -répression catabolique élevée en présence de glucose	- incapacité à assimiler les pentoses - peu de données omiques et d'outils génétiques comparé aux organisme modèles

La bactérie *Escherichia coli* (*E.coli*) a été choisie pour la réalisation de ces travaux de thèses. La connaissance de ses caractéristiques génétiques, biologiques et fermentaires et son absence de pathogénicité sont des atouts majeurs en biotechnologie. En étant l'organisme procaryote le mieux étudié *E.coli* sert souvent d'organisme test pour la construction et la validation de la conception d'une voie métabolique. L'absence de compartimentation facilite la démonstration de la fonctionnalité d'une nouvelle voie : si les enzymes de la voie s'expriment correctement dans *E.coli*, presque n'importe quelle molécule peut être produite par cet organisme à condition que les réactions soient thermodynamiquement favorables. De plus elle assimile naturellement les pentoses et les hexoses lignocellulosiques, substrats utilisés dans cette thèse.

3.2 *Escherichia coli*

La bactérie *Escherichia coli* est une des bactéries les plus étudiées, et certainement un des organismes vivant les mieux caractérisés. C'est un bacille Gram négatif appartement à la division des Proteobacterie et à la famille des Enterobacteriaceae, qui doit son nom à leur présence dans la flore intestinale aérobie des mammifères (Eckburg et al., 2005; Tenaillon et al., 2010) . Elle a été isolée pour la première fois en 1885 par le biologiste allemand Theodor Escherich. Plus d'un siècle de travaux ont permis d'accroître les connaissances physiologiques, biochimiques, génétiques et moléculaires de cet organisme.

La souche *E.coli* K-12-MG1655, souche d'étude de ces travaux, est une souche non pathogène dérivée de la souche ancestrale K-12. Cette dernière, isolée en 1922, a subi plusieurs modifications (perte de l'antigène O, du bactériophage λ et du plasmide F) ce qui a donné la souche MG1655 (Bachmann, 1996). Cette souche a été la première souche d'*E.coli* séquencée (Blattner, 1997) : la taille de son génome est de 4.64 Mb. D'après la dernière mise à jour de la base de données d'Ecocyc (30 juillet 2014), son génome contient 4499 gènes et environ 300 d'entre eux sont essentiels (Baba et al., 2006). Les connaissances du génome, transcriptome, protéome, métabolome et fluxome d'*E.coli* ont permis de reconstituer le réseau métabolique de cette bactérie modèle, ainsi que les systèmes de régulation qui assurent l'homéostasie cellulaire. A ce jour 484 souches d'*E.coli* ont été séquencées (NCBI 2018). Parmi celles-ci 7 représentent un intérêt majeur en laboratoire et en industrie : BL21, C, Crooks, DH5aplha, MG1655, W, W3100 (**Figure 1.5**) (Monk et al., 2016).

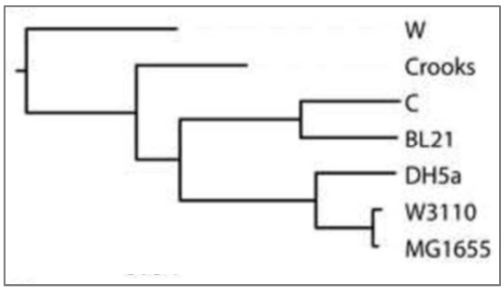


Figure 1.5. Arbre phylogénique à partir des séquences génomiques de 7 souches d'Escherichia coli *(Monk et al., 2016)*

Les souches MG1655, W3100, DH5alpha appartiennent au même clade et sont génétiquement très proches. La souche K-12 MG1655 et W3100 présentent le plus grand potentiel de production de produits potentiels en aérobie et anaérobie, respectivement, d'après des données de transcriptomique et de modélisation. La souche MG1655 a une vaste gamme de gènes dont le niveau d'expression est élevé en condition aérobie (Monk et al., 2016).

Conditions de culture - Le genre *Escherichia* est aéro-anaérobie facultatif. *E.coli* est mésophile avec une température optimale de 37°C et neutrophile, elle se développe à des pH compris entre 5,5 et 8,5 avec un optimum de 7. Son taux de croissance maximum peut aller jusqu'à 1,68 h^{-1} (Baev et al., 2006) ce qui a accéléré la recherche dans un premier temps, et aujourd'hui profite également aux procédés industriels. La souche *E.coli* MG1655 peut être cultivée en milieux riches ou sur milieux minimums supplémentés d'une source de carbone et de thiamine, en fonction du contexte d'utilisation. Les milieux riches tels que le milieu « Lysogenic Broth » (LB) (Bertani et al., 2006) sont des milieux complexes qui permettent une croissance rapide, ils sont largement utilisés pour des applications industrielles. Les milieux minimums n'autorisent pas des croissances aussi rapides qu'en milieu riche (taux de croissance maximum de 0,6 h^{-1} en présence de glucose) (Volkmer et Heinemann, 2011).

Métabolisme - *E.coli* est capable de métaboliser naturellement une grande variété de sources de carbone monomérique, incluant des sucres (hexoses et pentoses), les polyols (sucre-alcools), les acides organiques (Yoon et al., 2012). Sa capacité à croître sur le glucose, fructose, mannose, galactose, ribose, xylose, arabinose, lactate, glycérol présente un avantage en bioraffinerie (Buschke et al., 2013). Cependant *E.coli* n'est pas capable d'utiliser des polymères carbonés et la souche MG1655 le sucrose. Parmi les 3 souches présentées dans cette partie, *E.coli* est l'organisme qui possède naturellement le plus large spectre de substrats. *E.coli* utilise le système phosphotransferase (PTS) pour le transport actif et la phosphorylation de sucres tels que le glucose, le mannose (Kundig, Ghosh et Roseman, 1964; Hunter et Kornberg, 1979) ou encore des transport du galactose (GalP) (McDonald, Walmsley et Henderson, 1997) . Le phosphoenolpyruvate (PEP) fournit le phosphate nécessaire à l'étape de phosphorylation. Ce système irréversible convertit le PEP en pyruvate. Lorsque le glucose est présent dans le milieu une forte répression catabolique est exercée sur les autres sucres ce qui peut nuire à la productivité des procédés de valorisation de la biomasse.

En condition aérobie l'oxygène permet une production efficace d'ATP en utilisant les équivalents réducteurs pour la phosphorylation oxydative. Cependant les équivalents réducteurs utilisés pour la croissance peuvent être nécessaires pour la production de composés réduits. La croissance anaérobie chez *E.coli* se caractérise par une énergie cellulaire réduite, la production d'ATP est limitée. Le cycle de Krebs est alors sous-régulé, ce qui aboutit à une oxydation incomplète de la source de carbone : des sous-produits de fermentation sont excrétés : succinate, formate, acétate, lactate et éthanol (Clark, 1989).

Naturellement *E.coli* ne produit pas un métabolite d'intérêt industriel à des concentrations élevées, la flexibilité de son métabolisme et les nombreux outils génétiques développés pour cet organisme en ont fait le microorganisme le plus utilisé en biotechnologie industrielle.

L'ingénierie métabolique est couramment employée dans le développement des procédés de production microbienne. Quelles approches d'ingénierie métabolique sont développées pour dépasser le potentiel naturel des microorganismes ?

4. APPROCHES D'INGENIERIE METABOLIQUE UTILISEES POUR DEPASSER LE POTENTIEL NATUREL DES MICROORGANISMES

4.1 Les cinq niveaux d'ingénierie métabolique

Une voie métabolique est une succession de réactions enzymatiques qui permet la synthèse d'un composé en plusieurs étapes, à partir de sucres captés dans le milieu naturel. L'ingénierie métabolique peut s'appliquer à :

1) L'optimisation de voies existantes
2) Le transfert de voies existantes dans un nouvel hôte
3) La création de nouvelles voies à partir de réactions connues
4) La création de nouvelles voies à partir de nouvelles réactions basées sur des mécanismes catalytiques connus
5) La création de nouvelles voies à partir de nouvelles réactions basées sur de nouveaux mécanismes catalytiques

Ces cinq niveaux d'ingénieries sont décrits par Thomas Schwander *et al.* pour la présentation des travaux de fixation du CO_2 *in vitro* (Schwander et al., 2016).

4.1.1 Niveau 1 : optimisation de voies existantes

Les efforts de l'ingénierie métabolique de base se concentrent sur les voies existantes au sein de leur hôte naturel. L'efficacité de la voie métabolique est alors simplement améliorée par des délétions et surexpression. L'ingénierie d'*E.coli* pour rediriger le flux de la fermentation du glucose vers la production de succinate illustre cette approche (Zhang et al., 2009). En l'absence d'oxygène *E.coli* produit naturellement du succinate, du lactate, de l'éthanol et du formate. Le succinate est une production minoritaire, la majorité du carbone est converti en éthanol et en lactate. La stratégie développée selon cette approche par Zhang *et al.*. a permis de multiplier le rendement en succinate par 5 au détriment des autres productions.

4.1.2 Niveau 2 : transfert de voies existantes dans un nouvel hôte

L'exemple le plus marquant est celui de la production d'acide artémisinique par *S.cerevisiae*. Le paludisme est engendré par des parasites du genre *Plasmodium,* transmis par les piqures de moustique. Un traitement à base d'artéminisine est administré afin d'éviter que la maladie devienne mortelle. Cette molécule est extraite à partir de la plante *Artemisia annua* dont l'exploitation mondiale permet de récupérer 30 à 40 tonnes annuellement. Or pour traiter les centaines de millions de cas recensés, la capacité de production devait être augmentée à 300 tonnes par an. (Fiaroso, 2012).

La production d'un précurseur de artémisine hémi-synthétique par la levure *Saccharomyces cerevisiae* développée par Paddon *et al.* a apporté une réponse au problème posé (Paddon et al., 2013). *Saccharomyces* ne produit pas naturellement cette molécule, l'expression de voies de synthèse hétérologue provenant de l'*A.annua* a été nécessaire pour permettre sa production. La levure produit naturellement du FPP (farnésyl diphosphate), intermédiaire de l'acide artémisinique. Le FPP est assimilé dans la souche ingénieriée par la voie hétérologue provenant d'*A.annua* pour la synthèse d'acide artémisinique. Celui-ci est ensuite converti chimiquement en artémisinine.

Auparavant la production d'artémisinine à partir de la plante *A. annua* prenait dix-huit mois, l'ingénierie de *S. cerevisiae* a permis d'industrialiser la production de l'hémi-artémisinine en réduisant les délais de production et en sécurisant l'approvisionnement. Sanofi annonçait dans son communiqué du 12 aout 2014 l'expédition des premiers lots d'antipaludiques fabriqués par voie microbienne (Sanofi, 2012).

4.1.3 Niveau 3 : création de nouvelles voies à partir de réactions connues

La forme la plus simple de la biologie de synthèse repose sur le concept de « mix and match », dans lequel les voies non-naturelles sont créées à partir de la combinaison d'expression d'enzymes existantes (**Figure 1.6**). L'hôte produit les précurseurs naturels à partir de son métabolisme primaire ou secondaire, pris en charge par les enzymes de la voie non-naturelle.

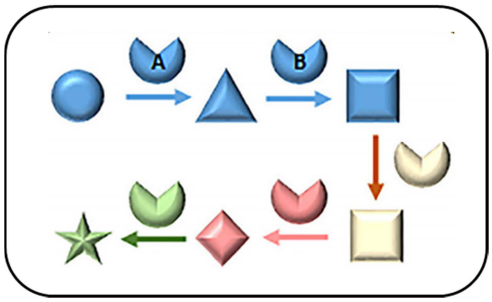


Figure 1.6. Le procédé d'ingénierie d'une voie non-naturellle basée sur le concept de 'mix and match' (Schwander et al., 2016). *En bleu : le métabolisme naturel qui permet la formation du précurseur (carré bleu), l'expression d'enzymes appartenant à d'autres microorganismes ou à l'organisme hôte mais n'appartenant à la voie métabolique naturelle (une couleur autre que bleue par enzyme).*

La biosynthèse d'arbutine, un glycoside utilisé comme agent éclaircissant de la peau, à partir de glucose illustre bien ce concept (Shen et al., 2017). En exprimant la 4-hydroxybenzoate-1-hydroxylase de *Candida parapsilopsis* (MNX1) et l'arbutine synthase from *Rauvolfia serpentina* (AS) dans l'organisme hôte *E.coli*. Shen *et al.* ont été capables de connecter l'acide précurseur p-hydroxybenzoique endogène au métabolite désiré.

4.1.4 Niveau 4 : création de nouvelles voies à partir de nouvelles réactions basées sur des mécanismes catalytiques connus

La conception de nouvelles voies *in silico* conduit fréquemment à l'identification de voies impliquant des étapes de réaction que les enzymes naturelles sont incapables de catalyser. Concevoir des voies métaboliques impliquant de nouvelles réactions repose la capacité de certains biocatalyseurs à convertir de nouveaux substrats tout en gardant inchangé leurs mécanismes catalytiques. Explorer la gamme de substrat des enzymes naturelles est une approche répandue sachant que 10% des enzymes bactériennes sont promiscuitaires (Martínez-Núñez et al., 2017). Les possibilités offertes par l'ingénierie enzymatique (i.e redirection de la spécificité de substrat, acceptation de co-facteurs artificiels) permettent également d'étendre la gamme de susbstrat d'enzymes connues. Cette dernière approche ne sera pas détaillée dans cette thèse.

Un exemple basé sur la promiscuité enzymatique évoquée dans cette thèse (**Chapitre III**) est celui de la voie xylulose-1P (X1P) qui est une voie alternative d'utilisation du xylose pour la production d'acide glycolique (Cam *et al.*, 2016). L'aldolase *aldoB* de mammifère est une fructose 1,6biphosphate aldolase exprimée chez *E.coli*, elle est capable de convertir le xylulose-1P en glycolaldéhyde et dihydroxyacétone phosphate (DHAP).

4.1.5 Niveau 5 : création de nouvelles voies à partir de nouvelles réactions basées sur de nouveaux mécanismes catalytiques

De nouvelles réactions biochimiques permettent la réaction de voies métaboliques artificielles illustrant le potentiel maximal de la biologie de synthèse. L'exemple le plus remarquable est celui du développement de l'activité formolase qui permet la réaction jamais décrite auparavant de carboligation de trois molécules de formate en une seule molécule de DHAP, entrant dans la voie de la glycolyse (Siegel *et al.*, 2015).

4.2 Optimisation du flux carboné bactérien

L'implémentation de nouvelles voies métaboliques entraine fréquemment des perturbations métaboliques et une production sous-optimale du produit cible. Dans l'approche 'mix and match' appliquée dans cette thèse (**Chapitre III**, **Chapitre IV**), l'intégration d'une nouvelle voie au métabolisme naturel peut entrainer une déplétion d'un des métabolites précurseurs, l'accumulation d'intermédiaires (toxiques), la formation de produits secondaires, conduisant à l'inhibition de la croissance et de faible rendement en produit (Pitera *et al.*, 2007; Dai et Nielsen, 2015; Xu *et al.*, 2017). La modulation du flux métabolique minimise la perte de productivité ou de rendement (**Figure 1.7**).

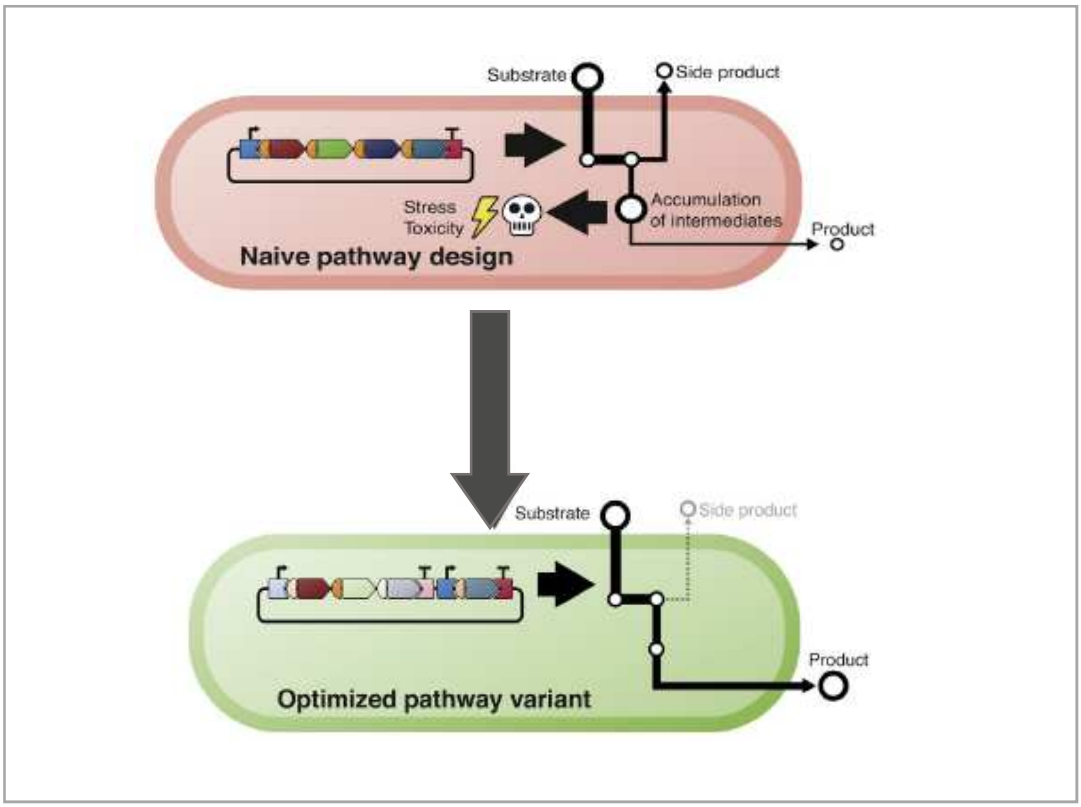


Figure 1.7 : représentation schématique du processus d'optimisation du flux métabolique après intégration d'une voie synthétique. *Figure adaptée de Jeschek, Gerngross et Panke, 2017.*

4.2.1 Les enzymes : catalyseurs clefs

Si l'on considère un microorganisme comme une micro-usine à chaque étape de production est associé une enzyme. Les enzymes sont des protéines possédant une activité catalytique. Elles catalysent des réactions où les substrats sont convertis en produits dans une structure que l'on appelle site actif. Celui-ci se compose d'une cavité hydrophobe, avec des résidus permettant d'orienter le substrat pour sa conversion en produit.

La caractérisation d'une enzyme repose sur l'appréciation de sa (i) spécificité (ii) sélectivité et de ses (ii) constantes cinétiques.

La **spécificité** est un paramètre qui évalue le degré de préférence d'une enzyme pour des substrats. La constante de spécificité (k_{cat}/K_m) est utilisée pour évaluer l'efficacité catalytique. Les enzymes du métabolisme carboné central sont spécifiques d'un substrat, tandis que les enzymes impliquées dans la détoxification cellulaire agissent sur une large gamme de substrat, on parle d'enzyme promiscuitaire. Bien que la promiscuité puisse s'appliquer à des enzymes acceptant plusieurs co-facteurs, ou catalysant plusieurs réactions biochimiquement différentes, le terme promiscuité sera appliqué dans cette thèse aux enzymes acceptant un large spectre de substrats.

La **sélectivité** est la capacité d'une enzyme de discriminer des substrats très proches. Il existe différents types de sélectivité chemiosélectivité, regioselectivité et stéréosélectivité. Ce paramètre ne sera pas exploité au cours de cette thèse.

Les **constantes cinétiques** sont dans la majorité des cas les mieux décrites par le modèle Michaelis-Menten. C'est l'un des premiers modèles d'enzymologie proposés et le plus utilisé encore aujourd'hui. D'après le modèle Michealis-Menten, l'activité enzymatique augmente de façon linéaire proportionnellement à la concentration de substrat jusqu'à une certaine concentration où l'enzyme atteint la saturation. L'équation est la suivante :

$$\boldsymbol{v} = \frac{\boldsymbol{Vmax} \times [\boldsymbol{S}]}{\boldsymbol{Km} + [\boldsymbol{S}]}$$

Où v est l'activité enzymatique (µmol.min^{-1}), vmax la vitesse enzymatique maximale (µmol.min^{-1}), la concentration de substrat et le K_m la constante de Michaelis Menten (mM), [S] la concentration en substrat (mM).

La performance d'une enzyme est évaluée sur la base de ses paramètres cinétiques (i) le nombre de 'turnover' (k_{cat}) qui est un dérivé de v_{max} et indique le nombre maximum de molécules de substrat converti par unité de temps (s^{-1}), (ii) la constante de Michaelis Menten (K_m) qui renseigne sur l'affinité d'une enzyme pour son substrat et (iii) l'efficacité catalytique (k_{cat}/K_m) qui mesure l'efficacité avec laquelle l'enzyme converti le substrat en produit, référant également à la spécificité de substrat. Des facteurs extérieurs influent les paramètres cinétiques tels que le pH, la température ou la présence d'ions. Les enzymes sont caractérisées *in vitro* après avoir été purifiées. Bien que les valeurs *in vitro* soient utilisées pour comprendre l'i*n vivo*, le changement d'environnement influe l'activité enzymatique ce qui fait encore débat.

La connaissance de la performance des enzymes de la voie est primordiale pour la conception d'une nouvelle voie métabolique. Des enzymes promiscuitaires pourront engendrer la formation de sous-produits indésirables tandis que des activités trop faibles pourront limiter la production et conduire à des accumulations de métabolites, en ralentissant le flux métabolique. La modulation de l'expression génique des gènes codant pour des enzymes cibles est une stratégie couramment utilisée pour moduler le flux carboné et maximiser la formation du produit final. Nous traiterons dans la prochaine sous-partie des moyens de moduler l'expression génique.

4.2.2 Modulation de l'expression génique

Le niveau d'expression d'un gène dépend de plusieurs paramètres: le nombre de copie du gène (origine (ori) de réplication), le niveau de transcription (promoteur) et le niveau de traduction (RBS) (**Figure 1.8**).

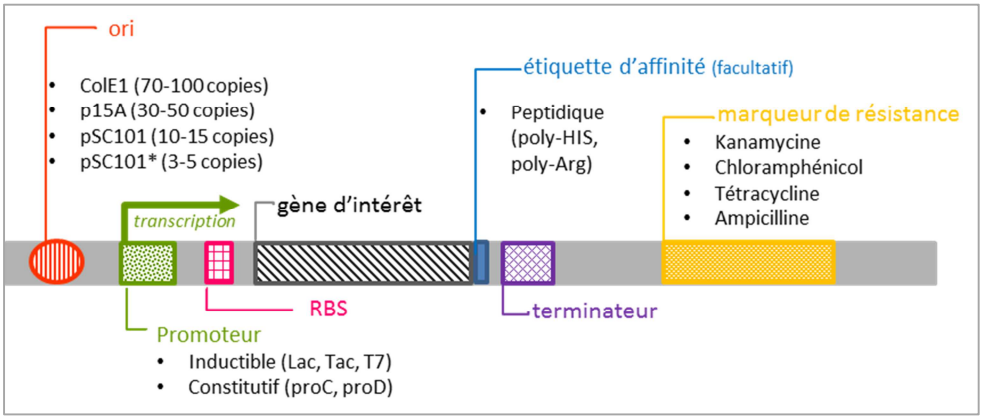


Figure 1.8 : Représentation schématique de la structure d'un plasmide d'expression

Variation du dosage génique - Le nombre de copies d'un plasmide est déterminé par son origine de réplication. Selon l'origine de réplication le nombre de copie peut varier de 3 copies à plus de 100 copies par cellule. La corrélation entre l'efficacité du système d'expression et le nombre de copies n'est pas cependant linéaire. Un nombre élevé de plasmides par cellule représente une charge métabolique cellulaire importante pour leur réplication au détriment de la production et de la croissance. (Ajikumar et al., 2010 ; Kittelson et al., 2011 ; Xu et al., 2013).

Régulation transcriptionelle - De nombreux promoteurs ont été développés pour les systèmes bactériens, un petit nombre est couramment utilisé. La plupart d'entre eux sont dérivés sur promoteur *lac* (Polisky, Bishop et Gelfand, 1976) promoteur faible qui n'est pas utilisé pour une production élevée de proteines et est avantageux pour la production de composés toxiques. Le promoteur synthétique tac (de Boer, Comstock et Vasser, 1983) est au moins 5 fois plus efficace que le promoteur lacUV5 (Amann, Brosius et Ptashne, 1983) et permet une accumulation de protéines qui représente entre 15 et 30% de la quantité totale de protéines cellulaires. L'induction des promoteurs lac et tac est réalisé par addition d'un analogue non hydrolysable du lactose l'isopropyl-betal-D-l-thiolgalactopyranoside (IPTG). D'autres systèmes d'expression comme le système T7 ARN polymerase, le promoteur phagique PL, le ptomoteur pBAD inductible au L-arabinose, le promteur rha PBAD inductible au L-rhamnose sont disponibles. Les systèmes d'expression plasmidiques bactérien peuvent être modulables comme la série de vecteur pZ (Expressys ®) les promoteurs, origine de réplication et cassette de résistance sont interchangeable entre les vecteurs de la série par restriction ligation ce qui facilite la conception de nouveaux plasmides.

Plusieurs bibliothèques de promoteurs ont été créées pour offrir des niveaux de transcription variés (Alper, Miyaoku et Stephanopoulos, 2005; Cox, Surette et Elowitz, 2007; Lee et al., 2013; Mutalik et al., 2013; Yuan, Du et Zhao, 2013; Redden et Alper, 2015; Yu et al., 2016). De nombreuses études d'ingénierie métabolique ont utilisé des banques de promoteurs chez *E.coli*. Selon le promoteur la transcription peut être inductible ou constitutive. Les promoteurs inductibles nécessitent l'expression de répresseur (exemple : lacI pour le promoteur plac) ce qui impose une charge métabolique supplémentaire importante pour la cellule hôte, de plus l'ajout d'activateur tel que l'IPTG représente un coût trop important à échelle industrielle. Les promoteurs constitutifs sont donc fréquemment utilisés mais ces derniers peuvent conduire à une inhibition de la croissance car ils ne permettent pas de découpler la croissance et la production.

Régulation traductionnelle - Chez les bactéries l'ingénierie du RBS (Ribosome Binding Site), qui contient la séquence de Shine-Delgarno en amont du codon d'initiation, présente plusieurs avantages. Un large éventail de niveaux d'expression est possible mais surtout c'est la seule approche permettant l'ajustement individuel de l'expression des gènes dans les opérons polycistroniques. Plusieurs études récentes utilisent des banques de RBS pour l'optimisation de l'expression des génique dans *E.coli* (Xu *et al.*, 2013; Smanski *et al.*, 2014; Jeschek, Gerngross et Panke, 2016).Un autre avantage de l'ingénierie traductionnelle est la disponibilité de modèles prédictifs qui permettent une conception sur mesure selon le niveau d'expression désiré, un certains nombres d'études utilisent des RBS conçus in silico.

Modification génomiques - Les modifications génomiques peuvent être réalisées avec le système de recombinaison Red du phage lambda qui rend possible la recombinaison de fragments linéaires et s'est montré très efficace pour l'introduction de modifications génomiques chez *E.coli* (Datsenko et Wanner, 2000) . Les gènes cibles peuvent facilement remplacés par une cassette de résistance à la kanamycine par recombinaison, la présence de site de reconnaissance de la flippase (FLP) de part et d'autre de la cassette assure le retrait de la cassette par flippase recombinase pour une délétion propre du gène. Seuls les sites de reconnaissance de la FLP demeurent insérés dans le génome, ce qui est non problématique dans la majorité des cas. La collection Keio, qui contient des souches d'*E.coli* portant chacune une unique délétion d'un gène non essentiel, a été construite en utilisant cette méthode (Baba et al., 2006). Le système de recombinaison Red du phage lambda a été utilisé pour la construction de nouvelles souches : délétions de gènes (Shen et al., 2011), introduction de mutations ponctuelles (Heermann, Zeppenfeld et Jung, 2008) et insertions au niveau génomique (Kuhlman et Cox, 2010).

4.3 Applications biotechnologiques chez *Escherichia coli*

E.coli a été utilisée dès les premières productions microbiennes pour la surproduction de l'acide aminé L-thréonine et pour la première synthèse recombinante d'insuline humaine (Goeddel et al., 1979). *E.coli* a depuis été ingénieriée pour la production d'arômes et de parfums (vanillin, benzaldehyde, cinnamaldéhyde et 2-phenylethanol), de di-acides pouvant être polymérisé pour faire du bioplastique tel que l'acide succinique (Thakker *et al.*, 2012; Cheng *et al.*, 2013; Zhu et Tang, 2017) ; de diols tel que le 1,2-propanediol (Zeng et Sabra, 2011), de lactames utilisés pour la synthèse de polymères de nylon polyamide (Chae et al., 2017), d'acides hydroxylé comme l'acide glycolique (Cam et al., 2016; Pereira et al., 2016) ou encore de biocarburants (Pontrelli et al., 2018).Nous pouvons nous demander face à l'éventail de possibilité offertes par l'ingénierie métabolique, quels sont les critères de choix pour déterminer la molécule cible à produire ?

5. CHOIX DE LA MOLECULE D'ETUDE : L'ACIDE GLYCOLIQUE

5.1 Molécules candidates

Le 'Department of Energy' (DOE) américain a publié en 2004 un rapport devenu une référence dans le domaine de la production de molécules biosourcées (Werpy et Petersen, 2004). Celui-ci présente les molécules cibles les plus prometteuses selon des critères propres aux procédés de bioraffinage : classification des molécules (molécule de remplacement entrant directement en compétition avec une/des molécule(s)produite(s) à partir de carbone fossile ; molécule nouvelle sans compétition avec la pétrochimie ; synthon chimique pour la production d'un large panel de molécules dérivées) ; coût de la matière première végétale, coût du procédé de conversion, estimation du prix de vente, marché potentiel, complexité technique du procédé. Les auteurs se sont assurés que la production à partir des glucides de la lignocellulose soit techniquement réalisable et économiquement viable. Les procédés biologiques (fermentation aérobie/anaérobie) et chimiques (catalyses, biotransformation) ont été considérés dans l'étude des productions actuelles ou potentielles. A l'origine 300 molécules pouvant être synthétisées à partir de la biomasse ont été recensées, puis la sélection s'est affinée jusqu'à ne retenir que 12 molécules : les 1,4-diacides (acides succinique, fumarique et malique), le 2-5 furane dicarboxylique, l'acide 3-hydroxypropanoïque, l'acide aspartique, l'acide glucarique, l'acide glutamique, l'acide itaconique, l'acide levulinique, la 3-hydroxybutyrolactone, le glycerol, le sorbitol et le xylitol/arabinitol. Ces 12 molécules candidates sont toutes 1) des molécules plateforme pour la synthèse d'une large variété de produits chimiques et matériaux à haute valeur ajoutée 2) des C1 à C6, constituées de 1 à 6 atomes de carbone 3) synthétisables à partir de lignocellulose et d'amidon 4) non dérivées de la lignine (aucune n'appartient à la classe des composés aromatiques).

Depuis le rapport original du DOE, des progrès considérables ont été réalisés dans l'utilisation des glucides comme matière premières pour la bioproduction. En 2010, Bozell et Petersen ont fourni une liste actualisée des principales molécules plateformes pouvant être produites à partir de glucides (Bozell et Petersen, 2010). Par rapport à la liste d'origine établie par le DOE les acides fumarique et malique, les acides aspartiques, glucarique, itaconique, le 3-hydroxylactone et l'arabitol n'ont pas été repris tandis que des composés ont été ajoutés : éthanol, furanes, hydrocarbures, acide lactique et aldéhyde 3-hydroxyproprionique. Cette sélection s'est faite entre autres sur la base d'études bibliographiques afin de déterminer l'état d'avancement des technologies nécessaires à la production et selon la taille du marché en considérant les applications possibles de chaque molécule candidate.

5.1.1 Les acides organiques

Les acides organiques tels que l'acide lactique, le 3-hydroxypropionique sont largement représentés parmi les molécules d'intérêt certain. Dans le rapport de la DOE 9 des 12 molécules les plus prometteuses étaient des acides organiques. Les acides organiques sont produits à l'échelle de millions de tonnes annuelle à destination des secteurs alimentaire, pharmaceutique et chimique et pour beaucoup à partir de ressources fossiles. Leur production via des procédés respectant les principes de la chimie verte aurait un impact immédiat et évident sur l'environnement. Ces dernières années l'attrait pour les acides carboxyliques produits par fermentation microbienne a considérablement augmenté que ce soit pour la production d'acide glycérique, glutarique, succinique, butirique, xylonique, fumarique, malique, itaconique, propionique, ou glycolique (Sauer *et al.*, 2008; Alonso, Rendueles et Díaz, 2015; Naraian et Kumari, 2017). Les productions microbiennes d'acide succinique, lactique, citrique gluconique et acétique ont pu être amené jusqu'à l'étape finale de commercialisation, montrant la voie pour la production des nombreux autres acides (Alonso, Rendueles et Díaz, 2015)

5.2 L'acide glycolique

L'acide glycolique (AG) est un acide organique à deux carbones arborant une fonction alcool et une fonction carboxylique, il est utilisé pour la synthèse de polyglycolate (PGA) qui présente des propriétés de barrière au gaz et une force mécanique remarquable ce qui en fait un polymère adapté à l'emballage. Sous forme non polymérisée l'AG est utilisé dans les cosmétiques et les produits nettoyants, 40 000 tonnes sont commercialisées annuellement. Le marché de cet α-hydroxy-acide devrait atteindre 300 millions d'euros d'ici 2020 avec un taux de croissance annuel moyen de 6% entre 2017 et 2023 (https://www.grandviewresearch.com/press-release/global-glycolic-acid-market). La production microbienne d'AG a attiré l'intérêt des acteurs de l'industrie chimique ce qui a conduit notamment au développement de procédés de production microbienne d'AG présentées dans la partie 6) ci-dessous.

6. ETAT DE L'ART DE LA PRODUCTION MICROBIENNE D'ACIDE GLYCOLIQUE

A l'échelle industrielle l'AG est produit par la carbonatation du formaldéhyde à haute température et haute pression (Loder, 1939) ou par la conversion enzymatique de glycolitrile par des nitrilases microbiennes (He et al., 2010; Panova et al., 2007) comme décrit par Dupont de Nemours et Co dans WO2006 / 069110 et dans le brevet U.S. Pat. 7,445,917 qui utilise le formaldéhyde et le cyanure d'hydrogène comme précurseurs pour la synthèse du glycolonitrile. Diverses nitrilases natives d'origine microbienne (*Rhodococcus rhodochrous*, *Acidovorax facilis 72W*, *Brevibacterium casei*, *Ainobacter sp. AK226)* ont été testées pour la synthèse d'AG *:* ainsi que des nitrilases mutantes avec l'activité nitrilase améliorée. L'AG obtenu par cette méthode a une pureté supérieure à 70%, et une productivité élevée de 17 g /L/h. Le principal inconvénient de ce procédé est la présence de glycolonitrile, une substance chimique qui peut polymériser, sous l'influence de traces d'acide ou de base, avec un risque d'incendie ou d'explosion. Cette substance se décompose en chauffant, produisant des fumées toxiques comprenant du cyanure d'hydrogène et des oxydes d'azote. Par conséquent, le glycolonitrile est répertorié comme une substance extrêmement dangereuse.

Par ailleurs l'utilisation du formaldéhyde (ou formol) est reconnu aujourd'hui comme fort allergène (classe A selon DIMDI) et qui fait l'objet de restrictions et d'interdictions dans les produits cosmétiques (règlement européen (CE) N° 1223/2009). Il est recommandé aux fabricants de substituer le formaldéhyde par un produit non dangereux (http://travail-emploi.gouv.fr/sante-au-travail/prevention-des-risques-pour-la-sante-au-travail/autres-dangers-et-risques/article/formaldehyde) et à la population d'éviter les produits contenant du formaldéhyde. Les préoccupations environnementales et de santé publique, soutenues par les gouvernements au travers d'une règlementation de plus en plus stricte, sont à l'origine d'une demande en AG biosourcé.

De nombreuses équipes de recherche se sont intéressées à la production d'acide glycolique par voie microbienne. Le **Tableau 1.4** reprend l'ensemble des projets d'ingénierie métabolique à la fin de ce chapitre.

6.1 Production naturelle par bioconversion de l'éthylène glycol

Les premières études ont consisté à identifier les microorganismes naturellement producteurs d'AG. En 1964, il a été montré que des levures et des bactéries pouvaient être capables de convertir naturellement l'éthylène glycol en AG (Brevet US3141033A). Cette méthode de bioconversion a été appliquée par Michihiko Kataoka pour la production d'AG (Kataoka *et al.*, 2001). En 2001, un criblage a été effectué sur plus de 3000 microorganismes : 2 000 microorganismes comprenant des bactéries, des actinomycètes, des levures, des molds et basidiomycètes connues provenant de stock et plus de 1000 autres organismes isolés à partir d'échantillon de sols ont été étudiés et classés selon leur capacité à oxyder l'éthylène glycol en AG. Seule une centaine de microorganismes (majoritairement bactéries, levures et actinomycètes), et 76 microorganismes provenant du sol, ont pu convertir au minimum 20% molaire d'EG en AG. Parmi ces candidats, les levures *Pichia naganishii* (provenant d'un stock glycérol) et *Rhodotorula sp. 3Pr-126* (isolée d'un échantillon du sol) ont le plus efficacement converti l'EG, produisant 105 et 110 g/L d'AG respectivement en 120 h (soit une conversion molaire de 88% et 92% à partir d'EG). L'éthylène glycol est peu onéreux, le procédé économiquement viable, mais une fermentation de 5 jours est trop longue pour une application industrielle. En 2009, *Gluconobacter oxydans* DSM 2003 a été à son tour considérée pour la biotransformation de l'EG en AG (Wei et al., 2009). Les conditions de cultures ont été optimisées pour la production d'AG, 70.2 g/l d'AG ont été obtenus à partir d'éthylène glycol en 48h, avec un rendement de conservation molaire de 86,7%.

6.2 <u>Niveau 1 d'ingénierie métabolique</u> : production d'AG à partir de glucose par ingénierie du cycle glyoxylate naturel

L'ingénierie de micro-organismes capables de convertir le glucose en AG possède des débouchés certains. Il a donc fallu répondre à la problématique suivante : comment un mircroorganisme peut-il convertir le glucose en AG ? La voie métabolique de conversion de l'EG en AG n'était pas adaptée, de nouvelles voies ont alors été considérées. Aucun organisme connu n'est capable de convertir naturellement le glucose en AG. Le glycolate est un produit de l'oxydation du glycoxylate, intermédiaire métabolique du cycle qui porte son nom.

La compréhension du fonctionnement et de la régulation du cycle glycoxylate a été nécessaire pour ouvrir cette voie métabolique au glucose pour la production d'AG.

6.2.1 Fonction et régulation du cycle glyoxylate

A) Fonction

La fonction principale du cycle glyoxylate dans certaines bactéries, plantes et champignons est de convertir les sources de carbones en C2 et C3 (acétate, acides gras) en précurseurs nécessaires à la cellule pour sa croissance en C4-C5-C6. Kornberg a décrit le cycle glycoxylate comme un cycle de Krebs modifié où les deux étapes de décarboxylation (isocitrate déhydrogénase et alpha cétoglutarate déhydrogénase) sont court circuités (Kornberg, 1957) La conservation du carbone rend possible la croissance à partir de C2.

Le produit net du cycle glyoxylate est une molécule de succinate. L'équation globale est la suivante :

2 acétyl-CoA (C2) + NAD+ + 2 H20 —> succinate (C2)+ 2 CoA-SH + NADH + H^+

Le succinate peut ensuite être utilisé pour la biosynthèse de glucides : il est converti en oxaloacétate utilisé pour la production de glucose par la voie de la néoglucogénèse. C'est ainsi que les organismes possédant le cycle glycoxylate peuvent utiliser des molécules carbonées simples pour la production des précurseurs biosynthétiques (Kornberg, 1966; Lorenz et Fink, 2002) (**Figure 1.9**).

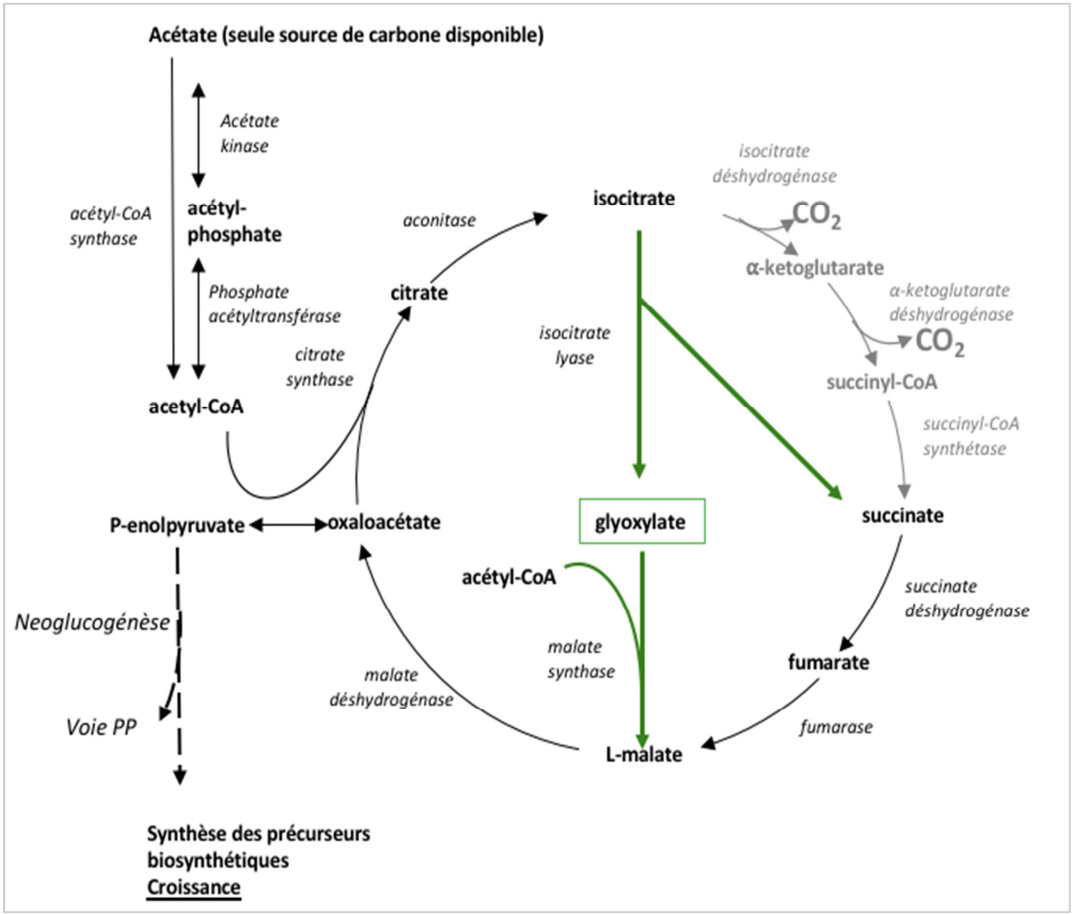


Figure 1.9 : Représentation de la voie glyoxylate en présence d'acétate – *En vert la voie glyoxylate, exprimée en présence d'acétate pour permettre la croissance.*

B) Régulation

Le cycle glyoxylate est induit en présence d'acétate. Lors de la transition glucose-acétate, comment s'organise la régulation de cette voie pour que les microorganismes puissent rapidement répondre à un changement de source de carbone ?

Présentation de l'opéron aceBAK - L'opéron 'acetate (ace) ' aceBAK, code pour trois protéines : la malate synthase (aceB), l'isocitrate lyase (aceA), et l'isocitrate déshydrogénase kinase/phosphatase (aceK) (Figure 1.12B). Ces 3 enzymes sont impliquées dans le cycle glyoxylate (**Figure 1.12A**). (Maloy and Nunn, 1982).

Dans le cycle complet de Krebs, l'isositrate est converti en alpha-ketoglutarate et en dioxyde de carbone par l'isocitrate déshydrogénase (IDH). Quand l'acétate est la seule source de carbone, l'aceK phosphoryle l'IDH ce qui entraine sa désactivation. L'inactivation de l'IDH engendre une accumulation d'isocitrate et favorise la conversion de l'isocitrate en glyoxylate par l'isocitrate lyase. Le glyoxylate combiné à l'acétyl-CoA forme du malate, un intermédiaire du cycle de Krebs. (**Figure 1.12A)** L'IDH kinase/phosphatase codée par *aceK* catalyse ce changement d'état (Zheng, Yates et Jia, 2012).

Le contrôle de l'expression de l'opéron aceBAK s'exerce à deux niveaux :

Régulation de l'opéron aceBAK au niveau transcriptionnel – L'expression de l'opéron aceBAK en présence de glucose est inhibée : le répresseur IclR empêche la fixation de l'ARN polymérase dans la région promotrice. Ce n'est qu'en présence d'acétate et en l'absence de glucose que la transcription peut avoir lieu. (**Figure 1.12B**). Lorsque les microorganismes sont cultivés sur acétate, le répresseur IclR est libéré de la région du promoteur, permettant l'expression de l'opéron aceBAK. Le glyoxylate et le pyruvate se lient au même site de fixation de la protéine IclR mais ont des effets opposés. Le glyoxylate empêche la formation du complexe IclR-promoteur, activant l'expression de l'opéron acétate 'aceBAK', alors que le pyruvate favorise la liaison de IclR au promoteur aceBAK lors de la croissance sur glucose et stabilise la forme tétramère active de IclR réprimant l'opéron (Lorca et al., 2007)

Régulation de l'activité de la protéine AceK au niveau post-traductionnel La régulation de l'événement de phosphorylation / déphosphorylation de l'IDH par AceK est un mécanisme post-traductionnel. La forme non phosphorylée de l'IDH (forme active) est dominante lorsque les produits des réactions enzymatiques génératrices d'ATP sont abondants (Laporte et Koshland, 1982). L'activité d'AceK dépend du ratio énergétique ATP/ADP, l'ATP active la fonction phosphatase.

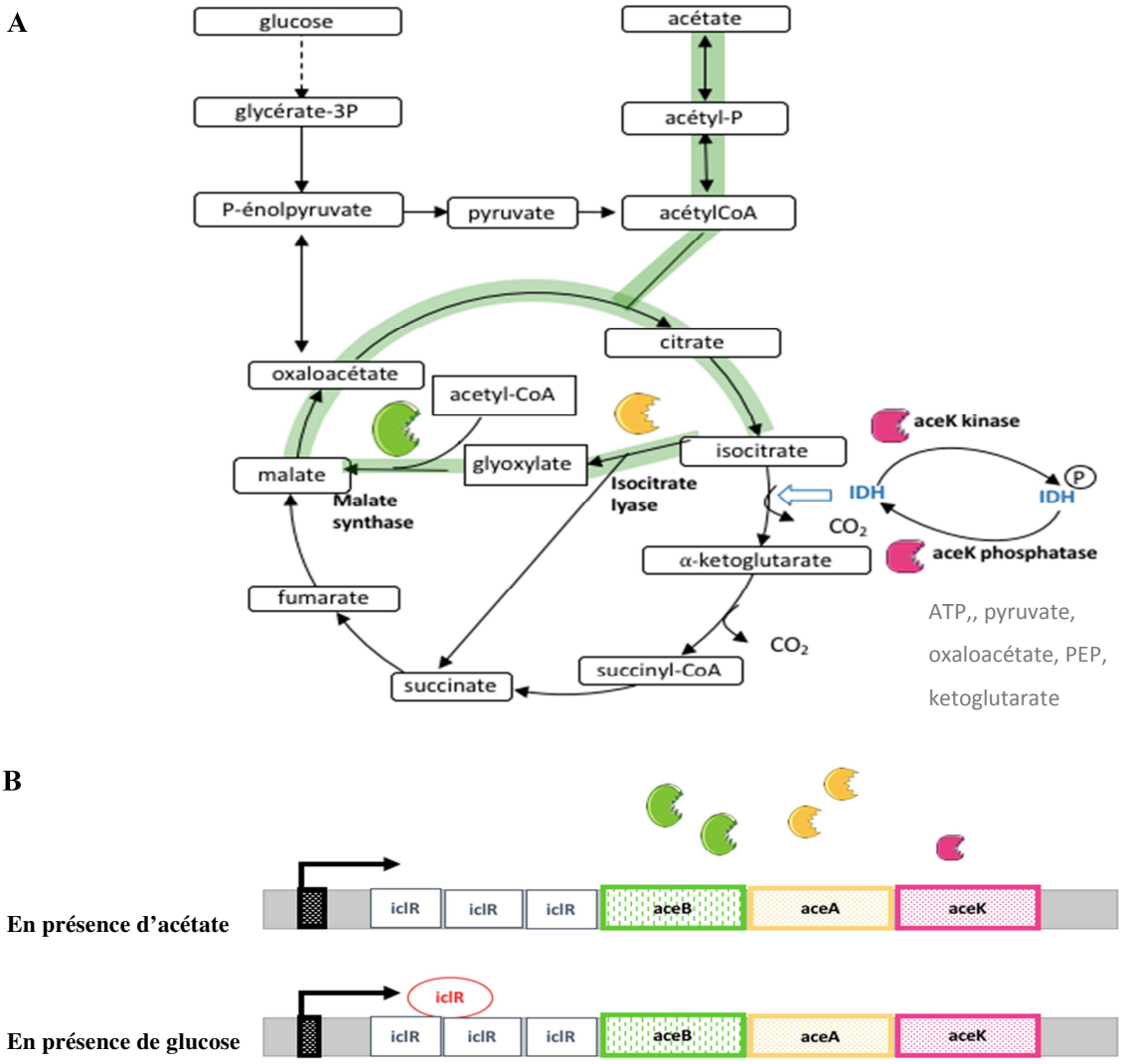


Figure 1.10 : Cycle de Krebs et du passage vers la voie glyoxylate médié par l'activité kinase d'aceK. A. Représentation du cycle de Krebs et de la voie glyxoylate, shéma adapté de celui de *(Zheng, Yates et Jia, 2012)****B La transcription de l'opéron aceBAK et iclR.*** *L'opéron 'acetate (ace) ' aceBAK, code pour trois protéines : la malate synthase (aceB) en vert, l'isocitrate lyase (aceA) en jaune, et l'isocitrate déshydrogénase kinase/phosphatase aceK en rose*

Pour aboutir à la conversion de glucose en AG par voie microbienne deux défis ont été relevés : (i) lever l'inhibition du cycle glyoxylate en présence de glucose (ii) favoriser l'accumulation de glyoxylate et sa conversion en AG. La production d'AG à partir d'acétate n'a pas été étudiée car utiliser le glucose comme substrat augmente le rendement théorique de production d'AG. La production d'AG à partir d'acétate a un rendement théorique de 0,57 g d'AG g^{-1} d'acétate (g/g), tandis que la production d'AG via la glycolyse et le cycle glyoxylate permet un rendement théorique maximal de 0,85 g/g à partir de glucose.

Cinq grands projets de recherche ont été menés sur l'ingénierie du cycle glyoxylate avec l'objectif commun de concevoir la voie de conversion du glucose en AG la plus performante, impliquant des laboratoires en Chine, aux Etats-Unis, en Belgique, en France, en Allemagne, en Suisse, en Corée du Sud et en Finlande. Ces études sont représentées de façon schématique dans la **Figure 1.11**.

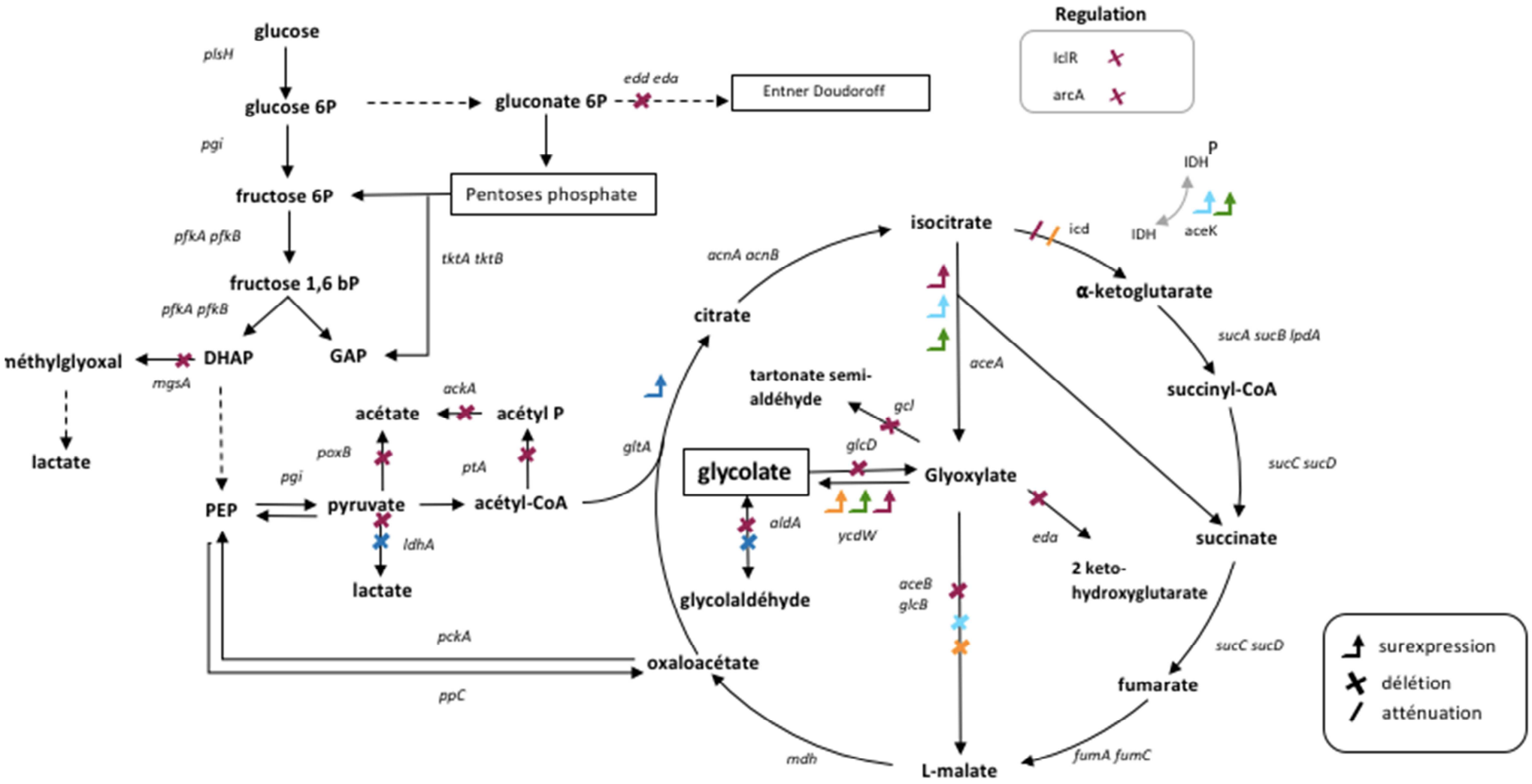


Figure 1.11 : Représentation des projets existants d'ingénierie du cycle glyoxylate procaryote pour la production d'acide glycolique : Dishert et al, 2012 US2012/0315682, Martin et al. (2013), Zahoor et al. (2013), Deng et al. (2015, 2018)

6.2.3 Le projet METEX : ingénierie du cycle glyoxylate d'*Escherichia coli*

En 2006 la société METEX (METabolic Explorer) basée à Clermont-Ferrand (France) a initié le projet de développement d'un procédé de production à fort rendement de l'AG à partir de glucose, pour l'industrie des plastiques biosourcés. Un accord a été signé avec Roquette, un des leaders mondiaux de la transformation d'amidon, pour l'industrialisation d'un bioprocédé breveté et la commercialisation de l'AG Ainsi de 2007 à 2012, plusieurs brevets ont été déposés par la société METEX (WO 2007/140816, WO2007/141316, WO2010004032, WO2011036213A2, WO2011577281A1, US2012/0315682) sur la production d'AG par fermentation de ressources renouvelables. Le glucose est utilisé comme substrat modèle, par *E.coli*. La version la plus aboutie (en **rose** sur la **Figure 1.11 :**US2012/0315682) propose une ingénierie métabolique de 'niveau 1' qui exploite les connaissances fondamentales de régulation du métabolisme carboné central et du cycle du glyoxylate.

A) Redirection du flux de carbone vers le cycle glyoxylate

Comme nous l'avons vu précédemment la régulation du cycle glyoxylate s'exerce au niveau de l'isocitrate déhydrogénase. Le gène *icd* codant pour l'IDH a été modifié afin d'atténuer l'activité de l'IDH, le gène *aceK* codant pour l'isocitrate déshydrogénase kinase/phosphatase a été délété afin de pouvoir entièrement contrôler l'activité résiduelle d'*icd.* Le gène codant pour répresseur IclR de l'opéron aceBAK a également été délété. Le gène *aceA* codant pour l'isocitrate lyase, première enzyme du cycle glyoxylate qui clive l'isocitrate en glyoxylate et succinate a été surexprimé.

Enfin le système Arc (composé de ArcA et ArcB) régule le métabolisme aérobie d'*E.coli.* Il module au niveau transcriptionel l'expression de 100 à 150 opérons impliqués dans le transport, le métabolisme énergétique, le catabolisme et l'état redox de la cellule (Georgellis, Kwon et Lin, 1999; Liu et De Wulf, 2004). Perrenoud et Sauer ont montré qu'ArcA régule négativement les gènes du cycle de Krebs indépendamment d'ArcB (Perrenoud et Sauer, 2005).. L'effet de la délétion d'*arcA* combinée avec les modifications génétiques préalablement effectuées (réduction de l'expression d'*icd*, surexpression d'*aceA*) augmente le flux dans le cycle glyoxylate.

B) Elimination des voies de consommation des intermédiaires entre glucose et acétyl-CoA –

Le glucose est assimilé par la voie de la glycolyse pour la synthèse d'acétyl-CoA qui alimente le cycle glyoxylate ingénieré. La conversion du glucose en acétyl-CoA comporte 10 réactions dont certains intermédiaires comme le glucose 6P, dihydroxyacétone phosphate (DHAP), le pyruvate et l'acétyl-CoA, sont impliqués dans des réactions secondaires. Les réactions non essentielles à la viabilité cellulaires ont été supprimées. Ainsi les gènes *mgsA*, *poxB*, *ldh*, *pta* ont été délétés pour minimiser la perte de carbone lors de l'assimilation du glucose.

C) Augmentation de la conversion du glyoxylate en AG dans *Escherichia coli*

Le glyoxylate ne s'accumule pas naturellement (**Figure 1.12**). Il peut être converti en (i) malate par la malate synthase codée par les gènes *aceB* et *glcB* (Chung, Klumpp et LaPorte, 1988) (ii) tartronate sema-aldéhyde par la glyoxylate carboligase *gcl* (Ornston et Ornston, 1969) (iii) 2-keto-4hydroxy-glutarate par la 2-keto-3-deoxyglutarate-6P aldolase codée par *eda* (Nishihara et Dekker, 1972) (iv) en glycolate par une réductase NADPH dépendante codée par *ycdW* (Nuñez et al., 2001). Pour favoriser la réduction du glyoxylate en glycolate les gènes *aceB, glcB, glc et eda* ont été délétés et l'expression du gène *ycdw* a été augmentée.

Le glycolate peut être réduit en glycolaldéhyde par une glycolaldéhyde déshydrogénase NAD+ dépendante, codée par *aldA*, ou oxydé en glyoxylate par une glycolate oxydase codée par *glcD*. Pour favoriser la production de glycolate *aldA* a été délété, tout comme *glcD* afin d'éviter un cycle futile glyoxylate-glycolate.

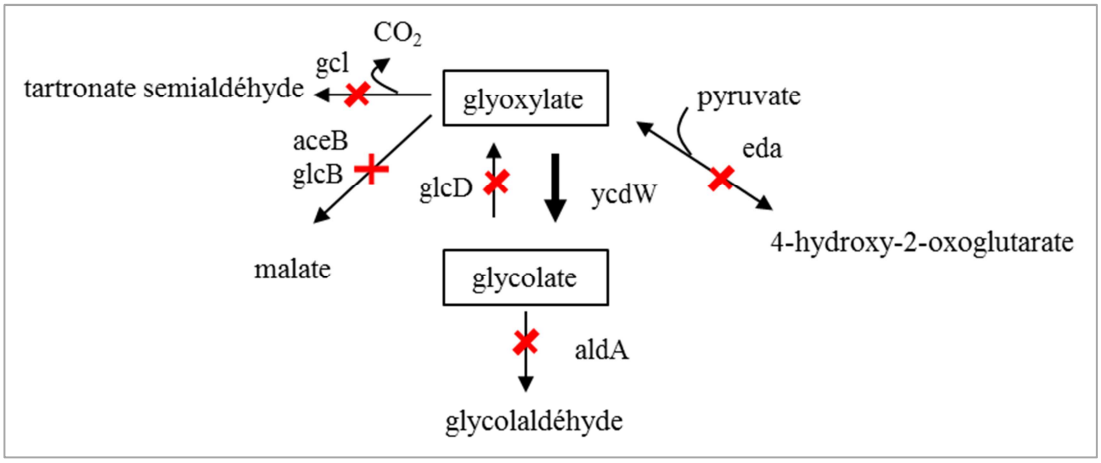


Figure 1.12 : Réactions de consommation du glyoxylate ou glycolate dans Escherichia coli. *Afin de favoriser la formation de glyoxylate et l'accumulation de glycolate, les malates synthases aceB et glcB, la 2-keto-3 deoxygluconate-6-P aldolase eda, ainsi que la glycolate oxydase glcD et l'aldéhyde déhydrogénase aldA ont été délétées.*

D) Obtention d'une souche optimisée pour la production d'AG à partir de glucose.

La souche finale dérivée d'*E.coli* K12-MG1655 comporte 16 délétions *: aceB, gcl, glcDEFGB, aldA, iclR, edd-eda ,poxB, ackA-pta, ldhA et mgsA*, le gène *icd* atténué et 2 gènes *ycdW* et *aceA* surexprimés. Cette souche a permis la production de 52,2 g/L d'AG, à un rendement de 0,38 g/g et une productivité de 1,33 g/L/h à partir de glucose.

6.2.4 Projets d'optimisation de l'ingénierie métabolique de la voie glyoxylate pour une production maximale d'acide glycolique chez *Escherichia coli*

Les études d'ingénierie métabolique succédant au projet METEX ont eu pour objectif d'apporter des éléments d'optimisation. Dans ce sens, le travail de Deng Yu (2015), en **bleu clair** sur la **Figure 1.11**,a consisté à faire évolué une souche d'*E.coli* dont les gènes des malates synthases *aceB* et *glcB* étaient délétés, le gène *glcB* avait été remplacé par une cassette surexprimant les gènes *ycdW*, *aceA et aceK* au niveau chromosomique (Deng, Mao et Zhang, 2015). L'évolution adaptative de la souche a entrainé une augmentation du taux de croissance et de la production d'AG à 56,44 g/L, 61,2% du rendement théorique a été atteint. La caractérisation de la souche finale avec la meilleure croissance a montré une activité isocitrate déhydrogénase augmentée comparée à la souche initiale. Cette étude a révélé que l'IDH doit être atténuée avec précaution : suffisamment réduite pour autoriser l'entrée de l'isocitrate dans la voie glyoxylate mais une inhibition trop élevée bloque le cycle de Krebs et la croissance. En 2018, la même équipe a démontré que la surexpression de la citrate synthase en complément de la surexpression d'*aceA*, *aceK* et *ycdW* et de la délétion d'*aceB* et *glcB* dans *E.coli* permettait d'atteindre 92,9% du rendement maximal fixé à 65,5g/g (les modifications apportées par rapport à l'étude de 2015 sont en **bleu foncé** sur la **Figure 1.11**) (Deng et al., 2018). Comment est-ce que la surexpression de la citrate synthase a permis d'atteindre un rendement aussi haut ?

Certaines des réactions de la glycolyse et du cycle de Krebs libèrent du NADH qui doit être en permanence recyclé en NAD+ pour permettre à ces voies la poursuite de la transformation du glucose (Castaño-Cerezo et al., 2009). L'oxydation du NADH dépend de la vitesse à laquelle la cellule peut consommer l'oxygène disponible dans le milieu. Lors d'une croissance exponentielle sur glucose, les capacités respiratoires de la cellule saturent (Eiteman and Altman, 2006) et tout le NADH généré par l'influx important de glucose ne peut être oxydé. Le NADH s'accumule. Or ce dernier inhibe, entre autres, l'enzyme citrate synthase qui catalyse la conversion de l'acétyl-CoA en citrate. Cette réaction est la première étape du cycle de Krebs. L'effet inhibiteur du NADH réduit donc le flux empruntant cette voie métabolique et un « goulot d'étranglement » se crée à l'entrée du cycle de Krebs ce qui entraine l'accumulation d'acétyl-CoA et de pyruvate (Majewski R A et Majewski R A, Domach, 1990; Vemuri *et al.*, 2006). Ces deux métabolites sont au carrefour de plusieurs voies métaboliques et ils servent de précurseurs à de nombreuses voies de biosynthèse. C'est pourquoi la bactérie doit

réguler leur concentration pour éviter toute perturbation de son fonctionnement. Normalement la solution optimale adoptée par la bactérie est de détourner le flux de matière arrivant à l'entrée du cycle de Krebs vers la production d'acétate mais ici en surexprimant la citrate synthase le flux est détourné vers la production de citrate et vers la voie glyoxylate (**Figure 1.11**).

Le rendement maximal de conversion du glucose en AG via le cycle glyoxylate ingénieré a été atteint dans *E.coli.* Les travaux se sont étendus à d'autres microorganismes aux caractéristiques intéressantes pour un procédé industriel de production d'AG.

6.3 Niveau 2 d'ingénierie métabolique : production d'AG *via* l'expression de voies hétérologues

6.3.1 Ingénierie du cycle glyoxylate d'autres microorganismes qu'*Escherichia coli*

La production d'AG via le cycle glyoxylate a été démontré chez *Corynebacterium glutamincum* en **orange** sur la **Figure 1.11** (Zahoor, Otten et Wendisch, 2014). *C. glutamicum* a l'avantage de résister à des concentrations d'acide élevées. La croissance est affectée seulement de moitié à un concentration de 35 g/L d'AG. La souche a été ingéniérée en atténuant d'*icd*, délétant *aceB* et surexprimant d'Y*cdw (*hétérologue provenant d'*E.coli*) *et d'aceA.* La production de glycolate rapportée est plus faible : 5,3 g/L, avec un rendement de 0,18 g/g.

Une seule étude publiée rapporte la production d'AG via le cycle glyoxylate chez les eucaryotes(Koivistoinen et al., 2013). Les micro-organismes eucaryotes ont l'avantage d'être tolérant aux inhibiteurs de croissance présents dans les hydrolysats lignocellulosiques et de croître à pH acide. O. Koivistoinen et al. (2013) a choisi *S.cerevisiae* et *Kluyveromyces lactis* pour la production d'AG. Ces deux espèces de levures sont capables de croître à des concentrations d'AG allant jusqu'à 30-50 g/L. L'originalité de cette étude vient du fait d'étudier la production d'AG à pH acide, ce qui s'est déjà montré avantageux pour la production d'acide lactique et d'exprimer une glyoxylate réductase hétérologue provenant d'*Arabidopsis thaliana*, qui présente une meilleure affinité pour le glyoxylate que l'enzyme native de *S.cerevisiae*.. Bien que la fermentation à pH acide soit avantageuse, la capacité de la souche à produire de l'AG est limitante. La meilleure production avec *K. lactis* est de 15 g/L.

Les études d'ingénierie du cycle glyoxylate pour la biosynthèse d'AG à partir de glucose chez *Corynebacterium*, chez la levure et *E.coli* (Koivistoinen *et al.*, 2013; Zahoor, Otten et Wendisch, 2014) démontrent que la suppression des activités malate synthase et isocitrate dehydrogénase et la surexpression de la glycolate réductase sont essentielles et suffisantes pour conférer le phénotype producteur d'AG.

6.3.2 Ingénierie métabolique pour la conversion du xylose en acide glycolique

La conversion optimale du glucose permise par le cycle ingénierié du glyoxylate n'est pas suffisante pour valoriser les hydrolysats lignocellulosiques qui se composent d'un mélange de D-glucose et D-xylose. Certains microorganismes tels que *Pseudomonasfragi Burkholderiaxenovorans et Caulobacter crescentus* possèdent une voie d'assimilation oxydative dépendante du NAD^+ du xylose complètement différente de celle des pentoses phosphate Le D-xylose est converti en 3 étapes en 2-keto-3-deoxy-xylonate (Stephen Dahms, 1974). Le 2-keto-3-deoxy-xylonate est alors converti à travers 2 réactions couplées en α-ketoglutarate (intermédiaire du cycle de Krebs) ou consommé par la voie de Dahms qui produit du glycolaldéhyde et du pyruvate. Cette dernière voie présente un intérêt pour la production d'AG. Le glycolaldéhyde peut être soit oxydé en AG par une aldéhyde déshydrogénase ou réduit en éthylène glycol par une réductase. Le pyruvate destiné à entrer dans le cycle de Krebs peut être redirigé vers la synthèse d'autres molécules à valeurs ajoutée telle que l'acide lactique. *S. cerevisiae* a ainsi été ingénierée comme plateforme de production d'AG et d'éthylène glycol avec des rendements théoriques de 1 mol d'AG par mol de xylose (Cabulong et al., 2017). L'expression hétérologue des gènes codant pour la D-xylose deshydrogénase (XylB) et la D-xylonate déshydratase (XylD) de *Caulobacter crescentus*, l'aldolase (YagE) et l'aldéhyde déshydrogénase (AldA) d'*E.coli* a permis la synthèse de 150 mg/L d'AG par *S.cerevesiae*. L'assimilation incomplète de l'acide D-xylonique sécrété, qui avait une concentration résiduelle de 3,21 g/L limite le rendement. C'est un phénotype observé pour les souches employant la voie de Dahms et un défi d'ingénierie métabolique (Cabulong et al., 2017) .

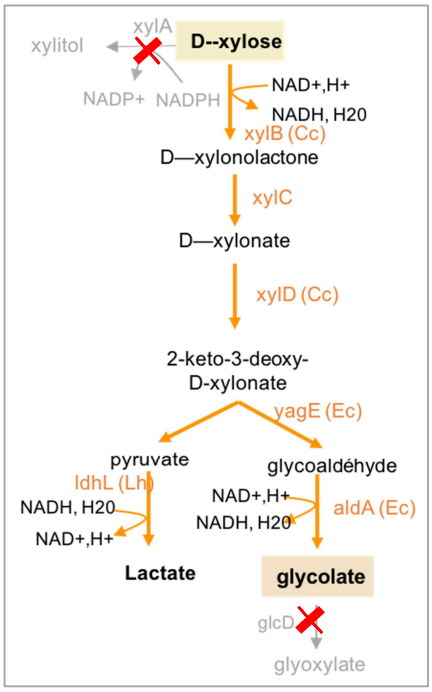


Figure 1.13 : Voie métabolique de synthèse d'oxydation du D-xylose pour la production d'acide glycolique (et de lactate) dans Saccharomyces cereviase. *En noir : voie métabolique de S.cerevisiae non modifiée , xylC : lactonase. En gris : réactions supprimées : xylA : aldo-keto reductase, glcD : glycolase oxydase. En orange : voies surexprimées synthétiques pour la mise en place de la voie de Dahms xylB : D-xylose déshydrogénase de Caulobacter crescentus, XylD : Dihydroxy-acid dehydratase de Caulobacter crescentus, yagE : aldolase d'Escherichia coli, aldA : aldéhyde déshydrogénase d'Escherichia coli et ldhL : lactate déshydrogénase de Lactobacillus helveticus.*

6.4 Niveau 3 d'ingénierie métabolique : production d'AG à partir de xylose via l'expression de voies non-naturelles

Les trois projets de biologie de synthèse présentés ci-dessous ont pour objectif commun d'ouvrir la synthèse d'AG par voie microbienne aux pentoses.

6.4.1 Voie synthétique du xylulose 1P

Le premier projet a été mené au sein de l'équipe du prof. Jean Marie François dans le but d'améliorer l'utilisation d'hydrolysats hémi-cellulosique. (WO2016079440,Alkim *et al.*, 2015, 2016; Cam *et al.*, 2016). Cette voie synthétique d'assimilation du D-xylose (pentose majoritaire de l'hémicellulose) développée chez *Escherichia coli* phosphoryle le xylulose sur le premier carbone (C1) au lieu du carbone 5 (C5) comme c'est le cas dans la voie des pentoses phosphate. Cette réaction de phosphorylation, catalysée par une ketohexokinase hétérologue issue de mammifère, est suivie d'une réaction de clivage par une aldolase conduisant à une molécule en C3 (dihydroxyacétone-P) et C2 (glycolaldéhyde). Ce dernier par une réaction d'oxydation catalysée par une glycoladéhyde déshydrogénase conduit à la synthèse d'AG (ou par une réaction de réduction catalysée par une réductase endogène produit à l'éthylène glycol, non-discuté ici) (Figure 1.16).

Cette voie synthétique permet d'obtenir une production théorique molaire en AG (1 mole par mole de C5), atteint à 90% avec un rendement de 0,46 g/g xylose, conférant un rendement théorique maximum d'AG supérieur de 20% par rapport au métabolisme naturel de ce microorganisme.

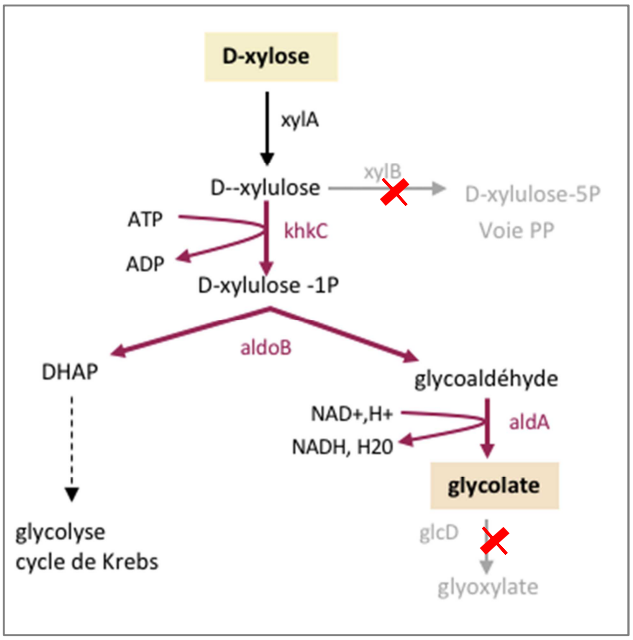


Figure 1.14 : Voie synthétique 'xylulose-1P' d'assimilation du D-xylose*. En rose framboise la voie synthétique, en noir réaction naturelle et en gris : délétion. xylA : xylose épimease, xylB : xylulose kinase, voie 'PP' : voie des pentoses phosphate, khkC : xylulose kinase, aldoB : glucose1,6bP aldolase (Homo sapiens), aldA : glycolate oxydas (E.coli)e.*

6.4.2 Voie synthétique du ribulose-1P

Un ensemble de voies synthétiques spécifiques aux pentoses pour la production de composés à deux carbones a été développée par l'équipe de Prof. Gregory Stephanopoulos, et un brevet a été déposé par le Massachusetts Institute of Technology (MIT) (WO2017059236, Pereira *et al.*, 2016). Ce réseau métabolique implique la phosphorylation en position 1 du D-ribulose clivé par une aldolase en glycolaldéhyde et DHAP. Le glycolaldéhyde est directement oxydé en glycolate (C2) (**Figure 1.15**)La démonstration a été faite à partir de D-xylose avec un rendement de 0,44 g/g.

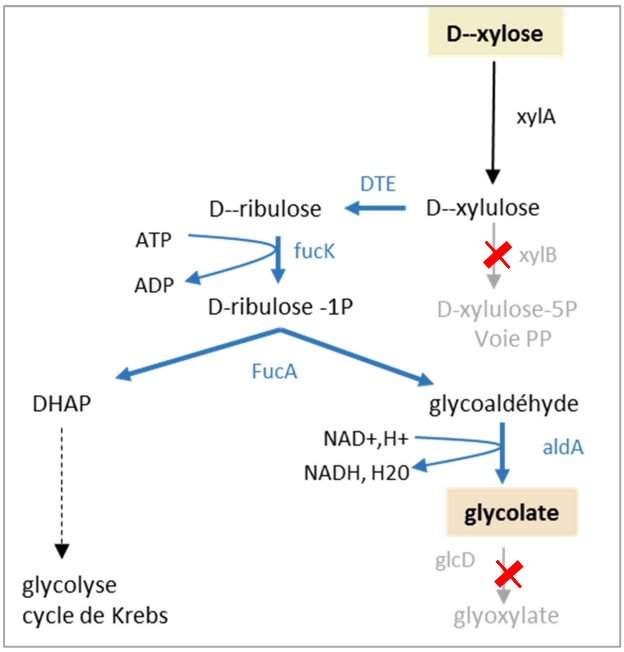


***Figure 1.15: Voie synthétique 'ribulose-1P' d'assimilation des pentoses dans Escherichia coli** . En noir les voies naturelles d'E.coli : xylA ; L-xylose isomérase, en bleu les gènes surexprimés pour la mise en place de la voie synthétique : DTE : D-tagatose 3- épimérase de Pseudomonas cichorii , rhaB (E.coli) : rhamnulokinase, fucK (E.coli) : Lfuculokinase, fucA (E.coli) : L-fuculose-P aldolase, En gris ; les gènes délétés. xylB : xylulokinase. glcD : glycolate oxidase.*

6.5 Combinaison d'approches d'ingénierie métabolique pour une valorisation des pentoses et hexoses en AG

Les premières études se sont basées sur l'ingénierie du cycle glyoxylate, pour la conversion du glucose contenu dans l'amidon en AG (**Figure 1.16** : **vert**) (Discher et al., 2007). Le rendement maximal autorisé est de 2 moles d'AG par mole de glucose, ce rendement a été atteint à 92,9% (Deng *et al.*, 2018). Pour les procédés dits de 2eme génération utilisant la lignocellulose - dont la portion de pentose peut atteindre 40% - une bioconversion restreinte au glucose est limitante. De nouvelles voies synthétiques de synthèse d'AG à partir de pentoses ont été décrites (Alkim et al., 2015, 2016b; Cabulong et al., 2017; Cam et al., 2016; Pereira et al., 2016). L'ensemble de ces voies clivent un C5 pour générer un C2 valorisé en AG, et C3 (DHAP ou pyruvate) avec un rendement théorique de 1 mole d'AG par mole de pentose. Afin d'éviter la perte de C3 dans le métabolisme carboné central, ces voies ont été couplées à la voie glyoxylate ingénierée ce qui conduit à une augmentation du rendement en AG à 2 mole par mole de pentose. (**Tableau 1.4**).

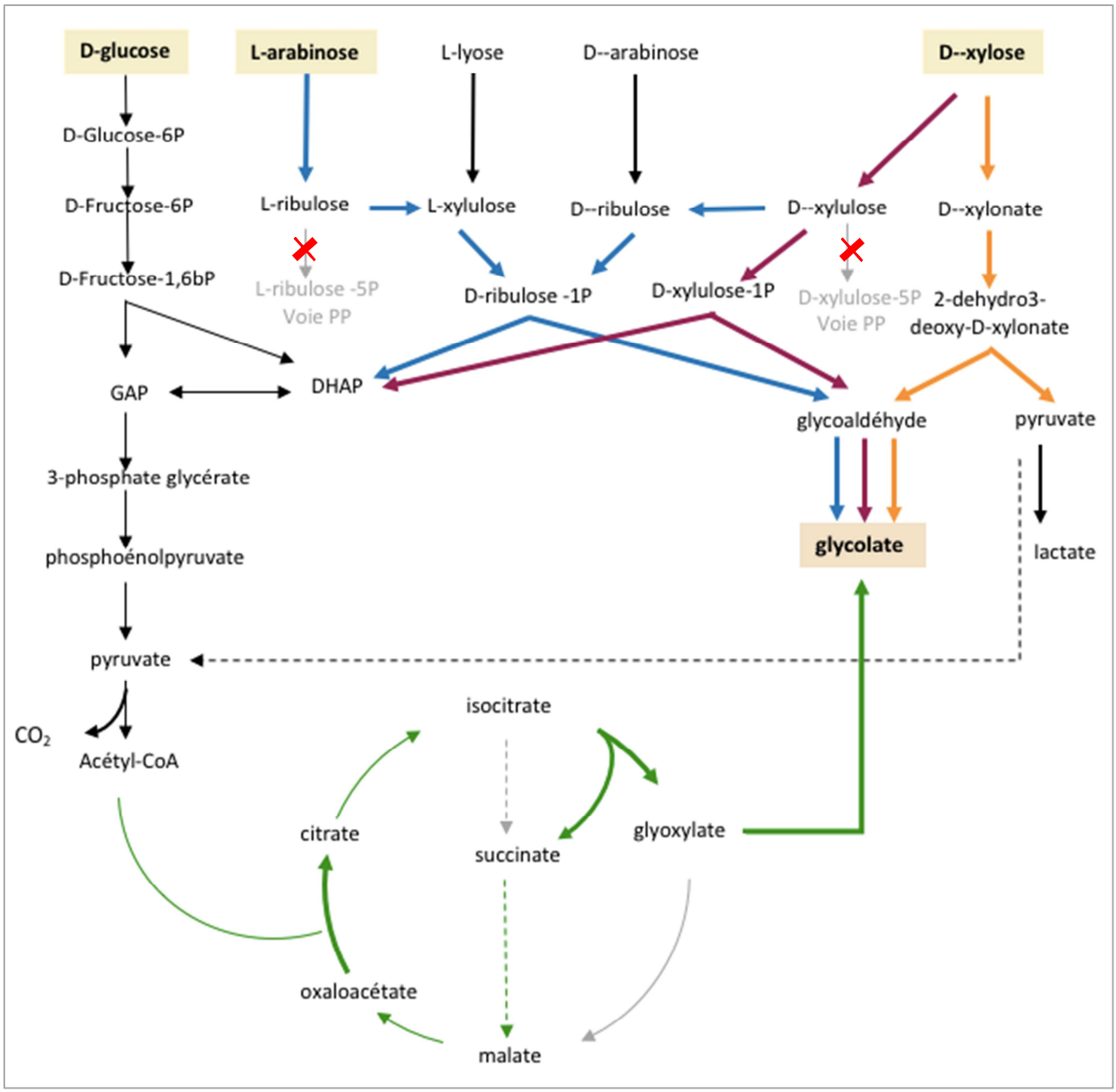


Figure 1.16 : Représentation de l'ensemble des voies de production microbienne d'AG existantes dans Escherichia coli. *En rouge : délétions. En vert la voie du glyoxylate ingéniérée de conversion du glucose, en bordeau la voie synthétique du Xylulose 1P d'asismilation du xylose, en bleu la voie synthétique 'épimérase' d'assimilation des pentoses, en orange la voie de Dahms d'oxydation du xylose.*

La combinaison de la synthétique D-xylulose-1P avec la voie du glyoxylate ingénierée pour valoriser la perte DHAP (C3) a abouti à un rendement de 0, 43g AG par g de xylose, et de 0,63 g/g par de sucres (mélange de xylose et glucose). (Alkim *et al.*, 2015, 2016; Cam *et al.*, 2016, **Figure 1.16** en **rose** et **vert**). De la même manière le couplage de la voie synthétique D-ribulose-1P a permis un rendement de 0,63 g/g à partir du xylose (**Figure 1.16** en **bleu** et **vert**).

L'approche la plus récente développée par Cabulong *et al.*, (2018)chez *E.coli* repose sur la production d'AG à partir de D-xylose par combinaison de la voie de Dahms avec le cycle glyoxylate (**Figure 1.16**, en **orange** et **vert**). La production autorisée par cette approche est de 4,57 g/L avec un rendement de 0,46 g/g à partir du D-xylose. Enfin, la production de poly(lactate-co-glycolate) (polymères contenant du glycolate en tant que co-monomère du lactate) à partir du D-xylose par la voie de Dahms a été mise en œuvre dans E. coli (Choi et al., 2017). Le glycoaldéhyde et le pyruvate issus de la voie de Dahms, sont alors onvertis en AG et lactique, respectivement.

Nous pouvons constater que la perte de CO_2 lors de l'étape de décarboxylation du pyruvate est commune à l'ensemble des voies (**Figure 1.16**). Pour dépasser le rendement théorique de 2 moles d'AG par mole de pentose et/ou hexose, il est nécessaire de redessiner les voies métaboliques existantes afin de s'affranchir de l'étape de catalysée par le complexe pyruvate déshydrogénase (PDH) qui cause une perte de CO_2 majeure.

Tableau 1.3 : Rendement en acide glycolique des voies de synthèse d'acide glycolique en fonction de la source carbonnée utilisée comme substrat.

Voie(s) métabolique(s) de production d'AG	Xylose	Glucose
Cycle glyoxylate ingénieré	/	2 mol/mol
Voie synthétique, clivage de C5 en C2 et C3	1 mol/mol	/
Voie de valorisation de C5 couplée au cycle glyoxylate ingénieré	2 mol/mol	2 mol/mol

7. Exemples de voies microbiennes de conservation du carbone

Le CO_2 libéré lors des étapes de décarboxylation peut être considéré comme un sous-produit, générant une perte de carbone. Il a longtemps été perçu comme un co-produit de réaction, et la perte de CO_2 n'était pas intégrée dans le processus d'optimisation.

La conversion du pyruvate en acétyl-CoA implique une décarboxylation et donc une perte de carbone sous forme de CO_2. Cette perte de carbone diminue sensiblement le rendement théorique des produits désirés. Pour obtenir une conservation totale du carbone trois alternatives existent: la voie naturelle Wood-Ljungdahl (Schuchmann et Müller, 2014); et deux voies de synthèse, la glycolyse non oxydative (NOG) (Bogorad, Lin et Liao, 2013) et le cycle glyoxylate inversé (rGS) (Mainguet et al., 2013) (**Figure 1.17**).

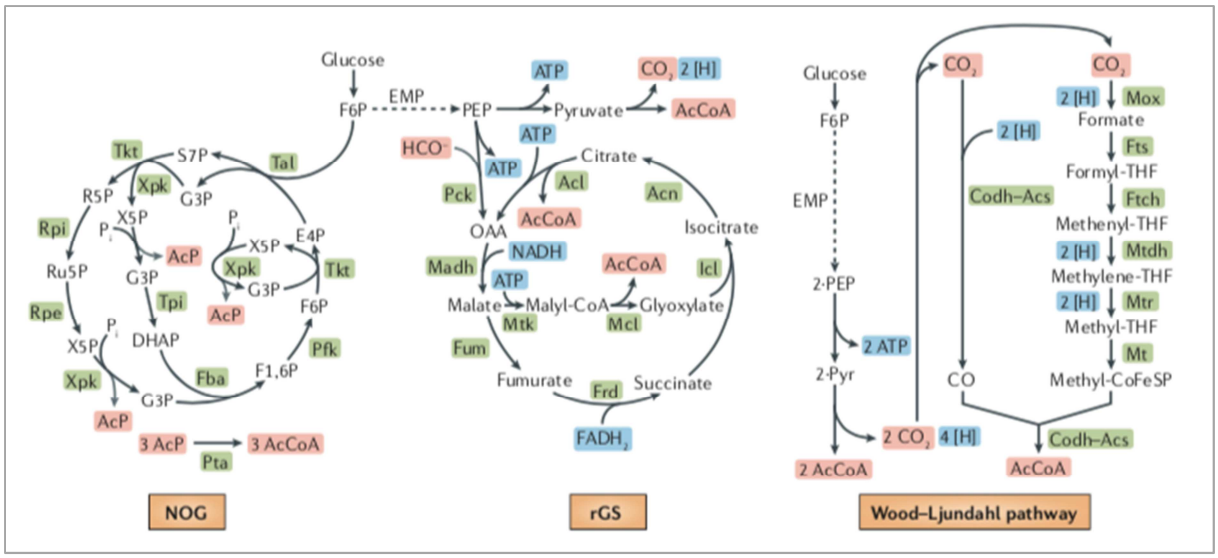


Figure 1.17 : Réarrangement du métabolisme carboné central pour une conservation totale du carbone (Liao et al., 2016) *AcCoA, acétyl CoA ; Acl, ATP-citrate lyase, Acn : aconitase, Codh-Acs : carbon monoxide dehydrogenase acetyl CoA synthase, CoFeSP : corrinoid iron-sulfur protein, DHAP : dihydroxyacétone phosphate, EMP : voie Embden-Meyerhof-Parnas (glycolyse), F1,6P : fructose 1,6 bisphosphate, fba : F1,6P aldolase, Frd : fumarate réductase, Ftch : formyl-THF cyclohydrolase, Fts : formyl THF synthase, Fum, fumarase, Icl : isocitrate lyase, Madh : malate déshydrogénase, Mox : methanol oxidase, Mt : méthyltransférase, Mtdh : méthylène-THF déshydrogénase, Mtr : méthylène-THF réductase, Pfk : phosphofructokinase, Pi : phosphate inorganique, Pta : phosphate acétyltransférase, R5P : ribose 5 phosphate, Rpe : R5P épimérase, Rpi : R5P isomérase, Ru5P : ribulose-5-phosphate, S7P : sédohéptulose-7-phosphate, THF : tétrahydrofolate, Tpi : triosephosphate isomérase.*

Voie naturelle de Wood-Ljungdahl - Les bactéries acétogènes et les méthanogènes possèdent une voie de fixation des 2 CO_2 libérés dans le cycle de Krebs appelée voie Wood-Ljungdahl (Schuchmann et Müller, 2014). Cependant, la voie implique plusieurs enzymes complexes qui sont sensibles à l'oxygène. Par conséquent, l'ingénierie de cette voie est un défi.

Voie non-naturelle 'reverse glyoxylate shunt (rGS) **-** Le rGS fournit une approche différente pour atteindre la conservation complète du carbone. Le schéma général de rGS commence par convertir le phosphoénolpyruvate (PEP) en oxaloacétate (OAA) en utilisant la PEP carboxykinase (Pck) tout en fixant une molécule de CO2. Le métabolite C4 résultant, OAA, peut être converti en deux molécules d'acétyl-CoA par rGS et les réactions associées. L'inversion de la dérivation de glyoxylate est rendue possible par une malate-thiokinase ATP-dépendante (Mtk) et une malyl-CoA-lyase (Mcl) (**Figure 1.17**). L'autre PEP généré à partir du glucose traverse la partie inférieure de la glycolyse pour générer un autre acétyl-CoA. Les deux voies parallèles combinées peuvent atteindre une production nette de trois molécules d'acétyl-CoA (Mainguet et al., 2013).

Voie synthétique NOG – Le travail de l'équipe du Prof. Liao (UCLA) publié dans Nature (Bogorad et al., 2013) a construit une voie glycolytique remaniée évitant la décarboxylation du pyruvate et conduisant à une conversion stoechiométrique de C6 ou C5 en C2 (acétate) sans perte de carbone. Nous pouvons nous demander si il est possible de remanier une partie de ces réseaux métaboliques sans porter atteinte à l'intégrité cellulaire, c'est-à-dire lui permettre au moins de survivre. Une conversion quasi stoechiométrique des glucides en acétate par NOG a été démontrée expérimentalement *in vivo*. Le rendement maximal de 2,5 moles d'acétate par mole de xylose a en effet presque été atteint avec un rendement de 2,2 moles d'acetate par mole de xylose, dépassant largement le rendement de 1,67 moles d'acétate par mole de xylose autorisé par la voie de la glycolyse.

Ces travaux montrent qu'il est possible de remanier le métabolisme carboné central bactérien pour une conservation optimale du carbone tout en assurant le maintien de la viabilité cellulaire. Un défi est celui de construire des systèmes de conversion de ressources renouvelables, avec un rendement carbone optimisé et de ce fait tendre vers une réduction significative de l'émission de CO_2.

8. Objectifs principaux de la thèse

L'acide glycolique est produit à partir du formaldéhyde qui est un composé irritant (conjonctivites, maux de tête), et cancérogène, ce qui exclut cette substance dans des préparations bénéficiant d'une Autorisation de Mise sur le Marché (AMM). Par conséquent, l'AG biosourcé présente un intérêt certain dans les domaines pharmaceutiques, cométiques mais aussi textiles. Les voies microbiennes de production d'AG existantes ont un rendement théorique limité par la perte de carbone due à la décarboxylation du pyruvate en acétyl-CoA.

En conséquence, cette thèse a pour objectif de répondre à deux défis biotechnologiques :

(i) le développement d'un procédé de production d'AG respectueux de l'environnement qui valorise les principaux sucres de la biomasse végétale

(ii) le remaniement du métabolisme carboné central microbien pour la production de biomolécules d'intérêt avec une conservation optimale du carbone, ce qui a pour effet de réduire les émissions de CO2 et d'augmenter le rendement en produit.

Pour répondre à ces objectifs *E.coli* a été choisi comme organisme d'étude. Nous avons développé deux approches (**Chapitre III ; IV**) afin de réaliser la synthèse d'AG à partir de pentoses et d'hexoses *via* une seule voie synthétique. Une production d'AG indépendamment de la voie glyoxylate lève la contrainte stœchiométrique imposée par le métabolisme naturel *d'E.coli* et permet le remaniement du métabolisme carboné central.

Tableau 1.4 : Tableau récapitulatif des projets d'ingénierie métabolique pour la production d'acide glycolique

Micro-organisme	Niveau d'ingénierie métabolique	Modifications génétiques	Production AG (g/L)	Rendement AG (g/g)	Conditions de culture	Référence
Escherichia coli	Niveau 1	Cycle du glycoxylate ingénieré : •Délétions : *aceB, glc, glcDEFGB, aldA, iclR, edd-eda, pgi* • Surexpression : *ycdW*	2,33	0,25	*Culture en erlenmyer bafflé de 250 m*L : 50 ml de milieu M9 avec 10 g/L de **glucose**, pH 6,8 OD : 0,3, 30°C, 400 rpm, 0,1 mM IPTG	Souaille, US2009/0155867 A1
			31	0,22	*Culture en fed-batch,fermenteur de 600 mL.* 200 mL de M9 avec 40 g/L **glucose**, 0,1 mM IPTG, DOi = 2, 30°C, p02 : 30%	Souaille, US2009/0155867 A1
Escherichia coli	Niveau 1	Cycle du glycoxylate ingénieré : • Atténuation : *icD* • Délétions : *aceB, gcl, glcDEFGB, aldA, iclR, edd-eda, poxB, ackA-pta, arcA, mgsA* • Surexpressions : *ycdW, aceA*	5,14	0,52	*Culture en erlenmyer bafflé de 250 ml* : 50 ml de milieu M9 avec 10 g/L de **glucose**, pH 6,8 OD : 0,3, 30°C, 200 rpm, 0,1 mM IPTG	Dishert *et al*., US2012/0315682
Escherichia coli	Niveau 1	Cycle du glycoxylate ingénieré : • Atténuation : *icD* • Délétions : *aceB, gcl, glcDEFGB, aldA, iclR, edd-eda, poxB, ackA-pta, arcA* • Surexpressions : *ycdW, aceA*	4,36	0,4		Dishert *et al.,* US2012/0315682
			52,2	0,38	*Culture en fed-batch, fermenteur de 600 mL.* 200 mL de M9 avec 20 g/L **glucose**, 0,1 mM IPTG, DOi = 2, 30°C, p02 : 30%, 5 fois pulse glucose	Dishert *et al.,* US2012/0315682
Escherichia coli	Niveau 1	Cycle du glyoxylate ingénieré : • Délétions : *endA, recA* • Surexpressions : *YcdW, aceA, aceK*	1,4	nd	*Culture en erlenmyer de 250 mL* : M9 10g/L **glucose**, 0,1 mM IPTG, 30°C, 250 rpm, pH non indiqué - 96h de culture	Martin *et al*., 2013
Escherichia coli	Niveau 1	Cycle du glyoxylate ingéniéré : • Délétions : *aceB, glcB* • Surexpressions : *ycdW, aceA-aceK* (au niveau génomique - évolution)	2,66	0,33	*Culture en erlenmeyer de 500 mL* : milieu M9 8g/L **glucose**	Deng *et al*., 2015
			8,96	0,47	*Culture en batch dans un fermenteur de 5L*. : 32,2 g/L **glucose**, pH 6,8, 37°C, 24h	Deng *et al*., 2015
			56,44	0,52	*Culture en fed-batch dans un fermenteur de 5L*. : 32,2 g/L **glucose**, pH 6,8, 37°C, 120h (125,6 g/L de glucose consommé)	Deng *et al*., 2015

Tableau 1.4 (suite): Tableau récapitulatif des projets d'ingénierie métabolique pour la production d'acide glycolique

Micro-organisme	Niveau d'ingénierie métabolique	Modifications génétiques	Production AG (g/L)	Rendement AG (g/g)	Conditions de culture	Référence
Saccharomyces cerevisae	Niveau 1	Voie de Dahms : • Délétion : *fra2* • Surexpression : *xylB, xylD, yagE, aldA*	0,14	0,15	*Culture en erlenmeyer 250mL* : 50 mL de milieu synthétique complet, 10g/L glucose, 20g/L **xylose** 250 rpm, 30°C, 168h	Salusjärji *et al*., 2017
Escherichia coli	Niveau 1	Cycle du glyoxylate ingénieré : • Délétion : *ldhA, glcB, aceB, aldA* • Surexpression : *ycdW, aceA, aceK, gltA*	2,82	0,79	*Culture en erlenmeyer* milieu M9, 20g/L **glucose**, 37°C jusqu'à induction avec 1mM IPTG ou 250 ng/mL anhydrotétracyline puis 30°C	Deng *et al*., 2018
			65,5		*Culture en fed-batch* : milieu M9, 5g/L **glucose**	Deng e*t al*., 2018
Saccharomyces cerevisae	Niveau 2	Cycle du glyoxylate ingénieré : •Délétions : *msl1, dal7, idp2, reg1* • Surexpression : *GLYR1*	0,91	nd	*Culture en erlenmeyer de 250 mL* : milieu synthétique complet (SC) 20g/L **xylose**, 20 g/L ethanol DOi = 2, 30°C;, 250 rpm	Koivistoinen *et al.,* 2013
			2	nd		Koivistoinen *et al*., 2013
Kluyveromyces lactis	Niveau 2	Cycle du glyoxylate ingénieré : • Délétions : *mls1, ipd2* • Surexpression : *GLYR1*	15	nd	*Culture en fed-batch dans un fermenteur de 5000 ml :* milieu SC sans uracil avec 20 g/L **xylose** et 15 g/L ethanol, DO600 initiale= 1, pH5, 30°C, 500 rpm, 1 vvm dans un	Koivistoinen *et al*., 2013
Corynebacterium glutamicum	Niveau 2	Cycle du glyoxylate ingénieré : • Atténuation : *icd* • Délétion *: aceB* • Surexpression *: ycdW*	5,3	0,18	*Culture en erlenmeyer bafflés de 500 mL* : 50 mL de milieu GXII 12g/L **glucose**, 17,6 g/L potassium acetate, DO initiatiale : 1, 30°C	Zahoor *et al*., 2014
Escherichia coli	Niveau 3	Voie synthétique du Xylulose 1P : • Surexpression de *khkC, aldoB, aldA* • Délétion de xylB, glcD	4,3	0,46	*Culture en erlenmeyer 250 mL :* 50 mL de milieu minimal M9 avec 10g/L **xylose**, pH 7, 200 rmp, 37°C	Cam *et al*., 2015

Tableau 1.4 (suite): Tableau récapitulatif des projets d'ingénierie métabolique pour la production d'acide glycolique

Micro-organisme	Niveau d'ingénierie métabolique	Modifications génétiques	Production AG (g/L)	Rendement AG (g/g)	Conditions de culture	Référence
Escherichia coli	Niveau 3	Voie synthétique du Xylulose 1P : • Surexpression : *KhkC, aldoB, aldA* • Délétions : *xylB, glcD* Cycle du glyoxylate ingénieré : • Délétion aceB, gcl, glcEFGB, iclR, edd-eda, arcA, icd, xylB • Surexpression: ycdW, aceA	2,24	0,43	*Culture en erlenmeyer 250 mL :* 50 mL de milieu minimal M9 avec 10g/L **xylose** (2g/L tryptone, 1g/L extrait de levure), pH 7, 200 rmp, 37°C	Alkim *et al*., 2016
			3,73	0,63	*Culture en erlenmeyer 250 mL :* 50 mL de milieu minimal M9 avec 2,5g/L glucose, 5g/L **xylose** (2g/L tryptone, 1g/L extrait de levure), pH 7, 200 rmp, 37°C	Alkim *et al*., 2016
Escherichia coli	Niveau 3	Voie synthétique du 'ribulose-1P" : • Délétion : *glcD, xylB* • Surexpression : *Pc-dte, fucK, fucA, aldA*	3,43	0,46	*Culture en tube 50mL (test)* : 10 mL milieu minimal MM1 avec 10g/L **xylose,** 37°C, 250 rpm	Peirera *et al*.2016
			44	0,44	*Culture en batch dans un fermenteur de 3L* : 2L en milieu minimal MM2 avec 110 g/L **xylose**, 37°C, pO2 : 20%	Peirera et al.2016
Escherichia coli	Niveau 3	Voie synthétique du "ribulose-1P" : • Délétion : *glcD, xylB* • Surexpressions : *Pc-dte, fucK, fucA, aldA* Cycle du glyoxylate ingénieré : • Surexpressions : *ycdW, aeA, aceK* • Délétion : aceB, glcB, gcl	40	0,63	*Culture en fed-batch dans un fermenteur de 3*L : 2 L milieu minimal MM2 avec 65 g/L **xylose**,	Peirera *et al*.2016

Tableau 1.4 (suite): Tableau récapitulatif des projets d'ingénierie métabolique pour la production d'acide glycolique

Micro-organisme	Niveau d'ingénierie métabolique	Modifications génétiques	Production AG (g/L)	Rendement AG (g/g)	Conditions de culture	Référence
Escherichia coli	Niveau 3	Voie de Dahms et voie glyoxylate ingénierées : • Délétion : *xylAB, glcD* • Surexpression : *xdh(Cc), yagE, yagF, aldA, ycdW, aceA, aceK*	4,57	0,46	*Culture en erlenmeyer* : 100 mL de milileu minimal M9, 10g/L **xylose**, (1 g/L peptone, 0,5g/L yeast extract), 0,5 mM IPTG (à DO600nm=0,3), 37°C, 250 rpm	Cabulong et al., 2018

CHAPITRE II

MATERIELS ET METHODES

1. Construction des plasmides

La liste des différents plasmides bactériens communément utilisés en laboratoire, à partir dequels les plasmides de l'étude sont dérivés, ainsi que leurs caractéristiques sont présentées dans le **Tableau 2.1**.

Tableau 2.1 : Caractéristiques des plasmides commerciaux utilisés

Plasmides	Résistance	Promoteur	Origine de réplication	Utilisation	Source
pGEM-T easy	Amp	pT7 – inductible à l'IPTG	F1 ori / moyen nombre de copies	clonage/séquençage	Promega
pET28a	Kan	pT7 – inductible à l'IPTG	ColE1 ori / haut nombre de copies	Production et purification de protéines étiquetées polyhistidine	Novagen
pACT3	Chm	pTac – inductible à l'IPTG	F1 ori/ moyen nombre de copies	Expression des protéines *in vivo*	Dykxhoorn, Pierre et Linn, 1996
pEXT20	Amp	pTac – inductible à l'IPTG	ColE1 ori / haut nombre de copies	Expression des protéines *in vivo*	Dykxhoorn, Pierre et Linn, 1996
pZE23	Kan (2)	pLac - inductible à l'IPTG	ColE1 ori (E) / haut nombre de copies	Expression des protéines *in vivo*	Expressys
pZA23	Kan (2)	pLac - inductible à l'IPTG	p15A ori (A) / moyen nombre de copies	Expression des protéines *in vivo*	Expressys
pZA33	Chm (3)	pLac - inductible à l'IPTG	p15A ori (A) / moyen nombre de copies	Expression des protéines *in vivo*	Expressys
pZS23	Kan (2)	pLac - inductible à l'IPTG	pSC101ori (S) / moyen nombre de copies	Expression des protéines *in vivo*	Expressys

La liste complète des vecteurs dérivés est présentée dans le **Tableau 2.2**

1.1. Plasmides utilisés pour les études *in vitro* dérivés de pET28

Le vecteur d'expression pET28a, vecteur classiquement utilisé pour l'expression et la purification de protéines a été choisi pour la production de protéines étiquetées poly-histidines.

A) Vecteurs pET28-*phosphatase*

Les gènes codants pour les phosphatases candidates AraL, PHO13, YbiV, YniC, Dog1 et TM1254 ont été clonés dans le vecteur d'expression pET28a. Les gènes de phosphatases ont été amplifiés par PCR, à partir d'ADN génomique, avec les amorces correspondantes aux extrémités des séquences codantes des phosphatases additionnées des sites de restrictions HindIII et BamHI. Les fragments de PCR ont ensuite été sous-clonés dans le plasmide pGEM-T easy où ils ont été extraits par double digestion HindIII/BamHI, afin d'être cloné dans le vecteur pET28a.

B) Vecteurs pET28-*phosphoglucomutase*

Les vecteurs pET28-PGM1, pET28-PGM2, pET28-PGM3 avait été construits dans le cadre du travail de Walther et al. (*2012*) Ces vecteurs ont été utilisés dans cette étude.

C) Vecteurs pET28 pour la production des enzymes de la voie KdsD-Fsa-AldA et du métabolisme carboné central

Pour construire un système *in vitro*, ORFs (pour « Open Reading Frames ») des arabinose isomérase (AraA), ribulose kinase (AraB), L-ribulose-5-phosphate 4-epimérase (AraD), ribulose-5-phosphate 3-epimérase (RPE), D-arabinose-5-phosphate isomérase (KdsD), fructose-6-phosphate aldolase (FSA) et aldéhyde déshydrogénase (AldA) d'*E. coli* K12 MG1655 ont été amplifiées. Les plasmides pour l'expression des protéines étiquetées polyhistidine en C-terminal XylA, XylB, AraA, AraB, SgbE, AraB, Rpe, KdsD, FSA, AldA ont été construits par HiFi assembly® (NEB) avec pET28a comme vecteur receveur. pET28a a été préalablement linéarisé avec les enzymes de restriction HindIII et BamHI (NEB)

1.2. Plasmides utilisés pour l'étude de surexpression génique *in vivo* dans E.coli

A) Surexpression de phosphatases et mutases par des vecteurs dérivés de pACT3

Le Kit In-Fusion HD (clontech) a été utilisé pour le clonage des phosphatases dans le vecteur d'expression pACT3. Le vecteur pACT3 a été linéarisé par les enzymes de restriction SalI/BamHI. Les phosphatases et mutases ont été amplifiées par PCR avec des amorces dont les queues flottantes contiennent chacune une région d'homologie avec une extrémité du vecteur pACT3 linéarisé.

B) Surexpression des enzymes de la voie KdsD-Fsa-AldA

Pour la surexpression d'enzymes dans *Escherichia coli* la série de vecteurs pZ (Expressys) a l'avantage d'être modulable : il est facile de changer l'origine de réplication, le marqueur de résistance et le promoteur des vecteurs par restriction/ligation. Les vecteurs pZA23, pZA33, pZE23 et pZS23 possèdent le promoteur *PA1lac0-1* qui est un promoteur dérivé du promoteur de l'opéron lactose inductible à l'IPTG.Des modifications ont été apportées au vecteur pZA33, le promoteur *PA1lac0-1* a été remplacé par le promoteur constitutif *proD* et par le promoteur inductible *Ptac*, générant les vecteurs pZA36 et pZA38 respectivement. Le promoteur *PA1lac0-1* de pZS23 a été modifié pour le promoteur *proD* générant les vecteurs pZS27. Les gènes de la voie synthétique de production de l'acide glycolique *kdsD*, *fsa* et *aldA* ont été amplifiés par PCR à partir de l'ADN génome extrait d'*E. coli K12 MG1655* et insérés par HiFi assembly® dans les vecteurs pZ, préalablement linéarisés par PCR avec des amorces s'hybridant de part et d'autre du MCS. Tous les plasmides ont été vérifiés par séquençage avant utilisation.

Tableau 2.2 : Plasmides utilisés dans les chapitres III et IV

	Vecteur	Caractéristiques	Source
Vecteur vides	**pET28**	*Ori colE1, pT7, Amp*^{R}	Promega ®
	pACT3	*ori* p15A, Ptac, Cm^R	Dykxhoorn, Pierre et Linn, 1996)
	pEXT20	*ori* colE1, Ptac, Amp^R	Dykxhoorn, Pierre et Linn, 1996
	pZA23	*ori* p15A, $P_{A1lacO-1}$, Kan^R	Expressys
	pZA36	*ori* p15A, Ptac, Cm^R	Cette étude
	pZA37	*ori* p15A, proC, Cm^R	Cette étude
	pZA38	*ori* p15A, proD, Cm^R	Cette étude
	pZE27	*ori* colE1, proC, Kan^R	Cette étude
	pZE28	*ori* colE1, proD, Kan^R	Cette étude
	pZS23	*ori* pSC101, $P_{A1lacO-1}$, Kan^R	Expressys
	pZS27	*ori* pSC101, proC, Kan^R	Cette étude
	pZS28	*ori* pSC101, proD, Kan^R	Cette étude

	Vecteur	Caractéristiques	Source
Vecteurs pour la production de phosphatases et mutases	**pVT-ARAL**	pET28 contenant *araL* de *B.subtilis*	Cette étude
	pVT-TM254	pET28 contenant *TM1254* de *T.maritima*	Cette étude
	pVT-YBIV	pET28 contenant YbiV de *E.coli*	Cette étude
	pVT-YNIC	pET28 contenant YniC de *E.coli*	Cette étude
	pVT-PHO13	pET28 contenant Pho13 de *S.cerevisae*	Cette étude
	pVT-DOG1	pET28 contenant Dog1 de *S.cerevisae*	Cette étude
	pVT-YKR070W	pET28 contenant YKR070W de *S.cerevisae*	Cette étude
	pVT-PGM1	pET28 contenant PGM1 de *S.cerevisae*	Walther *et al.*, 2012
	pVT-PGM2	pET28 contenant PGM2 de *S.cerevisae*	Walther *et al.*, 2012
	pVT-PGM3	pET28 contenant PGM3 de *S.cerevisae*	Walther *et al.*, 2012
Vecteurs pour l'expression de phosphatases et mutases dans *Escherichia coli*	**pX1P**	pEXT20 portant les genes humains *khk-c* et *aldoB*,et *aldA* de *E. coli*	Alkim *et al.*, 2016
	pVV1-ARAL	pACT3 contenant *araL* de *B.subtilis*	Cette étude
	pVV1-TM1254	pACT3 contenant *TM1254* de *T.maritima*	Cette étude
	pVV1-YBIV	pACT3 contenant YbiV de *E.coli*	Cette étude
	pVV1-YNIC	pACT3 contenant YniC de *E.coli*	Cette étude
	pVV1-PHO13	pACT3 contenant Pho13 de *E.coli*	Cette étude
	pVV1-DOG1	pACT3 contenant Dog1 de *E.coli*	Cette étude
	PVV1-YKR070W	pACT3 contenant YKR070W de *S.cerevisae*	Cette étude
	pVV1-PGM1	pACT3 contenant PGM1 de *S.cerevisae*	Cette étude
	pVV1-PGM2	pACT3 contenant PGM2 de *S.cerevisae*	Cette étude
	pVV1-PGM3	pACT3 contenant PGM3 de *S.cerevisae*	Cette étude
Vecteurs pour la production d'enzyme de la voie cyclique (KdsD-Fsa-AldA)	**pVT-FSA**	pET28 portant *fsa* d'*E.coli*	Cette étude
	pVT-KDSD	pET28 portant *kdsD* d'*E.coli*	Cette étude
	pVT-ALDA	pET28 portant *aldA* d'*E.coli*	Cette étude
	pVT-RPE	pET28 portant *rpe* d'*E.coli*	Cette étude
	pVT-XYLA	pET28 portant *xylA* d'*E.coli*	Cette étude
	pVT-XYLB	pET28 portant *xylB* d'*E.coli*	Cette étude
	pVT-ARAA	pET28 portant *araA* d'*E.coli*	Cette étude
	pVT-ARAB	pET28 portant *araB* d'*E.coli*	Cette étude
	pVT-ARAD	pET28 portant *araD* d'*E.coli*	Cette étude

	Vecteur	Caractéristiques	Source
L'expression des gènes kdsD,fsa, aldA	**pKFA**	pEXT20 portant *kdsD, fsa* and *aldA* genes d'*E.coli*	Cette étude
	pKF1	pEXT20 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pKF2	pZA23 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pKF3	pZA36 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pKF4	pZS23 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pKF5	pZA37 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pKF6	pZA38 portant *kdsD* and *fsa* d'*E.coli*	Cette étude
	pK	pZA23 portant *kdsD* d'*E.coli*	Cette étude
	pF	pZA23 portant *fsa* g d'*E.coli*	Cette étude
	pA1	pACT3 portant *aldA* d'*E.coli*	Cette étude
	pA2	pEXT20 portant *aldA* d'*E.coli*	Cette étude
	pA3	pZA23 portant *aldA* d'*E.coli*	Cette étude
	pA4	pZS23 portant *aldA* d'*E.coli*	Cette étude
	pA5	pZE27 portant *aldA* d'*E.coli*	Cette étude
	pA6	pZE28 portant *aldA* d'*E.coli*	Cette étude
	pA7	pZS27 portant *aldA* d'*E.coli*	Cette étude
	pA8	pZS28 portant *aldA* d'*E.coli*	Cette étude

2. Construction de souches d'*Escherichia coli*

La souche DH5α compétente (NEB) a été utilisée pour le clonage et l'amplification de plasmides tandis que la souche BL21(DE3) compétente (Invitrogen) a été utilisée pour la production de protéines. Escherichia MG1655 (ATCC 47076) et W3CG *gapA* dérivée de W3110-a été utilisée comme souche parentale pour les souches construites dans cette étude. Les souches construites sont listées dans les **Tableau 2.3** et **Tableau 2.4.**

Tableau 2.3 : Souches d'Escherichia coli utilisées dans le chapitre III

Strain name	Genotype	Reference
MG1655	*F− λ−ilvG-rfb-50 rph-1*	ATCC 47076
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	Invitrogen
NEB5-alpha	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NEB
JW3536-2	*F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514 ΔxylB747 ::kan*	Baba *et al.*, 2006
JW5478	*F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514 ΔtktA ::kan*	Baba *et al.*, 2006
JW2449	*F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514 ΔtktB ::kan*	Baba *et al.*, 2006
JW2946	*F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514 ΔglcD ::kan*	Baba *et al.*, 2006
ProdARAL	NEB5-alpha contenant pVT-ARAL	Cette étude
ProdTM1254	BL21 (DE3) contenant pVT-TM1254	Cette étude
ProdYBIV	BL21 (DE3) contenant pVT-YBIV	Cette étude
ProdYNIC	BL21 (DE3) contenant pVT-YNIC	Cette étude
ProdPHO13	BL21 (DE3) contenant pVT-PHO13	Cette étude
ProdDOG1	BL21 (DE3) contenant pVT-DOG1	Cette étude
ProdYKR070W	BL21 (DE3) contenant pVT-YKR070W	Cette étude
ProdPGM1	BL21 (DE3) contenant pVT-PGM1	Cette étude
ProdPGM2	BL21 (DE3) contenant pVT-PGM2	Cette étude
ProdPGM3	BL21 (DE3) contenant pVT-PGM3	Cette étude
Screen ΔxylB	MG1655 *ΔxylB ΔtktA ΔtktB ΔglcD*	Cette étude
ScreenX1P	ScreenΔxylB contenant pX1P	Cette étude
ScreenARAL	ScreenX1P contenant pVV1-ARAL	Cette étude
ScreenTM1254	ScreenX1P contenant pVV1-TM1254	Cette étude
ScreenYBIV	ScreenX1P contenant pVV1-YBIV	Cette étude
ScreenYNIC	ScreenX1P contenant pVV1-YNIC	Cette étude
ScreenPHO13	ScreenX1P contenant pVV1-PHO13	Cette étude
ScreenDOG1	ScreenX1P contenant pVV1-DOG1	Cette étude
ScreenYKR070W	ScreenX1P contenant pVV1-YKR070W	Cette étude
ScreenPGM1	ScreenX1P contenant pVV1-PGM1	Cette étude
ScreenPGM2	ScreenX1P contenant pVV1-PGM2	Cette étude
ScreenPGM3	ScreenX1P contenant pVV1-PGM3	Cette étude

Tableau 2.4 : Souches d'Escherichia coli utilisées dans le chapitre IV

Souche	Genotype	Reference
MG1655	*F− λ−ilvG-rfb-50 rph-1*	ATCC 47076
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	Invitrogen
prodFSA	BL21 contenant pVT-FSA	Cette étude
prodKDSD	BL21 contenant pVT-KDSD	Cette étude
prodALDA	BL21 contenant pVT-ALDA	Cette étude
prodRPE	BL21 contenant pVT-RPE	Cette étude
prodXYLA	BL21 contenant pVT-XYLA	Cette étude
prodXYLB	BL21 contenant pVT-XYLB	Cette étude
prodARA	BL21 contenant pVT-ARAA	Cette étude
prodARAB	BL21 contenant pVT-ARAB	Cette étude
prodARAC	BL21 contenant pVT-ARAC	Cette étude
Screen	MG1655 ΔtktA ΔtktB ΔglcD	Cette étude
Screen00	Screen contenant pZA36	Cette étude
Screen01	Screen contenant pKFA	Cette étude
Screen02	Screen contenant pKF1	Cette étude
Screen03	Screen02 contenant pA1	Cette étude
Screen04	Screen02 contenant pA4	Cette étude
Screen05	Screen02 contenant pA3	Cette étude
Screen06	Screen contenant pKF2	Cette étude
Screen07	Screen06 contenant pA2	Cette étude
Screen08	Screen contenant pKF3	Cette étude
Screen09	Screen08 contenant pA4	Cette étude
Screen10	Screen08 contenant pA2	Cette étude
Screen11	Screen containing pKF4	Cette étude
Screen12	Screen11 contenant pA1	Cette étude
Screen13	Screen11 contenant pA2	Cette étude
Screen14	Screen contenant pZA27	Cette étude
Screen15	Screen contenant pKF5	Cette étude
Screen16	Screen15 contenant pA5	Cette étude
Screen17	Screen15 contenant pA6	Cette étude
Screen18	Screen15 contenant pA7	Cette étude
Screen19	Screen15 contenant pA8	Cette étude
Screen20	Screen contenant pFK6	Cette étude

Tableau 2.4 (suite 2) : Souches d'Escherichia coli utilisées dans le chapitre IV

Souche	Genotype	Reference
Screen21	Screen20 contenant pA5	Cette étude
Screen22	Screen20 contenant pA6	Cette étude
Screen23	Screen20 contenant pA7	Cette étude
Screen24	Screen20 contenant pA8	Cette étude
Cont01	Screen contenant pK	Cette étude
Cont02	Screen contenant pF	Cette étude
Cont03	Screen contenant pA2	Cette étude
Cont04	Con01 contenant pA2	Cette étude
Cont05	Con02 contenant pA2	Cette étude
W3CG gapA-	*F-, λ-, gapA 10 ::Tn10, IN(rrnD-rrnE), rph-1*	Ganter and Pluckthun, 1990
GA00	W3CG gapA- ΔglcD ΔarcA ΔmgsA ΔfucA ΔpfkA $galP^{proD}$	Cette étude
GA09	GA00 contenant pKF3 and pA4	Cette étude
GA23	GA00 contenant pKF6 and pA7	Cette étude

2.1. Délétions des gènes d'*E. coli*

Les délétions ont été réalisées par transduction, en utilisant le phage P1vir. La préparation des lysats P1vir et les procédures de transduction ont été réalisées comme décrit par Bremer *et al.* (1984) avec peu de modifications.

Ainsi, les souches KEIO portant une délétion unique et une cassette de résistance à l'antibiotique Kanamycine (souche donneuse) ont été infectées avec P1vir et des lysats λ de titre élevé ont été obtenus (Baba et al., 2006). Les souches donneuses (KEIO) ont été cultivées sur une nuit dans du LB à 37°C. Le lendemain, 5 ml de LB contenant 0,2% de glucose et 5 mM de $CaCl_2$ ont été inoculés avec 200 µl de la souche donneuse et cultivés pendant 30 min à 37°C. Ensuite, 100 µl de lysat de P1vir (~ 5 x 10^8 phages/ml) ont été ajoutés à chaque culture de donneur et incubés de nouveau à 37°C pendant 2 à 3 heures jusqu'à ce que la culture soit claire et les cellules soient complètement lysées. Les lysats ont été filtrés en utilisant des filtres à seringue stériles de 25 mm avec une membrane de 0,2 µm et conservés à 4°C.

La souche à déléter (souche réceptrice) a été infectée avec P1vir contenant une cassette de délétion de gène donneur avec une résistance à la kanamycine. Pour cela, la souche réceptrice a été cultivée dans 5 ml de milieu LB à 37°C. Les cellules ont été centrifugées à 1 500 g pendant 10 min et remises en suspension dans 1,5 ml de $MgSO_4$ à 10 mM et de $CaCl_2$ à 5 mM. Du lysat de la souche donneuse (0,1 ml) est ajouté à la suspension cellulaire qui est incubée pendant 30 min. Ensuite, 0,1 ml de citrate

de sodium à 1 M est ajouté au mélange cellules et P1vir. Ensuite, 1 mL de LB est ajouté à la suspension homogénéisée avant une incubation de 1h à 37°C, 200 rpm. Les suspensions cellulaires sont étalées sur un milieu LB solide avec l'antibiotique approprié puis les colonies sont criblées par PCR pour mettre en évidence des événements de transduction réussis. Pour éliminer la cassette à l'antibiotique, les cellules ont été transformées avec un plasmide pCP20 portant la FLP recombinase. Chaque étape a été vérifiée par PCR. Lorsque la souche délétée est sensible à la kanamycine après le retrait de la cassette, elle peut de nouveau être utilisée comme souche réceptrice afin d'ajouter une nouvelle délétion à partir d'un nouveau lysat de phage.

2.2. Préparation de bactéries compétentes et transformation

Les souches non commerciales compétentes sont préparées selon le protocole de Chung et al, 1989 légèrement modifié (Chung, Niemela et Miller, 1989). Une pré-culture est réalisée en LB sur la nuit pour inoculer le lendemain une culture fraîche de LB à une DO_{600} de 0,1. Lorsque la DO_{600} atteint 0,3-0,5 (les bactéries peuvent être rendues compétentes jusqu'à une DO_{600} de 1), une quantité de culture cellulaire équivalente à une unite DO_{600} est prélevée et centrifugée (8000 rpm, 2 min). Le surnageant est retiré tandis que le culot est repris dans 300 µl de tampon TSS (2,5 $\%_{(wt/vol)}$ PEG 3350, 1 M MgCl2, 5% $_{(vol/vol)}$ DMSO. Le mélange est incubé pendant 10 min sur glace. Le plasmide peut alors être ajouté aux cellules compétentes. Après 30 min supplémentaires d'incubation sur glace, un choc thermique est réalisé à 42°C pendant 90 secondes. Les cellules transformées sont mises dans la glace 10 min. 400 µl de LB sont ajoutés et la culture est incubée à 200 rpm pendant 1h, à une température adaptée (la température ne peut excéder 30°C dans le cas d'une transformation avec un plasmide dont l'origine de réplication est thermosensible). La culture bactérienne est centrifugée à 8000 rpm pendant 2,5 min. 600 µl de surnageant sont retirés, le volume restant est inoculé sur une boîte de LB solide avec l'antibiotique approprié.

3. Milieux et conditions de culture

3.1. Compositions des milieux

Les cellules sont cultivées sur le milieu LB pour les étapes de biologie moléculaire (clonage, délétions, transformation). Ce milieu riche est composé de 10 g/L de tryptone, 5 g/L d'extraits de levure et 5 g/L de NaCl. 15 g/L d'agar sont ajoutés pour l'obtention d'un milieu solide. Le LB, avec ou sans agar, est stérilisé par autoclavage 20 min à 110°C avant utilisation.

Les cultures pour la production d'acide glycolique sont réalisés en milieu minéral M9 (Tableau 2.5) contenant une source de carbone (glucose, xylose ou arabinose) à une concentration de 10 g/L ou 20 g/L et des traces de LB afin de réduire la phase de latence (2 g/L tryptone, 1 g/L d'extraits de levure et 1 g/L de NaCl).

Tableau 2.5 : Composition du milieu minéral M9

Composé	Concentration finale (g/L)
Na2HPO4 * 12 H2O	18
KH2PO4	3
NaCl	0,5
NH4Cl	2
MgSO4 * 7 H2O	0,5
CaCl2 * 2 H2O	0,015
FeCl3	0,010
Thiamine HCl	6.10-3
NaEDTA	0,49.10-3
CoCl2 * 6 H2O	1,8.10-3
ZnCl2SO4 * 7 H2O	1,8.10-3
Na2MoO4 * 2 H2O	0,4 .10-3
H3BO3	0,1.10-3
MnSO4 * H2O	1,2.10-3
CuCl2 * 2 H2O	1,2.10-3

Les cultures des souches MG1655 ΔxylB ΔtktA ΔtktB ΔglcD pour le criblage des D-xylulose 5P phosphatases/mutases candidates *in vivo* sont réalisées en milieu minéral M9 contenant une source de carbone (xylose ou arabinose ou glucose) à une concentration de 10 g/L et 10 à 20% (v/v) de LB afin de réduire la phase de latence (2 g/L tryptone, 1 g/L d'extraits de levure et 1 g/L de NaCl). Les souches dépourvues de transcétolase (TktA, TktB) ne peuvent pas synthétiser l'érythrose 4P qui est un précurseur essentiel de vitamines telle que la pyridoxine et d'acides aminés aromatiques, le milieu est complémenté avec 500 µM L-phénylalanine, 250 µM L-tyrosine, 200 µM L-tryptophane, 6 µM p-aminobenzoate, 6 µM p-hydroxydenzoate et 280 µM de shikimate pour permettre la croissance. Le milieu est tamponné par l'addition de 20 g/L d'acide 3-(N-morpholino)propanesulfonique (MOPS) à pH 7 puis filtré sur des membranes à 0,2 µm pour obtenir un milieu stérile. Si besoin, les antibiotiques appropriés ont été ajoutés au milieu (ampicilline 100 µg/mL, kanamycine 50 µg/mL). Pour les souches

contenant un vecteur inductible à l'IPTG, une concentration finale de 0,1 mM d'IPTG est utilisée. Tous les produits ont été commandés chez Sigma. Les cultures sont placées dans un agitateur (Infors) à 200 rpm, 37°C le temps de l'expérience. La croissance des cultures est suivie par la mesure de la densité optique à 600 nm (DO_{600}) au spectrophotomètre (Biochrom Libra S11).

3.2. Conditions de culture

A) <u>En erlenmeyer</u>

Souches contenant des promoteurs inductibles - Les souches sont reprises à partir d'un stock glycérol conservé à -80°C dans 10 mL de LB, à 37°C sur la nuit. Les précultures en LB des souches contenant un vecteur avec promoteur inductible sont mises en culture à une DO_{600} initiale de 0,1 dans 30 mL de M9 glucose 1%. Lorsque la DO_{600} atteint 0,6, 0,1 mM d'IPTG sont ajoutés pour induire l'expression des gènes sous le contrôle du promoteur inductible à l'IPTG (ptac ou plac). 24 h après, les cellules sont centrifugées, un lavage à l'eau stérile permet d'enlever les traces de glucose, et resuspendues dans 20 mL du milieu choisi pour l'étude (M9 xylose ou arabinose 1%, 0,1 mM IPTG) pour une phase d'adaptation de 24h permet aux souches de s'adapter à l'utilisation de pentose. Les cellules sont ensuite centrifugées (4000 rpm, 5 min) et reprises dans les conditions de culture finales : à une DO_{600} initiale de 0,5 dans 50 mL du milieu de composition identique à celui utilisé pendant la phase d'adaptation (M9 xylose ou arabinose 1%), 0,1 mM IPTG). Des erlenmeyer de culture à baffles de 250 mL sont utilisés pour une oxygénation optimale. Les cultures sont agitées à 200 rpm, à 37°C.

Souches contenant des promoteurs constitutifs - Les souches sont reprises à partir d'un stock glycérol conservé à -80°C dans 10 mL de LB, à 37°C sur la nuit. Les pré-cultures sont centrifugées le lendemain à 4000 rpm pendant 5 min et resuspendues dans 20 mL de milieu M9 xylose ou arabinose 1% en flacons d'Erlenmeyer de 100 mL. Une phase d'adaptation d'une durée de 24h permet aux souches de s'adapter à l'utilisation de pentose. Les cellules sont ensuite centrifugées (4000 rpm, 5 min) et reprises dans les conditions de culture finales : à une DO_{600} initiale de 0,5 dans 50 mL du milieu de composition identique à celui utilisé pendant la phase d'adaptation (M9 xylose ou arabinose 1%). Des flacons de culture à baffles de 250 mL sont utilisés pour une oxygénation optimale. Les cultures sont agitées à 200 rpm, à 37°C.

B) En bioréacteur BIOSTAT® B Startorius Stedim Biotech(2000L)

Une première préculture de 3mL en LB additionné de kanamycine, chloramphénicol et tétracycline de la souche GA23 a été réalisée en tube Falcon® 50 mL sur la nuit, celle-ci a servi à l'inoculation d'une deuxième préculture en fiole erlenmeyer 100mL contenant 15 mL de LB avec antibiotiques, mise en culture à 37°C, 200 rpm pendant 24h. Celle-ci a servi à l'inoculation d'une troisième et dernière préculture en erlenmeyer à baffles de 1000 mL contenant 150 mL de LB et antibiotiques mise à 37°C, 150 rpm pendant 8h. En phase exponentielle la préculture 3 a servi à inoculer la culture en fermenteur à une DO de 0,3. Le milieu de culture M9 (**Tableau 2.5**) a été complété par une source de carbone type hexose (glucose) ou pentose (xylose) à une concentration de 10 g/l ainsi que de 0,4 g/L d'acide malique préalablement ajusté à pH 7 avec KOH afin de complémenter la délétion de la glycéraldéhyde 3-phosphate déshydrogénase et 10% (v/v) de LB afin de réduire la phase de latence (2 g/L tryptone, 1 g/L d'extraits de levure et 1 g/L de NaCl). Les cultures ont été réalisées à 37°C, avec une agitation (300-1500 rpm) et une aération maintenant l'oxygène dissous au-dessus de 20% du débit d'air. Le pH a été maintenu à 7 avec une solution de base KOH. La culture a été conduite en mode batch pendant 90h.

4. Extraction et quantification des métabolites

4.1 Métabolites extracellulaires

La consommation de sucre et la production d'acide glycolique sont suivies en prélevant régulièrement des échantillons en conditions stériles qui sont centrifugés à 13,000 rpm pendant 5 min dans une centrifugeuse de paillasse (Eppendorf 5415D), filtrés à travers un filtre de 0,2 µm et stockés à -20°C. La quantification des métabolites extracellulaires est réalisée par chromatographie liquide haute performance (CLHP) avec un chromatographe Ultimate 3000 (Dionex, Sunnyvale, USA). Le système de CLHP est équipé d'une colonne à échange de cations (Aminex, HPX87H – 300 x 7,8 mm, 9 µm, BioRad), d'un injecteur automatique (WPS-3000RS, Dionex), d'un détecteur IR (RID 10A, Shimadzu) et d'un détecteur UV (SPD-20A, Shimadzu). La phase mobile est une solution de d'acide sulfurique à 1,25 mM à un débit de 0,5 mL/min. Les échantillons sont maintenus à 4°C et l'injection de 20 µl se fait dans la colonne à 35°C.

4.2 Métabolites intracellulaires

Les méthodes d'extraction et de quantifications utilisées sont celles décrites par Walther *et al.*(2012).5 mL de culture ont été filtré sur une membrane polyamide (0,2 µm, Sartorius), les cellules sont filtrées en utilisant 10 ml d'eau, la membrane est trempée dans de l'éthanol 80% à 75°C. Le traitement des échantillons a été effectué comme décrit plus tôt. Les métabolites ont été quantifiés par chromatographie en phase liquide couplée à la spectrométrie de masse (LC-MS) a été réalisée sur un système ICS-3000 de Dionex (Sunnyvale, USA) équipé d'un système de générateur d'éluant automatique (KOH) (RFIC, Dionex) et d'un échantillonneur automatique (AS50, Dionex), où les

échantillons sont à 4°C. Les analytes ont été séparés sur une colonne IonPac AS11 (250 x 2 mm, Dionex) colonne protégée par une pré-colonne AG11 (50 x 2 mm, Dionex). La température de la colonne a été maintenue à 25°C, avec un flux fixé à 0,25 ml/min, les analytes ont été élués en appliquant un gradient de KOH (. Le volume d'échantillon injecté est de 15 µl. Pour éliminer le bruit, le système ASR ultra II (2 mm, 75 mA) a été utilisé. Les analytes ont été quantifiés en utilisant un détecteur à barrettes de diodes (Ultimate 3000, Dionex), un capteur de conductivité (partie de ICS-3000, Dionex) and un détecteur de masse (MSQ Plus, Thermo) en mode ESI (division de 1/3, pression en nitrogène à 90 psi, la tension capillaire de 3,5 kV et température de sonde de 450 °C).

5. Expression et purification des protéines étiquetées polyhistidine.

Toutes les protéines exprimées sont solubles et ont été produites à partir du vecteur d'expression pET28a transformé dans une souche d'*E. coli BL21 (DE3)*. Une pré-culture en milieu LB (pour « Luria-Bertani ») additionnée de l'antibiotique kanamycine est réalisée sur la nuit à 37°C. La pré-culture est utilisée pour ensemencée une culture fraîche de 200 mL de LB-Kanamycine à une densité optique à 600 nm (DO_{600}) de 0,1 (37°C, 200 rpm). Lorsque la DO_{600} est comprise entre 0,6 et 0,8, l'expression de la protéine d'intérêt est induite par l'ajout d'IPTG à 1 mM final. Les protéines sont exprimées sur la nuit à 16°C. Les cellules sont collectées par fractions de 50 ml et centrifugées à 4800 rpm pendant 15 min à 4°C. Les culots cellulaires sont conservés à -20°C.

La purification des protéines est réalisée à partir des culots cellulaires obtenus lors de l'étape de production. Toutes les étapes sont faites à froid pour éviter la dégradation des protéines par des protéases. Les culots cellulaires sont remis en suspension dans 1,5 mL de tampon de lavage (50 mM HEPES, pH 7,5 ; 0,3 M NaCl) puis soniqués sur glace. Une étape de centrifugation à 13000 rpm, pendant 15 min à 4°C permet de séparer les débris cellulaires du liquide cytoplasmique. Les lysats clarifiés sont déposés sur 600 µl de résine cobalt (Clontech) préalablement équilibrée avec le tampon de lavage. Après 20 min à température ambiante au contact de la résine, les tubes sont centrifugés (700 rcf, 3 min, 4°C). Le surnageant est retiré, la résine est mise en contact avec 3 mL de tampon de lavage pendant 10 min afin d'éliminer les interactions non spécifiques. Après centrifugation (700 rcf, 3 min, 4°C) et élimination du surnageant, 3 mL d'une solution à 15 mM d'imidazole sont mis en contact avec la résine pendant 5 min. Le surnageant est séparé de la résine par centrifugation, remplacé par 500 µl d'imidazole à 200 mM. L'imidazole entraîne l'élution des protéines porteuses d'une étiquette polyhistidine. Pour favoriser la stabilité des protéines à leur pH optimal, le tampon a été modifié.

La méthode utilisée pour mesurer la concentration des protéines en solution est basée sur la méthode de Bradford. Le réactif Protein assay® commercialisé par BioRad est dilué au ¼, le mélange réactionnel comprend 160 µl de réactif et 40 µl d'éluat dilué (au 10ème et au 20ème). Une gamme standard est réalisé avec de la BSA de 12,5 à 100 µg/ml.

6. Essais enzymatiques.

6.1. Activité phosphatasique

A) Sur substrat artificiel pNPP

Les phosphatases hydrolysent des liaisons monoester-phosphorique sur un large panel de petits substrats phosphorylés. Il a été démontré que la majorité des phosphatases ont une activité significative sur un substrat artificiel chromogène de petite taille, le pNPP (p-nitrophenyl phophate). Même à des températures élevées (70 °C), le pNPP a un taux d'hydrolyse non-enzymatique très faible, ce qui rend possible de longues incubations (2-4 heures) afin de détecter une activité enzymatique pour les enzymes avec un turnover (Kcat) très faible. La plupart des phosphatases ont une forte affinité pour le pNPP (Km 0.005-5 mM) (BRENDA). De plus les phosphatases acides et alcalines ont une activité significative sur pNPP à pH neutre (20-80% de l'activité à pH optimal). Ainsi le pNPP est couramment utilisé pour mettre en évidence une activité « phosphatasique » lors de ciblages enzymatiques. Dans cette étude nous avons utilisés le pNPP comme contrôle positif. Une activité sur pNPP indique que l'enzyme purifiée est fonctionnelle. Le p-nitrophényl phosphate (pNPP) est un composé incolore en solution. L'hydrolyse du pNPP par une phosphatase libère un ion monophosphate et du para-nitrophénol (pNP), composé de couleur jaune qui absorbe à 405 nm. Le p-Nitrophényl Phosphate Liquid Substrate System (SIGMA) contient du pNPP à 3mM, du tampon TrisHCl, et des ions Mg2+. La réaction est réalisée dans des conditions physiologiques : à pH 7.5 à 37°C. Sur microplaques, 190µl de la solution de pNPP sont ajoutés à 10µl d'éluat. L'absorbance est mesurée à 405nm, à 10, 20 30 et 60 minutes.

B) Sur substrats biologiques

La réaction de déphosphorylation catalysée par les phosphatases génère du phosphate libre, quel que soit le substrat. Nous avons utilisé le kit colorimétrique de Sigma contenant du Vert de Malachite pour mesurer l'apparition de phosphate libre dans le milieu réactionnel (Baykov, Evtushenko et Avaeva, 1988). Le dosage du phosphate par la méthode au Vert de Malachite est la méthode la plus répandue pour caractériser des phosphatases sur des substrats biologiques. Son coefficient d'extinction molaire est de 90000 M-1.cm-1(Baykov, Evtushenko et Avaeva, 1988). Elle est 5 fois plus sensible que la méthode de caractérisation de pNPP (17514 M-1 cm-1).Pour maintenir le bruit de fond bas lors des tests enzymatiques et se rapprocher des conditions physiologiques, les tests d'activité phosphatasique sur substrat naturels ont été réalisés avec des concentrations en susbtrat de 0,5 mM qui sont non saturantes pour la plupart des HADs. Par conséquence les vitesses observées ne représentent pas les activités maximales de ces enzymes. L'activité phosphatase des HADs vis-à-vis de substrats naturels à pH neutres dépend de la présence de cations métalliques divalents (Mg2+, Mn2+, Zn2+). Les concentrations utilisées sont celles utilisées par E.kuznetsova pour le screening de HADs

phosphatases, le mélange réactionnel contient les éléments nécessaires à l'activité de l'ensemble des phosphatases (Kuznetsova et al., 2006).

Les dosages sont réalisés en plaque de 96 puits. 200 µl de réaction contenant 50 mM HEPES-K (pH 7,5), 0,5 mM de substrat, 5 mM MgCl2, 0,5 mM MnCl2, 0,5 mM NiCl2 et 1-10 µg de protéine est mis à incuber 30 min à 37°C. 30µl de réactif sont ajoutés à la réaction et la quantité de Pi libéré est calculé à partir de l'absorbance à 630 nm. Nous avons choisi pour notre étude des sucres phosphorylés qui sont présents dans le réseau métabolique qui nous intéresse, tels que le G6P, F6P, R5P, Ru5P, Xu5P,E4P et GAP.

6.2. Mesure d'activité mutasique

A) Sur D-glucose 1P, substrat de référence

Les phosphoglucomutases ont été testées dans un mix réactionnel contenant 60 mM HEPES, 60 mM KCl, 3 mM MgCl2, 0,5 mM NADP, 1U/ml de glucose 6P déshydrogénase et 5 µM de glucose 1,6-bis-phosphate. La réaction est initiée par l'ajout de 8 mM de glucose 1P. Les réactions sont faites dans des plaques 96 puits, l'apparition de NADPH est mesurée à partir de l'absorbance à 340 nm.

B) Sur D-xylulose- 5P, substrat d'étude

Les phosphoglucomutases ont été testées dans un mix réactionnel contenant 60 mM HEPES, 60 mM KCl, 3 mM MgCl2, 0,5 mM NADH, 1U/ml de fructose 1,6bP aldolase (aldoB), purifiée à partir de culot bactérien) et 5 µM de glucose 1,6-bis-phosphate. La réaction est initiée par l'ajout de 4 mM de D-xylulose 5P. Les réactions sont faites dans des plaques 96 puits, la consommation de NADH est mesurée à partir de l'absorbance à 340 nm.

6.3. Mesure d'activité des enzymes clefs de la voie KdsD-Fsa-AldA

Toutes les enzymes ont été testées dans 100 mM Tris, pH 7.5, 10 mM $MgCl_2$, à 37°C. Les co-facteurs et activateurs ont été ajoutés si nécessaire. Tous les substrats enzymatiques ont été achetés chez Sigma-Aldrich. Les tests enzymatiques sont couplés à une réaction redox avec pour co-facteur, le NADH qui absorbe l'ultraviolet avec un pic à 340 nm avec un coefficient d'extinction molaire de 6 220 $M^{-1}.cm^{-1}$. Le spectrophotomètre Epoch 2 de Bio Tek a été utilisé pour le suivi des réactions en UV. La **Figure 2.1** schématise les essais enzymatiques réalisés, décrits ci-dessous.

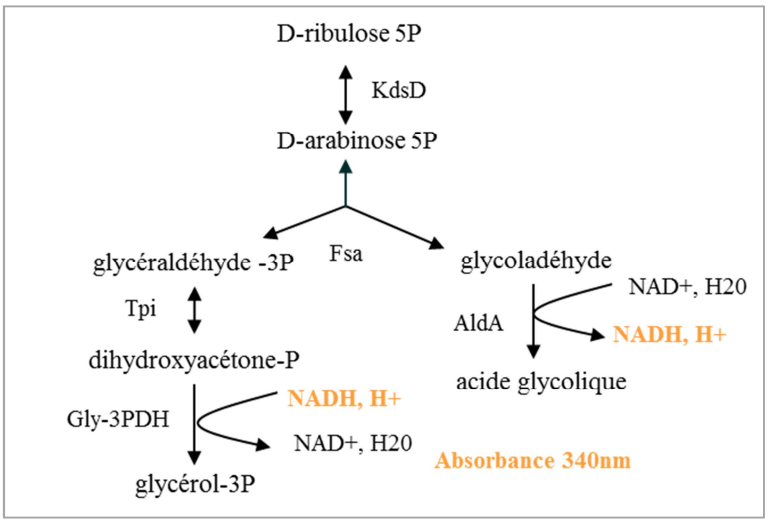


Figure 2.1 : Représentation schématique de l'essai enzymatique de vérification de l'activité de l'arabinose-5P isomérase (KdsD), la fructose-6P aldolase (F6P) et l'aldéhyde déshydrogénase (AldA).

A) Mesure d'activité de l'aldéhyde déshydrogénase (AldA)

Cette activité a été mesurée en présence de 5 mM de glycoaldéhyde, 3 mM de NAD^+ et 3 mM d'ATP. Pour une molécule de glycoaldéhyde oxydée en glycolate, une molécule de NAD^+ est réduite en NADH.

B) Mesure d'activité de la Fructose-6-phosphate aldolase (Fsa)

L'enzyme FSA clive le D-arabinose-5-phosphate en glycolaldéhyde et glycéraldéhyde-3-phosphate (GAP). La vitesse d'apparition de glycolaldéhyde a été mesurée en présence d'AldA avec 2,5 mM de NADH. Pour suivre l'apparition de GAP, la Tpi et la Gly-3PDH ont été ajoutées, en présence de 0,4 mM de NADH.

C) Mesure d'activité de la D-arabinose-5-phosphate isomérase (KdsD)

La D-arabinose-5-phosphate isomérase catalyse l'interconversion du D-ribulose-5-phosphate en D-arabinose-5-phosphate. L'activité de KdsD sur D-ribulose-5-phosphate a été déterminée en ajoutant FSA, AldA en excès et 3 mM de NAD^+.

Remarque : Les tests enzymatiques sur extraits bruts ont été réalisé à partir d'une quantité de cellules en phase exponentielle correspondant à 40 unités de DO_{600nm} (exemple :pour une culture avec une DO_{600nm} de 2, 20mL ont été prélevés). Les cellules sont récupérées dans un falcon 50 mL, centrifugées 10 min à 4000 rpm. Le culot est ensuite suspendu dans 2 mL du tampon Tris-HCl 100 mM à pH 7.5 et soniqué sur glace. Les essais enzymatiques sont ensuite réalisés comme décrit ci-dessus. La quantité de protéines totales mesurées par Bradford est utilisée pour le calcul de l'activité enzymatique spécifique ($\mu mol.min^{-1}.mg^{-1}$)

6.4. Conversion du D-Xylose, L-arabinose et D-glucose en acide glycolique *in vitro*

Tout comme dans la partie 6.3) les enzymes ont été testées dans 100 mM Tris, pH 7.5, 10 mM $MgCl_2$, à 37°C. Les co-facteurs et activateurs ont été ajoutés si nécessaire. Les enzymes Hexokinase de *S.cerevisiae* (HxK), phosphoglucose isomérase d'*E.coli* (Pgi) et transcétolase d'*E.coli* (Tkt) ont été commandé chez Sigma. Le reste des enzymes : xylulose isomérase (xylA), xylulokinase (XylB), Ribulose-5-phosphate-3-épiméras (Rpe), arabinose isomérase (AraA), ribulokinase (AraB), L-ribulose-5-phosphate -4-épimérase (AraD), D-arabinose-5P isomérase (KdsD), Fructose-6-phosphate aldolase (Fsa), aldéhyde déshydrogénase (AldA) d'*E.coli* MG1655 ont été produites et purifiées. Les voies de conversion *in vitro* du D-xylose, L-arabinose et D-glucose en acide glycolique sont présentées dans la **Figure 2.2**.

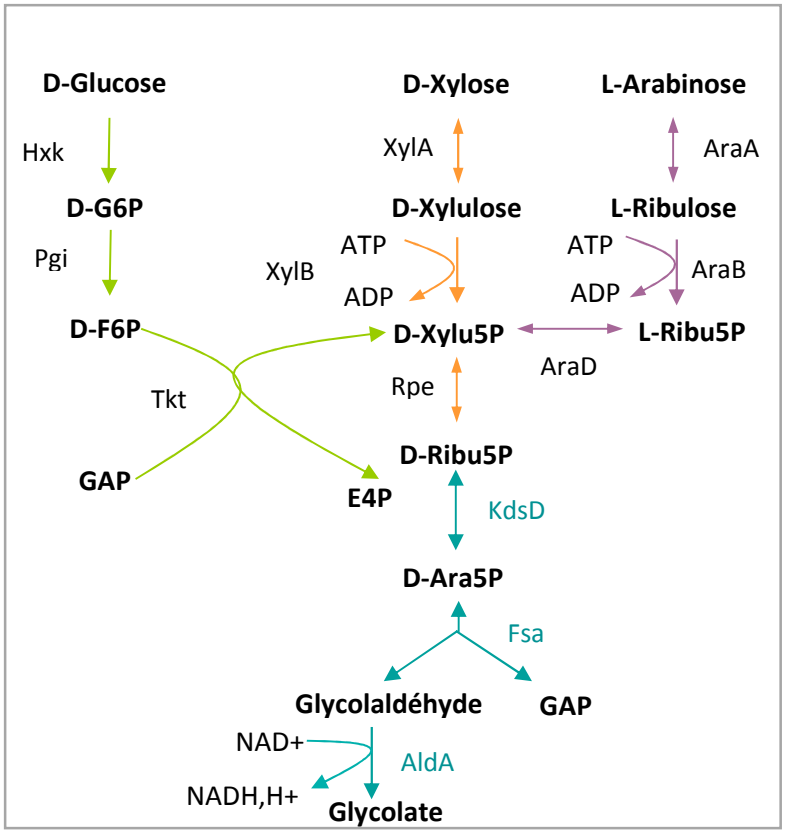


Figure 2.2 : Représentation des voies de conversion du D-glucose, D-xylose et L-arabinose en acide glycolique in vitro. En bleu la voie non-naturelle KdsD-Fsa-AldA. *En vert : voie d'assimilation du glucose, en orange : voie d'assimilation du xylose, en violet : voie d'assimilation du L-arabinose, en bleu-vert : voie non-naturelle de production d'acide glycolique. Hxk : Hexokinase de S.cerevisae (Sigma), Pgi : phosphoglucose isomérased'E.coli (sigma), Tkt : transcétolase d'E.coli (Sigma) xylA : xylulose isomérase, XylB : xylulokinase, Rpe : Ribulose-5-phosphate-3-épimérase, AraA : arabinose isomérase, AraB : ribulokinase, AraD : L-ribulose-5-phosphate -4-épimérase, KdsD : D-arabinose-5P isomérase, Fsa : Fructose-6-phosphate aldolase, AldA : aldéhyde déshydrogénase*

A) Conversion du D-xylose en AG *in vitro*

La conversion du D-xylose en acide glycolique requiert XylA, XylB, Rpe, KdsD, FSA et aldA (**Figure 3.1,** orange et bleu). La conversion du glycolaldéhyde, D-arabinose-5-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, D-xylulose et D-xylose en acide glycolique est démontrée dans 100 mM Tris HCl pH 7,5, 10 mM $MgCl_2$, 0,5 mM $MnCl_2$, 3 mM ATP et 2,5 mM NAD.

B) Conversion du L-arabinose en AG *in vitro*

La conversion du L-arabinose en acide glycolique requiert AraA, AraB, AraD, Rpe, KdsD, FSA et aldA (**Figure 3.1**, violet et bleu). La conversion du glycolaldéhyde, D-arabinose-5-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, L-ribulose et L-arabinose en acide glycolique est démontrée dans 100 mM Tris HCl, pH7,5, 10 mM $MgCl_2$, 0,5 mM $MnCl_2$, 3 mM ATP et 2,5 mM NAD.

C) Conversion du D-glucose en AG *in vitro*

La conversion du glycolaldéhyde, D-arabinose-5-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, D-fructose-6-phosphate, D-glycéraldéhyde-3-phosphate, D-glucose-6-phosphate et D-glucose en acide glycolique est vérifiée dans 100 mM Tris HCl, pH 7,5, 10 mM $MgCl_2$, 0,5 mM $MnCl_2$, 1 mM Thiamine Pyrophosphate, 3 mM ATP et 2,5 mM NAD (**Figure 3.1**, vert et bleu)..

6.5 Etude de la faisabilité thermodynamique

Les valeurs d'énergie libre de Gibbs ΔG proviennent de Metacyc. Les estimations du l'énergie libre de Gibbs des réactions et de formation des composés, $\Delta_r G'^{\circ}$ et $\Delta_f G'^{\circ}$ respectivement, ont été fait à pH 7,3 et une force ionique de 0,25. Les calculs des $\Delta_r G'^{\circ}$ et $\Delta_f G'^{\circ}$ se basent sur la technique développée par Robert Alberty (*Thermodynamics of Biochemical Reactions*. Wiley-InterScience, John Wiley & Sons, Hoboken, New Jersey, 2003)

La relation suivante a été utilisée pour calculer les énergies libres de Gibbs des réactions, à partir des énergies libres de Gibbs de formation ($\Delta_f G'^{\circ}$) des produits et des substrats :

$$\Delta_r G'^{\circ} = \sum(\Delta_f G'^{\circ}\text{Produit}) - (\Delta_f G'^{\circ}\ \text{Substrat})$$

CHAPITRE III

EVALUATION DE L'ACCESSIBILITE DE LA VOIE 'XYLULOSE-1P' AUX HEXOSES ET PENTOSES PAR EXPRESSION D'UNE PHOSPHATASE OU MUTASE POUR LA PRODUCTION D'ACIDE GLYCOLIQUE

1. INTRODUCTION

1.1 La voie synthétique de production d'acide glycolique 'X1P'

Dans un contexte de valorisation de la biomasse, l'équipe du Prof. Jean Marie François a mis au point un procédé de conversion de D-xylose, qui peut représenter jusqu'à 70% de sucres totale de la fraction hémicellulosique, en acide glycolique (AG) (**Figure 3.1**) (Alkim *et al.*, 2015, 2016; Cam *et al.*, 2016). Cependant l'utilisation de xylose pur n'est pas un scénario réaliste pour la production industrielle d'acide glycolique et l'hémicellulose n'est qu'une des trois fractions de la lignocellulose représentant 20 à 40% de la biomasse en poids. La fraction majoritaire est la cellulose, composée de molécules de glucose (Chen et Cong, 1985; Hespell *et al.*, 1997). Bien que le glucose soit le plus abondant hexose, et le xylose le principal pentose de la lignocellulose, la fraction d'arabinose présente dans les hydrolysats peut aussi être importante (Wenzl, 1970). Pour qu'un procédé de production soit économiquement viable, il est préférable que tous les sucres présents dans la lignocellulose puissent être assimilés et convertis en acide glycolique.

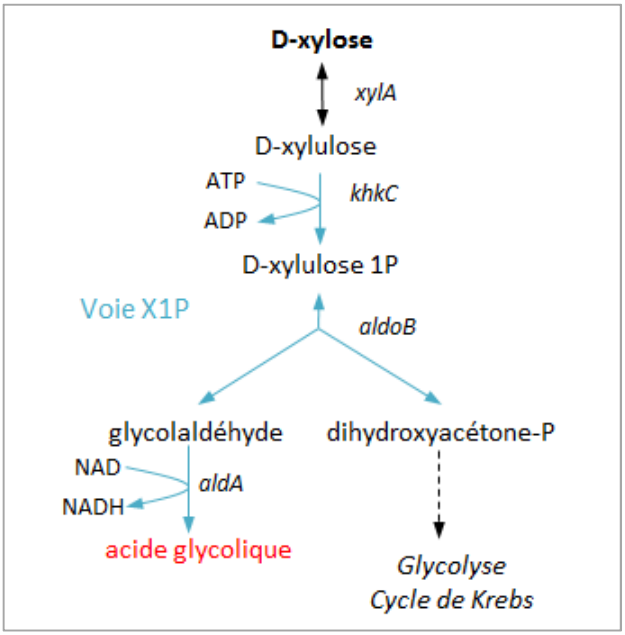


Figure 3.1. Voies naturelle et synthétique d'assimilation du D-xylose. *Les réactions de la voie synthétique X1P sont en bleu, les réactions du métabolisme naturel sont en noir et le produit désiré en rouge. xylA : xylose isomérase , khkC : xylulose-1-kinase, aldoB : (D)-xylulose-1P aldolase, aldA : glycolaldéhyde déshydrogénase.*

La première approche pour convertir le D-xylose et D-glucose en acide glycolique (AG) a été de coupler la voie 'Xylulose-1P' (X1P) avec le cycle glyoxylate ingénieré (Alkim *et al.*, 2016). Cette voie naturelle de production d'acide glycolique dérive du cycle de Krebs à partir de l'isocitrate, comme présenté dans le premier chapitre. Le résultat de cette combinaison a conduit à un rendement pratique de 0,63 g AG par g de sucres C6 (hexoes) et C5 (pentoses), soit 35% de plus que dans les travaux antérieurs utilisant la voie naturelle, et seulement 26 % de moins que le rendement théorique (Alkim *et al.*, 2016). Dans cette approche, le rendement théorique est limité par la stœchiométrie du métabolisme naturel. Une des étapes de synthèse de l'acide glycolique via le métabolisme du

glyoxylate est la décarboxylation du pyruvate, qui génère la perte de CO_2 et représente une perte de carbone pour la formation d'AG.

L'objectif de cette étude est de rendre la voie synthétique X1P accessible aux autres sucres de la lignocellulose en C5 et C6. Une fois cet objectif atteint, le métabolisme carboné central bactérien devra être remanié afin de s'affranchir de l'étape de décarboxylation du pyruvate et augmenter le rendement de la conversion de carbone à son maximum théorique.

1.2 L'assimilation des sucres lignocellulosiques dans *Escherichia coli*

Les bactéries assimilent naturellement les pentoses tels que le xylose et l'arabinose par la voie des pentoses phosphate via le D-xylulose 5P (Xu5P). Le glucose est un aldohexose, principalement catabolisé par la voie de la glycolyse mais le catabolisme du glucose-6P peut également emprunter la voie des pentoses phosphate en débutant par la formation de 6-phosphogluconate, converti en D-gluconate-6P (Gumaa et McLean, 1969). Celui-ci est converti en D-ribulose-5P puis D-Xu5P. Le D-Xu5P est un métabolite commun au catabolisme de l'arabinose, du xylose et du glucose chez *Escherichia coli* (**Figure 3.2**).

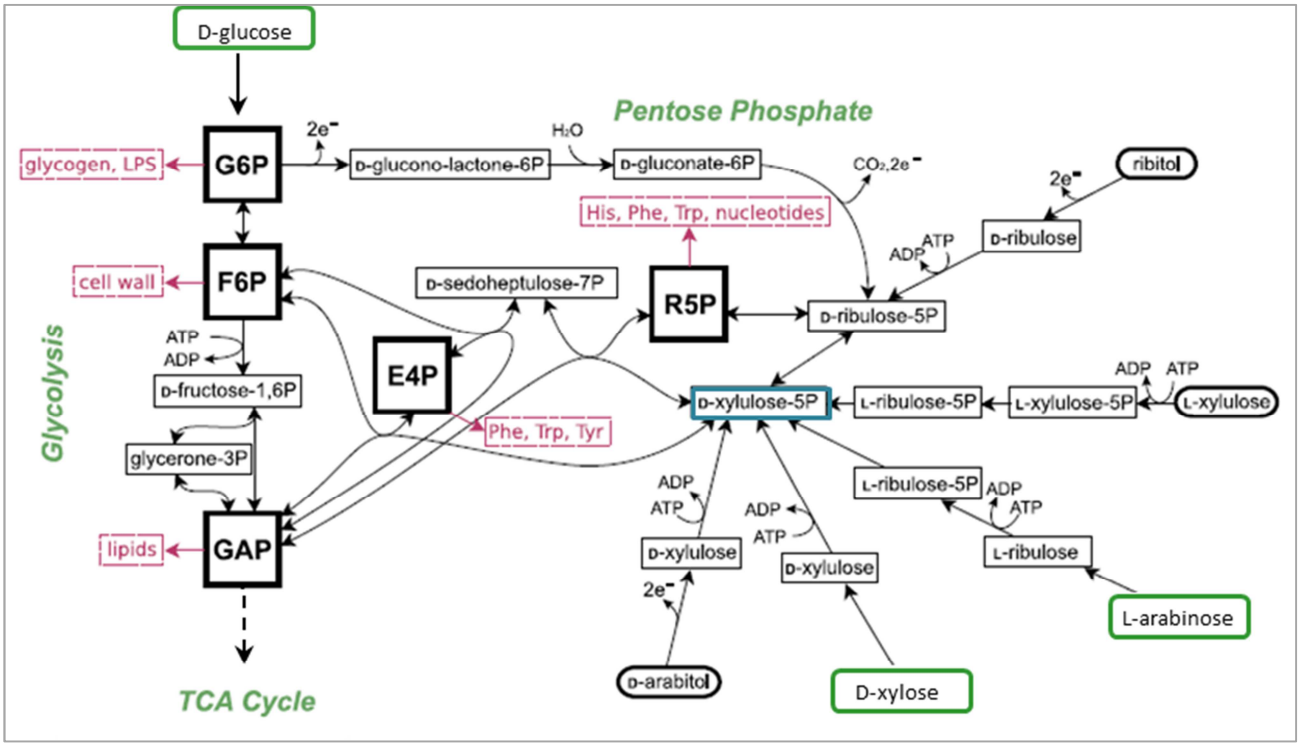


Figure 3.2 . Métabolisme carboné central d'Escherichia coli et assimilation des sucres lignocellulosiques principaux : D-glucose, D-xylose, L-arabinose. *Figure adaptée de Noor et al., 2014. En vert : le nom des principales voies du métabolisme carboné central – Encadré noir : quelques précurseurs métaboliques essentiels : D-glucose-6-phosphate (G6P), D-fructose-6-phosphate (F6P), D-ribose-5-phosphate (R5P), D-erythrose-4-phosphate (E4P), D-glyceraldehyde-3-phosphate (GAP)– En violet : les produits finaux des voies de biosynthèse en aval du métabolisme carboné central - Encadré vert : les sucres d'intérêt pour l'étude.*

1.3 L'intégration d'une D-xylulose-5P mutase ou phosphatase à la voie X1P

L'introduction d'une enzyme de type phosphatase ou mutase active sur D-xylulose 5P rendrait possible la conversion de ces sucres en molécules d'intérêt par la voie synthétique X1P comme indique la **Figure 3.3.**

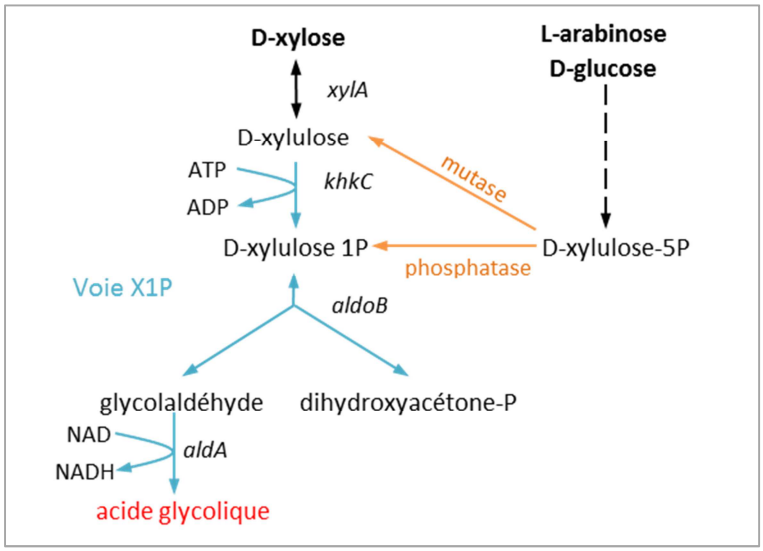


***Figure 3.3 : Voies naturelles et synthétiques d'assimilation des pentoses et hexoses.**. En noir : le métabolisme naturel, en bleu : la voie synthétique 'X1P', en orange : D-xylulose-5P phosphatase ou mutase putative. xylA : xylose isomérase , khkC : xylulose-1-kinase, aldoB : (D)-xylulose-1P aldolase, aldA : glycolaldéhyde déshydrogénase.*

La première étape a consisté à rechercher dans la littérature l'existence de phosphatases et/ou mutases qui pourraient être potentiellement actives sur le D-Xu5P. Bien que l'introduction d'une activité 'mutase' apparaisse à première vue la meilleure solution d'un point de vue énergétique, évitant la consommation d'un ATP, il n'y a ce jour qu'une seule phosphoglucomutase identifiée et qui requiert dans le mécanisme catalytique la présence d'un intermédiaire di-phosphorylé (le glucose $1,6P_2$ pour conversion du glucose-1P en glucose-6P) (Walther *et al.*, 2012). En revanche, les phosphatases agissant sur des sucres phosphorylées appartiennent à la superfamille des Haloacide Déshalogénases (HADs) , offrant ainsi un large choix de phosphatases candidates potentiellement actives sur D-Xu5P.

1.4 La superfamille des Haloacide Déshalogénases

Bien que nommée d'après le premier membre dont la structure a été caractérisée, la 2-haloacid dehalogenase, la famille des HADs comporte une majorité de membres catalysant le transfert de groupement phosphorylé. Elle comprend des phosphatases, phosphonatases et phosphomutases (**Figure 3.4**). Ce sont des protéines ubiquitaires, présentes dans le protéome de tous les organismes, allant des archées, bactéries aux eucaryotes (Burroughs *et al.*, 2006). Les membres de la famille des HADs partagent des motifs conservés et utilisent un mécanisme commun pour catalyser différents types de réactions enzymatiques (Lahiri *et al.*, 2006).

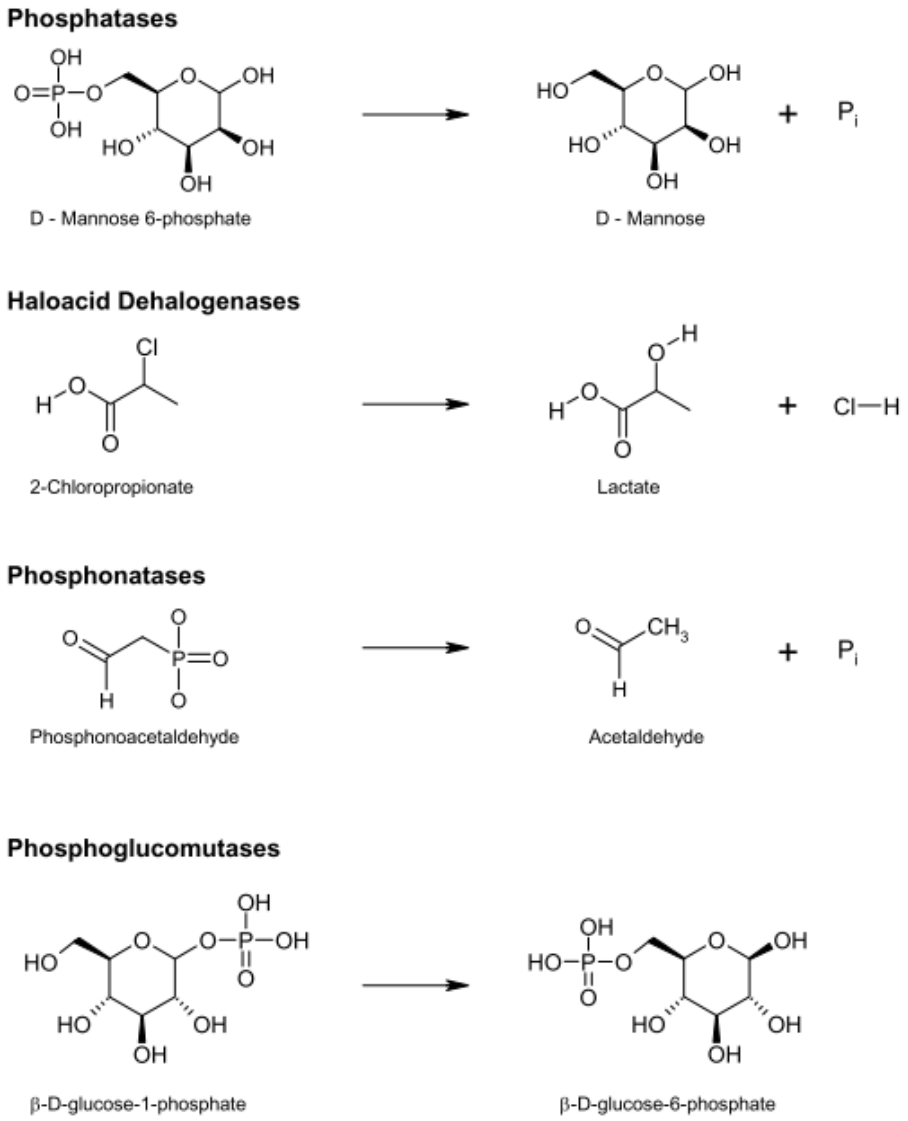


Figure 3.4 : Réactions représentatives des 4 activités catalysées par les membres de la superfamille des HADs. *(Kuznetsova, 2009)*

Koonin et Tatusov ont conduit une analyse bioinformatique à partir des séquences des HADs connues (Koonin et Tatusov, 1994). De cette manière quatre motifs conservés ont été découverts au sein de la superfamille enzymatique. Cette découverte a permis de rapidement identifier de nouveaux membres de la famille des HADs à partir d'études *in silico*. En 1995, 50 protéines étaient identifiées, aujourd'hui elles sont plus de 30 000 (Allen et Dunaway-mariano, 2010).

1.4.1 Structure des HADs

A) Le site actif

Actuellement, plus de 70 structures à rayons X de membres de la superfamille des HADs sont disponibles (Banque de données sur les protéines, www.rcsb.org, Juin 2018). La majorité des hydrolases HAD ont une organisation structurelle illustrée dans la **Figure 3.5**.

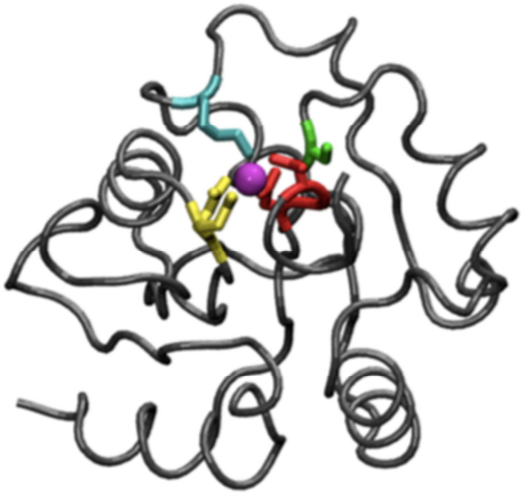


Figure 3.5 : Les quatre motifs catalytiques conservés des HADSF phosphotransferases représentés dans le pli Rossman *La coiffe n'a pas été représentée pour une meilleure visibilité du site actif. Le motif I (DXD) est en rouge, le motif II (S/T) en vert, le motif III (K/R), le motif IV (DD, GDxxxD, ou GDxxxxD) en jaune. Le cofacteur Mg2+ est réprésenté sous la forme d'une sphère magenta. (Allen et Dunaway-mariano, 2010)*

Tous les membres de la HADSF possèdent un domaine central qui consiste en un pli Rossmann qui abrite leur site actif. Chaque motif est situé sur une boucle, la région du site actif responsable de l'activité catalytique est constituée de quatre boucles. Sur la boucle 1 réside à la fois l'aspartate nucléophile et l'aspartate catalyseur acide/base qui font partie intégrante du mécanisme de transfert du groupement phosphorylé et constitue un premier motif. La boucle 2 arbore une serine/threonine qui se lie au substrat possédant à cette étape encore son groupement phosphorylé, alors que la boucle 3 qui porte la lysine oriente la charge négative du groupement phosphorylé vers l'aspartate nucléophile. Pour finir la boucle 4 expose 2 ou 3 résidus d'asparagine qui se lient à des ions Mg2+, co-facteur qui a pour rôle de neutraliser les charges négatives du groupement phosphorylé pendant son transfert pour stabiliser le complexe (Seifried, Schultz et Gohla, 2013 ; Allen and Dunaway-Mariano, 2009 ; Lu, Dunaway-Mariano and Allen, 2005) Le rôle du domaine central est de se lier au substrat et d'orienter le groupement phosphorylé pour le transfert du groupement phosphate vers l'accepteur final.

Le domaine central des HADs constitue l'équipement standard requis pour catalyser le transfert de groupement phosphate, il reste inchangé au sein de la famille des HADs et n'est pas lié à la spécificité de substrat. La compréhension des caractéristiques structurelles qui contribuent à la spécificité du substrat est un objectif principal de recherche sur les membres de la superfamille HAD (Srinivasan *et al.*, 2015).

B) La coiffe et la spécificité de substrat

La structure impliquée dans la reconnaissance du substrat au sein de la famille des HADs est une structure secondaire. Ce second domaine appelé le « cap domain » ou coiffe est à proximité du site actif. Sa nature et sa localisation sont à l'origine de la classification des HADs en trois sous familles (I, IIA-IIB-III) (Zhang *et al.*, 2002; Morais *et al.*, 2004; Tremblay, Dunaway-Mariano et Allen, 2006; Caparrós-Martín, McCarthy-Suárez et Culiáñez-Macià, 2013). La sous-famille I possède une petite coiffe entre les motifs I et II. La sous famille II a une coiffe plus importante, située entre les motifs II et III tandis que la sous famille III ne possède pas de coiffe. La sous famille II est elle-même divisée en deux sous-classes IIA et IIB, possédant respectivement des 'cap domains' de type αβ αββ αβ αβ et αββ (αβ αβ) αββ (**Figure 3.6**). Il n'existe qu'une seule enzyme appartenant au type A, la ribonucléotide monophosphatase NagD d'*Escherichia coli*, le type B a donc été étudié permettant la compréhension d'un plus large groupe d'enzymes.

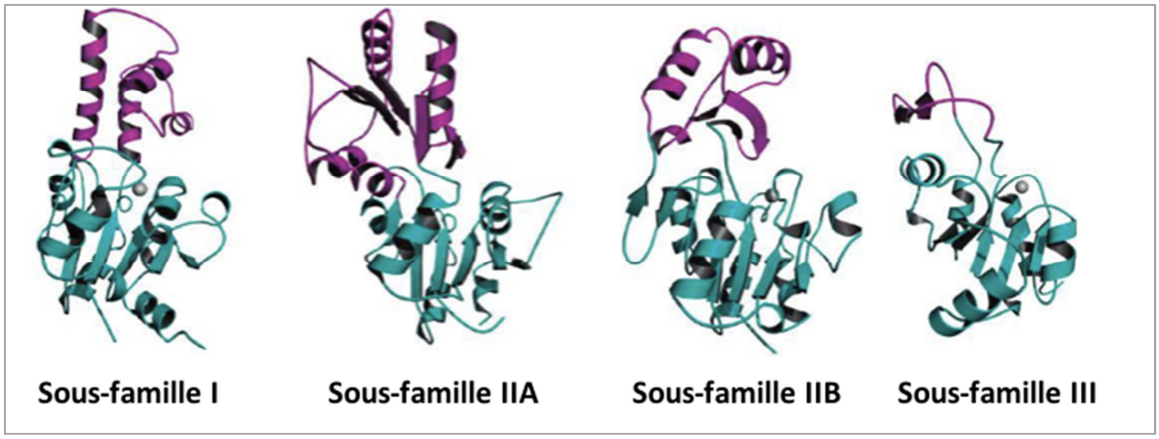


Figure 3.6 : Structure de représentants de la sous-famille I (B-PGM), sous-famille II de type A (NagD), subfamily II de type B (phosphoglycolate phosphatase), et la sous-famille III (phosphatase 1 magnesum-dépendante) *(Lu, Dunaway-Mariano and Allen, 2005)*

La coiffe agit structurellement comme un « couvercle » au-dessus du site catalytique en limitant du l'accès au site actif, autorisant le passage aux molécules de poids moléculaire adapté pour se frayer un chemin jusqu'au site catalytique et retenant les autres.

Les aspects essentiels qui ont été soulignés dans les études menées sur les HADs sont : (i) l'alternance des états ouverts et fermés de la coiffe (ii) une réaction préliminaire favorisée par l'exclusion du solvant et une étape terminale favorisée par le contact avec le solvant.

Compte tenu de l'importance de ces étapes, la sélection naturelle semble avoir favorisé l'émergence de coiffes qui ont rendu les processus d'exclusion des solvants plus efficace. En plus de favoriser la catalyse, l'émergence de diverses coiffes a fourni un moyen de reconnaissance du substrat avec de nouvelles surfaces d'intéraction, ce qui n'était pas le cas du site actif « ancestral ».

L'étude de la conformation de l'enzyme permet à partir de la taille de la cavité d'évaluer quels métabolites phosphorylés pourraient occuper au mieux la cavité (Allen et Dunaway-mariano, 2010). Pour les phosphatases hydrolysant les sucres phosphorylés appelées 'sugar phosphatases', les sucres phosphorylés en position 1 ont une conformation cyclique contrairement aux sucres phosphorylés en position terminale phosphorylé sur le carbone 4 pour les C4, 5 pour les C5,.. de ce fait les enzymes pouvant accueillir les sucres linéaires n'ont pas le site actif adapté pour les sucres cycliques. Elles ne sont pas spécifiques d'un sucre donné mais elles ont de par leur conformation une affinité pour les métabolites phosphorylés selon leur nombre de carbone et le positionnement du groupement phosphate. Le Xu5P est un pentose linéaire, sa structure est très proche du ribulose-5P dont il est le diastéréoisomère (**Figure 3.7**). Les autres pentoses-P de la voie des PP couramment utilisés lors des caractérisations enzymatiques sont le D-xylose-5P et D-ribose-5P, leur conformation et leur taille sont similaires à celle du D-Xu5P. Une HAD active sur ces pentoses-5P sera probablement active sur Xu5P.

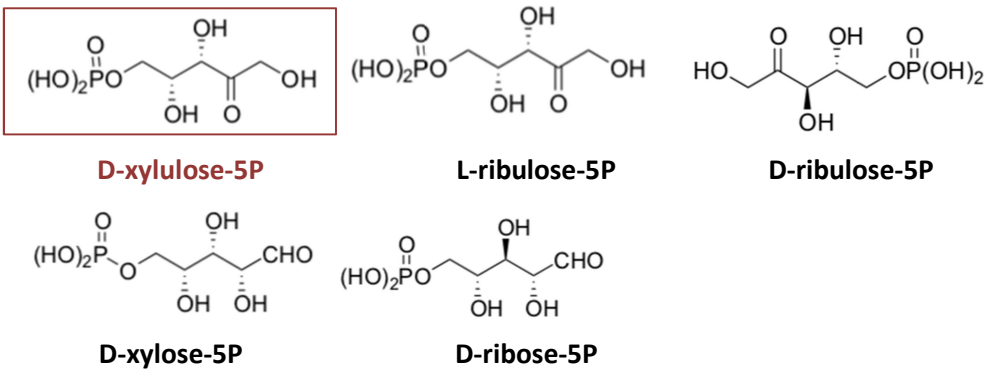


Figure 3.7 : Structure des pentoses-P de la voie des PP.
En rouge : le D-xylulose-5P, substrat cible de cette étude.

1.4.2 Phosphatases

Les phosphatases sont les enzymes majoritaires de la superfamille des HADs (79%) (Allen et Dunaway-Mariano, 2009). Bien que de nombreuses protéines aient été annotées comme étant des phosphatases, relativement peu ont été caractérisées. Une enzyme est définie comme phosphatase si elle est capable d'hydrolyser le substrat artificiel pNPP (para-nitrophényl-phosphate). Ce substrat est utilisé dans des études de criblages enzymatiques, le pH optimal d'hydrolyse du pNPP permet de déterminer si la phosphatase incriminée est alcaline ou acide.

Chez les microorganismes, les HAD phosphatases sont des protéines possédant un unique domaine fonctionnel et hydrolysant de petites molécules phosphorylées, alors que chez les eucaryotes supérieurs de nombreuses HADS ont fusionnés avec d'autres domaines fonctionnels, conférant une plus grande spécialisation. (Allen *et al.,* 2004; Burroughs *et al*. 2006).

HAD phosphatases monomériques- Les HADS possédant un seul domaine fonctionnel peuvent hydrolyser un très large panel de petits métabolites phosphorylés tels que les sucres, les nucléotides, et les intermédiaires de plusieurs voies métaboliques : cofacteurs, acides aminés, Les enzymes varient dans leur degré de spécificité pour leur substrat, de très spécifique, comme la 3-Deoxy-d-manno-octulosonate 8-phosphate (KO8P) phosphatase (Wu *et al.,* 2003), à non spécifique, comme la phosphatase levurienne YKR070W, qui n'a d'affinité marquée pour aucun substrat. Le D-xylulose 5P étant un métabolite de petite taille, les HADs monomérique, levuriennes ou bactériennes semblent être les plus adaptées à cette étude.

HAD phosphatases multimériques - Présentes chez les eucaryotes supérieurs, ces protéines seraient issues de la fusion de plusieurs domaines fonctionnels (Burroughs *et al.* 2006), ces fusions apportent en outre une plus grande spécificité et une localisation définie permettant le positionnement de deux étapes enzymatique côte à côte, ce qui serait bénéfique pour cellules eucaryotes plus complexes. Ces phosphatases hydrolysent des polymères de grande taille tels que des protéines phosphorylées.

Nous pouvons à partir de ces connaissances définir les critères de sélection des enzymes candidates : i) actives sur pentoses-P linéaires ii) monomériques (provenant de procaryote ou eucaryote inférieur)

1.4.2 Phosphoglucomutases

Dans la cellule la conversion du glucose 1-phosphate (G1P) en glucose 6-phosphate (G6P) est catalysée par 2 différentes enzymes : les α-phosphoglucomutases (αPGM) et les β-phosphoglucomutase (βPGM). Les αPGM sont spécifiques des α-glucose phosphate, on les trouve chez les eucaryotes et les procaryotes, tandis que les βPGMs, sont spécifiques des β-glucose phosphate (Dai *et al.*, 2006).Les αPGMs sont des protéines comportant 4 domaines et appartiennent à la superfamille des phosphohexomutases. Les βPGM appartiennent à la famille des HADs, elles sont plus petites et comportent 2 domaines. Ces 2 types d'enzymes requièrent dans le mécanisme catalytique la présence d'un intermédiaire di-phosphorylé, le glucose 1,6 bisphosphate. L'activité des PGMs est impliquée dans l'équilibre entre synthèse et dégradation des polysaccharides de la paroi cellulaire et la glycolyse chez les bactéries. (Mesak and Dahl 2000 ; Lahiri, Zhang et al. 2002). Le rôle et l'importance des βPGM dépend de l'organisme : chez *Lactococcus lactis* βPGM occupe un rôle central dans la catabolisme du maltose et du tréhalose, tandis que chez *E.coli* cette enzyme n'est pas importante. (Levander, Andersson et al. 2001). Walther *et al.*(2012) ont testé la capacité des phosphoglucomutase Pgmi1 Pgm2 et Pgm3 de la levure *S. cerevisiae* à convertir le ribose-1P en ribose-5P. Les 3 enzymes sont actives sur glucose-1P et ribose-1P, avec Pgm3 ne possédant qu'une activité résiduelle sur ribose-1P. La délétion de Pgm3 a entrainé chez les mutants une accumulation de rbose-1P mettant en evideence la role majeur de phosphoribomutase de Pgm3 *in vivo*. Le site actif de Pgm1, Pgm2, Pgm3 est donc capable d'accepter un pentose, nous testerons dans cette étude ces 3 isoenzymes sur D-Xu5P.

Les connaissances structurales des HADs indiquent qu'un site actif capable de reconnaitre un pentose linéaire phosphorylé aura très probablement une affinité pour le D-Xu5P. De ce fait les D-Xu5P phosphatases/mutases candidates ont été sélectionnées d'après leur capacité à accepter le D-Xu5P ou des sucres à 5 ou 6 carbone phosphorylés en position terminale, comme substrat. Cette approche a pour avantage d'élargir le choix de phosphatases candidates, cependant une faible spécificité de substrat pourrait engendrer des perturbations métaboliques. Le métabolisme carboné central comprend des sucres phosphorylés que ce soit dans la glycolyse (glucose-6P, fructose 6P, fructose 1,6bP,..) ou la voie des pentoses phosphate (ribose-5P, xylulose-5P, ribulose-5P,..). Il sera important d'évaluer l'activité sur D-xylulose-5P ainsi que la spécificité des phosphatases candidates sur des métabolites phosphorylés essentiels.

1.5 Sélection de D-xylulose 5P phosphatases candidates

Seule une phosphatase a été caractérisée sur D-Xu5P dans la littérature (Godinho et De Sá-Nogueira, 2011). Deux autres protéines ont montré une activité sur D-Xu5P mais n'ont pas été caractérisées (Mervi H Toivari *et al.*, 2007; Kim *et al.*, 2013). Par conséquent nous avons élargi nos recherches aux phosphatases capables d'hydrolyser des sucres phosphorylés de taille comparable au D-Xu5P (C5), en position terminale. Le travail de caractérisation des HADs d'Ekatarina Kuznetsova a été une précieuse source d'informations. Les HADs d'*E.coli* et *S.cerevisiae* ont été étudiées sur 80 substrats phosphorylés. Parmi les substrats des pentoses-P en position terminale : ribose-5P, ribulose-5P ; des hexoses-P phosphorylés sur le carbone n°6 : glucose-6P, fructose-6P, mannose-6P mais également sur le carbone n°1 : glucose-1P, fructose-1P, mannose-1P. Le D-Xu5P n'a pas été testé car c'est un substrat très onéreux, nous considérons l'activité sur D-ribulose-5P comme prédicatrice de l'activité sur D-Xu5P car ce sont des épimères. Les activités sur hexoses-1P dénotent de la capacité à déphosphoryler des sucres en position 1, comme le D-xylulose-1P est un métabolite clef de la synthèse d'AG c'est un paramètre important à considérer. Ainsi, 7 phosphatases de bactéries et de levures ont été sélectionnées pour l'étude de la déphosphorylation du D-xylulose-5P en D-xylulose.

1.5.1 AraL de *Bacillus subtilis*

La protéine AraL de *Bacillus subtilis* appartient à la famille des HADs. Son activité phosphatasique est optimale à pH7, à une température 65°C. Lia M.Godinho, a mis en évidence une activité de la phosphatase AraL pour les sucres phosphorylés dont le D-Xu5P (**Tableau 3.1**, Godinho et De Sá-Nogueira, 2011). Il s'agit de la seule étude montrant une activité enzymatique sur le substrat D-xylulose 5P.

Tableau 3.1: Constantes cinétiques d'AraL pour différents substrats phosphorylés (pH7, 37 °C) (Godinho et De Sá-Nogueira, 2011)

Substrat	K_m (mM)	k_{cat} (s^{-1})	k_{cat} / K_m (s^{-1} M^{-1})
D-xylulose 5P	29.14 ± 4.87	2.75± 0.26	0.943 x 10^2
Glucose 6P	26.96 ± 4.08	2.46 ± 0.26	0.998 x 10^2
D-arabinose 5P	27.36 ± 1.8	2.92 ± 0.10	1.06 x 10^2
Fructose 6P	34.89 ± 4.51	2.817 ± 0.22	0.807 x 10^2
Fructose 1,6bisP	40.78 ± 11.40	1.49 ± 0.26	0.365 x 10^2
Galactose 1P	40.74 ± 6.03	4.28 ± 0.40	1.02 x 10^2
pNPP	50.00 ± 23.32	0.012 ± 0.0006	0.24

La constante d'affinité, K_m, correspond à la concentration en substrat pour laquelle la vitesse enzymatique est de $V_{max}/2$. Ainsi plus le K_m est faible plus l'affinité pour le substrat est grande. Le K_m pour le D-xylulose 5P est de 29,14 mM, ce qui est élevé sachant que la concentration de D-xylulose 5P intracellulaire est de 1mM.

La constante k_{cat} est le 'turnover' de l'enzyme, plus le k_{cat} est élevé plus la réaction est rapide. Le k_{cat} (D-Xu5P) est de 2,75 s^{-1}. Enfin, la spécificité, renseignée par le rapport k_{cat} / K_m, est de $0{,}943.10^2$ s^{-1} M^{-1}. Un rapport élevé indique que la spécificité de l'enzyme pour le substrat est grande. AraL a montré une activité sur X5P ainsi que G6P, A5P, F6P, F1,6bisP et Gal1P . Le rapport k_{cat} /K_m étant faible pour ces sucres phosphorylés, la capacité d'AraL de distinguer ces différents métabolites nous apparait très limitée.

1.5.2 Pho13 de *Saccharomyces cerevisiae*

Le gène *pho13* code pour la protéine monomérique Pho13 de *Saccharomyces cerevisiae* qui a été annotée comme faisant partie des HADs en 1998 et plus particulièrement comme phosphatase car elle montre une activité sur le substrat pNPP (Tuleva, Vasileva-Tonkova et Galabova, 1998). Ce n'est que récemment que Pho13 a été étudiée pour son activité sur le D-xylulose 5P. *S. cerevisiae* est une levure couramment utilisée pour la production de biocarburant à partir de biomasse végétale mais elle n'assimile pas naturellement le xylose. Deux voies d'assimilation du xylose ont été identifiés chez des champignons et bactéries (Zaldivar, Nielsen et Olsson, 2001). L'une ou l'autre de ces voies est souvent combinées avec la surexpression de la xylulose kinase dans *S.cerevisiae* ce qui facilite le flux de la voie hétérologue vers la voie des PP native (Toivari *et al.*, 2001; Madhavan *et al.*, 2009; Kim *et al.*, 2013). La xylulose kinase hétérologue phosphoryle le D-xylose en D-xylulose 5P qui alimente la voie des PP. Il a été démontré que la délétion du gène *pho13* améliorait la croissance et la production d'éthanol à partir de D-xylose (Ni, Laplaza and Jeffries, 2007; Van Vleet, Jeffries and Olsson, 2008; Fujitomi *et al.*, 2012; Huang *et al.*, 2012; Kim *et al.*, 2013, 2015; Sakihama, Hasunuma et Kondo, 2015). Au moment de la sélection des enzymes candidates l'hypothèse principale était que Pho13 déphosphorylait le D-xylulose 5P créant ainsi un cycle futile diminuant de la production d'éthanol. Sur la base de cette hypothèse, l'activité de Pho13 sur D-xylulose 5P a été mise en évidence sur extraits bruts (Kim *et al.*, 2013). Nous avons souhaité valider l'activité de Pho13 purifiée sur D-Xu5P et évaluer son activité sur D-Xu5P *in vivo*.

L'activité de la phosphatase de Pho13 a été récemment testée *in vitro* sur des sucres phosphorylés intermédiaires de la voie des PP (xylulose 5P, ribulose 5P, ribose5P, desoheptulose 7P, E4P) et de la glycolyse (glucose 6P, fructose-6P) (Xu *et al.*, 2016). Pho13 déphosphoryle non-spécifiquement les intermédiaires de la voie des PP dont le sedoheptulose 7P mais n'a pas d'activité sur ceux de la glycolyse (**Figure 3.8**).

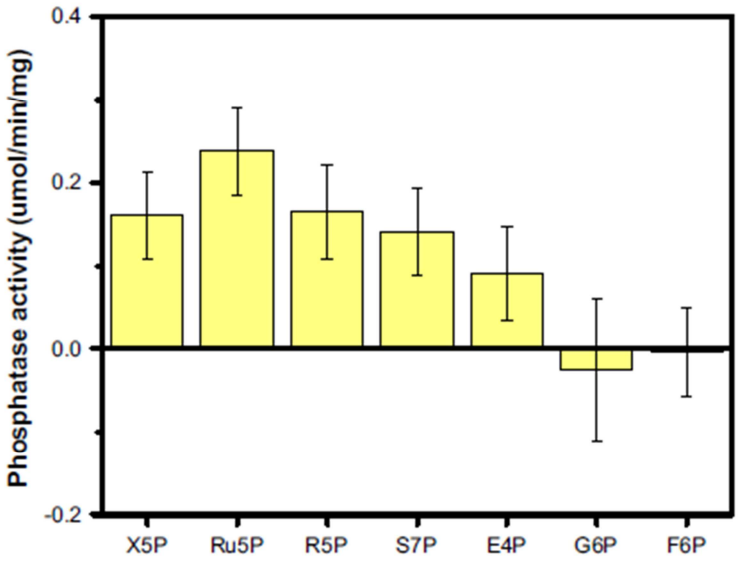


Figure 3.8 : Activité phosphatasique de Pho13 sur des des sucres phosphorylés variés. *X5P : xylulose-5P, Ru5P : ribulose-5P, R5P : ribose-5P, S7P : sedoheptulose-7P, E4P : erythrose-4P, G6P : glucose-6P, F6P : fructose-6P.Substrats concentrés à 0,1 mM.*(Xu *et al.*, 2016).

1.5.3 TM1254 de *Thermogota maritima*

La protéine TM1254 de *Thermogota maritima* est une phosphatase découverte en 2006 (Kuznetsova *et al.*, 2006). Cette phosphatase est principalement active sur l'erythrose 4P, fructose 6P, 2-deosyglucose 6P et mannose-6P. Cette phosphatase a l'affinité la plus élevée pour l'erythrose 4P (K_m = 152.6 ± 55.µM) et bien que la concentration intracellulaire de l'erythrose 4P ne soit pas connue, ce K_m est parmi les plus faibles des enzymes ayant une activité sur erythrose 4P (BRENDA). L'intérêt dans l'étude de TM1254 est d'évaluer sa spécificité en la caractérisant sur un panel de sucres phosphorylés.

1.5.4 YbiV d'*Escherichia coli*

La protéine YbiV d'*Escherichia coli* a été annotée comme étant une phosphatase HAD en 2005. Cette PPase est capable d'hydrolyser des pentoses et des hexoses dont le fructose 1P, le ribose 5P, le glucose 6P. YbiV a la plus forte affinité pour le fructose 1P et ribose 5P (**Tableau 3.2**, Roberts *et al.*, 2005). Le ribose est un pentose, tout comme le xylulose 5P. Bien qu'YbiV n'ait jamais été caractérisé sur xylulose 5P, elle pourrait avoir une activité sur ce métabolite. Si YbiV est retenue pour son activité sur D-xylulose-5P il faudra cependant s'assurer qu'elle n'ait pas d'activité sur Xylulose 1P car cette enzyme est aussi capable de déphosphoryler des sucres en position 1 (fructose 1P).

Tableau 3.2 : Constantes cinétiques de YbiV pour différents substrats phosphorylés (pH7, 37 °C) *(Roberts et al., 2005)*

Substrat	K_m (mM)	K_{cat} (s^{-1})	K_{cat} / K_m (s^{-1} M^{-1})
Fructose 1P	1.4 ± 0.3	111± 6	8.0 x 10^4
Ribose 5P	2.4 ± 0.2	28 ± 1	1.2 x 10^4
Glucose 6P	3.1 ± 0.3	22 ± 1	6.9 x 10^3
Acetyl P	4.9 ± 0.7	101 ± 4	21 x 10^3
Imido diP	0.12 ± 0.02	7.5 ± 0.6	62 x 10^3

1.5.5 Dog1 de *Saccharomyces cerevisiae*

Dog1P est une phosphatase de *Saccharomyces cerevisiae* qui a une activité élevée sur 2-Desoxyglucose 6P (100%) mais aussi sur D-ribose 5P (47%), xylulose-5P (15%) et D-ribose-5P (7%) (Mervi H. Toivari *et al.*, 2007). Les constantes cinétiques K_m, K_{cat} et K_{cat} /K_m n'ont pas été calculées pour le métabolite D-xylulose 5P. Dog1 a montré une activité sur D-xylulose-5P in *vitro* mais n'a pas montré d'activité significative sur les pentoses phosphorylés *in vivo*.

1.5.6 YniC d'*Escherichia coli*

La protéine YniC est l'homologue de Dog1 chez *Escherichia coli.* Cette phosphatase a montré une activité sur 2-Desoxyglucose 6P, mannose-6P, ribose-5P, glucose-6P, avec une forte affinité pour le 2-desxyglucose-6P*,* tout comme Dog1 (**Tableau 3.3**, Kuznetsova *et al.*, 2006). Dog1 ayant une activité sur D-xylulose 5P, il est fort probable qu'YniC soit active sur ce métabolite.

Tableau 3.3 : Constantes cinétiques de YniC pour différents substrats phosphorylés (pH7, 37 °C) (Kuznetsova et al., 2006)

Substrat	K_m (mM)	k_{cat} (s^{-1})	k_{cat} / K_m (s^{-1} M^{-1})
2deoxyglucose 6P	0.61 ± 0.05	33 ± 1	5.4 x 10^4
Mannose 6P	4.7 ± 0.4	11 ± 0.4	2.2 x 10^3
2 deoxyribose 5P	2.5 ± 0.1	9.4 ± 0.1	3.7 x 10^3
Ribose 5P	2.6 ± 0.3	2.8 ± 0.1	1.1 x 10^3
Glucose 6P	3.6 ± 0.5	25 ± 2	6.7 x 10^3

1.5.7 YKR070W de *Saccharomyces cerevisiae*

YKR070W est une HAD phosphatase mitochondriale peu spécifique mais montrant une activité sur pentoses et hexoses phosphorylés ainsi que sur le glycérol-P (Kuznetsova *et al.*, 2006). D'après les valeurs de K_m relativement faibles (inférieures à 1 mM) observées sur la majorité des substrat testés (glucose-6P, glycérol-3P, glycérol-1P, 2-deoxy-glucose-6P, 6P-gluconate, 2-deoxy-ribose-5P, Fructose-1,6bisP, Fructose-6P, Mannose-6P, Ribose-5P, Glycérol-2P) cette phosphatase peut être responsable de la réduction des taux sucres phosphorylés en présence d'oxygène afin de diminuer leur taux d'utilisation (effet pasteur). Les efficacités catalytiques ne varient que d'un ordre de grandeur parmi les substrats cités allant de $1{,}6.10^5$ à $5{,}1.10^4$. YRKR070W est très promiscuitaire, il serait intéressant de connaitre son activité sur Xu5P. Si cette activité est suffisante pour la production d'acide glycolique nous pourrons envisager d'accroitre la spécificité de YKR070W par ingénierie.

1.5.8 Synthèse

7 phosphatases provenant de bactéries (*Escherichia coli, Bacillus subtilis, Thermogota maritima*) et de S.cerevisiae, ainsi que 3 isoenzyme de phosphoglucomutase de *S.cerevisiae* ont été sélectionnées pour cette étude. Le **Tableau 3.4** synthétise les principales données d'activités enzymatiques connues pour ces enzymes.

Tableau 3.4 : Constantes cinétiques des phosphatases et mutases sélectionnées

Enzyme	subtsrat	K_m (mM)	k_{cat} (s^{-1})	k_{cat} / K_m ($s^{-1} M^{-1}$)	Référence
AraL	**D-xylulose-5P**	**29,14**	**2,75**	**$0{,}943.10^2$**	Godinho et De Sá-Nogueira, 2011
	D-glucose-6P	24,96	2,49	$0{,}998.10^2$	Godinho et De Sá-Nogueira, 2011
	D-arabinose-5P	27,36	2,92	$1{,}06.10^2$	Godinho et De Sá-Nogueira, 2011
	D-fructose-6P	34,89	2,817	$0{,}807.10^2$	Godinho et De Sá-Nogueira, 2011
	D-fructose 1,6-biP	40,78	1,49	$0{,}365.10^2$	Godinho et De Sá-Nogueira, 2011
	galactose-1P	40,74	4,28	$41{,}02.10^2$	Godinho et De Sá-Nogueira, 2011
	pNPP	50	0,012	0,24	Godinho et De Sá-Nogueira, 2011
TM1254	D-erythrose-4P	0,152	1,24	8130	Kuznetsova *et al.*, 2006
	D-fructose-6P	0,2	0,62	3089	Kuznetsova *et al.*, 2006

Tableau 3.4 (suite) : Constantes cinétiques des phosphatases et mutases sélectionnées

Enzyme	subtsrat	K_m (mM)	k_{cat} (s^{-1})	k_{cat} / K_m ($s^{-1} M^{-1}$)	Référence
Dog1	2-deoxy-glucose-6P	1,39	99,60	$7,1.10^4$	Kuznetsova *et al.*, 2006
	Mannose 6P	5,7	46,4	$8,1.10^4$	Kuznetsova *et al.*, 2006
	2-deoxyglucose-6P	17	ns	ns	Randez-Gil *et al.*, 1995
	D-ribose-5P	30	ns	ns	Randez-Gil *et al.*, 1995
	2-deoxyglucose-6P	ns	ns	ns	Mervi H Toivari *et al.*, 2007
	D-ribose-5P	ns	ns	ns	Mervi H Toivari *et al.*, 2007
	D-ribulose-5P	ns	ns	ns	Mervi H Toivari *et al.*, 2007
	D-xylulose-5P	ns	ns	ns	Mervi H Toivari *et al.*, 2007
Pho13	phosphoglycolate	0,061	67,4	$1,1.10^6$	Kuznetsova *et al.*, 2015
	pNPP	0,036	292,6	$8,12.10^6$	Tuleva, Vasileva-Tonkova et Galabova, 1998
	D-xylulose-5P	nd	ns	nd	Kim *et al.*, 2013
YKR070W	Glucose-6P	0,22	32,1	$1,6.10^5$	Kuznetsova *et al.*, 2015
	2-deoxyglucose-6P	0,59	46,2	$7,8.10^4$	Kuznetsova *et al.*, 2015
	2-deoxyribose-5P	0,31	27,3	$1,4.10^5$	Kuznetsova *et al.*, 2015
	fuctose-1,6bisP	0,65	40,1	$7,8.10^4$	Kuznetsova *et al.*, 2015
	fructose-6P	1,04	67,4	$8,8.10^4$	Kuznetsova *et al.*, 2015
	Mannose 6P	0,67	39,6	$6,2.10^4$	Kuznetsova *et al.*, 2015
	ribose-5P	0,75	49	$6,5.10^4$	Kuznetsova *et al.*, 2015
	Erythrose-4P	0,33	29,5	$5,9.10^4$	Kuznetsova *et al.*, 2015
	Glycerol-3P	0,28	44,1	$6,5.10^4$	Kuznetsova *et al.*, 2015
	Glycérol-1P	0,39	35,6	$9,2.10^4$	Kuznetsova *et al.*, 2015
	Glycérol-2P	1,22	12,2	$1,0.10^4$	Kuznetsova *et al.*, 2015
YbiV	Ribose-5P	6	nd	nd	Kuznetsova *et al.*, 2006
	pNPP	2,43	nd	nd	Kuznetsova *et al.*, 2006
	Fructose-1P	1,4	111	8.10^4	Kuznetsova *et al.*, 2006
	Ribose-5P	2,4	28	$1,2.10^4$	Kuznetsova *et al.*, 2006
	Glucose-6P	3,1	22	$6,9.10^3$	Kuznetsova *et al.*, 2006
	Acetyl-P	4,9	101	21.10^3	Kuznetsova *et al.*, 2006
	Imido-di-P	0,12	7,5	62.10^3	Kuznetsova *et al.*, 2006
YniC	2-deoxyglucose-6P	0,61	33	$5,4.10^4$	Kuznetsova *et al.*, 2006
	Mannose-6P	4,7	11	$2,2.10^3$	Kuznetsova *et al.*, 2006
	2-deoxyribose-5P	2,5	9,4	$3,7.10^3$	Kuznetsova *et al.*, 2006
	ribose-5P	2,6	2,8	$1,1.10^3$	Kuznetsova *et al.*, 2006
	glucose-6P	3,6	25	$6,7.10^3$	Kuznetsova *et al.*, 2006
PGM1	glucose-1P	0,06	0,29	$4,21.10^3$	Walther *et al.*, 2012
	ribose-1P	nd	3.10^-5	nd	Walther *et al.*, 2012
PGM2	glucose-1P	0,026	35,4	$1,36.10^6$	Walther *et al.*, 2012
	ribose-1P	0,53	0,33	$6,35.10^2$	Walther *et al.*, 2012
PGM3	glucose-1P	0,112	0,13	$1,16.10^3$	Walther *et al.*, 2012
	ribose-1P	0,75	0,34	$4,58.10^2$	Walther *et al.*, 2012

2. RESULTATS

2.1 Expression et purification des phosphatases

L'expression des protéines a été induite par l'ajout d'IPTG au milieu de culture LB en phase exponentielle de croissance. Les protéines fusion 6His-*phosphatase* ont été purifiées sur une résine cobalt. A chaque étape de la purification des aliquotes ont été prélevés, ils ont ensuite été analysés sur gel SDS-PAGE. La **Figure 3.9** présente le gel SDS PAGE de la purification de *PHO13*. Les profils de purification sont similaires pour toutes les phosphatases (résultats non montrés). La présence d'une bande protéique à la taille attendue (*PHO13* : 35 kDa) sur la piste de migration de l'éluat confirme l'expression de la protéine d'intérêt. Seule la protéine TM1254 présente 2 bandes, mettant en évidence une co-élution. Les phosphatases purifiées sont présentés sur un même gel SDS PAGE, elles ont toutes la taille attendue (**Figure 3.9**). Le tag Histidine, en N terminal des protéines, permet une bonne efficacité de fixation de la protéine sur la colonne (comparaison des piste 'BB' et 'AB' : avant fixation et après fixation) qui ne s'élue pas au cours des deux lavages successifs (piste W1 et W2). La présence d'une seule bande au niveau des pistes d'élution permet de conclure en une pureté satisfaisante des protéines (**Figure 3.9, Figure *3.10* 3.10**) .

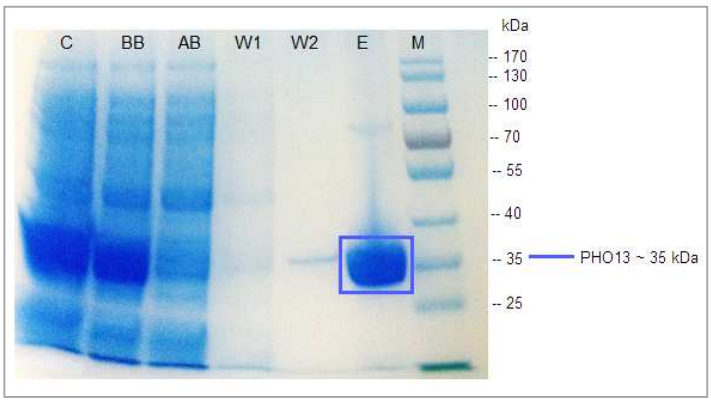


Figure 3.9 : SDS PAGE de la purification de Pho13

***C** : Culot cellulaire, **BB** : « before binding », avant mis en contact avec les billes de cobalt, **AB** : « after binding », après la mise en contact avec les billes de Cobalt, **W1 :** « wash 1 » , 1er lavage, **W2 : :** « wash 2 » , deuxième lavage, **E** : Elution, **M** : Marqueur de poids moléculaire 10-170 kDa*

Les phosphatases purifiées ont toutes les tailles attendues. Le dosage par la méthode de Bradford des protéines ont démontré que les protéines de fusion AraL-6His, PHO13-6His, Dog1-6His, YbiV-6His, YniC-6His, YKR070W-6His et TM1254-6His ont été produites en quantité satisfaisante par la souche BL21(DE3) d'*Escherichia coli*. Les élutions ont permis de récupérer entre 0,8 et 2 mg de protéines purifiées par ml d'éluat, ce qui équivaut à un rendement de 8 à 21 mg par litre de culture.

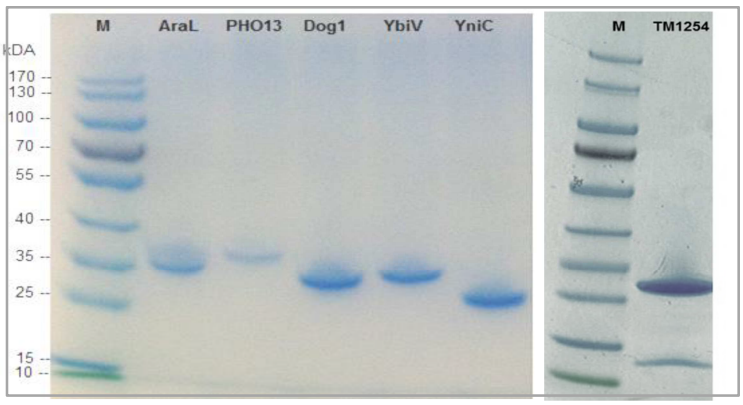


Figure 3.10 : SDS PAGE des élutats des phosphatases candidates*. M = Marqueur de taille protéique (10-170 kDa)*

2.2 Activité enzymatique des phosphatases purifiées sur pNPP *in vitro*

L'expression hétérologue de protéines dans *Escherichia coli* a de nombreux avantages : la croissance bactérienne est rapide et peu onéreuse, *E.coli* est une bactérie non pathogène, son système génétique est bien connu, de nombreuses souches ont pu être améliorées (ex : BL21(DE3)) et les rendements de production de protéines sont élevés. Cependant un des inconvénients de ce système de production est qu'il n'est pas capable de modification post-traductionnelles contrairement aux systèmes eucaryotes comme *Saccharomyces cerevisiae*. De ce fait l'activité biologique de la protéine hétérologue peut différer de la protéine native. De plus, l'étiquette polyhistidine en N-terminal peut affecter la conformation de la protéine et son activité. Afin de s'assurer de la fonctionnalité des phosphatases exprimées dans *Escherichia coli* nous avons mesuré leur activité 'phosphatase'sur pNPP. Une activité sur pNPP indique que l'enzyme purifiée est fonctionnelle. Les résultats d'activités obtenus sont présentés dans le **Tableau 3.5**.

Tableau 3.5 : Activité spécifique des phosphatases sur pNPP (µmol/min/mg)– 3mM pNPP, pH7.5, 37°C

Phosphatase	pNPP
AraL	0,25 ± 0,08
TM1254	0,43 ± 0,12
YbiV	0,98 ± 0,12
YniC	0,21 ± 0,7
Pho13	22 ,4± 1,8
Dog1	0,51 ± 0,11
YKR070W	0,42 ± 0,16

PHO13 de *Saccharomyces cerevisiae* est classée sur la base de données BRENDA comme étant une p-nitrophényl phosphatase. Cette phosphatase est connue pour avoir une activité élevée sur pNPP. (200

µmol.min^{-1}.mg^{-1}). L'activité a été mesurée dans cette étude à pH 7.5, à 37°C sur une solution de pNPP à 3 mM. L'activité de PHO13 sur pNPP est de 22,4 µmol. min^{-1}.mg^{-1} ce qui correspond à l'activité la plus élevée mesurée parmi les phosphatases candidates.

Toutes les enzymes ont montré une activité 'phosphatase' sur pNPP (3 mM, pH 7.5, 37°C). La protéine YbiV a une activité de 0,98 µmol.min^{-1}.mg^{-1}sur pNPP (**Tableau 3.6**). L'étude d'Ekatarina Kuznetsova sur les HADs d'Escherichia coli et *Saccharomyces cerevisiae* indique une activité de 2,6 µmol.min-1.mg-1 pour YbiV, sur pNPP à 10 mM, à pH 7,5 et 37°C (Kuznetsova, 2009). Selon cette même étude l'activité de YniC sur pNPP est de 0,6 µmol.min-1.mg-1. Dans notre étude l'activité de YniC sur pNPP est également plus faible que celle de YbiV, elle est de 0,02 µmol.min-1.mg-1. Nos résultats sont cohérents avec les publications, le fait que les conditions expérimentales diffèrent de celle utilisées par Ekatarina Kuznetsova a pour effet de diviser d'un facteur 3 les activités obtenues.
La faible activité de AraL sur pNPP (0,3 µmol.min^{-1}.mg^{-1}) s'explique par une faible affinité pour le pNPP (K_m = 50 mM). Enfin, l'activité de Dog1 sur pNPP est de 0,51 µmol.min-1.mg-1, celle de YKR070W est de 0,42 µmol.min-1.mg-1 ces activités n'ont pu être comparé à d'autres études, mais permettent de mettre en évidence une activité 'phosphatase' et donc la fonctionnalité de cette enzyme. TM1254 avait montré une activité sur pNPP à 65°C (Kuznetsova *et al.*, 2005a), l'activité à 37°C est faible (0,43 U/mg). En réalisant l'essai sur pNPP à 65°C, l'activité est de 3,28 U/mg. Cette enzyme provenant de l'hyperthermophile *Thermogota maritima* perd donc 86% de son activité phosphatasique à 37°C.

2.3 Profils d'activité des phosphatases et mutases sur substrats biologiques *in vitro*

2.3.1 Activité enzymatique des phosphatases

L'activité des phosphatases a été testée sur des métabolites de la voie des pentoses phosphate dont le D-Xu5P (D-ribose-5P, D-ribulose-5P, D-xylulose-5P) et de la glycolyse (D-glucose-6P, D-fructose-6P et D-glycéraldéhyde-3P) à 0,5 mM, pH 7 et à 37°C. Les résultats sont présentés dans la Figure 3.11.

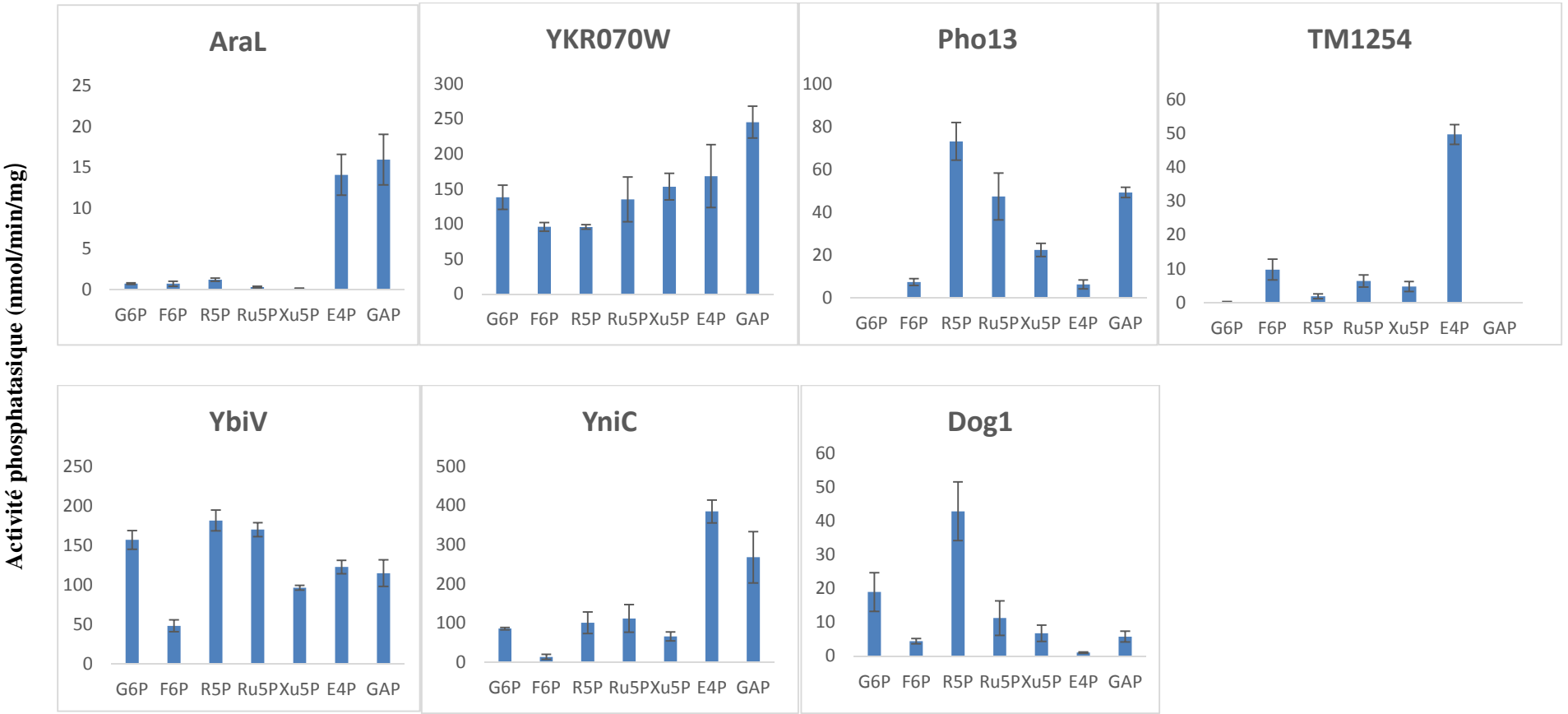


Figure 3.11 : Profils d'activité enzymatique des HADs phosphatases de B.subtilis (AraL) , S.cerevisiae (YKR070, Dog1, Pho13), T.maritima (TM1254) et E.coli (YbiV,YniC) sur des métabolites phosphorylés du métabolisme carboné central dont le D-xylulose-5P. *D-glucose-6P (G6P), D-fructose-6P (F6P), D-ribose-5P (R5P), D-ribulose-5P (Ru5P), D-xylulose-5P (Xu5P), D-erythrose-4P (E4P), Glycéraldéhyde-3P (GAP). Les mélanges réactionnels contiennent 0,5 mM de substrat et 1-10 µg d'enzyme purifié, à 37°C, pH 7.5, la détection de phosphate inorganique a été réalisé en utilisant le réactif malachite green (Sigma).*

A) Profils catalytiques des HADs candidates

AraL de *Bacillus subtilis* est la seule phosphatase dont les constantes cinétiques sont connues sur D-xylulose-5P, son efficacité catalytique sur ce substrat est faible (0,9.102 s-1.M-1) et le K_m élevé (29,14 mM). AraL n'a pas montré d'activité sur D-Xu5P à 0,5 mM (0,1 nmol.min^{-1}.mg^{-1}), l'affinité pour ce substrat n'est pas assez élevée pour permettre la déphosphorylation dans des conditions proches des conditions intracellulaires. (à 0,5 mM). Il en est de même pour le fructose-6P et le glucose-6P, pour lesquels les efficacités catalytiques sont similaires à celle du D-Xu5P (Godinho and De Sá-Nogueira, 2011). AraL est active sur le D-erythrose-4P et le glycéraldéhyde-3P, plus faiblement dur D-ribose-5P, ce qui laisse à penser que le K_m pour ces métabolites est plus bas que les autres métabolites testés.

Dog1 possède une activité sur ribose-5P et glucose-6P, ce profil catalytique est cohérent avec les résultats de l'étude de Kuznetsova mais diffèrent de ceux rapportés par Mervi H Toivari *et al.*, (2007). Lors du criblage sur différents substrats à 0,1 mM décrit par Kuznetsova une activité sur 2-deoxyglucose-6P (substrat de référence), ribose-5P, 2-desoxyribose-5P, glucose-6P et Fructose-1P (Kuznetsova *et al.*, 2005). Aucune activité à cette concentration en substrat n'a été détectée sur D-ribulose-5P tout comme dans notre étude. Toivari *et al* ont utilisés 20 mM pour déterminer l'activité enzymatique de Dog1 sur ce substrat (Mervi H Toivari *et al.*, 2007). Notre résultat et celui de Kuznetsova indique que les affinités pour le D-xylulose-5P et ribulose-5P sont faibles, le K_m doit être supérieur à 0,5 mM (Kuznetsova *et al.*, 2005). Cela explique pourquoi la surexpression de Dog1 chez *Saccharomyces cerevisiae* n'a pas permis la déphosphorylation du D-xylulose-5P dans un mutant transkétolase. Dans une telle souche la concentration en D-Xu5P varie de 0,2 mM à 5 mM selon le milieu de culture. Le K_m pourrait être supérieur à 5 mM (Teleman *et al.*, 1999).

YniC est une HAD monomérique considérée comme homologue de Dog1 chez *Escherichia coli*, tous les deux ont comme substrat de référence le 2 deoxyglucose-6P. Les profils catalytiques observés dans cette étude entre ces deux HADs sont différents. YniC a une activité sur E4P et GAP, absente chez Dog1, de plus elle est active sur Ru5P, R5P et Xu5P avec une meilleure affinité. L'activité sur D-Xu5P est dix fois plus élevé que celle de Dog1 bien qu'elle soit inférieure à celle des autres pentoses-P ce qui révèle un manque de spécificité. Kuznetsova a également décrit une activité décroissante sur les substrats E4P, GAP, G6P, R5P et F6P testés à 0,1 mM (Kuznetsova *et al.*, 2006). Nous avons confirmé ces résultats, étonnement le Ru5P avait été testé mais aucune activité n'a été rapportée par ces auteurs.

Pho13 est annotée comme 4-nitrophenyl phosphatase (UniProt P19881), ce substrat n'étant pas naturel son rôle biologique était inconnu lors de sa sélection. D'après nos résultats Pho13 a une activité sur les pentoses-P et sur GAP qui est un intermédiaire entre la voie des PP et la glycolyse, ce qui est confirmé par (Xu *et al.*, 2016). Parmi les PP, Pho13 montre une meilleure activité sur R5P (73,2 nmol.min^{-1}.mg^{-1}) et l'activité la plus faible sur Xu5P (22,4 nmol.min^{-1}.mg^{-1}). Pho13 n'est pas actif sur G6P, et a une très faible activité sur F6P, E4P.

Les autres phosphatases (TM1254, YbiV, TM1254) ont été sélectionnées d'après les études d'Ekatarina Kuznetsova pour être actives sur des sucres phosphorylés C4, C5 et/ou C6 qu'elles déphosphorylent respectivement en position 4, 5 et 6 terminale.

TM1254 est une enzyme de *Thermogota maritima*. Nous avions montré lors de l'essai enzymatique sur pNPP que cette phosphatase perdait 86% de son activité à 37°C comparé à sa température optimale (65°C). Lors de sa découverte par E. Kuznetsova, TM1254 a été décrite comme E4P phosphatase ce qui est confirmé d'après les résultats obtenus. A 37°C, TM1254 présente une activité réduite mais conservée sur E4P et F6P. Elle n'est pas active sur les autres substrats testés dont le D-Xu5P.

YbiV d'*E.coli* et **YKR070W** de *S.cerevisiae* sont des enzymes promiscuitaires, elles sont actives sur l'ensemble des substrats testés. YKR070W présente dans la littérature une affinité élevée pour les sucres phosphorylés, son activité n'avait pas été évaluée sur Xu5P. Comme attendu YKR070W est active sur Xu5P (153,4 nmol.min^{-1}.mg^{-1}), l'activité dépasse celle de R5P et est similaire à celle sur Ru5P. L'activité sur Ru5P n'est pas rapportée par Kuzntesova, tout comme pour YniC. Il est possible que le screening réalisé par Kuznetsova *et al.* ne soit pas valable pour ce susbtrat. Les activités de YKR070W sont élevées comparés aux autres HADs, comprises entre 95,9 nmol/min/mg (F6P) et 245,6 nmol.min^{-1}.mg^{-1} (GAP). YKR070W montre l'activité la plus élevée sur GAP ce qui peut s'avérer problématique *in vivo*. YbiV a une activité comprise entre 48,5 nmol.min^{-1}.mg^{-1} (F6P) et 181,8 nmol.min^{-1}.mg^{-1}(R5P), elle n'a pas de préférence pour le Xu5P (96,7 nnmol.min^{-1}.mg^{-1}). YKR070W est une phosphaatse mitochondriale qui a un profil singulier parmi les HAD phosphatases levuriennes qui ont été décrites comme plus spécifiques que les HADs bactériennes. Cette observation est en accord avec l'origine des mitochondries qui pourraient être le fruit de l'évolution de protobactérie endosymbiotiques.

B) <u>Évaluation de l'activité des phosphatases sur D-xylulose-5P</u>

AraL, TM1254, Dog1 ne sont pas actives sur D-Xu5P dans les conditions testées. Une activité a été sur D-Xu5P a été démontré pour les HADs YKR070W, YbiV, YniC et Pho13. YKR070W est la protéine avec l'activité la plus élevée : 2,3 fois plus active sur Xu5P que YniC, 1,5 fois plus que YbiV. Pho13 a une activité très faible activité, elle est 7 fois moins active qu'YKR070W.

C) Évaluation de la spécificité des phosphatases

On estime que le métabolome bactérien est composé à 35-40% de molécules phosphorylés. Il est essentiel d'identifier ou de développer une phosphatase spécifique pour le D-Xu5P pour ne pas perturber le métabolisme central d'*Escherichia coli.* Cinq des métabolites testés sont des métabolites essentiels (**Tableau 3.6**).

Number	Metabolite	Abbreviation	Building Blocks Produced
1	D-glucose-6-phosphate	G6P	glycogen, LPS
2	D-fructose-6-phosphate	F6P	cell wall
3	D-ribose-5-phosphate	R5P	His, Phe, Trp, nucleotides
4	D-erythrose-4-phosphate	E4P	Phe, Trp, Tyr
5	D-glyceraldehyde-3-phosphate	GAP	lipids
6	glycerate-3-phosphate	3PG	Cys, Gly, Ser
7	phosphoenolpyruvate	PEP	Tyr, Trp
8	pyruvate	PYR	Ala, Ile, Lys, Leu, Val
9	acetyl-CoA	ACA	Leu, lipids
10	2-ketoglutarate	2KG	Glu, Gln, Arg, Pro
11	succinyl-CoA	SCA	Met, Lys, tetrapyrroles (e.g., heme)
12	oxaloacetate	OXA	Asn, Asp, Ile, Lys, Met, Thr, nucleotides

Tableau 3.6 : Les 12 métabolites précurseurs pour la production de biomasse dans *Escherichia coli* (Noor *et al.*, 2010)**.** Encadré en bleu : les métabolites précurseurs testés lors de l'évaluation de la spécificité des HADs *in vitro* : D-glucose-6P,D- fructose-6P, D-ribose-5P, D-erythrose-4P, D-glycéraldéhyde-3P.

Le D-glucose-6P est nécessaire chez *Escherichia coli*, et les bactéries Gram négatif, à la synthèse de lipopolysaccharide (LPS), c'est également le précurseur du glycogène. Le D-fructose-6P est essentiel à la synthèse de la paroi cellulaire. Le D-ribose-5P est le précurseur de synthèse des nucléotides et d'acides aminés (histidine, phénylalanine, tryptophane), tout comme l'érythrose-4P impliqué également dans la synthèse d'acides aminés aromatique et de vitamines. Enfin le glycéraldéhyde-3P participe à la formation de lipides. La déphosphorylation d'un ou plusieurs de ces métabolites précurseurs pourrait perturber l'ensemble du métabolisme d'Escherichia coli.

Bien qu'YKR070W soit la plus active sur D-Xu5P des phosphatases candidates, du fait de son activité sur GAP, G6P, F6P et R5P sa surexpression dans *Escherichia coli* peut s'avérer problématique pour la croissance et la production d'acide glycolique. Il en est de même pour YbiV. Pho13 est active sur une gamme de substrat plus réduite, ce qui limite les risques de perturbations métaboliques.

2.3.2 Activité enzymatique des mutases

Les phosphoglucomutases purifiées PGM1, PGM2 et PGM3 ont été testées sur leur substrat de référence, le Glucose 1P à une concentration finale de 8 mM. Toutes ont montré une activité sur ce substrat (PGM1 : 0,05 U/mg, PGM2 : 3 U/mg, PGM3 : 0,7 U/mg) comme attendu, mais les phosphoglucomutases n'ont pas montré d'activité sur le Xu5P

2.3.3 Conclusion

La phosphatase ou mutase idéale doit être active et assez spécifique du D-Xu5P pour ne pas déphosphoryler d'autres métabolites phosphorylés important du métabolisme de la cellule. L'activité des phosphatases a été mesurée sur G6P, F6P, R5P, Ru5P, E4P, X5P et GAP. Aucune des enzymes testées n'a le profil idéal recherché, un travail d'ingénierie pour augmenter l'activité de Pho13 ou rendre YKR070W spécifique pourrait être envisagé. Il est nécessaire d'exprimer les phosphatases *in vivo* dans *Escherichia coli* afin d'évaluer l'activité sur D-Xu5P dans les conditions intracellulaires. Les activités mesurées *in vitro* permettent d'évaluer l'affinité et la spécificité des phosphatases sur une gamme de substrats choisis dans des conditions se rapprochant des conditions intracellulaires mais elles ne peuvent pas tout à fait reproduire les conditions *in vivo* (concentration en ions, disponibilité de substrat, compétition de susbtrats,*etc*…). Il nous apparait donc nécessaire d'évaluer si ces HADs peuvent remplir l'objectif initial de déphosphoryler le D-Xu5P *in vivo*.

2.4 Evaluation de l'activité des phosphatases et mutases sur D-xylulose-5P *in vivo*

Le deuxième volet de ce chapitre consiste à tester les phosphatases et mutases *in vivo*. Pour déterminer si les activités *in vitro* observées sont pertinentes *in vivo* un crible fonctionnel a été développé où la croissance est directement liée à l'activité de l'enzyme étudiée sur D-xylulose-5P.

2.4.1 Principe du criblage fonctionnel *in vivo*

Le double mutant Δ*tktA* Δ*tktB*, dépourvu de trancétolases, est incapable de croitre sur pentoses. Les trancétolases sont des enzymes essentielles de la voie des pentoses phosphate, elles font le lien entre la voie des pentoses phosphate et la glycolyse et sont impliquées dans le catabolisme des pentoses et la formation de D-ribose-5P et de D-erythrose-4P, un précurseur des acides aminés aromatiques et de la vitamine B6 (pyridoxine). *Escherichia coli* contient 2 isoenzymes de transcétolases : TktA et TktB. Une souche dépourvue des trancétolases complémentée en acides aminés aromatiques et en pyridoxine peut croître sur D-glucose, mais est incapable de pousser sur D-xylose ou L-arabinose. Un triple mutant Δ*xylB* Δ*tktA* Δ*tktB* exprimant la voie synthétique X1P est capable de croitre sur xylose via la voie X1P qui génère du DHAP alimentant le métabolisme carboné central (glycolyse et cycle de Krebs) tandis qu'il sera incapable de croitre sur L-arabinose (

Figure 3.*12*). Sur L-arabinose en l'absence de trancétolase, le D-Xu5P s'accumule avec le D-ribulose-5P, D-ribose-5P mais le carbone ne peut accéder à la glycolyse. L'expression d'une phosphatase ou

mutase capable de déphosphoryler le D-Xu5P rendrait la croissance possible en générant respectivement du D-xylulose ou D-xylulose-1P, qui alimente la voie X1P. Le DHAP formé par la voie synthétique est directement utilisé par la glycolyse afin de satisfaire le besoin de la cellule en précurseurs, équivalents réducteurs et énergie pour la croissance.

Le mutant Δ*xylB* Δ*tktA* Δ*tktB* Δ*glcD* exprimant la voie X1P (ScreenX1P) a été construit pour le criblage des phosphatases et mutases candidates par test de croissance (

Figure 3.*12*). Une phosphatase ou mutase active sur Xu5P pourra produire de la biomasse et de l'acide glycolique à partir de L-arabinose. Afin que l'acide glycolique puisse s'accumuler le gène *glcD*, codant pour l'aldéhyde déshydrogénase qui catalyse l'oxydation de l'acide glycolique en glyoxylate a été délété.

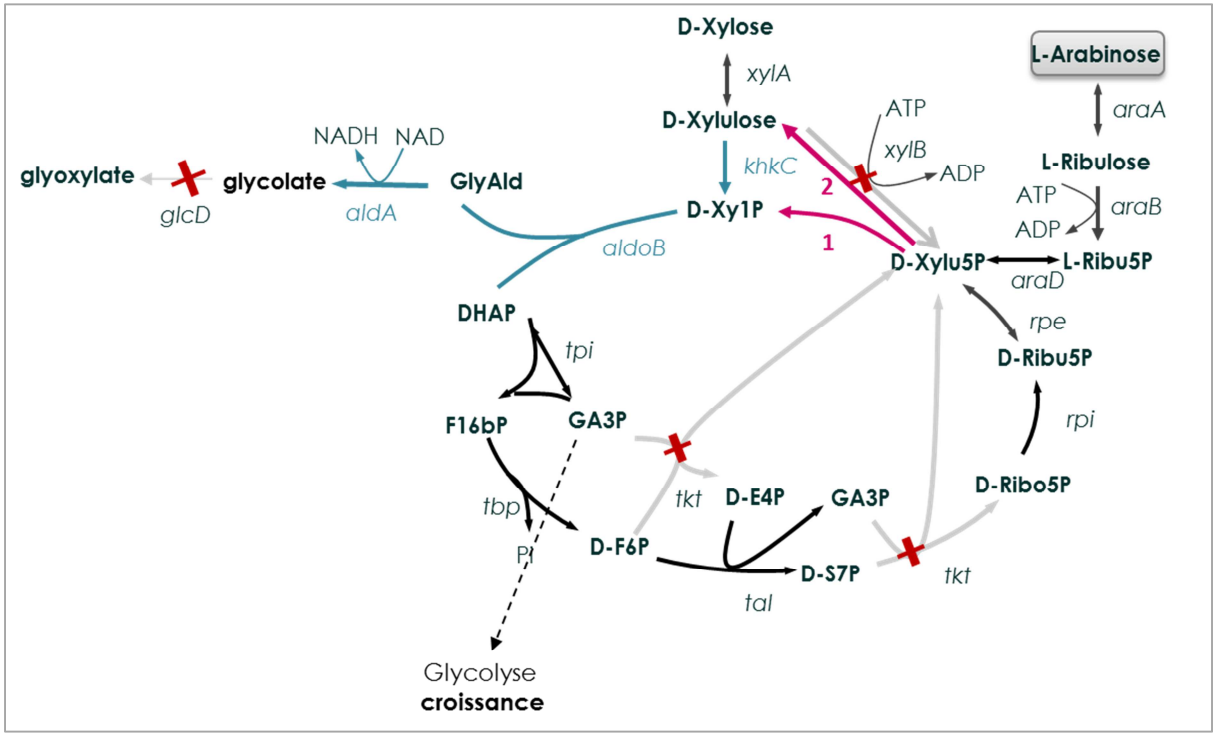


Figure 3.12: Mutant xylB tktA tktB glcD exprimant la voie X1P et une mutase (1) ou phosphatase (2). *La voie X1P est en bleu, les enzymes criblés en violet, (1) : D-xylulose-5P mutase, (2) D-xylulose-5P phosphatase. Les croix rouges situent les délétions des trancétolases tktA, tktB ; de la xylulokinase xylB et la glycolate déshydrogénase glcD. Ribu-5P : ribulose-5P, Ribo5P : ribose-5P, S7P : sedoheptulose-7P, G6P : glucose-6P, F6P : fructose-6P, E4P : erythrose-4P, F16bP : Fructose 1,6bisP, GA3P : glycéraldéhyde-3P, Xu5P : xylulose-5P.*

2.4.2 Évaluation de l'effet de l'expression des phosphatases et mutases sur la croissance et la production d'AG

La souche MG1655 Δ*xylB* Δ*tktA* Δ*tktB* Δ*glcD* exprimant la voie X1P (ScreenX1P) a été transformée avec les vecteurs pACT3 portant les phosphatases ou mutases candidates générant ainsi 10 souches. L'étude *in vitro* avait montré que la majorité des phosphatases étaient peu spécifiques avec des activités sur les précurseurs essentiels d'*Escherichia coli.* YKR070W avait en particulier montré une activité sur D-xylulose-5P mais une absence de spécificité susceptible de perturber le métabolisme carboné centrale. YbiV, YniC et Pho13 avaient monté une activité sur D-Xu5P plus réduite. Pour évaluer l'impact de la surexpression des phosphatases ou mutases sur la production d'acide glycolique et la croissance dans *Escherichia coli* les souches dérivées de ScreenX1P ont été mises en culture sur xylose. Les résultats sont présentés dans le **Tableau 3.8**.

	phosphatase ou mutase	xylose consommé (g/L)	Biomasse		Acide glycolique	
			concentration finale (g/L)	rendement (g/g)	concentration finale (g/L)	rendement (g/g)
ScreenX1P	-	**3,89**	**0,40**	**0,10**	**0,65**	**0,19**
ScreenARAL	araL	2,15	0,16	0,08	0,24	0,12
ScreenTM1254	TM1254	3,20	0,29	0,09	0,51	0,16
ScreenYNIC	yniC	2,10	0,16	0,10	0,34	0,22
ScreenYBIV	YbiV	2,87	0,26	0,09	0,38	0,14
ScreenPHO13	**pho13**	**5,00**	**0,80**	**0,18**	**1,75**	**0,37**
ScreenDOG1	Dog1	2,47	0,21	0,10	0,54	0,26
ScreenYKR070W	YKR070W	2,60	0,28	0,12	0,54	0,24
ScreenPGM1	pgm1	2,92	0,29	0,10	0,50	0,18
ScreenPGM2	pgm2	1,99	0,17	0,12	0,34	0,22
ScreenPGM3	pgm3	3,85	0,23	0,06	0,73	0,17

Tableau 3.7 : Production en acide glycolique de différentes souches d'E.coli ΔxylB ΔtktA ΔtktB ΔglcD dans un milieu contenant du xylose comme source de carbonée. *En vert : le contrôle négatif ScreenX1P n'exprimant ni phosphatase ni mutase. En orange : la souche ScreenPHO13 exprimant Pho13 avec la croissance et la production d'AG les plus élevées. Conditions de croissance : M9 10 g/L xylose, 500 µM L-phénylalanine, 250 µM L-tyrosine, 200 µM L-tryptophane, 6 µM p-aminobenzoate, 6 µM p-hydroxydenzoate et 280 µM de shikimate pH 7. 100h à 37°C, 200 rpm.*

A) Impact de l'expression des phosphatases et mutases sur la croissance

La perturbation du métabolisme cellulaire peut être appréciée par la mesure de la densité optique en cours de culture. La biomasse produite à 100h de culture a été mesurée pour chacune des souches (**Tableau 3.7, Figure 3.13**)

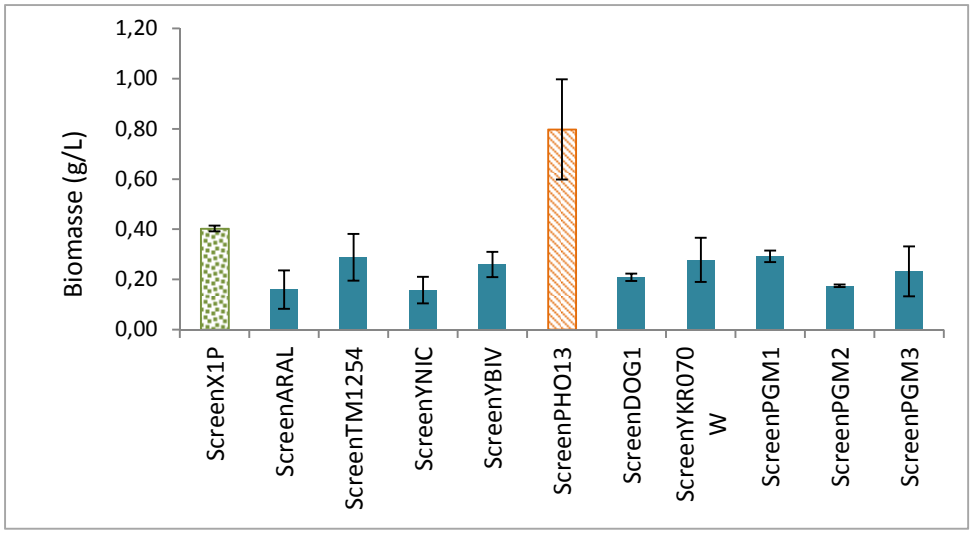


Figure 3.13 : Production de biomasse des souches d'E.coli ΔxylB ΔtktA ΔtktB ΔglcD. *En vert : le contrôle négatif ScreenX1P n'exprimant ni phosphatase ni mutase. En orange : la souche ScreenPHO13 exprimant Pho13 avec la croissance et la production d'AG les plus élevées. Conditions de croissance : M9 10 g/L xylose, 500 µM L-phénylalanine, 250 µM L-tyrosine, 200 µM L-tryptophane, 6 µM p-aminobenzoate, 6 µM p-hydroxydenzoate et 280 µM de shikimate pH 7. 100h à 37°C, 200 rpm. Production de biomasse calculée à partir de la DO600nm.*

D'après les profils d'activité *in vitro*, YKR070W est la HAD au spectre d'activité le plus large et aux activités les plus élevées. A l'inverse, nous avions démontré que TM1254 perdait 86% de son activité à 37°C. Étonnement les souches ScreenYKR070W et ScreenTM1254 présentent la même production de biomasse. Toutes les souches, mis à part ScreenPHO13, ont une production de biomasse inférieure à la souche contrôle ScreenX1P. L'effet de l'expression de ces phosphatases et mutases *in vivo* n'est pas corrélé aux activités catalytiques mesurées *in vitro*. L'impact négatif sur la croissance est homogène parmi les souches, cela indique que la charge métabolique que représente l'expression d'un deuxième plasmide porteur d'une phosphatase ou mutase est à l'origine du ralentissement de la croissance observée pour ces souches, indépendamment de l'enzyme exprimée. L'expression de Pho13 permet de compenser le 'metabolic burden' et améliore la croissance. Chez *S. cerevisiae* l'expression de Pho13 a un impact sur la régulation des gènes du métabolisme carboné central

B) Effet de l'expression des phosphatases ou mutases sur la production d'acide glycolique

La production d'acide glycolique a été mesurée dans le milieu extracellulaire par HPLC au cours des cultures (**Tableau 3.7**). La Figure 3.14 présente la production finale (A) et le rendement (B) à 100h de culture pour chacune des souches.

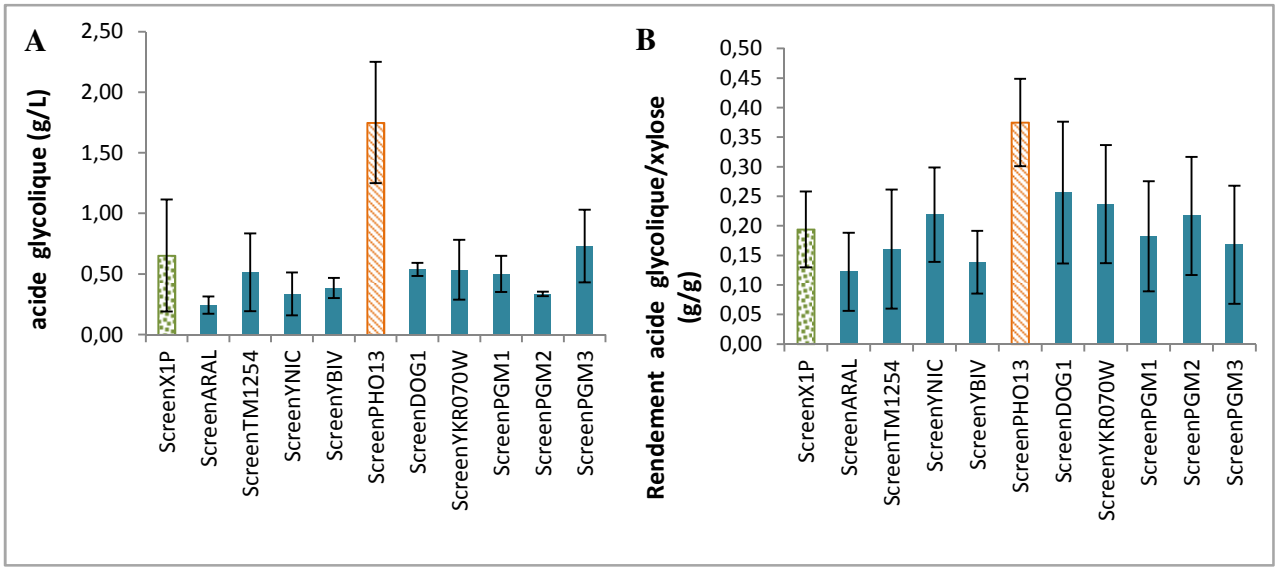


Figure 3.14 : A. Production d'acide glycolique B. Rendement en acide glycolique des souches ΔxylB ΔtktA ΔtktB ΔglcD à partir de xylose. *En vert : le contrôle négatif ScreenX1P n'exprimant ni phosphatase ni mutase. En orange : la souche ScreenPHO13 exprimant Pho13 avec la croissance et la production d'AG les plus élevées. 100h de croissance, 50 mL M9 10 g/L xylose, 500 µM L-phénylalanine, 250 µM L-tyrosine, 200 µM L-tryptophane, 6 µM p-aminobenzoate, 6 µM p-hydroxydenzoate et 280 µM de shikimate pH 7, 37°C, 200 rpm. Production d'acide glycolique mesurée dans le milieu extracellulaire par HPLC.*

Aucune différence significative de production d'acide glycolique n'a été observée par rapport au témoin ScreenX1P hormis pour ScreenPHO13. La souche ScreenPHO13 a une croissance augmentée et présente une production d'AG plus élevée que les autres souches. La souche ScreenPHO13 a également un rendement en biomasse plus élevé que ScreenX1P, le xylose consommé est donc plus efficacement catabolisé du fait la surexpression de Pho13 (Tableau 3.13, Figure 3.12)

La production et le rendement plus élevée d'AG s'explique-t-elle par un plus grand nombre de cellules productrices consommant plus efficacement le xylose? La production d'AG ramené à la biomasse soit le rendement AG/biomasse de ScreenX1P est de 1,625 g/g, tandis qu'il est de 2,18 g/g pour ScreenPHO13. La souche ScreenPHO13 a un métabolisme amélioré pour l'assimilation du xylose et la production d'acide glycolique, ce qui pourrait provenir d'une meilleure efficacité de la voie synthétique X1P.

La voie synthétique X1P convertit le xylose en glycolaldéhyde et DHAP en quantité équimolaire. Le DHAP entre dans la voie de basse de la glycolyse pour fournir les éléments nécessaires à la croissance tandis que le glycolaldéhyde est oxydé en acide glycolique. Si la voie synthétique X1P est plus efficace en présence de Pho13 le rendement en glycolaldéhyde et donc en acide glycolique est augmenté tout comme la production de biomasse. Un argument vient appuyer cette hypothèse. Nous avons pu également observer une excrétion de xylulose dans le milieu lors des cultures. Le xylulose est un intermédiaire de la voie X1P à l'interface entre le métabolisme naturel et la voie synthétique X1P. En présence de Pho13 la sécrétion de xylulose est réduite comparé au contrôle (données non présentées). Il est possible que Pho13 ait une action globale sur le métabolisme central, améliorant la consommation de DHAP ce qui 'tirerait' le carbone à travers la voie X1P. En 2016, une étude publiée dans Nature Chemical Biology a menée à une avancée majeure dans la compréhension du rôle de Pho13 (Collard *et al.*, 2016). Pho13 a une activité sur le 4-P-erythronate, produit de l'activité secondaire de la glycéraldéhyde-3P déshydrogénase (GAPDH). Le 4-P-erythronate un inhibiteur de la gluconate-6P déshydrogénase (Zwf), une enzyme de la voie des pentoses-P. Si cette étude est applicable à *Escherichia coli*, la surexpression de Pho13 permettrait un flux augmenté dans le métabolisme carboné central.

2.4.3 Evaluation de l'activité des phosphatases et mutases sur xylulose-5P *in vivo*

L'étude la croissance des souches de criblage sur L-arabinose a montré qu'elles étaient capables d'assimiler l'arabinose, mais aucune production de biomasse ou d'acide glycolique n'a été observée. Le criblage par test de croissance a révélé que les mutases et phosphatases ne sont pas assez actives sur D-Xu5P *in vivo* pour permettre une croissance sur arabinose. Le contrôle de la croissance sur xylose a montré que la voie X1P était fonctionnelle en présence des phosphatases et mutases. Les phosphatases et mutases testées ne sont donc pas suffisamment actives et spécifiques du D-Xu5P pour permettre la croissance et la synthèse d'acide glycolique à partir de L-arabinose *via* la voie X1P. Ce résultat est cohérant avec les activités des enzymes candidates sur Xu5P mesurées *in vitro*, les phosphatases ont montré un manque de spécificité et une faible activité sur le D-Xu5P tandis les phosphoglucomutases n'ont montré aucune activité sur le D-Xu5P, ce qui explique l'absence de croissance sur l'arabinose.

3. CONCLUSION

La voie conçue repose sur l'expression d'une D-Xu5P phosphatase ou mutase afin d'ouvrir la voie X1P, initialement conçue pour la valorisation du xylose, au glucose et à l'arabinose. L'évaluation *in vitro* et *in vivo* de l'activité des phosphatases et mutases candidates sur D-Xu5P a révélé que leur activité n'était pas suffisante pour rendre la voie X1P accessible à d'autres sucres. Toutefois YKR070W testé dans cette étude, possède un profil intéressant. Son activité sur Xu5P est la plus élevée et l'étendue de la gamme de substrat dont il catalyse la déphosphorylation pourrait être un avantage d'un point de vue évolutif. Les HADs non spécifiques fourniraient un pool de spécificité potentielle pouvant être sélectionné au cours des mutations spontanées pour réguler le métabolisme bactérien. En considérant YKR070W comme un « réservoir de spécificité », il est envisageable de conserver ses capacités catalytiques et la rendre spécifique du D-Xu5P par des approches évolution dirigée et mutagénèse aléatoire. L'identification des mutations ayant un impact sur la spécificité de substrat alimentera les travaux sur la compréhension de la spécificité des HADs.

Par ailleurs nous avons pu constater que la surexpression de l'enzyme Pho13 de *Saccharomyces cerevisiae* semblait avoir un effet positif sur le métabolisme carboné central bactérien. L'effet de Pho13 sur la croissance d'*Escherichia coli* MG1655 devra être évalué, ainsi que l'effet sur la production d'acide glycolique par la voie X1P de la souche MG1655 Δ*xylB* Δ*glcD khkC-aldoB-aldA* (Alkim *et al.*2015,2016) Si l'effet de Pho13 est confirmé, des analyses globales de type '-omiques' seront employées pour comprendre comment Pho13 agit sur le métabolisme bactérien. Chez *S.cerevisiae* Pho13 est impliqué dans la détoxification cellulaire en limitant l'accumulation de 2 inhibiteurs du métabolisme carboné central, l'érythronate-4P et le fructose 2,6-bisphosphate, via leur déphosphorylation en érythronate et fructose 6 phosphate. Le fructose 2,6-bisphosphate est un régulateur du métabolisme carboné présent chez les eucaryotes mais absent du métabolisme procaryote ((Michels and Rigden, 2006). L'érythronate-4P présent chez *E.coli* est un inhibiteur de la glucose-6-phosphate-1-deshydrogénase, enzyme qui catalyse l'entrée du glucose dans la voie des PP. Dans le cas de la croissance sur D-xylose cette enzyme n'est pas requise, le niveau d'érythronate-4P intracellulaire n'est donc probablement pas lié à l'amélioration de la croissance. Pho13 régule également la transcription de gènes de la voie des PP tels que Tal1 codant pour la trancétolase, la délétion de Pho13 augmente l'activité de ces gènes de façon bénéfique pour la production d'éthanol par *S.cerevisiae*. Il serait intéressant d'évaluer l'effet de Pho13 sur le niveau de transcription des gènes de la voie des PP d'*E.coli.*

Les phosphatases n'ayant pas permis la production d'AG à partir de l'arabinose via la voie X1P, une nouvelle approche a été développée, qui fait l'objet du prochain chapitre (**Chapitre IV**).

CHAPITRE IV

DEVELOPPEMENT D'UNE VOIE CYCLIQUE DE PRODUCTION D'ACIDE GLYCOLIQUE A PARTIR DE GLUCOSE, XYLOSE, ARABINOSE DANS *ESCHERICHIA COLI*

1. INTRODUCTION

1.1 Synthèse d'acide glycolique par voie microbienne

La biomasse lignocellulosique est une source de carbone renouvelable abondante utilisée comme matière première pour la production de produits chimiques et de carburants. L'hydrolyse de la cellulose et de l'hémicellulose génère du glucose, du xylose et de l'arabinose. Les études de production d'acide glycolique (AG) par voie microbienne ont abouti au développement de voies optimisées pour la conversion du D-glucose ou du D-xylose chez la bactérie et la levure (Cam *et al.*, 2016). Il n'existe à l'heure actuelle aucune voie non-naturelle capable de convertir à elle seule ces deux substrats en AG. La valorisation des pentoses et des hexoses requièrent la combinaison de voies synthétiques optimisées pour les pentoses et du cycle du glyoxylate naturel ingénieré pour la production d'AG à partir d'hexose. Le rendement de ces voies est limité par la décarboxylation du pyruvate en l'acétyl-CoA.

Les voies métaboliques d'assimilation et d'utilisation des sucres ont été façonnées par l'évolution pour permettre une croissance cellulaire optimale mais elles ne le sont pas pour des objectifs biotechnologiques, et en particulier ceux de produire des métabolites avec rendement en carbone maximum. Ces pertes de carbone sont non seulement pénalisantes pour le rendement des procédés biotechnologiques mais elles posent aussi un problème important sur le plan économique et écologique. Nous proposons ici de remanier le métabolisme carboné central d'*Escherichia coli* pour conserver le carbone et limiter sa libération.

1.2 Conversion des sucres lignocellulosiques en acide glycolique avec une conservation optimale du carbone

Une nouvelle voie de conversion des hexoses et pentoses en acide glycolique a été conçue, elle se compose d'une D-arabinose-5-phosphate isomérase (KdsD), une fructose-6-phosphate aldolase (Fsa) et d'une aldéhyde déhydrogénase (AldA). Ces enzymes sont naturellement présentes dans *Escherichia coli* mais une surexpression des gènes de *kdsD*, *fsa* et *aldA* est nécessaire pour permettre la conversion de D-ribulose-5-phosphate (Ru5P) en AG. Le Ru5P est un métabolite de la voie des pentoses phosphate (PP) commun au catabolisme du D-glucose, D-xylose et L-arabinose. Ainsi la surexpression de ces enzymes permet la conversion des sucres lignocellulosiques en AG à partir du Ru5P. Le Ru5P est converti en D-arabinose-5-phosphate (A5P) par KdsD, celui-ci est clivé par Fsa en glycéraldéhyde-3-phosphate (GAP) et glycolaldéhyde (GO), oxydé en glycolate par AldA.

Le GAP est naturellement utilisé pour alimenter la partie basse de la glycolyse, le cycle de Krebs et fournir à la cellule en précurseurs et co-facteurs réduits nécessaire à la croissance. Nous avons conçu une voie cyclique de production d'AG en inactivant la glycéraldéhyde-3-phosphate déshydrogénase A (GapA). Le GAP est réutilisé pour la production d'AG via la voie des PP et la voie KdsD-Fsa-AldA (KFA). Cette voie de synthèse cyclique permet la conversion du D-glucose, D-xylose et L-arabinose en AG avec un rendement maximum théorique est de 3 mol/mol à partir de glucose et 2,5 mol/mol à partir des pentoses (L-arabinose et D-xylose), ce qui est à notre connaissance le rendement théorique le plus élevé pour la synthèse d'AG. L'étude se divise en deux parties. Une première partie réalisée *in vitro* a permis de démontrer la faisabilité de la production d'AG par la voie KFA à partir d'enzymes purifiées. La deuxième partie apporte la preuve de concept *in vivo* de la production d'AG par la voie KFA à partir d'hexose et pentoses.

2. RESULTATS

2.1 Conception d'une nouvelle voie de production d'acide glycolique

2.1.1 Présentation de la voie KdsD-Fsa-AldA

La voie non naturelle présentée dans la **Figure 4.1** est composée de 3 enzymes naturellement présentes chez *Escherichia coli* : l'arabinose-5-phosphate isomérase (KdsD), la fructose-6-phosphate aldolase (Fsa) et l'aldéhyde déshydrogénase (AldA).

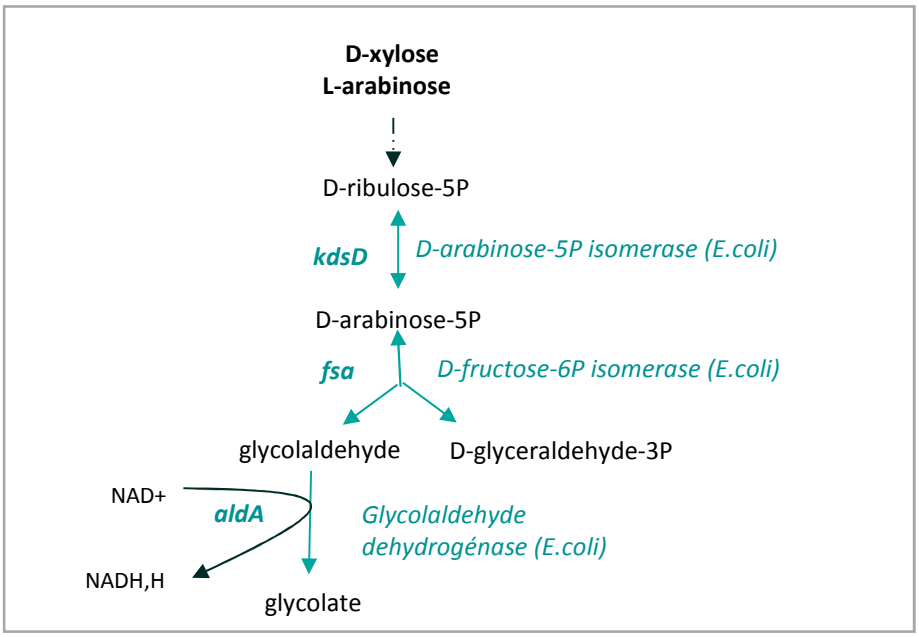


Figure 4.1 : Les trois réactions clefs pour la production d'acide glycolique, isolées du métabolisme carboné central d'Escherichia coli

Tableau 4.1 : Constantes catalytiques des enzymes KdsD, Fsa et AldA d'Escherichia coli sur les substrats d'intérêt de l'étude- *Ru5P – Ribulose-5P, Ara5P : Arabinose-5P, F6P : Fructose-6P, GO : glycolaldéhyde.*

Enzyme	Abb.	EC #	Org.	Vmax (U/mg)	Km (mM)	Reference
D-arabinose 5P isomerase	KdsD	5.3.1.13	*E.coli*	1,14	Ru-5P-1,34	Meredith et Woodard, 2003
Fructose 6P aldolase	Fsa	4.1.2-	*E.coli*	0,652	Ara5P-0,26	Garrabou *et al.*, 2009
				0,34	F6P-9	Garrabou *et al.*, 2009
Aldehyde dehydrogenase	AldA	1.2.1.3	*E.coli*	19	GO-0,38	Baldomà et Aguilar, 1987

Fsa est une aldolase qui a la particularité d'accepter des donneurs non phosphorylés tels que le dihydroxyacétone (DHA), l'hydroxyacétone, l'hydroxybutantone, ou le glycolaldéhyde (GO) au lieu de leurs (couteux) analogues phosphorylés ce qui en fait une cible attrayante pour la synthèse biocatalytique. La polyvalence de Fsa de type sauvage a été démontrée dans le cadre de synthèses conduisant à la production de sucres et d'analogues phosphorylés. Lors de sa découverte en 2001 par Schürmann et Sprenger, Fsa a été décrite pour sa capacité à cliver le Fructose-6-phosphate (F6P) en DHA et GAP. Garrabou *et al.*(2009) ont mis en évidence l'activité de Fsa sur D-arabinose-5P (A5P) . Ce dernier est clivé en GO et GAP avec une efficacité catalytique 65 fois plus élevée que sur F6P. Du fait de la faible affinité de Fsa pour le F6P (K_m : 9 mM) nous considèrerons cette activité négligeable (**Tableau 4.1**).

L'enzyme AldA catalyse l'oxydation du GO en AG issus du clivage de l'A5P. A ce stade de la conception de la voie, la synthèse d'AG à partir d'A5P via l'expression de Fsa et AldA était envisageable. Nous avons recherché dans la littérature une enzyme capable de convertir un intermédiaire du catabolisme des C5 (pentoses) et C6 en A5P. Chez *Escherichia coli* la D-arabinose-5P isomérase (KdsD) interconvertit l'A5P en Ru5P. Le xylose, l'arabinose et le glucose assimilés par la voie des PP en Ru5P pourraient être convertis par la voie non-naturelle KdsD-Fsa-AldA en AG.

2.1.2 Evaluation de la faisabilité de la production d'acide glycolique à partir de glucose, xylose et arabinose

Les enzymes KdsD, Fsa et AldA sont impliquées des voies métaboliques distinctes chez *Escherichia coli*. Le rôle biologique de Fsa n'a pas encore été élucidé tandis que KdsD est impliqué dans la voie de synthèse du LPS où elle interconvertit l'A5P en Ru5P, et AldA contribue essentiellement à la production de lactate à partir de méthylglyoxal. Cette déshydrogénase catalyse plus largement l'oxydation d'aldéhydes tels que le GO, le lactaldéhyde et le glyceraldéhyde. Nous pouvons nous demander si l''assemblage' de ces trois réactions pour la conception d'une nouvelle voie de métabolique est thermodynamiquement viable et si leurs efficacités catalytiques sont adaptées pour rediriger efficacement le flux de carbone vers l'acide glycolique (**Figure 4.2**).

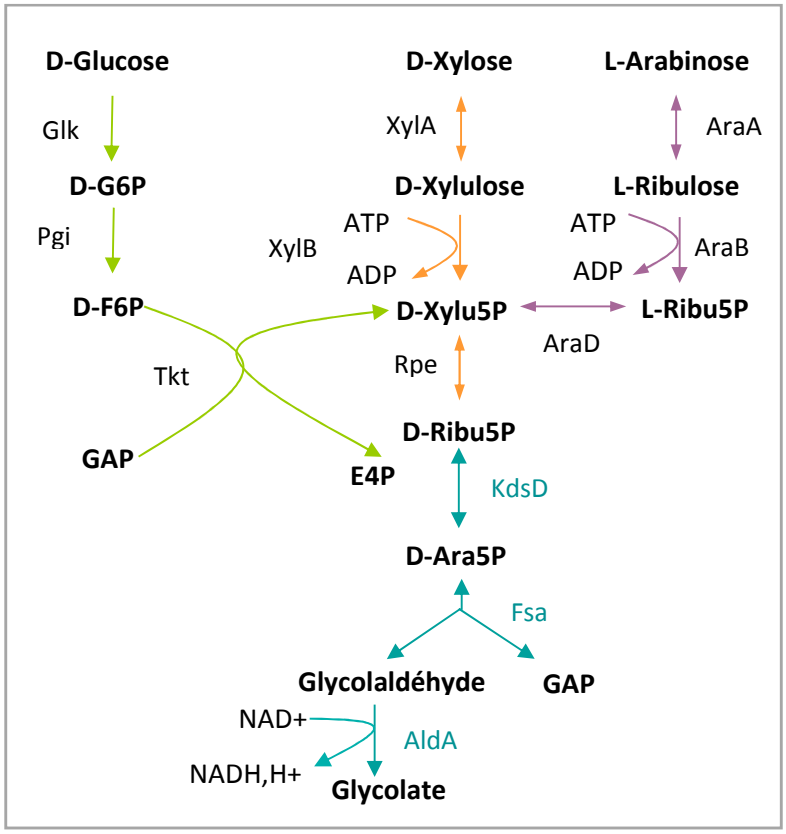


Figure 4.2 : Voies d'assimilation naturelles du D-glucose, D-xylose et L-arabinose combinées à la voie non-naturelle de production d'acide glycolique. *En vert : voie naturelle d'assimilation du glucose, en orange : voie d'assimilation du xylose, en violet : voie d'assimilation du L-arabinose, en bleu-vert : voie non-naturelle de production d'acide glycolique. Glk : glucokinase, Pgi : phosphoglucose isomérase, Tkt : transcétolase, xylA : xylulose isomérase, XylB : xylulokinase, Rpe : Ribulose-5-phosphate-3-épimérase, AraA : arabinose isomérase, AraB : ribulokinase, AraD : L-ribulose-5-phosphate -4-épimérase, KdsD : D-arabinose-5P isomérase, Fsa : Fructose-6-phosphate aldolase, AldA : aldéhyde déshydrogénase*

A) Thermodynamique

Pour qu'une voie métabolique soit thermodynamiquement favorable la somme des énergies libres de Gibbs (ΔrG'°) des réactions la composant doit être négative. En considérant les réactions de façon individuelle, nous constatons qu'une majorité d'entre elles sont thermodynamiquement favorables (ΔrG'° < 0) (**Figure 4.3**). Seules les réactions catalysées par Pgi, KdsD et Fsa ont un ΔrG'° positifs indiquant que la conversion du Ru5P en GAP et GO est défavorable. Autrement dit, la réaction a lieu spontanément dans le sens inverse, par conséquent le niveau de GO et GAP doit être maintenu bas pour favoriser le clivage de l'A5P. Pour cela une expression élevée d'AldA est nécessaire. L'oxydation du GO est la seule réaction irréversible de la voie KFA, elle seule peut tirer le flux efficacement vers l'AG.

De façon plus globale, les voies de conversion de glucose, arabinose et xylose en AG ont un ΔrG'° négatif, la synthèse d'AG est donc favorable à partir de ces trois substrats comme le présente la **Figure 4.3**. Deux réactions, sans lesquels la synthèse d'AG serait défavorable, sont essentielles : les réactions de phosphorylation catalysées par XylB (xylulose kinase), AraB (ribulose kinase) ou Glk (glucokinase), et l'oxydation du GO catalysée par AldA.

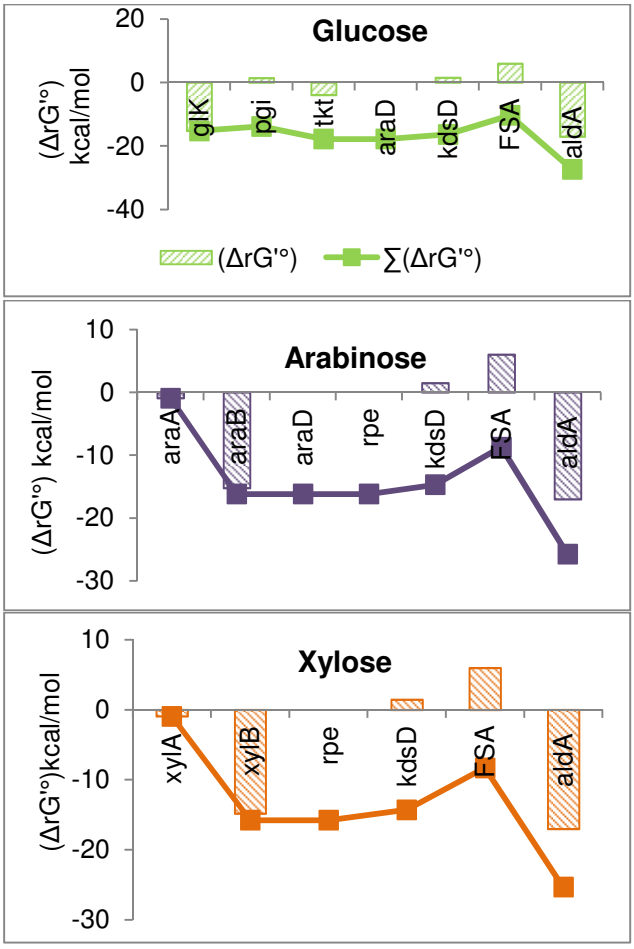


Figure 4.3 : Energie libre de Gibbs (ΔrG'°) des réactions des voies de conversion du glucose, arabinose et xylose en acide glycolique. *Les barres représentent l'énergie libre de Gibbs des réactions, les lignes de points représentent la somme des énergies libre de Gibbs de gauche vers la droite. Glk : glucokinase, Pgi : phosphoglucose isomérase, Tkt : transcétolase, xylA : xylulose isomérase, XylB : xylulokinase, Rpe : Ribulose-5-phosphate-3-épimérase, AraA : arabinose isomérase, AraB : ribulokinase, AraD : L-ribulose-5-phosphate -4-épimérase, KdsD : D-arabinose-5P isomérase, Fsa : Fructose-6-phosphate aldolase, AldA : aldéhyde déshydrogénase.*

B) Efficacité catalytique

L'efficacité catalytique (k_{cat}/K_m) des enzymes impliquées dans la conversion du glucose, xylose et arabinose en AG par la voie non-naturelle sont présentées dans les **Tableaux 4.2, 4.3** et **4.4**. L'efficacité catalytique renseigne sur le taux de catalyse d'une enzyme. Un k_{cat}/K_m élevé indique que le substrat est rapidement converti à son contact avec l'enzyme. A l'inverse une efficacité catalytique faible met en évidence un taux de catalyse bas. Le flux carboné se distribuera de façon optimale à travers une voie métabolique dont les efficacités catalytiques sont homogènes.

Tableau 4.2 : Efficacité catalytique des enzymes impliquées dans la conversion du glucose en acide glycolique dans Escherichia coli.

Etape	Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	Référence
1	Glk	92,16	0,78	1,18E+05	Meyer *et al.*, 1997
2	Pgi	-	-	-	-
3	Tkt	149	1,1	1,36E+05	Sprenger *et al.*, 1995
4	**Rpe**	67000	1,6	**4,19E+07**	Sobota et Imlay, 2011
5	KdsD	1,34	2,3	2,57E+05	Meredith et Woodard, 2006
6	**Fsa**	0,68	0,561	**1,22E+03**	Schürmann et Sprenger, 2001
7	AldA	18,33	0,14	1,31E+05	Rodríguez-Zavala *et al.*, 2006

Tableau 4.3 : Efficacité catalytique des enzymes impliquées dans la conversion de l'arabinose en acide glycolique dans Escherichia coli.

Etape	Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	Référence
1	AraA	-	-	-	-
2	AraB	80	0,14	5,71E+05	Lee *et a*l. 2001
3	AraD	36,5	0,415	8,80E+04	Lee *et al.*, 2000
4	Rpe	67000	1,6	**4,19E+07**	Sobota et Imlay, 2011
5	KdsD	1,34	2,3	2,57E+05	Meredith et Woodard, 2006
6	**Fsa**	0,68	0,561	**1,22E+03**	Schürmann et Sprenger, 2001
7	AldA	18,33	0,14	1,31E+05	Rodríguez-Zavala *et al.*, 2006

Tableau 4.4 : Efficacité catalytique des enzymes impliquées dans la conversion du xylose en acide glycolique dans Escherichia coli.

Etape	Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	Référence
1	**XylA**	5	10	**5,00E+02**	Batt *et al.*, 1990
2	XylB	255	0,29	8,79E+05	Di Luccio *et al.*, 2007
3	**Rpe**	67000	1,6	**4,19E+07**	Sobota et Imlay, 2011
4	KdsD	1,34	2,3	2,57E+05	Meredith et Woodard, 2006
5	**Fsa**	0,68	0,561	**1,22E+03**	Schürmann et Sprenger, 2001
6	AldA	18,33	0,14	1,31E+05	Rodríguez-Zavala *et al.*, 2006

L'ordre de grandeur des efficacités catalytique de la majorité des enzymes impliquées dans la conversion des sucres lignocellulosiques est de 10^5 $M^{-1}s^{-1}$ (*i.e* Glk, Tkt, AraB, AraD, XylB, KdsD, AldA). Les enzymes au k_{cat}/K_m éloigné de cette valeur peuvent être à l'origine d'une distribution de flux carboné suboptimale au sein de la voie concernée. La ribulose-phosphate-3-épimérase, Rpe, qui catalyse l'interconversion du D-Xu5P en Ru5P, a une efficacité catalytique très élevée, du fait d'un k_{cat} de 67000 s^{-1}. Le Ru5P plus rapidement synthétisé qu'il n'est pris en charge par KdsD pourrait s'accumuler. Enfin l'efficacité catalytique de Fsa est faible, du fait d'un 'turnover' ou k_{cat} bas, bien que l'affinité pour l'A5P soit élevée. Le ralentissement du flux carboné pourrait entrainer l'accumulation d'A5P (**Tableau 4.2**, **Tableau 4.3**, **Tableau 4.4**). Il est alors d'autant plus important de surexprimé AldA pour tirer l'A5P vers le GO et le GAP. La surexpression de Fsa devra être par ailleurs être suffisante pour compenser la faible activité catalytique et diriger de façon optimale le flux de carbone vers l'acide glycolique.

2.2 Preuve de concept *in vitro*

2.2.1 Conversion du D-ribulose-5P en acide glycolique

La voie non-naturelle est constituée des enzymes KdsD (D-arabinose 5P isomérase), Fsa (fructose-6P aldolase) et AldA (glycolaldéhyde déshydrogénase) et convertit le Ru5P naturellement synthétisé à partir d'hexose ou pentose en acide glycolique. La démonstration de faisabilité *in vitro* a consisté à reconstruire cette voie avec des enzymes purifiées dans un tampon adapté. Premièrement AldA a été caractérisée sur le GO, puis l'activité de Fsa sur l'A5P a été vérifiée par couplage avec AldA afin de mettre en évidence la formation de GO et par couplage avec la triose phosphate isomérase (Tpi) et la glycérol-3-phosphate déshydrogénase (Gly3PDH) pour mettre en évidence la synthèse de GAP (***Figure 4.4*** **4.4**). Les activités calculées par ces deux méthodes comme attendu ont été identiques. Enfin l'activité de KdsD sur D-Ru5P a été mesurée par couplage avec les enzymes Fsa et AldA en excès. Ce dernier essai enzymatique a permis de vérifier la fonctionnalité de KdsD et de reconstituer la voie de synthèse d'AG entière. Le NADH mesuré par absorbance à 340nm est produit en quantité équimolaire avec l'AG. La production d'AG est vérifiée en HPLC. Le principe de l'essai enzymatique est présenté dans la **Figure 4.4** et les résultats dans le **Tableau 4.5.**

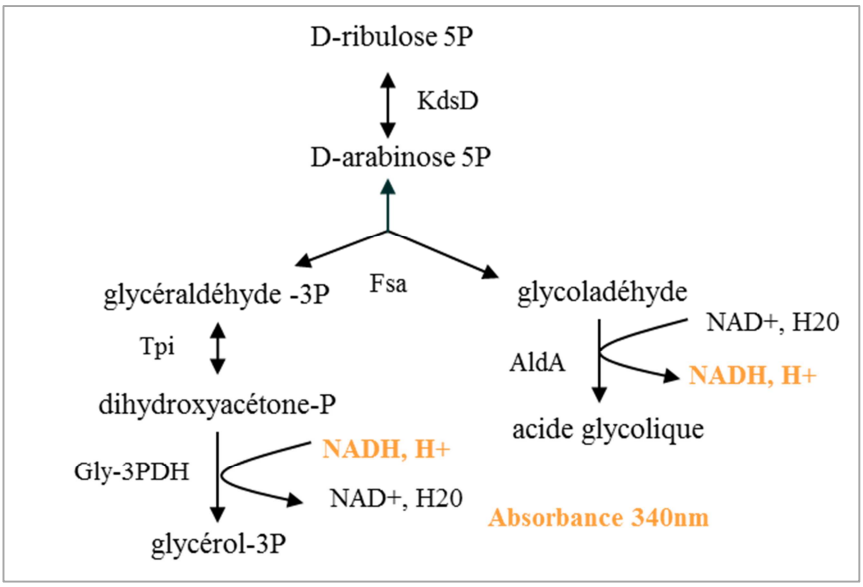


Figure 4.4 : Représentation schématique de l'essai enzymatique de vérification de l'activité de l'arabinose-5P isomérase (KdsD), la fructose-6P aldolase (F6P) et l'aldéhyde déshydrogénase (AldA).

Tableau 4.5 : Constantes catalytiques de l'arabinose-5P isomérase (kdsD), la fructose-6P aldolase (Fsa), l'aldéhyde déshydrogénase (AldA) étiquetées polyhistidine. *(37°C, pH 7)*

Enzyme	**Substrat**	**K_m (mM)**	**V_{max}(U mg^{-1})**	**V_{max}/K_m(.10^{-6}) (min^{-1} mg^{-1})**
KdsD-His	D-ribulose-5P	1,34 ± 0,12	1,14 ± 0,17	851
Fsa-His	D-arabinose-5P	0,652 ±0,15	0,26 ± 0,09	399
AldA-His	glycolaldéhyde	0,21 ±0,05	1,16 ± 0,31	5552

Le mélange réactionnel avec les trois enzymes de la voie en présence de Ru5P a été analysé, 1,5 mM d'AG ont été quantifiés par HPLC. Le NADH et l'AG ont bien été synthétisés en quantité équimolaire. Le Ru5P est converti en AG *in vitro* par l'activité couplée des enzymes KdsD, Fsa et AldA. La voie KFA est fonctionnelle *in vitro.*

La conversion des intermédiaires métabolique en AG a également été vérifiée, afin de s'assurer que les enzymes purifiées étaient fonctionnelles et identifier d'éventuelles activités secondaires. La première enzyme responsable de l'assimilation Ru5P, KdsD, a été retirée du mélange réactionnel. L'ajout de Ru5P en présence de Fsa et AldA sans KdsD induit une faible synthèse de NADH. Cela indique que que Fsa ou aldA a une faible activité secondaire sur Ru5P. L'essai enzymatique présenté dans la **Figure 4.5** démontre que Fsa est active sur Ru5P. Guérard-Hélaine *et al.*, 2014 ont rapporté que Fsa catalyse le clivage du xylulose-5-phosphate (Xu5P) en DHA et glycolaldéhyde-P. Le Ru5P est structurellement très proche du Xu5P, ce sont des diastéréoisomères. Il est possible que Fsa clive le

Ru5P de la même façon que le Xu5P. Le production de l'activité de Fsa sur le Ru5P est ensuite oxydé par AldA. A notre connaissance aucune activité sur glycolaldéhyde-P n'a été rapportée, il est possible que le glycolaldéhyde-P soit instable et se déphosphoryle spontanément en GO à 37°C. Celui-ci est alors oxydé par AldA et génère le NADH mesuré. Une autre hypothèse a été avancée. Les sucres phosphorylés tels que le Ru5P peuvent être dégradés lors du stockage à 4°C ou 37°C, générant du D-ribulose. L'activité de Fsa sur D-ribulose a été testée. L'activité est très faible, une inhibition par le substrat est observée, ce qui écarte cette hypothèse.

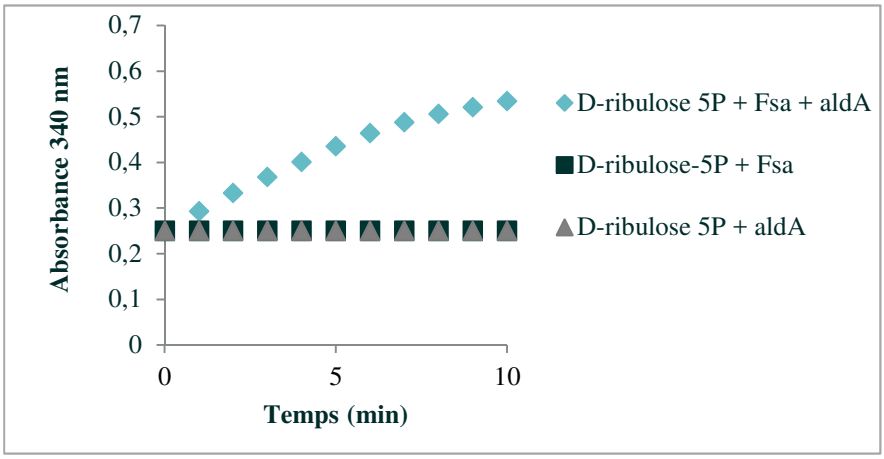


Figure 4.5: Mesure de la production de NADH à 340 nm à 37°C. *5 mM Ribulose-5P, 1-10 µg d'enzyme, pH7, 37°C. Losange bleu : l'essai enzymatique avec Fsa et AldA en présence de 5 mM de D-ribulose-5P, en carré noir*

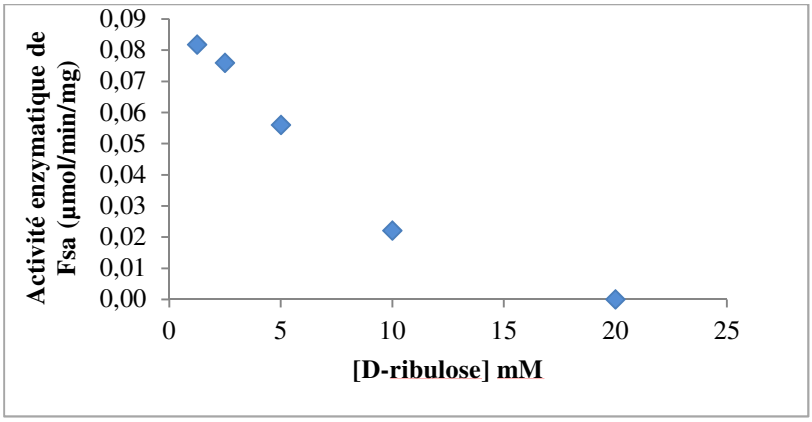


Figure 4.6 : Activité enzymatique de Fsa sur D-ribulose

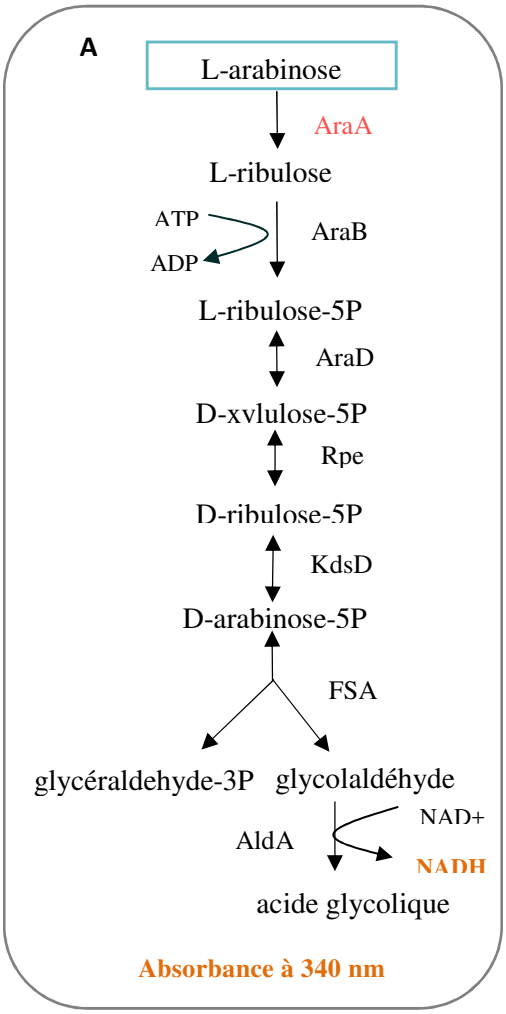
L'activité secondaire de Fsa sur Ru5P a été prise en compte dans le calcul de l'activité de KdsD et ne remet pas en cause la preuve de concept de la conversion du Ru5P en AG par la voie KFA. L'ajout d'A5P en présence d'AldA sans Fsa et l'ajout du GO sans AldA n'ont pas permis la synthèse de NADH et d'AG.

2.2.2 Conversion de L-arabinose ou D-xylose ou D-glucose en acide glycolique

Les enzymes catalysant la conversion du L-arabinose, D-xylose et D-glucose en acide glycolique ont été purifiées (AraA, AraB, AraD, XylA, XylB, Rpe, KdsD, Fsa, AldA) ou commandées (Tkt, Hxk, Pgi) afin de démontrer la faisabilité de ces voies *in vitro*.

A) Conversion du L-arabinose en acide glycolique

La conversion du L-arabinose en acide glycolique et glycéraldéhyde 3P est catalysée par AraA, AraB, AraD, Rpe, KdsD, Fsa et AldA (**Figure 4.7A**). Lorsque que ces enzymes sont mises en présence de L-arabinose, l'apparition de NADH est mesurée au spectrophotomètre (**Figure 4.7B**). La production de NADH est le témoin indirect de la production d'AG à partir de L-arabinose. La conversion du L-arabinose en AG a été ainsi démontrée *in vitro*, nous avons également vérifié la conversion de chacun des intermédiaires de la voie en AG et mis en évidence la présence d'activité secondaires. Ces essais ont permettent également de vérifier la fonctionnalité des enzymes purifiées.



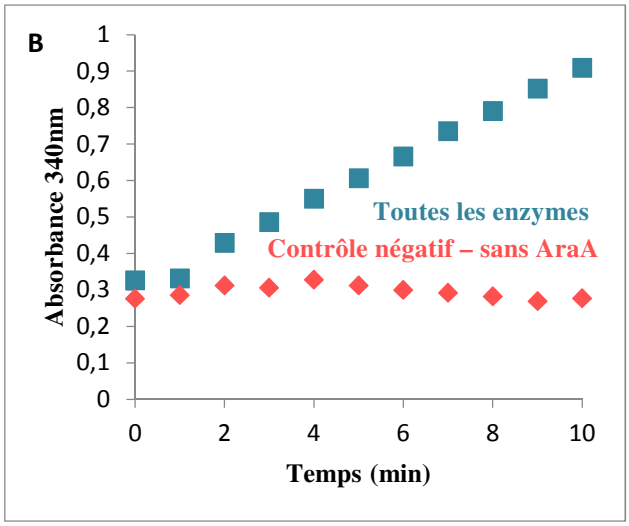


Figure 4.7: A) Représentation schématique de la conversion du L-arabinose en acide glycolique, B) Mesure de l'apparition de NADH à 340 nm *En bleu la réaction avec l'ensemble des enzymes représentées en A) en présence de L-arabinose. En rouge : la réaction contrôle avec toutes les enzymes sauf AraA en présence de L.-arabinose.*

AraA : arabinose isomérase, AraB : ribulose kinase, AraD : L-ribulose-5-phosphate-4 épimérase, Rpe :ribulose-3-phosphate 3 épimérase , KdsD : D-arabinose-5-phosphate isomérase, FSA : fructose-6-phoshate aldolase, aldA : aldéhyde déshydrogénase.

Nous avons dans la partie 2.2.1 étudié la conversion du Ru5P en AG. La démonstration débute donc directement par la conversion du D-xylulose-5P (Xu5P) en AG. Pour cela le Xu5P est mis en présence des enzymes Rpe KdsD,Fsa et AldA. Une activité de 11,93 U est mesurée. Le contrôle sans Rpe révèle une activité élevée de 9,32 U confirmant les données de Guérard-Hélaine *et al.*(2014) ***Tableau 4.6***. En effet le Xu5P est clivé par Fsa en DHA et glycolaldéhyde-P. Comme précédemment pour le Ru5P deux explications à l'activité observée sont possibles (i) le glycolaldéhyde-P est instable à 37°C, son produit de dégradation le glycolaldéhyde est pris en charge par aldA (ii) la déphosphorylation spontanée du Xu5P au cours du stockage à 4°C, ou lors de l'essai à 37°C, génère du D-xylulose que Fsa clivera en DHA et GA (Szekrenyi *et al.*, 2014) pris en charge par AldA. Cette activité est secondaire et ne remet pas en question l'efficacité de la voie de conversion du L-arabinose en AG. La même démarche a été suivie avec le L-ribulose, et le L-arabinose, aucune activité secondaire significative n'a été révélée sur ces substrats (***Tableau 4.6***). Nous n'avons pu nous procurer du L-Ru5P au moment de l'essai mais il serait intéressant d'évaluer l'activité de Fsa sur ce métabolite, Fsa possède une activité sur D-Ru5P il est possible qu'il en soit de même avec son énantiomère. Nous avons démontré dans cette partie la conversion de D-arabinose en AG, cette voie est thermodynamiquement favorable conformément à nos prédictions (**Figure 4.3**).

Tableau 4.6 : Activité enzymatique des enzymes impliquées dans la conversion du L-arabinose en acide glycolique.

Substrate (5mM)	Enzymes	Activité (µmol/min)	Remarques
L-arabinose	**AraA**, AraB, AraD, Rpe, KdsD, Fsa, AldA	11,59	L'apparition de NADH atteste de la synthèse d'acide glycolique, la voie complète est fonctionnelle.
	AraB, AraD, Rpe, KdsD, Fsa, AldA	0,00	
L-ribulose	**AraB**, AraD, Rpe, KdsD, Fsa, AldA	14,11	AraB-HIS est fonctionnelle
	AraD, Rpe, KdsD, Fsa, AldA	2,82	
D-xylulose-5P	**Rpe**, KdsD, Fsa, aldA	11,93	Rpe-HIS est fonctionnelle. L'activité du contrôle sans Rpe s'explique par l'activité de FSA sur D-xylulose-5P ou D-xylulose (produit de dégradation spontanée de D-xylulose 5P).
	KdsD, Fsa, aldA	9,32	

B) Conversion de D-xylose en acide glycolique

La conversion du D-xylose en AG est catalysée par les enzymes XylA, XylB, Rpe, KdsD, Fsa et AldA (**Figure 4.8A**). Lorsque que ces enzymes sont mises en présence de D-xylose, l'apparition de NADH est mesurée au spectrophotomètre (**Figure 4.8 B**). La production de NADH est le témoin indirect de la production d'AG à partir de D-xylose. Lors de l'ajout de XylA au mix réactionnel contenant l'ensemble des enzymes citées nous avons pu mesurer une production de NADH. Le contrôle négatif sans xylA ne montre aucune production de NADH.

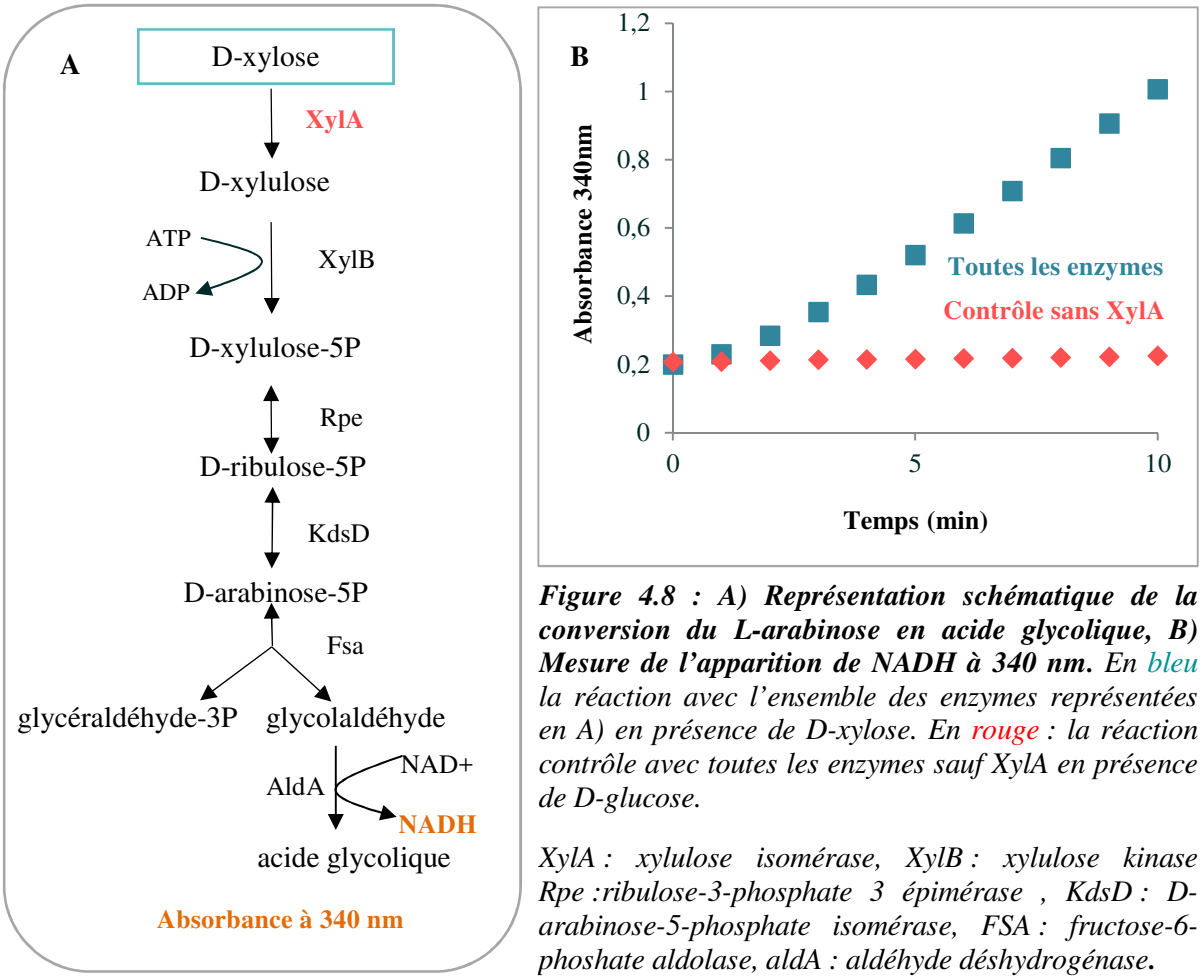


Figure 4.8 : A) Représentation schématique de la conversion du L-arabinose en acide glycolique, B) Mesure de l'apparition de NADH à 340 nm. *En bleu la réaction avec l'ensemble des enzymes représentées en A) en présence de D-xylose. En rouge : la réaction contrôle avec toutes les enzymes sauf XylA en présence de D-glucose.*

XylA : xylulose isomérase, XylB : xylulose kinase Rpe :ribulose-3-phosphate 3 épimérase , KdsD : D-arabinose-5-phosphate isomérase, FSA : fructose-6-phoshate aldolase, aldA : aldéhyde déshydrogénase.

Le xylose est assimilé tout comme l'arabinose par la voie des PP, tous deux sont convertis en Xu5P. Nous avons ici évalué l'activité des enzymes XylA, XylB, Rpe, KdsD, Fsa, AldA sur l'intermédiaire propre à l'assimilation du xylose : le D-xylulose. L'ajout du D-xylulose induit la production de NADH, l'activité est 21,47 U (**Tableau 4.7**). Le xylulose est pris en charge par XylB qui catalyse sa phosphorylation en Xu5P. En l'absence de XylB, une activité de 9,25 U est mesurée. En effet le D-xylulose, est clivé en GO et DHA. AldA oxyde le GO issu de cette réaction, ce qui expliquerait l'activité secondaire mesurée. Szekrenzi *et al*. ont démontré que Fsa pouvait former du xylulose à partir de GO et DHA, les constantes catalytiques sont cependant inconnues.

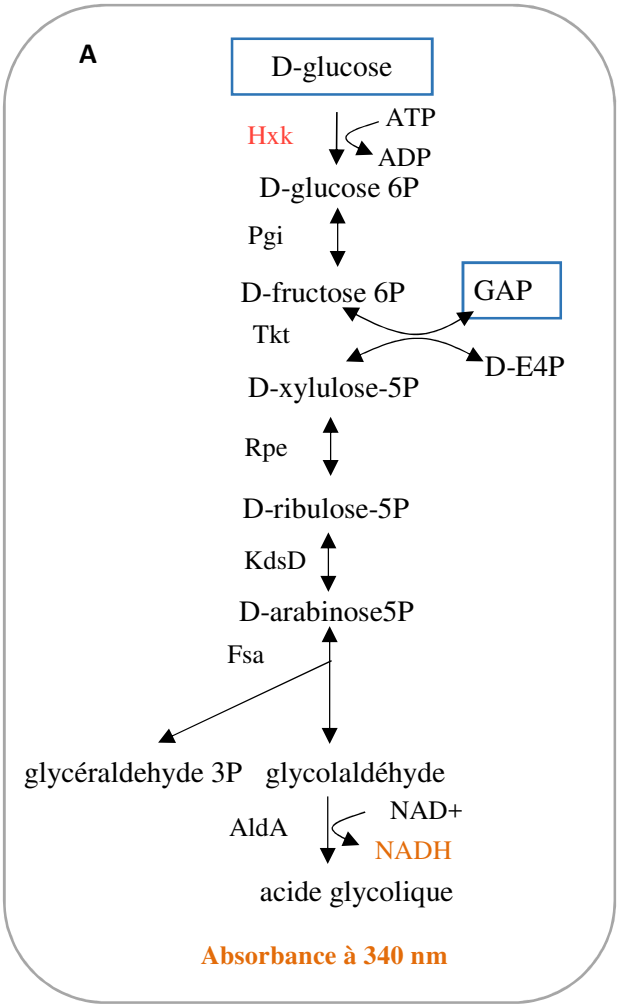
Tableau 4.7 : Activité enzymatique des enzymes impliquées dans la conversion du D-xylose en acide glycolique

Substrate (5mM)	Enzymes	Activité (µmol/min)	Remarques
D-xylose	**XylA,** XylB, Rpe, KdsD, FSA, aldA	15,07	Voie de conversion du D-xylose en acide glycolique fonctionnelle, l'apparition de NADH démontre la synthèse d'acide glycolique.
	XylB, Rpe, KdsD, FSA, AldA	0,00	
D-xylulose	**xylB**, Rpe, KdsD, FSA, AldA	21,47	XylB est fonctionnelle. L'activité secondaire observée est dûe au clivage du D-xylulose en glycolaldéhyde et dihydroxyacetone catalysé par FSA
	Rpe, KdsD, FSA, AldA	9,25	
	Fsa, AldA	9,1	

Le D-xylose est converti en AG, la voie est fonctionnelle. L'activité secondaire observée sur D-xylulose n'est pas majoritaire, mais elle participe à la production d'AG. L'efficacité catalytique de Fsa devra être déterminée sur ce substrat.

C) Conversion du D-glucose en acide glycolique

La conversion du D-glucose en AG a été démontrée en mettant en présence le D-glucose et le GAP avec Hxk, Pgi, Tkt, Rpe, KdsD, Fsa et AldA (**Figure 4.9A**). L'apparition de NADH atteste de la synthèse d'AG, vérifiée par HPLC (**Figure 4.9B**). Le contrôle négatif présente une activité importante.



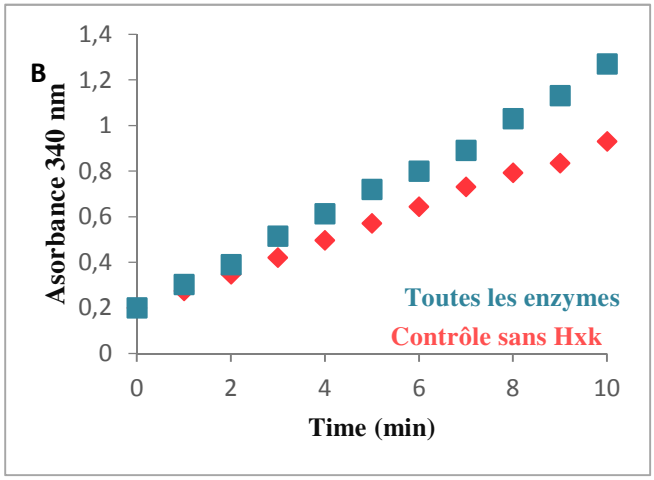


Figure 4.9 : A) Représentation schématique de la conversion du D-glucose en acide glycolique, B) Mesure de l'apparition de NADH à 340 nm. *En bleu la réaction avec l'ensemble des enzymes représentées en A) en présence de D-glucose et glycéraldéhyde-3P. En rouge : la réaction contrôle avec toutes les enzymes sauf Hxk en présence de D-glucose.*

Hxk : hexokinase, Pgi : phosphoglucose isomérase, Tkt : transcétolase, Rpe :ribulose-3-phosphate 3 épimérase , KdsD : D-arabinose-5-phosphate isomérase, FSA : fructose-6-phoshate aldolase, aldA : aldéhyde déshydrogénase.

L'activité secondaire mesurée n'apparait qu'en présence de GAP qui est dans l'essai, avec le F6P, le substrat de la trancétolase Nous avons observé la synthèse de NADH lorsque GAP était en présence des enzymes Rpe, KdsD, FSA et AldA sans trancétolase (**Tableau 4.8**). Parmi ces enzymes AldA est celle ayant une activité sur GAP. Baldoma *et al.* (1998) a démontré qu'AldA d'*Escherichia coli* était active sur glycolaldéhyde, lactaldéhyde, méthylglyoxal et L-glycéraldéhyde. Le GAP est une molécule relativement instable, il se déphosphoryle alors spontanément en glycéraldéhyde. AldA convertit alors le L-glycéraldéhyde en glycérate (Baldomà et Aguilar, 1987). La synthèse de NADH dans le contrôle négatif ne serait donc pas corrélée à la production de glycolate mais de glycérate provenant de l'oxydation du glycéraldéhyde, ce qui devra être vérifié par HPLC. Cette activité secondaire n'est pas pertinente *in vivo*, d'autant plus que le L-glycéraldéhyde n'est pas formé dans *Escherichia coli*.
D'après les résultats obtenus la voie de conversion du D-glucose en AG est fonctionnelle et thermodynamiquement favorable.

Tableau 4.8 : Activité enzymatique des enzymes impliquées dans la conversion du D-glucose en acide glycolique.

Substrate (5mM)	Enzymes	Activité (µmol/min)	Remarques
D-glucose, GAP	**Hxk**, Pgi, Tkt, Rpe, KdsD, Fsa, AldA	23,5	La voie de conversion du D-glucose en acide glycolique est fonctionnelle. L'activité secondaire est importante
	Pgi, Tkt, Rpe, KdsD, Fsa, AldA	17,3	
D-glucose-6P, GAP	**Pgi**, Tkt, Rpe, KdsD, Fsa, AldA	22,2	Pgi fonctionnelle, activité secondaire importante
	Tkt, Rpe, KdsD, KdsD, AldA	16,9	
D-fructose-6P, GAP	**Tkt**, Rpe, KdsD, Fsa, AldA	21,5	Tkt fonctionnelle, activité secondaire identique aux réactions ci-dessus
	Rpe, KdsD, Fsa, AldA	16,82	
D-fructose-6P	Tkt, Rpe, KdsD, Fsa, AldA	0	En l'absence de GAP, il n'y a pas d'apparition de NADH. L'activité secondaire résulte de la présence de GAP dans le milieu réactionnel.
GAP	Tkt, Rpe, KdsD, Fsa, AldA	17,5	Cette réaction confirme qu'une des enzymes a une activité sur GAP
	Rpe, KdsD, Fsa, AldA	17,3	Une de ces 4 enzymes est responsable de l'activité secondaire observée
	Fsa, AldA	17,3	
	AldA	16,9	AldA a une activité sur GAP.

2.2.3 Conclusion

Nous avons démontré la production d'AG à partir de glucose, xylose et arabinose via la voie non naturelle KFA *in vitro*. Cependant Fsa possède une activité sur plusieurs intermédiaires, représentée dans la **Figure 4.10**, dont les constantes catalytiques n'ont pas été rapportées dans la littérature. Il sera important de caractériser l'activité de Fsa sur D-xylulose-5P, D-ribulose-5P et D-xylulose. Nous pouvons nous demander si l'activité sur le D-A5P sera prépondérante *in vivo*. La production de DHA représenterait une perte de carbone, car le DHA n'entre pas dans le métabolisme carboné central. Un contrôle sans KdsD devra être réalisé, si l'activité est fortement réduite cela confirmera que l'activité de Fsa sur l'A5P pour la production d'AG.

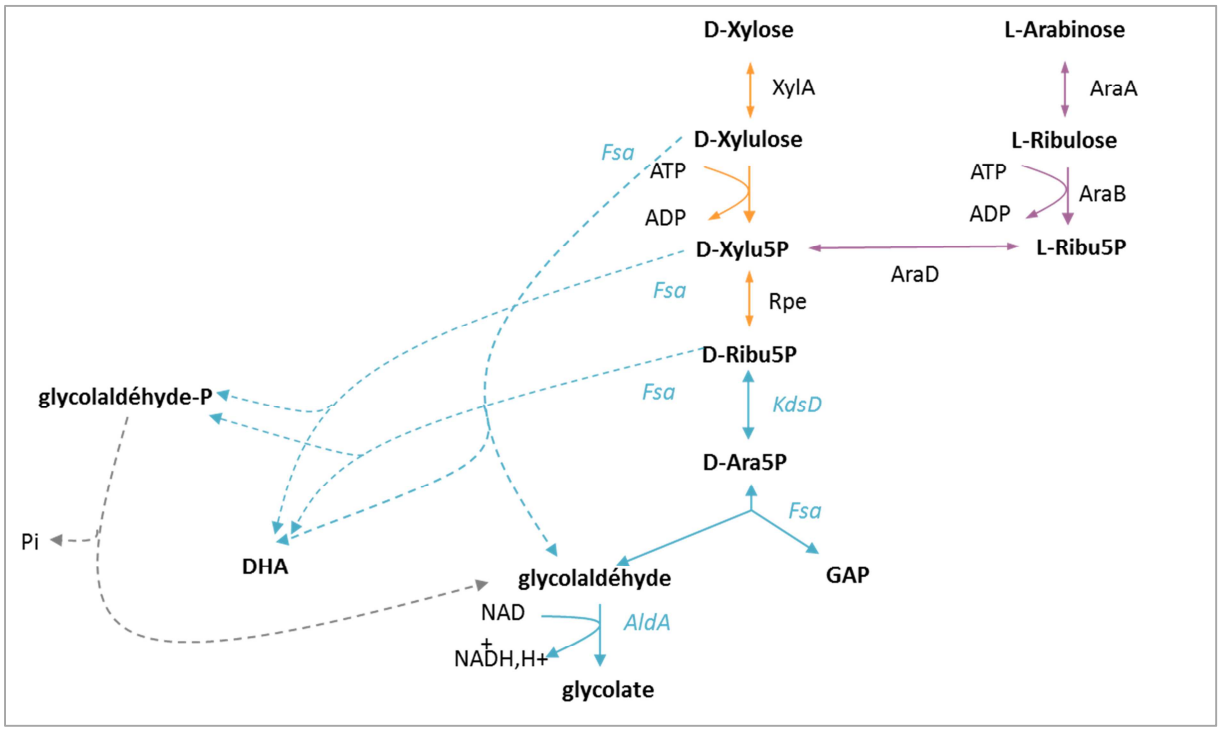


Figure 4.10 : Conversion du D-xylose et L-arabinose en AG in vitro en tenant compte des activités secondaires de la fructose-6P aldolase Fsa. *En pointillé bleu apparaissent les réactions secondaires catalysées par Fsa, les pointillés gris représentent la déphosphorylation spontanée du glycolaldéhyde-P in vitro. Trait plein bleu : la voie KdsD-Fsa-AldA, trait plein orange : vois d'assimilation du D-xylose, trait plein violet : voie d'assimilation du L-arabinose.*

2.3 Preuve de concept *in vivo*

Nous avons conçu une voie non naturelle capable de convertir l'ensemble des monosaccharides contenus dans la biomasse végétale en AG dans *Escherichia coli*. Pour cela deux actions ont été réalisées (i) surexprimer la voie KFA, (ii) développer une souche dont le métabolisme est optimisé pour la conversion du carbone en acide glycolique (**Figure 4.11**).

2.3.1 Ingénierie de la *souche Escherichia coli K12-W3CG*

A) Ingénierie du métabolisme d'Escherichia coli pour favoriser le flux vers la production d'acide glycolique

Choix de la souche W3CG - Nous avons choisi *Escherichia coli K-12 W3CG* pour l'implémenter de la voie de synthèse d'AG, KdsD-Fsa-AldA *in vivo*. Cette souche est dérivée d'*E.coli K12-W3110* ne transcrit pas la glycéraldéhyde 3P-déshydrogénase GapA. Le transposon Tn10 a été intégré dans l'ORF du gène *gapA*. La souche W3CG a donc une activité glycéraldéhyde 3P-déshydrogénase (GAPDH) réduite (Ganter et Plueckthun, 1990). Le GAP n'alimente pas la voie basse de la glycolyse et reste dans la voie des PP où il est utilisé pour la production d'AG par la voie KFA surexprimée. D'autres délétions ont été réalisées afin de diriger le flux vers la production d'AG présentées dans le **Tableau 4.9** et la **Figure 4.11**.

Délétions de glcD* et *arcA*-** Tout d'abord le gène ***glcD codant pour l'aldéhyde déshydrogénase, qui catalyse l'oxydation du glycolate en glyoxylate a été délété (Lord, 1972; Pellicer *et al.*, 1996) . L'AG n'est ainsi plus consommé par GlcD et s'accumule. Pour favoriser la production d'AG à partir de GO le gène ***arcA*** qui régule négativement la transcription du gène *aldA* au niveau génomique a été délété (Pellicer *et al.*, 1999).

Surexpression de galP -La galactose perméase **GalP** d'*Escherichia coli* a été surexprimée par remplacement du promoteur naturel par un promoteur fort constitutif (proD). Alkim *et al.* avait démontré la surexpression de GalP avait un effet bénéfique sur la production d'AG en améliorant l'assimilation du glucose, et xylose (Alkim *et al.*, 2016). Surexprimer ce système de transport permet de s'affranchir du système appelé phosphoenol-pyruvate-sucre phosphotransférase ou PTS qui assure la phosphorylation et le transport du glucose. L'énergie nécessaire et le phosphate proviennent du phospho-enol-pyruvate, un métabolite intermédiaire de la glycolyse, qui est converti en pyruvate (Postma, Lengeler et Jacobson, 1993; Deutscher *et al.*, 2006).

Délétion de mgsA - Par ailleurs, le gène ***mgsA*** codant pour la méthylglyoxal synthase a été délété. Cette enzyme catalyse la conversion de l'excédent de DHAP en méthylglyoxal (MG), un niveau élevé de MG est létal, entrainant des modifications covalentes des protéines, de l'ADN et des lipides, et représente une perte de carbone. (Krymkiewicz, 1973; Colanduoni and Villafranca, 1985; Ferguson *et al.*, 2000).

Délétion de fucA - Le GO, précurseur de l'AG, peut être utilisé par la L-fuculose-phosphate aldolase ***FucA*** pour la formation de D-ribulose-1P (Ghalambor et Heath, 1962). Nous avons délété le gène ***fucA*** pour que le glycolaldéhyde intracellulaire soit uniquement utilisée pour la production d'AG.

Délétion de pfkA- Lors de l'assimilation de glucose par la voie de la glycolyse du F6P est formé, celui est naturellement converti en fructose-1,6bP par la 6-phosphofructokinase I-(*pfkA*) (Atkinson et Walton, 1965; Hellinga et Evans, 1985). Nous avons délété le gène *pfkA* afin de favoriser la prise en charge du F6P par les trancétolases et transaldolases de la voie des PP, et alimenter le flux vers l'AG. Par ailleurs quel que soit la source carbonée GAP produit via le clivage de l'A5P par Fsa est converti en DHAP, le GAP et DHAP accumulés servent à la formation du Fructose-1,6bP, qui peut être déphosphoryler en F6P. Le F6P peut alors entrer dans la voie des PP. L'activité de PfkA créerait un cycle futile, la suppression de PfkA contribue donc à diriger le flux de carbone vers l'AG.

Tableau 4.9 : Gènes de la souche *E.coli* W3CG dont l'expression a été modifiée dans le but d'augmenter le rendement en acide glycolique à partir d'hexoses et pentoses.

Gène	Enzyme	Modification	Réaction	Référence
glcD	Aldéhyde déshydrogénase	délétion	Oxydation du glycolate en glyoxylate	Lord, 1972; Pellicer *et al.*, 1996
arcA	Régulateur transcriptionnel	délétion	régulation négative de l'expression d'aldA	Pellicer *et al.*, 1999
mgsA	Methylglyoxal synthase	délétion	Formation irréversible de méthylglyoxal à partir de DHAP	Krymkiewicz, 1973; Colanduoni et Villafranca, 1985; Ferguson *et al.*, 2000
fucA	L-fuculose-phosphate aldolase	délétion	Formation du D-ribulose-1P à partir de GO et DHAP	Ghalambor et Heath, 1962
pfkA	6-phosphofructokinase I	délétion	Phosphorylation du D-fructose-6P en D-fructose 1,6-bP	Atkinson et Walton, 1965; Hellinga et Evans, 1985
galP	galactose permermease	surexpression	Transport (symporteur :H+) de sucres (hexoses et pentoses)	Postma, Lengeler et Jacobson, 1993; Deutscher *et al.*, 2006

La souche W3CG Δ*glcD* Δ*arcA* Δ*mgsA* Δ*fucA* Δ*pfkA galP* ^{proD} a été utilisé comme microorganisme hôte pour l'expression de la voie cyclique de production d'acide glycolique (**Figure 4.11**).

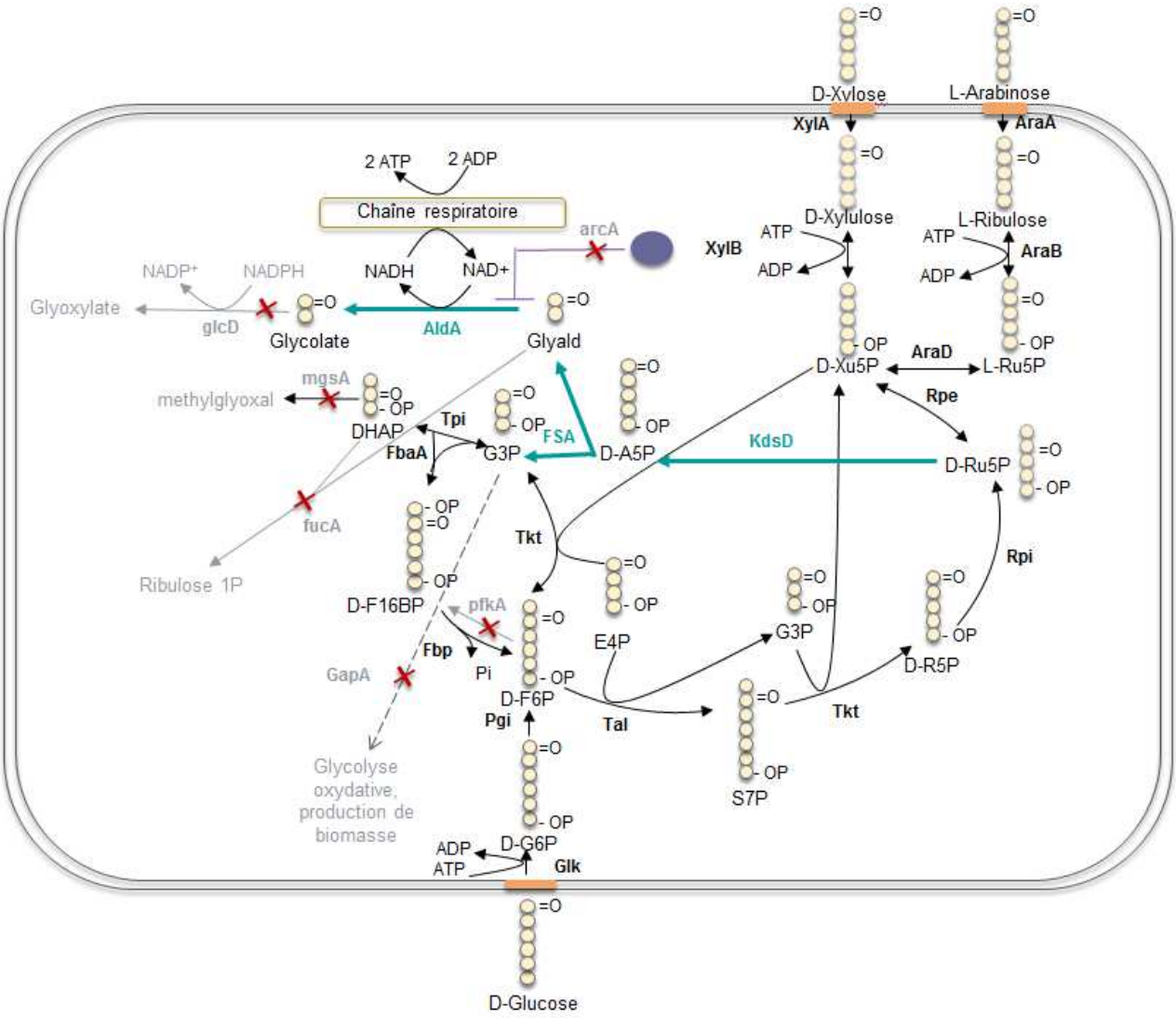


Figure 4.11 : Escherichia coli WC3G ΔglcD ΔarcA ΔmgsA ΔfucA Δpfk galP ^{proD} (GA00) exprimant la voie non naturelle dépendante de kdsD-fsa –aldA pour la production d'acide glycolique à partir de D-glucose, D-xylose et L-arabinose.

B) Expression de la voie non-naturelle KdsD-Fsa-AldA dans la souche W3G ingéniérée

Instabilité de l'opéron complet kdsD-fsa-aldA **-**Les gènes *kdsD*, *fsa* et *aldA* ont tout d'abord été clonés en un seul opéron qui a été intégré dans des vecteurs de la série pZ (Expressys ®). Le séquençage des vecteurs a mis en évidence une région manquante du promoteur lac (**Figure 4.12**). Kawe, Horn et Plückthun (2009) ont récemment décrit ce phénomène. Il s'agit d'un mécanisme de 'replication splipage', un réarrangement permis par la présence de motif répété décrit au nibeau génomique pour la première fois dans les années 60's par Georges Streisinger (Streisinger *et al.*, 1966).

```
Seq_1  2101  gaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgt  2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  123   GAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGT  182

Seq_1  2161  cttcacctcgagaaa-tttatcaaaaagagtgttgacttgtgagcggataacaatgatac  2219
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Seq_2  183   CTTCACCTCGAGAAAATTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAACAAT-----  237

Seq_1  2220  ttagattcaattgtgagcggataacaatttcacacagaattcGTTTAACTTTAAAGGAGG  2279
                                                       ||||||||||||||||||||||||||||||||||
Seq_2  238   ----------------------------TTCACACAGAATTCGTTTAACTTTAAAGGAGG  269

Seq_1  2280  TATATATGTCGCACGTAGAGTTACAACCGGGTTTTGACTTTCAGCAAGCAGGTAAAGAAG  2339
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Seq_2  270   TATATATGTCGCACGTGGAGTTACAACCGGGTTTTGACTTTCAGCAAGCAGGTAAAGAAG  329

Seq_1  2340  TCCTGGCGATTGAACGTGAATGCCTGGCGGAGCTTGATCAATACATCAATCAGAATTTCA  2399
              ..............................................................
```

Figure 4.12 : Analyse de séquence. *Seq 1 : séquence de référence du plasmide pZA23 kdsD-fsa-aldA, Seq2 : séquençage du plasmide pZA23 kdsD-fsa-aldA. En vert clair : la séquence du promoteur plac.en rose le RBS, en bleu une partie de la séquence de KdsD. Chaque pointillé dans la Seq 2 indique une paire de base manquante en comparaison avec la séquence de référence (Seq1).*

Ce mécanisme décrit dans le génome des procaryotes peut également survenir dans un vecteur en présence d'un motif répété. Le promoteur lac présent sur les vecteurs pZ possède un motif répété autorisant cet évènement comme le montre la **Figure 4.13**.

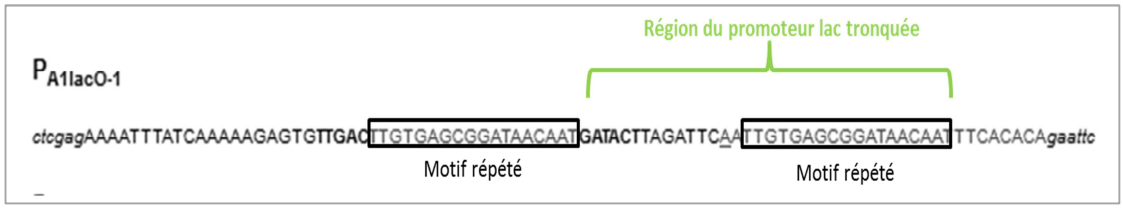


Figure 4.13 : Séquence du promoteur plac ou $P_{A1lacO-1}$. *Les régions répétées sont encadrée, indiqué en vert la région du promoteur qui apparait tronquée dans la figure 4.10.*

Nous avons donc modifié notre stratégie. La première étape a été de remplacer le promoteur lac par le promoteur tac qui ne contient pas de séquence répétée dans le plasmide pZA33, générant le plasmide pZA36. L'opéron kdsD-fsa-aldA a été intégré dans le vecteur pZA36 (vecteur à moyen nombre de copie, Ptac, sans copie de *lacI*) et dans le vecteur pEXT20 (vecteur à haut nombre copie, Ptac, une copie de *lacI*).

Ces constructions ont été exprimées dans la souche ingénieriée W3CG Δ*glcD* Δ*arcA* Δ*mgsA* Δ*fucA* Δ*pfkA* galP ^{proD}. Les souches obtenues ont été mises en croissance sur glucose, xylose et arabinose. Aucune production d'AG n'a été détectée par HPLC. Nous nous avons alors vérifié que les gènes de la voie soient correctement surexprimés. Lors de l'étude thermodynamique et des efficacités catalytiques des enzymes de la voie, nous avions conclu que la surexpression Fsa et AldA était essentielle pour tirer le carbone vers l'AG. Les activités enzymatiques de Fsa et AldA ont été mesurées sur extrait brut, elles sont présentées dans la **Figure 4.14**. Parallèlement les plasmides ont été extraits en fin de culture et séquencés. D'après les résultats d'activité le vecteur pEXT20, à haut nombre de copies, n'est pas adapté à la surexpression de l'opéron. En effet l'absence de production d'AG peut s'expliquer par une surexpression d'AldA et Fsa insuffisante (**Figure 4.14**). Le vecteur pZA36, à moyen nombre de copie, permet une augmentation d'un facteur 10 de la surexpression du gène AldA par rapport à son expression naturelle **Figure 4.14**). Les résultats du séquençage de pZA36 *kdsD-fsa-aldA* ont mis en évidence l'intégration du transposon Tn10 dans l'ORF du gène *kdsD*. L'isomérase KdsD ne peut alors être transcrite ce qui explique l'absence de production d'AG par ce vecteur. Face à la 'fuite' d'expression autorisée par le plasmide pZA36 qui ne contient pas de copie du *lacI,* la bactérie ne peut plus mettre en œuvre le mécanisme de 'replication splippage' en l'absence du promoteur lac, le transposon Tn10 serait alors intégré à l'opéron.

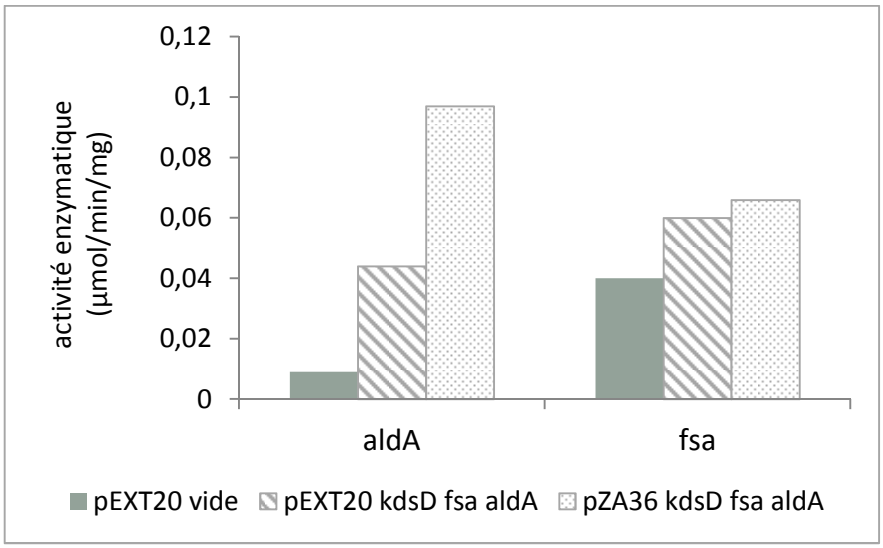


Figure 4.14 : Activité d'AldA et Fsa sur extraits bruts des souches dérivées de W3CG Δglcd ΔarcA ΔmgsA ΔfucA ΔpfkA galP ^{proD} (GA00). *En gris la souche GA00 exprimant pEXT20 vide, en rouge la souche GA00 exprimant le plasmide pEXT20 kdsD-fsa-aldA, en jaune : la souche GA00 exprimant le plasmide pZA36 kdsD-fsa-aldA.*

Construction de divers systèmes d'expression plasmidiques inductibles et constitutif pour la surexpression de kdsD, fsa et aldA - L'opéron *kdsD-fsa* sans *aldA* n'entraine pas de 'replication splippage' ou d'insertion du transposon Tn10. Nous avons donc choisi de coupler l'expression d'un vecteur contenant *kdsD* et *fsa* à un deuxième vecteur portant *aldA*. L'expression de vecteurs seuls contenant *kdsD* et *fsa* sans couplage avec *aldA* a également été étudiée comme contrôle. Plusieurs constructions ont été réalisées avec des promoteurs constitutifs (proD, proC) ou inductibles (lac, tac) (**Figure 4.15 ,Figure 4.16**). Le promoteur proD est plus fort que le promoteur proC. Ces deux promoteurs ont été conçus pour être très faiblement impactés par les conditions environnementales (Davis, Rubin et Sauer, 2011). Les promoteurs lac et tac sont tous deux des promoteurs forts inductibles à l'IPTG.

La co-expression de deux vecteurs implique l'emploi d'origines de réplication compatibles. Nous avons pour cela utilisé trois origines de réplication différentes : colE1 - haut nombre de copies (70-100), p15A - moyen nombre de copies (30-50) et pSC101 - faible nombre de copies(10-15), dont la compatibilité est présentée dans le **Tableau 4.10**.

Tableau 4.10 : Compatibilité entre les origines de réplication ColE1, p15A, pSC101

Ori	ColE1	p15A	pSC101
ColE1			
p15A			
pSC101			

La variation des promoteurs, des origines de réplication, ainsi que du nombre de gènes surexprimés *kdsD* et *aldA* ou *kdsD* et *aldA* et *fsa* a abouti à la construction de 23 systèmes d'expressions différents. Les RBS des gènes *kdsD*, *fsa* et *aldA* sont inchangés parmi les constructions. Les systèmes d'expression ont été représentés de façon schématique dans les **Figure 4.15** et **Figure 4.16**. La **Figure 4.15** représente les constructions avec des promoteurs inductibles type lac ou tac tandis que la **Figure 4.16** présente les constructions avec des promoteurs constitutifs type proD ou proC. La première construction (Screen01, pEXT20 kdsD-fsa-aldA), est la seule à posséder l'opéron kdsD-fsa-aldA complet. Le reste des constructions correspondent à un vecteur seul contenant *kdsD* et *fsa* (Sreen02, Screen06, Screen08, Screen11, Screen15, Screen20) ou à deux vecteurs dont un portant *kdsD-fsa* et l'autre portant aldA (Screen03, 04, 07, 09, 10, 12, 13, 16, 17, 18, 19, 21, 22, 23, 24).

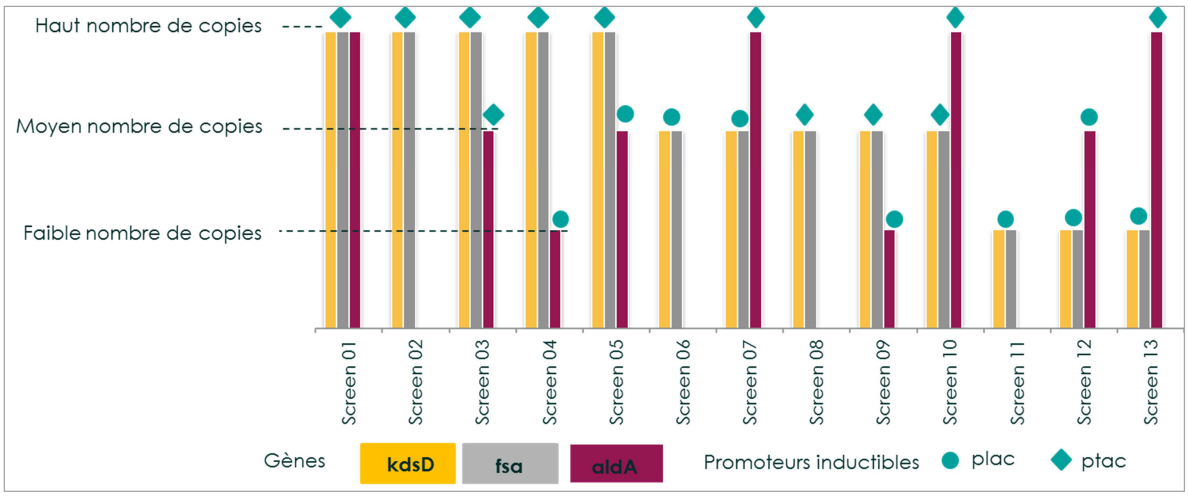


Figure 4.15: Représentation schématiques des systèmes d'expression à un ou deux vecteurs inductibles construits pour la surexpression des gènes kdsD, fsa et aldA dans Escherichia coli. *En abscisse les noms correspondent aux souches E.coli ΔtktA ΔtktB ΔglcD porteuses des systèmes d'expression n°01 à 13.*

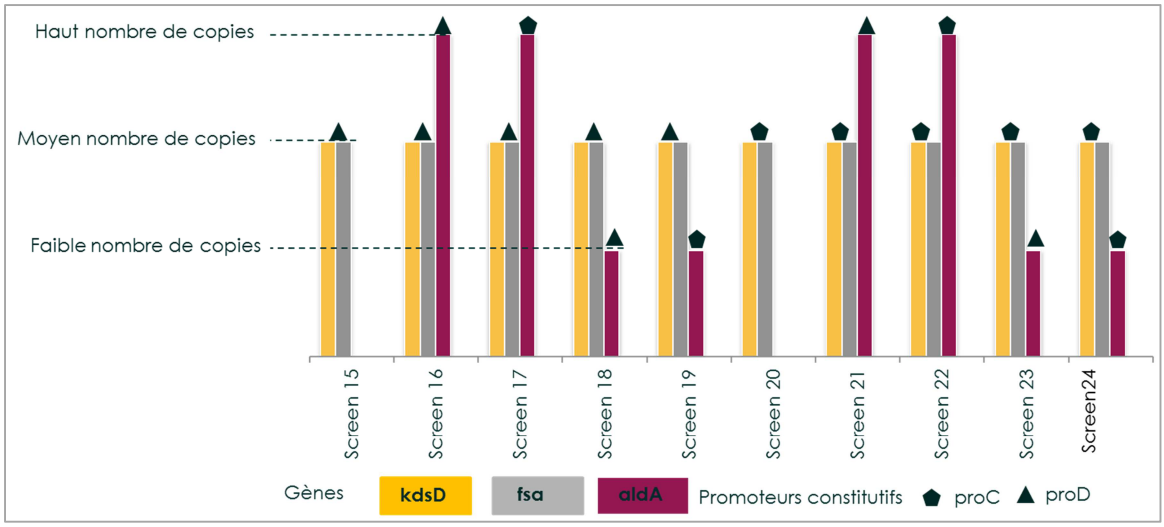


Figure 4.16 : Représentation schématiques des systèmes d'expression à un ou deux vecteurs constitutifs construits pour la surexpression des gènes kdsD, fsa et aldA dans Escherichia coli. *En abscisse les noms correspondent aux souches E.coli ΔtktA ΔtktB ΔglcD porteuses des systèmes d'expression n°15 à 16*

Afin de déterminer quelles constructions permettent la meilleure production d'AG un criblage fonctionnel a été développé.

2.3.2 Sélection des meilleurs systèmes d'expression pour la production d'acide glycolique par criblage fonctionnel *in vivo*

A) Principe du criblage fonctionnel

Le double mutant *E.coli* Δ*tktA* Δ*tktB* est incapable de pousser en présence des pentoses L-arabinose et D-xylose. Le principe du criblage est basé sur un test de croissance. L'expression fonctionnelle des gènes *kdsD*, *fsa* permet la synthèse de GAP et rend possible la croissance à partir de pentose. Quant à aldA, elle n'est pas nécessaire à la production de GAP mais est essentielle à la production d'AG. La croissance peut-elle être dans ce cas directement corrélée à la production d'acide glycolique ? La croissance est l'indicateur de la production de GAP et de ce fait de de la production de GO qui est produit en quantité équimolaire. Le gène *glcD* codant pour la glycolaldéhyde déshydrogénase a été délété pour permettre l'accumulation d'AG dans la souche *E.coli* Δ*tktA* Δ*tktB* Δ*glcD*. La conversion du glycolaldéhyde en AG n'est pas un défi technique au vue de l'efficacité catalytique d'AldA et de l'énergie libre de Gibbs très négative de cette réaction. Ainsi si KdsD et Fsa ont un niveau d'activité suffisant pour permettre la production de GAP et GO, nous devrions détecter une production d'AG proportionnelle à la croissance. Le suivi de la DO à 600nm et la quantification de l'AG produit en cours de croissance pour chacune des souches testées permettra de tester cette hypothèse. (**Figure 4.17**)

.

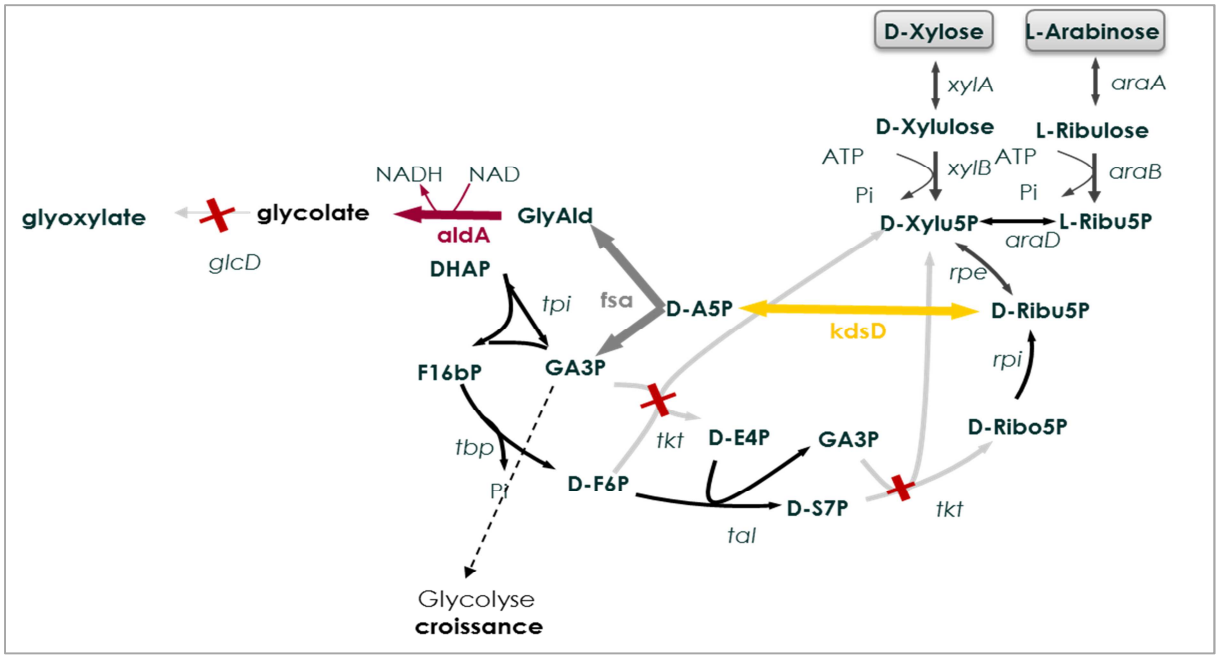


Figure 4.17 : Souche E.coli ΔtktA ΔtktB ΔglcD utilisée pour le criblage des systèmes d'expression des enzymes KdsD, Fsa et AldA.

B) Etude de la croissance et de la production d'acide glycolique des souches Screen00 à Screen24

E.coli Δ*tktA* Δ*tktB* Δ*glcD* a été transformé avec les systèmes d'expressions inductibles et constitutifs des gènes *kdsD*, *fsa* et *aldA* générant ainsi 25 souches (Screen00 à Screen24). Ces souches surexprimant les enzymes de la voie à différents niveaux d'expression ont été mises en culture sur D-xylose afin de démontrer la production d'acide glycolique à partir de pentoses *in vivo* et déterminer quelles constructions plasmidiques sont associées à la meilleure production d'AG. Le milieu a été complémenté par des traces de peptone, d'extrait de levure afin de réduire la phase de latence et lancer la croissance.

Évaluation de la faisabilité d'un crible au débit en microplaque - Dans le but de développer un criblage à haut débit, les souches dérivées de MG1655 Δ*tktA* Δ*tktB* Δ*glcD* ont été testées en plaque 96 puits, sous agitation à 37°C dans un lecteur microplaque relevant la $D0_{600nm}$ toute les 30 minutes sur 24h. Les souches testées fragilisées par de multiples délétions et l'expression de 2 plasmides, ont un taux de croissance faible incompatible avec les conditions de culture en microplaque. Les principales limitations sont le manque d'oxygène et l'évaporation. En effet la vitesse de transfert d'oxygène (ORT, oxygen transfer rate) dépend de la surface d'interface gaz-liquide. En microplaque très faible surface d'échange des puits en microplaque (0,32cm^2) a un impact négatif sur la vitesse de transfert d'oxygène. Les résultats obtenus n'étaient pas exploitables. Nous avons donc poursuivi l'étude en erlenmeyer à baffles pour permettre une oxygénation optimale.

Résultats de cultures en erlenmeyer - Nous avons réalisé les cultures des 24 souches de screening en erlenmeyer. Les résultats sont présentés dans le **Tableau 4.11 A, B,** les souches contrôle Screen00 et Screen14 portant des vecteurs vides ont été cultivés en triplicat, les autres résultats proviennent d'une seule expérience.

Tableau 4.11 : Production d'acide glycolique à partir de xylose des souches dérivées d'E.coli Δ*tktA* Δ*tktB* Δ*glcD* , surexprimant kdsD-fsa (KF) ou kdsD-fsa-aldA(KFA). *Conditions de culture : M9 xylose 1%(m/v), tryptone 0,2 %(m/v), extrait de levure 0,1%(m/v) à 37°C pendant 200h.*

Code couleur : expression des résultats en pourcentage en fonction du contrôle négatif.

Légende – xylose, xylulose, acide glycolique	
0 à 19 %	-0 à -19 %
20 à 39 %	-20 à -39 %
40 à 59 %	-40 à -59 %
60 à 79 %	-60 à -79 %
80 à 99%	-80 à -99%
>100 %	-100 % <

Légende-biomasse
0 à100 %
200 à 390 %
400 à 590 %
600 à 790 %
800 à 990%
>1000 %

A. ***Souches d'E.coli contenant des systèmes d'expression inductibles.*** *(*) contrôle négatif avec plasmide vide, utilisé comme référence (0%)*

Souche	Gènes	Conso. xylose (g/L)	Production de xylulose (g/L)	Acide glycolique		Biomasse	
				Conc.finale (g/L)	AG/xylose (g/g)	Prod.finale (g/L)	OD600nm finale
Screen00*	-	3,8	1,01	0,35	0,1	0,08	0,8
Screen01	KFA	5,36	1,6	0,64	0,12	0,18	1,00
Screen02	KF	1,9	0,37	0,28	0,15	0,25	1,24
Screen03	KFA	1,86	0,4	0,25	0,14	0,09	0,63
Screen04	KFA	3,8	1,02	0,53	0,14	0,17	0,91
Screen05	KFA	2,11	0,35	0,38	0,18	0,29	1,4
Screen06	KF	4,7	1,01	0,74	0,16	0,34	1,62
Screen07	KFA	2,91	0,59	0,41	0,14	0,28	1,51
Screen08	KF	5,44	1,03	0,88	0,17	0,58	2,50
Screen09	KFA	4,86	0,4	1,11	0,24	0,74	2,90
Screen10	KFA	2,87	nd	0,49	0,17	0,19	0,86
Screen11	KF	6,41	0,89	1,13	0,18	0,42	1,92
Screen12	KFA	3,06	0,71	0,34	0,11	0,26	1,38
Screen13	KFA	4,25	0,85	0,75	0,18	0,28	1,33

B. ***Souches d'E.coli contenant des systèmes d'expression inductibles*** *(*) contrôle négatif avec plasmide vide utilisée comme référence (0%)*

Souche	Gènes	Conso. xylose (g/L)	Production de xylulose (g/L)	Acide glycolique		Biomasse	
				Conc.finale (g/L)	AG/xylose (g/g)	Prod.finale (g/L)	OD600nm finale
Screen14*	-	3,04	0,75	0,426	0,12	0,10	0,62
Screen15	KF	3,84	0,84	0,55	0,15	0,32	1,66
Screen16	KFA	3,64	0,51	0,48	0,14	0,31	1,47
Screen17	KFA	3,53	0,54	0,48	0,14	0,35	1,55
Screen18	KFA	4,76	0,70	0,74	0,16	0,45	1,94
Screen19	KFA	3,24	0,53	0,62	0,20	0,36	1,73
Screen20	KF	3,29	0,74	0,42	0,13	0,37	1,68
Screen21	KFA	3,13	0,76	0,42	0,14	0,34	1,56
Screen22	KFA	3,14	0,69	0,44	0,14	0,37	1,66
Screen23	KFA	3,23	0,41	0,75	0,24	0,90	3,30
Screen24	KFA	3,20	0,70	0,46	0,15	0,41	1,88

Production d'AG des souches contrôle Screen00 et Screen14 - Les souches contrôle contenant des plasmides vides (Screen00, Screen14) utilisées comme références dans les **Tableau 4.11A ; 4.11B** n'ont pas montré de croissance une fois les traces de peptones et d'extrait de levure consommés. Toutefois une production d'acide glycolique a été observée, 0,35 g/L pour la souche Screen00 et 0,42g/L pour la souche Screen14. Les souches contrôle ne surexpriment pas les enzymes KdsD, Fsa et AldA et le cycle naturel du glyoxylate est inhibé en présence de source de carbone complexe telle que le xylose. Nous observé dans la partie '*in vitro'* 222.B) une activité de Fsa sur D-xylulose, qui serait clivé en DHA et GO. Le DHA ne permet pas la croissance et le GO est oxydé en AG par AldA exprimé au niveau génomique dans *E.coli*. De plus, le xylulose s'accumule particulièrement dans les souches *MG1655 ΔtktA ΔtktB ΔglcD* dépourvues de la voie KFA: 1 g/L et 0,75 g/L ont été quantifié dans le surnageant des cultures des souches Screen00 et Screen14, respectivement (**Tableau 4.1 A,B**).

Tableau 4.11Une autre explication pourrait être la production d'AG par l'activité combinée des enzymes KdsD, Fsa et AldA naturellement présentes dans *E.coli*. Ces activités ne permettent pas de production d'AG dans une souche sauvage mais pourraient être mises en évidence dans le triple mutant en raison de l'accumulation des intermédiaires de la voie des pentoses P (Teleman *et al.*, 1999). Le niveau élevé de Ru5P facilite peut être la prise en charge par KdsD, enzyme qui possède le K_m le plus élevé à 2,6 mM (**Tableau 4.1**, Meredith et Woodard, 2003). La voie KFA génère du GAP, si celle-ci était à l'origine de la production d'AG une croissance serait observée ce qui n'est pas le cas des souches Screen00 et Screen14. L'hypothèse de l'activité in *vivo* de Fsa sur D-xylulose est donc la plus probable, elle pourra être vérifiée en étudiant la production d'AG de la souche *E.coli* Δ*tktA* Δ*tktB* Δ*glcD* Δ*fsa* sur xylose.

Croissance et production d'AG des souches test Les souches Screen01, 06, 08, 09, 11, 13, 18, 19 et 23 produisent significativement plus d'AG que les souches contrôle Screen00, 14 (**Tableau 4.11**, **Figure 4.18**). Avec une concentration finale en produit de 1,14 g/L la souche Screen09 accumule jusqu'à 2,17 fois plus d'AG que la souche contrôle Screen00 (0,36 g/L) (**Tableau 4.11**, **Figure 4.18**). Ces résultats montrent que l'expression de la voie KdsD-Fsa-AldA permet la conversion le D-xylose en AG *in vivo*. Toutefois toutes les souches testées n'ont pas montré une production augmentée d'AG, l'efficacité de la voie synthétique KdsD-Fsa-AldA dépendrait du niveau d'expression de ces enzymes et du 'poids métabolique' du système d'expression utilisé. Aucune des souches contenant un plasmide à haut nombre de copies (Screen02, Screen 03, Screen04, Screen05, Screen07, Screen10, Screen16, Screen17, Screen21, Screen22, en rose dans la **Figure 4.18**) n'a permis une production significative d'AG comparé au contrôle. Un nombre élevé de copies de plasmide par cellule semble imposer une charge métabolique cellulaire trop importante (metabolic burden) au détriment de la production d'AG.

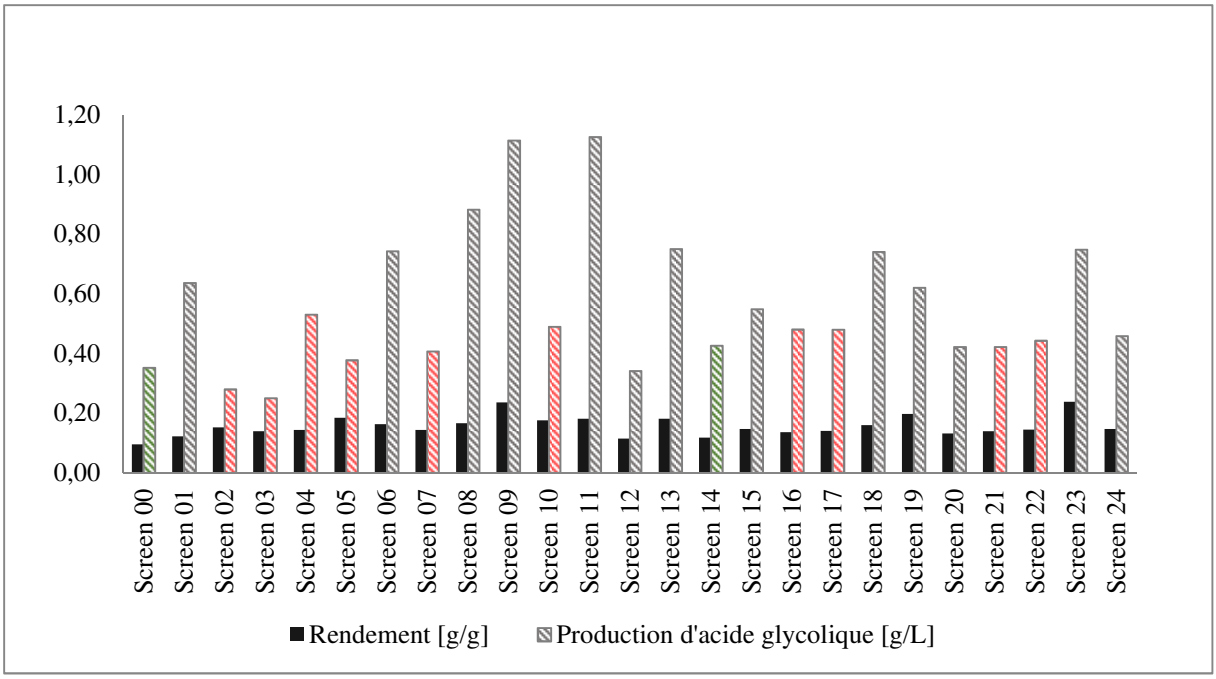


Figure 4.18 : Rendement et production d'acide glycolique des souches Screen00 à Screen24. Souches cultivées en erlenmeyer dans un milieu M9 xylose 1%,, *tryptone 0,2%, extrait de levure 0,1% à 37°C pendant 200h. En vert : la production d'acide glycolique des souches contrôle contenant des plasmides vides, en rouge la production d'acide glycolique des souches contenant au moins un plasmide à haut nombre de copie.*

En terme de rendement en acide glycolique la majorité des souches testées lors du criblage *in vivo* dépassent la souche contrôle. Le plus haut rendement atteint de 0,24 g/g a été obtenu avec les souches Screen09 et Screen23. Toutes deux surexpriment *aldA* en faible nombre de copie. La meilleure combinaison semble être *kdsD-fsa* sur un vecteur à moyen nombre de copie et *aldA* sur un vecteur à faible nombre de copie.

Evaluation de la nécessité de surexprimer de kdsD et fsa et aldA pour la synthèse d'AG dans Escherichia coli **-** Les souches Screen 06, 08 et 11 surexpriment uniquement les gènes *kdsD* et *fsa* et produisent également trois fois plus d'AG que la souche contrôle, à des rendements plus faibles que Screen09. Ce résultat questionne la nécessité de la surexpression d'*aldA*. Dans ce contexte une expérience a été réalisée. Les gènes *kdsD*, *fsa* et *aldA* ont été surexprimés séparément et de façon combinée dans une souche d'E.coli Δ*tktA* Δ*tktB* Δ*glcD* (non délétée de *kdsD*, *fsa* et *aldA*).

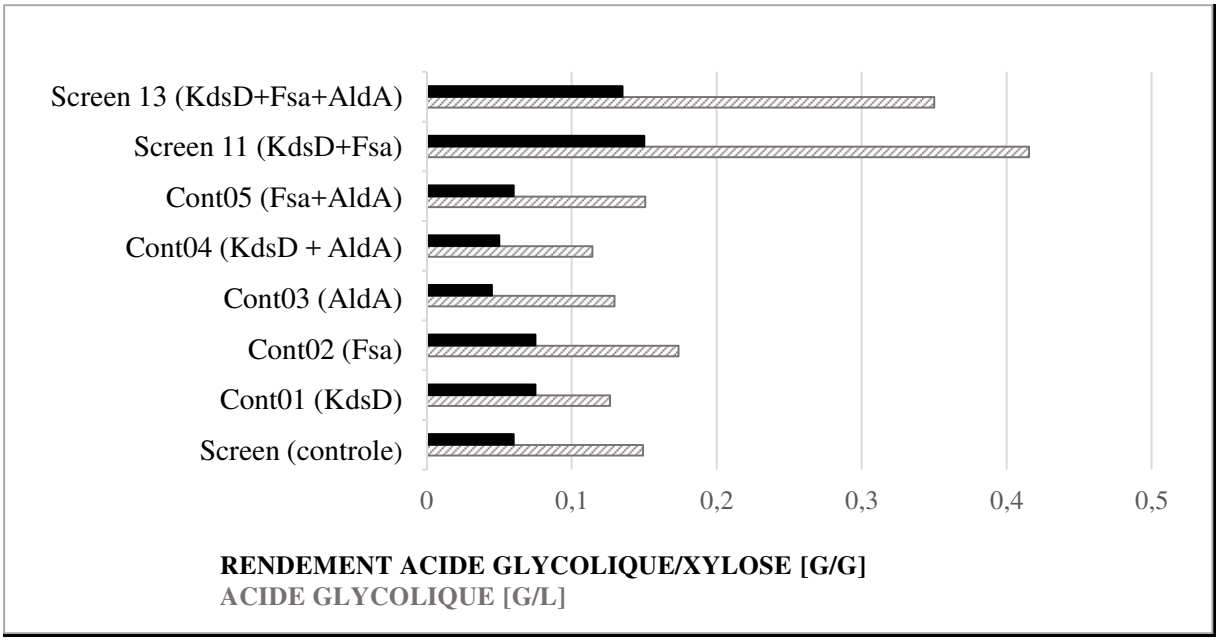


Figure 4.19: Etude de la production d'acide glycolique à partir de xylose par les souches E. coli MG1655 ΔtktA ΔtktB ΔglcD strains surexprimant aucune (Screen), seulement une (Cont02, Cont03), deux (Cont04, Cont05, Screen11) ou trois (Screen13) des enzymes clefs de la voie non-naturelle de production d'acide glycolique. *Conditions de culture : M9 xylose 1%, tryptone 0,2%, extrait de levure 0,1% à 37°C . pendant 100h.*

Les résultats de cette expérience indiquent que la surexpression de *kdsD*, *fsa* ou *aldA* seuls, avec les deux autres gènes exprimés à leur niveau natifs n'est pas suffisante pour une production d'AG significativement plus élevée que le contrôle (**Figure 4.19**). La surexpression de *kdsD* et *fsa*, avec *aldA* exprimé au niveau génomique, induit une augmentation de la production d'AG. La nécessité de surexprimer *kdsD* indiquerait que la production de GO est le produit de l'activité de Fsa sur A5P, si elle provenait du clivage du D-xylulose la surexpression de *kdsD* couplée à *fsa* n'aurait pas un tel effet.

La surexpression d'*aldA* n'apparait pas essentielle, l'expression génomique est suffisante pour oxyder le GO. Toutefois la surexpression d'*aldA* et la charge métabolique imposée par le système d'expression aurait un fort impact sur la production d'AG et apparaissent comme des paramètres centraux d'optimisation. Dans l'expérience *aldA* est porté par un vecteur à nombre élevé de copie qui s'accompagne d'un effet négatif sur la production d'acide glycolique, rejoignant les résultats obtenus lors du criblage. Lorsqu'*aldA* est exprimé sur un vecteur à faible nombre de copie sa surexpression permettrait d'atteindre un meilleur rendement et une meilleure production qu'avec la surexpression de *kdsD-fsa* seule comme le montre les souches Screen09 et Screen23.

L'expérience devra être répétée en délétant les gènes *kdsD*, *fsa* et *aldA* de la souche *MG1655 ΔtktA ΔtktB ΔglcD* pour évaluer l'effet de l'expression de ces 3 gènes sans le biais de leur expression naturelle.

Rendement en AG - La majorité des souches ont un rendement compris entre 0,14 g/g et 0,16 g/g alors que la production d'AG est bien plus variable, du fait de la variabilité de consommation de xylose (**Figure 4.20**). En effet les analyses HPLC des échantillons pris en cours de croissance révèlent une grande variabilité du niveau de xylose consommé entre les souches testées (1,90-6,41 g/L) (**Tableau 4.11**). Les souches convertissant le plus efficacement le xylose en AG sont les souches Screen09 et Screen23.

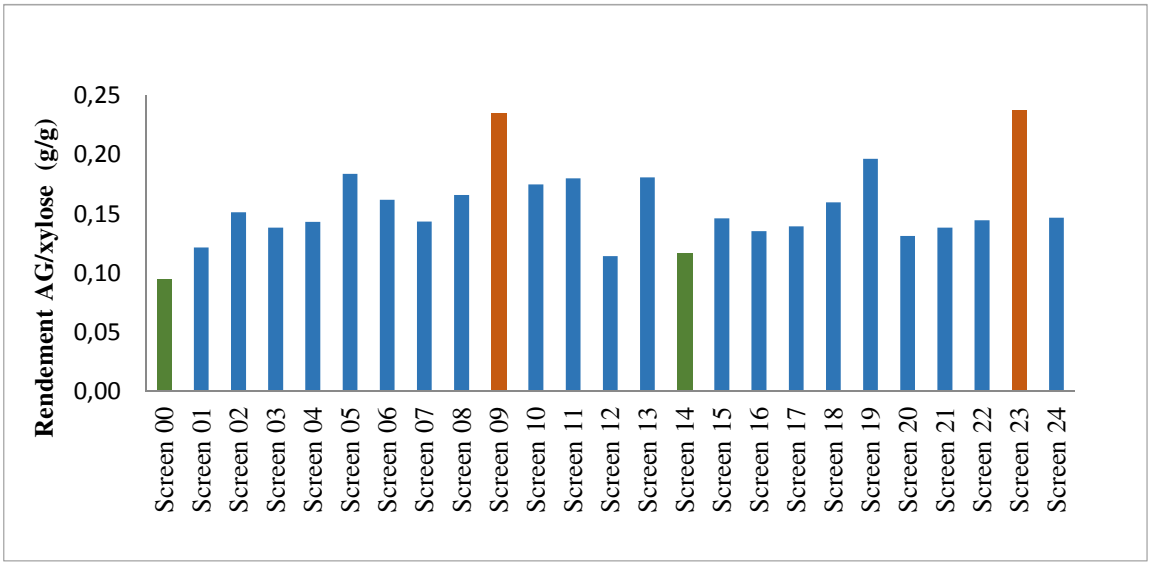


Figure 4.20 : Rendement en AG (g/g) des souches Screen00-24. *Figures réalisées à partir des données HPLC de consommation de xylose et de production d'AG à 200h de culture sur xylose 1%.(Tableau 4.11)*

Accumulation et sécrétion de xylulose –Nous avons constaté une excrétion de xylulose dans le milieu de culture des souches ne produisant pas ou peu d'AG, mettant en évidence une forte accumulation intracellulaire de ce métabolite. Lors de l'assimilation du xylose, l'ATP est utilisé comme co-facteur pour la conversion du xylulose en xylulose-5P, la régénération de l'ATP dépend de l'efficacité de l'oxydation en AG, catalysée par aldA, qui génère du NADH pris en charge par la chaine respiratoire pour la production d'ATP. Lorsque la voie KFA n'est pas efficace, l'ATP n'est pas régénéré et le déficit en ATP rend impossible la phosphorylation du xylulose qui s'accumule. Pour appuyer à cette hypothèse, la **Figure 4.21** met en évidence une corrélation négative entre la production de xylulose et d'AG à partir de xylose. Les souches Screen09 et Screen23 accumulent le moins de xylulose, et présentent le meilleur rendement de production.

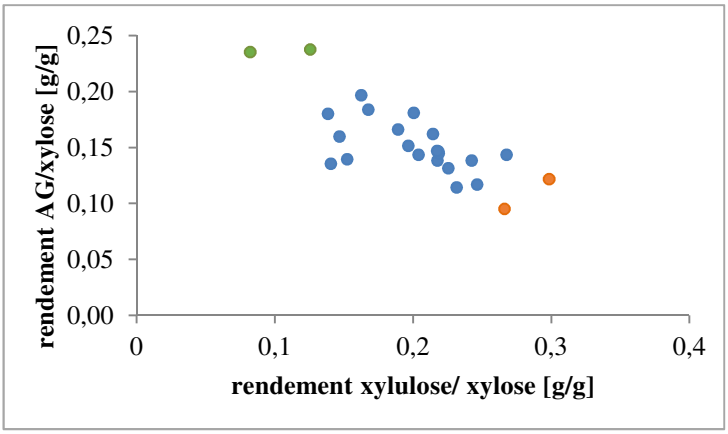


Figure 4.21 : Rendement AG/xylose exprimé en fonction du rendement xylulose/xylose des souches dérivées de MG1655 ΔtktA ΔtktB ΔglcD . *En vert : les souches Screen00 et Screen14 contrôle, contenant des plasmides vides. En orange les souches productrices sélectionnées Screen09 et Screen23. Figures réalisées à part des données HPLC de consommation de xylose et de production d'AG à 200h de culture sur xylose 1%.*

Relation entre croissance bactérienne et production d'AG - Un des objectifs était de développer un criblage fonctionnel basé sur la croissance. Le triple mutant Δ*tktA* Δ*tktB* Δ*glcD* n'est pas capable de croitre sur pentose, l'expression d'une voie synthétique fonctionnelle générant du GAP est nécessaire à la croissance. Cependant d'après les résultats obtenus il n'y a pas de corrélation entre le taux de croissance maximal des souches et la production d'AG (**Figure 4.22**).Toutes les souches ont un taux de croissance très faible compris entre 0 et 0,01 h^{-1}. Ce faible taux de croissance est en grande partie la conséquence du manque de trancétolases, enzymes importantes de la voie des pentoses phosphate, et d'un système d'expression à 2 plasmides qui représente une lourde charge métabolique (réplication, trancription, traduction). Il est difficile de calculer des taux de croissances si faibles et de comparer la variabilité de croissance entre les souches.

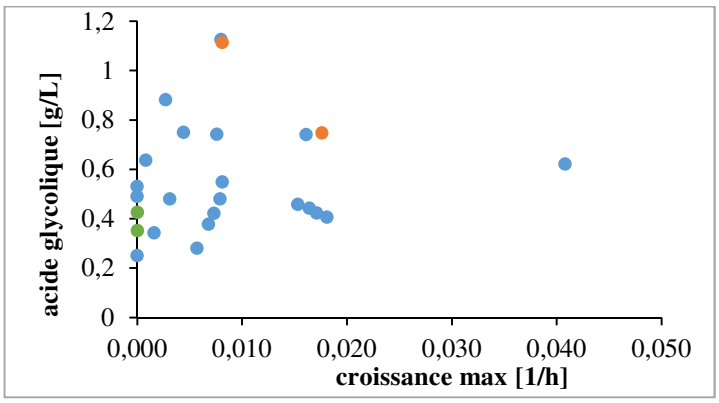


Figure 4.22 : Production finale d'acide glycolique (A) et rendement acide glycolique/xylose en fonction du taux de croissance maximal des souches Screen00 à Screen24. *En vert : les souches Screen00 et Screen14 contrôle, contenant des plasmides vides. En orange les souches productrices sélectionnées Screen09 et Screen23.*

La densité optique à 600nm mesure de la turbidité du milieu de culture, liée à la concentration en biomasse. La mesure de la DO_{600nm} en phase stationnaire permet d'estimer la variabilité de production de biomasse entre les souches. Aucune corrélation n'a été observée entre la DO_{600nm} en phase stationnaire et la production d'AG mais on peut constater que les souches contrôle ont la DO_{600nm} finale les plus faibles tandis que les souches productrices d'AG Screen09 et Screen23 ont la DO_{600nm} la plus élevée (**Figure 4.23**). Nous aurions pu sur la base de la mesure de la DO_{600nm} sélectionner les souches Screen09 et Screen23 comme meilleures productrices. Le criblage développé dans cette étude permet de sélectionner les souches les plus performantes à partir de la DO600. Seules les souches à la DO_{600nm} finale élevée nécessiteraient d'être caractérisée (production d'AG). Toutefois la relation entre DO_{600nm} et production n'est pas assez nette pour pouvoir déduite la production en AG à partir de la lecture de la DO_{600nm}.

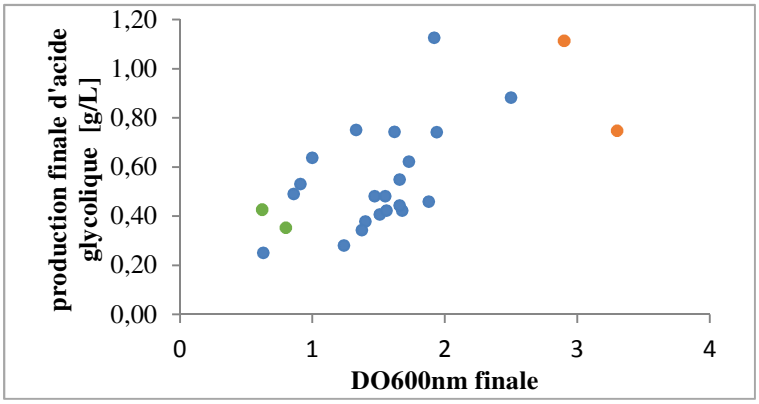


Figure 4.23 : Production finale d'acide glycolique (A) et rendement acide glycolique/xylose (B) en fonction de la D600 finale. . *En vert : les souches Screen00 et Screen14 contrôle, contenant des plasmides vides. En orange les souches productrices sélectionnées Screen09 et Screen23. Figures réalisées à part des données HPLC de consommation de xylose et de production d'AG à 200h de culture sur xylose 1%.*

2.3.3 Production des souches Screen09 et Screen23 sur xylose et arabinose

La caractérisation des souches Screen09 et Screen23 a permis de démontrer la fonctionnalité de la voie KFA sur arabinose et xylose. Les deux souches atteignent des productions et rendements comparables à partir de ces pentoses (

Figure 4.24 **4.24**). Nous en concluons que les systèmes d'expression portés par les souches Screen09 et Screen23 sont d'efficacité égale. La différence entre ces deux systèmes d'expression est leur caractère inductible ou constitutif, le système d'expression 09 est inductible à l'IPTG tandis que le système 23 est constitutif.

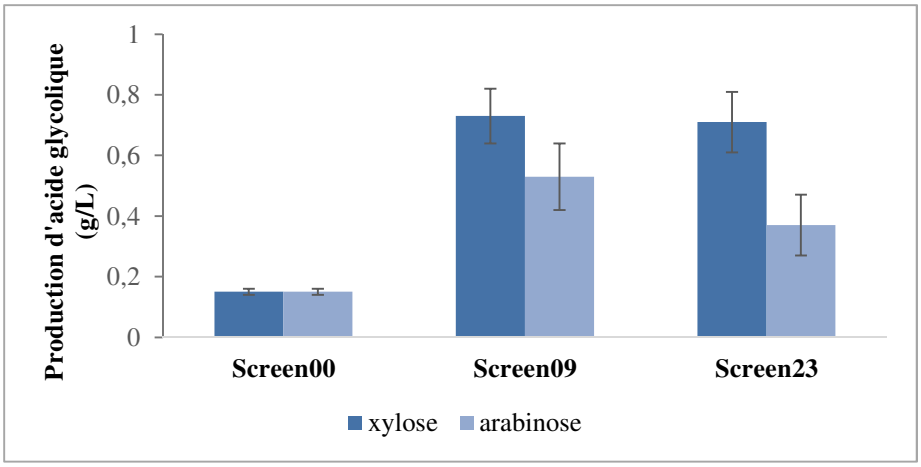


Figure 4.24 : Production d'acide glycolique par les souches MG1655 ΔtktA ΔtktB ΔglcD Screen00 (contrôle avec plasmide vide) , Screen09 et Screen23 exprimant la voie KdsD-Fsa-AldA *sur milieu minimal M9 (10g/L) sur xylose et arabinose , culture sur 100 h, 37 °C. Les barres d'erreur ont été calculées à partir de trois expériences indépendantes.*

2.3.4 Caractérisation des souches productrices GA23 et GA09 sur glucose, xylose et arabinose en erlenmeyer

Les constructions 09 et 23 sélectionnées lors du screening ont été transformées dans la souche W3CG ingéniérée, W3CG *ΔgapA ΔmgsA ΔarcA ΔglcD ΔfucA ΔpfkA* galP ^{proD}(GA00) générant ainsi les souches GA09 et GA23 **(Figure 4.25)**. La production d'AG de ces deux souches a été étudiée sur glucose, xylose et arabinose. L'inactivation *gapA* limite l'entrée du GAP, produit la voie des PP de à partir de pentoses ou d'hexose, dans la voie basse de la glycolyse pour alimenter le cycle de Krebs en acétyl-CoA. La souche W3CG ne peut pas croitre avec comme une seule source de carbone en C5 ou C6. L'ajout de malate (C4) alimentant le cycle de Krebs est nécessaire pour restaurer la croissance des mutants GapA$^-$. Du malate à 0,4 g/L a. été utilisé pour complémenter le milieu contenant 10 g/L d'hexose ou pentose.

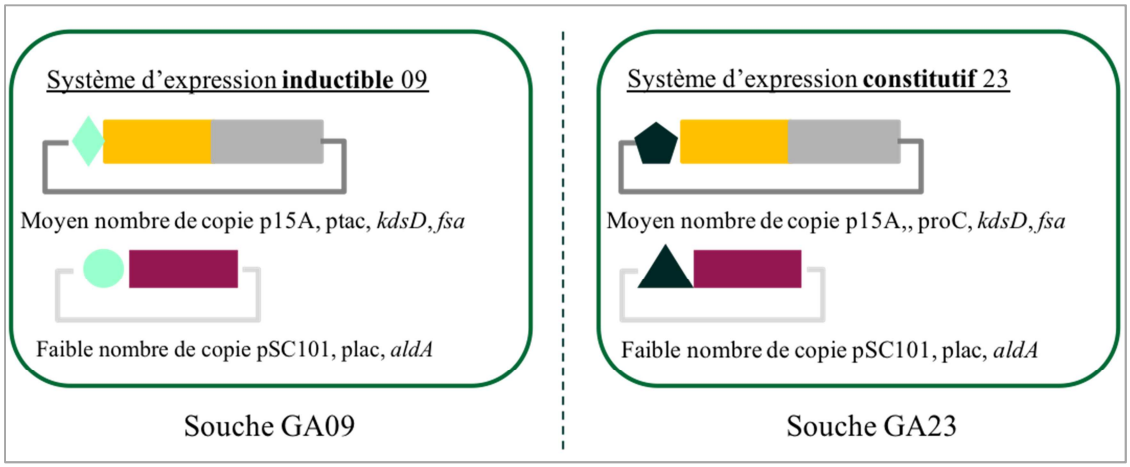


Figure 4.25 : Représentation schématique de la construction des souches GA09 et GA23 dérivées de GA00.

Production d'acide glycolique - GA09 et GA23 produisent de l'AG à partir de glucose, le xylose et plus faiblement l'arabinose ce qui démontre la fonctionnalité de la voie KdsD-Fsa-AldA *in vivo* sur l'ensemble des sucres lignocellulosiques. Les constructions 09 et 23 dans le contexte génétique du triple mutant Δ*tktA* Δ*tktB* Δ*glcD* avait une efficacité comparable tandis que dans la souche dérivée de W3CG la construction 23 se démarque avec une production d'AG supérieure sur l'ensemble des sucres (**Figure 4.26**). La production sur arabinose est faible du fait d'une assimilation très limitée. La souche GA23 a l'avantage de posséder un système d'expression constitutif, sa culture ne nécessite pas l'emploi d'IPTG pour l'induction de la voie.

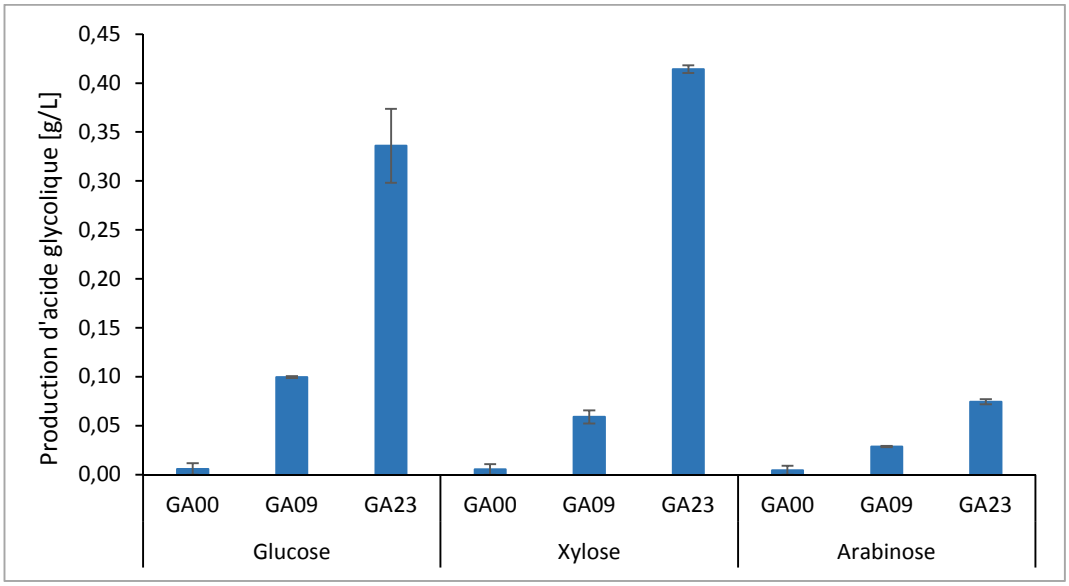


Figure 4.26 : Production d'acide glycolique à partir de glucose, xylose et arabinose par les souches GA00, GA09 et GA23 dérivés de WC3G ΔglcD ΔarcA ΔmgsA ΔfucA Δpfk proD-galP. *Culture sur 50h, à 37°C dans un milieu contenant 10g/L de xylose, arabinose ou glucose additionné de 0,4 g/L de malate. Résultats de trois expériences indépendantes.*

La performance de la souche GA23 a été évaluée en erlenmeyer sur glucose, xylose et arabinose. Le rendement a été calculé pour chacun des sucres, 11% du rendement théorique maximum a été atteint sur xylose (0,14 g/g) , 7% sur glucose (0,08 g/g) et 8 % sur arabinose (0,1 g/g) (**Tableau 3.12**), mettant en évidence une importante perte de carbone.

Tableau 3.12 : Production d'acide glycolique et rendement sur glucose, xylose et arabinose de la souche GA23. *Résultats de trois expériences indépendantes*

Source de carbone	Production d'AG		Rendement $_{GA/sucre}$		Rendement $_{GA/sucre}$ comparé au rendement théorique**
	[mM]	[g/L]	[mol/mol]	[g/g]	
glucose	4,42 ± 0,50	0,34 ± 0,04	0,20 ± 0,02	0,08 ± 0,01	7 %
xylose	5,45 ± 0,05	0,41 ± 0,00	0,27 ± 0,10	0,14 ± 0,05	11 %
arabinose	0,98 ± 0,03	0,07 ± 0,00	0,21 ± 0,02	0,10 ± 0,01	8 %

Analyse de l'extra-métabolome - Les concentrations extracellulaires des métabolites acétate, citrate, fumarate, pyruvate et xylulose ont été mesurés au cours des cultures de la souche contrôle W3CG *ΔgapA ΔmgsA ΔarcA ΔglcD ΔfucA ΔpfkA* galP ^{proD} (GA00) et la souche exprimant le système d'expression 23 (GA23). Le suivi de la consommation de malate et de sucre en cours de croissance révèle que la malate est rapidement consommé, en 20h, tandis que la consommation des sucres n'est pas totale. La consommation de sucre est indispensable à la synthèse d'acide glycolique ainsi que de D-ribose-5P et d'érythrose 4P qui sont les précurseurs des nucléotides, acides aminés aromatiques et vitamines, tandis que la malate est requis pour la croissance mais ne participe pas à la synthèse d'AG. Nous avons observé la présence de fumarate et citrate dans les milieux de cultures qui sont des intermédiaires du cycle de Krebs et dénote d'une assimilation massive de malate. Une production d'acétate est également mesurée, c'est un dérivé de l'acétyl-coA qui atteste de la saturation des capacités respiratoires cellulaires et représente une fuite de carbone à partir du cycle de Krebs qui aurait pu être utilisé pour la croissance. Aucune accumulation de xylulose n'a été rapportée lors de la croissance sur xylose.

2.3.5 Croissance en bioréacteur de la souche GA23 sur xylose et glucose

La souche GA23 a été cultivée en bioréacteur pendant 96h. Le milieu utilisé est le même que celui utilisé pour la caractérisation en erlenmeyer : milieu M9 complémenté de 10g/L de xylose ou glucose, 0,4 g/L de malate et 10% LB pour réduire la phase de latence. Une injection de malate pour relancer la croissance a été réalisée à 24h de culture. L'objectif est de pouvoir identifier les pertes de carbone mises en évidence en erlenmeyer afin de proposer des solutions adaptées, pour cela des prélèvements ont été effectués afin de quantifier les métabolites in et extra-cellulaires, le TOC (Total Organic Carbon), et la DO600nm. Les résultats du suivi de croissances sont présentés dans la **Figure 4.27**.

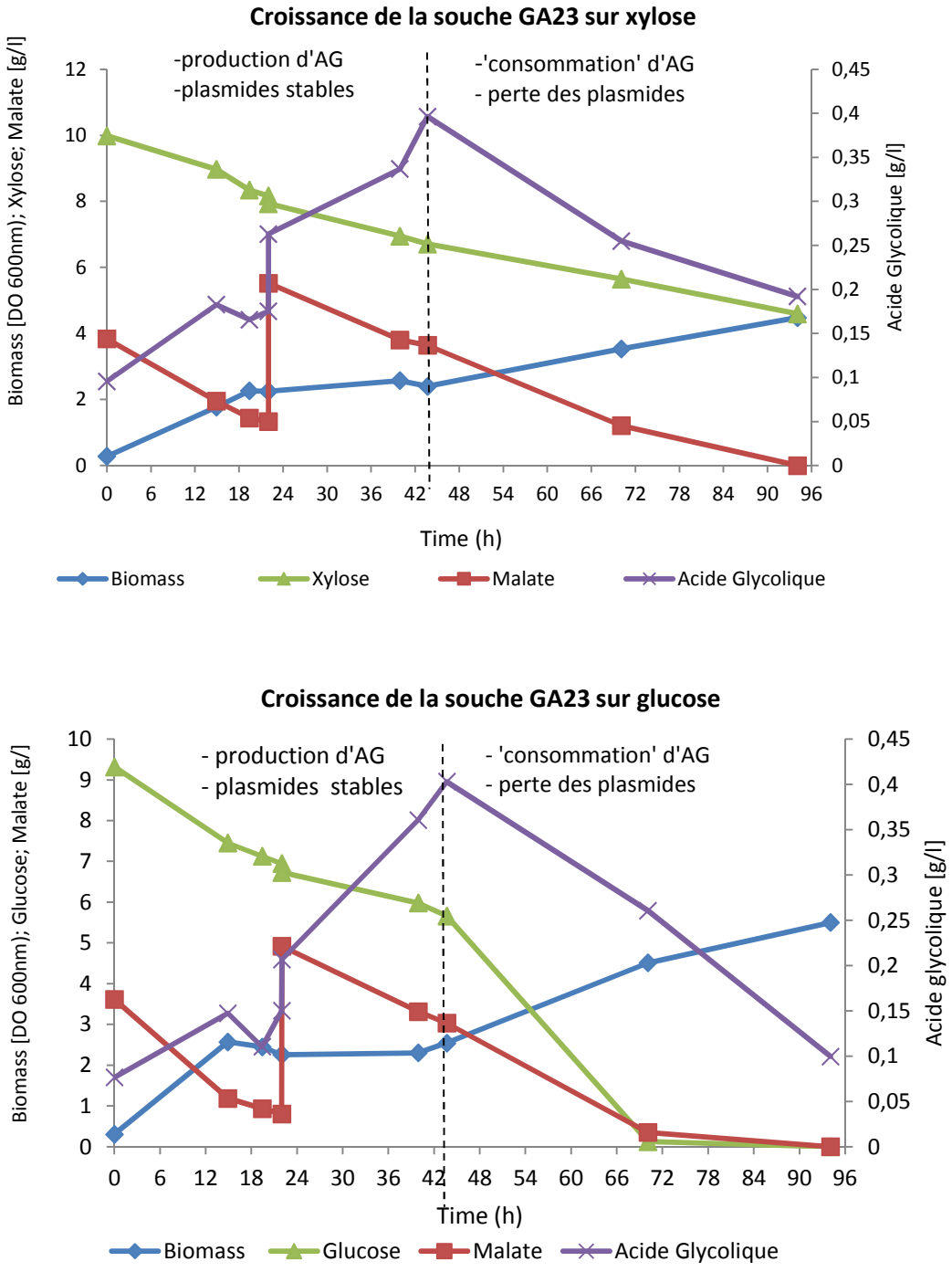


Figure 4.27 : Suivi de croissance de la souche GA23 sur xylose et glucose.

A) Etude de stabilité des plasmides –

Une étude de la stabilité des plasmides par étalement sur milieu sélectif a révélé qu'à T24h l'ensemble de la population bactérienne possédait le système d'expression 23, à T48h seulement 50% de la population contenait les deux plasmides. Les données de croissance, de production d'AG, et de consommation de malate et glucose ou xylose sont présentées sur 96h mais les analyses seront faites sur les 40h où au moins la moitié de la population bactérienne possède la voie KdsD-Fsa-AldA.

B) Analyse de l'exométabolome

Consommation de malate, des monosaccharide et production d'acide glycolique - D'après la **Figure 4.27**, le malate est très rapidement consommé par la souche GA23, le glucose est entièrement consommé, contrairement au xylose qui présente une assimilation incomplète plus lente. Le glucose est soit (i) transporté par la perméase GalP puis phosphorylé par la glucokinase Glk dans le cytoplasme soit (ii) pris en charge par le système PTS, qui transporte et phosphoryle le glucose en glucose-6P, le phosphate provient de la déphosphorylation du phosphoenol-pyruvate en pyruvate. L'absorption du xylose se produit principalement par l'intermédiaire d'un transporteur-ATP dépendant de haute affinité (XylFGH), bien qu'un second symporteur/H^+ de faible affinité (XylE) soit également présent. L'efficacité de l'utilisation du xylose dans *Escherichia coli* dépend du niveau d'ATP cellulaire, la capacité naturelle d'assimilation du xylose par *E.coli* n'est pas optimale (Lam *et al.*, 1980; Moussatova *et al.*, 2008).

Les profils de production d'acide glycolique sont identiques entre les 2 conditions testées, l'AG est produit de façon biphasique sur les 42h premières heures : la production s'arrête lors de la consommation totale du malate et reprend après le pulse. La reprise la synthèse d'AG après l'ajout de malate pourrait s'expliquer par la synthèse d'équivalent réducteur nécessaire à la production. Etonnement la concentration d'AG diminue après 42h de culture. *Escherichia coli* assimile habituellement les C2 lorsque les sources carbonées plus complexes sont absentes du milieu. Or, la consommation d'AG semble débuter en présence de malate et monosaccharides. Nous pouvons poser l'hypothèse que l'AG est dégradé ou consommé au cours de la culture, la production grâce à la voie KFA compensait cette perte jusqu'à la perte des plasmides. Dans le cas d'une consommation, trois transporteurs membranaires pourraient être impliqués dans l'importation d'AG chez *Escherichia coli.* Le symporteur glycolate/lactate :H^+ GlcA, qui appartient à la famille des lactate permease (Saier, 2000; Núñez *et al.*, 2002) contribue à l'absorption du lactate, le glycolate et l'hydroxybutanoate. Il a été démontré que la croissance sur glycolate était totalement inhibée par la délétion de *glcA* (Núñez *et al.*, 2001). La protéine membranaire Lldp est une perméase avec une affinité pour le L-lactate, D-lactate et glycolate (Dong *et al.*, 1993; Núñez *et al.*, 2001, 2002). Enfin l'Actp est une acétate/glycolate perméase impliquée dans le transport d'acétate et de glycolate (Gimenez *et al.*,

2003).Une souche mutée pour ces transporteur n'est plus capable d'importer l'AG et pourrait être capable de résister à des concentrations élevées en AG.

Production et rendement de la souche GA23 en AG - La production et le rendement en AG de la souche GA23 a été calculée à 40h de culture, avant la phase de consommation d'AG. La production d'AG est de 0,4 g/L sur glucose et xylose. Le rendement sur glucose est de 0,2 g/g (0,25 mol/mol) et de 0,13 g/g (0,26 mol/mol) sur xylose. Ces données sont identiques à celles obtenues en erlenmeyer.

C) Analyse de l'endo-métabolome

Les résultats de l'analyse de l'endo-métabolome de la souche GA23 sont présentés dans les **Tableaux 4.13** (xylose) et **4.14** (glucose), les accumulations de métabolites sont représentées sur la **Figure 4.28**

Assimilation du malate - L'analyse de l'endo-métabolome à 44h révèle une accumulation de malate, de fumarate, de succinate, d'acétate, indépendamment du monosaccharide utilisé pour la croissance. Ces métabolites sont des produits de l'assimilation de malate. La consommation rapide de malate mène à l'accumulation des intermédiaires du cycle de Krebs.

Assimilation du glucose - Du sedoheptulose 7P s'accumule sur glucose (0,18 mM à T44h) comparé à la culture sur xylose (0,03 mM à T44h), ce qui s'explique par la délétion de la phosphofructokinase (**Tableau 4.13, 4.14)** .En effet glucose est converti par la voie de la glycolyse en F6P qui ne peut être phosphorylé en F1,6bP. Le F6P est alors pris en charge par la transaldolase (Tal) pour la synthèse de S7P et de GAP (**Figure 4.28**). L'accumulation de S7P qui devrait être pris en charge par la trancétolase (Tkt) pourrait être le signe d'une accumulation de Xu5P et ribose-5P limitant la conversion du S7P et du GAP. Le Xu5P n'a pas été quantifié, mais étonnement le Ribose-5P ne présente pas d'accumulation.

Assimilation du xylose - Des traces de fructose 1,6bP ont été quantifiées particulièrement sur xylose (0,22 mM) mais également sur glucose (0,14 mM). Celui-ci ne peut être le produit de la phosphorylation du F6P car la souche ne possède pas d'activité phosphofructokinase. Le fructose 1,6bP est donc produit à partir de GAP et DHAP. Cette observation converge vers l'hypothèse de l'accumulation des substrats de la trancétolase (S7P, GAP, Xu5P, R5P).

Enfin lorsque GA23 pousse sur xylose, 12,06 à 2,52 mM d'inosine sont quantifiés. L'inosine est un métabolite de la voie des purines dérivé du D-ribose-5P, ce qui expliquerait pourquoi le R5P ne s'accumule pas (0,04 mM) **(Tableau 4.13**).

Synthèse d'acide gycolique - Dans les deux conditions testées de l'AG intracellulaire est détecté (2,46 mM sur xylose, 1,22 mM sur glucose) **(Tableaux 4.13, 4.14)**

Métabolisme énergétique **-** L'analyse met en évidence l'absence d'ATP et d'ADP intracellulaire. Une nouvelle méthode d'extraction devra être développée, la durée de filtration est trop importante et engendre la dégradation de l'ADP et ATP sous forme d'AMP. Nous ne pouvons pas exploiter ce résultat en l'état.

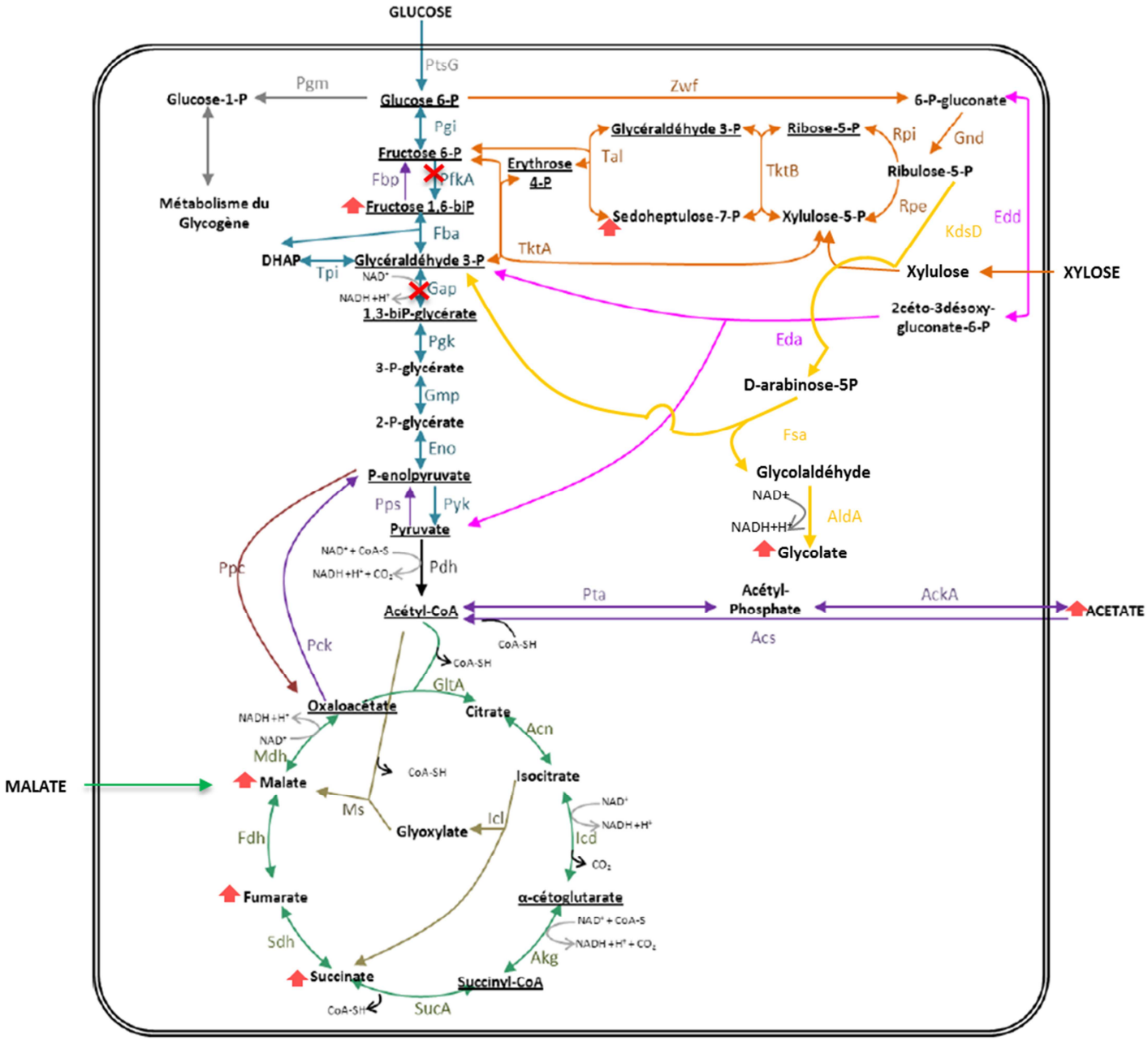


Figure 4.28 : Représentation du métabolisme carboné central de la souche GA23. *Bleu : glycolyse, Vert : cycle de Krebs, Violet : néoglucogénèse, Orange : voie des pentoses phosphate, Rose : voie d'Entner Doudoroff, Marron : cycle glycoxylate, Jaune : voie non naturelle KdsD-Fsa-AldA. Flèche rose : accumulation. Croix rouge : délétion.*

Tableau 4.13 : Analyse de l'endométabolome de la souche GA23 en bioréacteur avec comme source carboné 10g/L de xylose et 0,4 g/L de malate. *En rouge sont mises en évidence les accumulations. AG = acide glycolique, 3PG : 3P-glycérate, R5P : ribulose-5P, G6P : Glucose-6P, S7P : sédoheptulose-7P.*

Temps (h)	AG (mM)	fumarate (mM)	succinate (mM)	malate (mM)	3PG (mM)	citrate (mM)	acétate (mM)	R5P (mM)	G6P (mM)	F6P (mM)	S7P (mM)	F1,6-bisP (mM)	AMP (mM)	inosine (mM)
20	0,87	1,05	0,77	6,12	0,08	0,17	8,04	0,03	0,00	0,00	0,03	0,11	0,00	12,66
44	2,46	13,00	2,55	24,83	0,03	0,43	4,12	0,04	0,04	0,02	0,03	0,22	0,04	2,52

Tableau 4.14 : Analyse de l'endométabolome de la souche GA23 en bioréacteur avec comme source carboné 10g/L de glucose et 0,4 g/L de malate. *En rouge : les accumulations. AG = acide glycolique, 3PG : 3P-glycérate, R5P : ribulose-5P, G6P : Glucose-6P, S7P : sédoheptulose-7P.*

Temps (h)	AG (mM)	fumarate (mM)	succinate (mM)	malate (mM)	3PG (mM)	citrate (mM)	acétate (mM)	R5P (mM)	G6P (mM)	F6P (mM)	S7P (mM)	F1,6-bisP (mM)	AMP (mM)	inosine (mM)
20	0,78	2,25	1,32	8,42	0,01	0,09	8,79	0,01	0,03	0,03	0,03	0,08	0,00	3,48
43,7	1,22	9,41	2,60	22,53	0,07	0,35	6,20	0,02	0,04	0,06	0,18	0,14	0,00	0,65

D) Bilan carbone global

Le bilan carbone réalisé à T40h renseigne sur la distribution du carbone au cours de la croissance. Les monosaccharides et le malate sont convertis principalement en $C0_2$, biomasse, en fumarate, en AG et acétate (**Figure 4.29**).

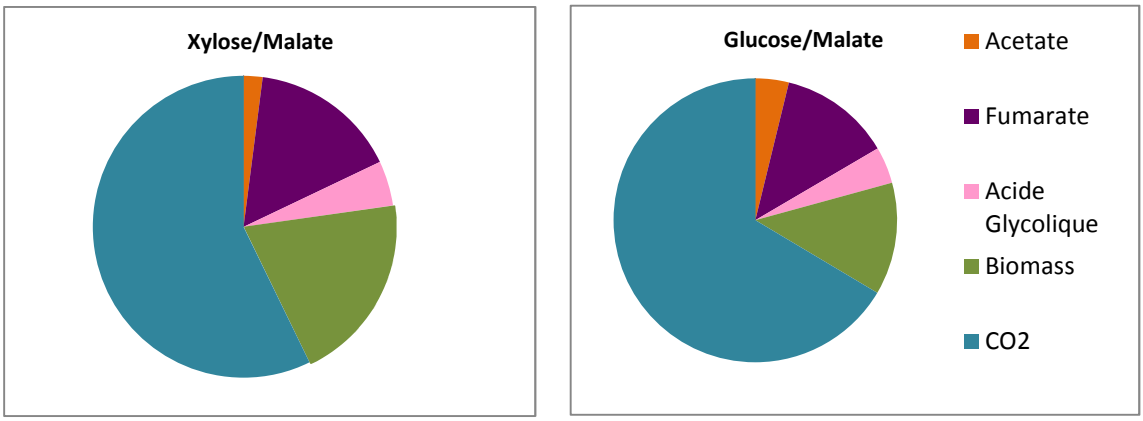


Figure 4.29 : Répartition du carbone dans la souche GA23 en croissance sur en xylose et malate et en glucose et malate à T40h.

Production d'acétate **-** L'acétate est un dérivé de l'acétyl-CoA, c'est un marqueur d'un niveau de NADH cellulaire élevé. La croissance génère du NADH, si les capacités respiratoires ne permettent pas l'oxydation du NADH la citrate synthase, qui catalyse la conversion de l'acétyl-CoA en citrate, est inhibée. L'acétyl-CoA et le pyruvate s'accumulent alors à l'entrée du cycle de Krebs. Afin de réguler la concentration de ces deux métabolites la cellule détourne le surplus de flux de matière vers la production d'acétate. L'acétate est produit à partir de l'acétyl-CoA par la voie Pta AckA. On constate une production d'acétate trois fois plus importante à partir du glucose. Lors du transport du glucose par le système PTS, la synthèse de pyruvate est suivie par une étape de décarboxylation en acétyl-CoA avec libération de CO_2. L'acétyl-CoA est converti en acétate.

Production de fumarate **-** La fumarase catalyse l'interconversion du malate en fumarate. La production de fumarate observée est la conséquence assimilation importante de malate.

Production de CO2 – La carbone assimilé est majoritairement libéré sous forme de de CO_2 Deux réactions du le cycle de Krebs génèrent du CO2 ce qui représente une perte majeure du carbone provenant du malate. Nous pouvons constater que la production de CO2 lorsque le glucose est utilisé comme source de carbone est plus important que sur xylose. En effet le transport du glucose par le système PTS génère du pyruvate qui est décarboxylation en acétyl-CoA. De plus le glucose-6P peut entrer dans la voie des PP, converti en gluconate-6P qui peut être décarboxyler en D-Ru5P.

Production d'acide glycolique - Une partie du carbone est utilisé pour la production d'AG ce qui démontre l'activité de la voie KFA sur xylose et glucose.

Production de biomasse - La biomasse est nécessaire à la production d'AG. La souche GA23 produit plus efficacement la biomasse à partir de xylose, bien que la quantité de biomasse produite soit plus importante à partir de glucose. En culture sur xylose moins de CO_2 est libéré comparé à la culture sur glucose, carbone qui n'est pas libéré sous forme de CO_2 est utilisé pour la biomasse.

3. CONCLUSION

Ce travail a permis la conception d'une voie non-naturelle capable de convertir les pentoses et hexoses en acide glycolique. Le métabolisme carboné central d'*Escherichia coli* a été ingénieré dans le but de minimiser la production de C02 et maximiser le rendement.

Preuve de concept in vitro - La reconstitution *in vitro* des voies métaboliques de conversion des hexoses et pentoses en AG a confirmé la faisabilité thermodynamique de la voie : les enzymes KdsD, Fsa et AldA forment une voie fonctionnelle de conversion du D-Ru5P en AG comme l'atteste la production d'AG mesurée en HPLC et l'apparition de NADH mesurée en UV. Cette étude a mis en évidence les activités secondaires de Fsa sur des métabolites intermédiaires tels que le D-Xu5P, le D-xylulose et le D-Ru5P. L'affinité de Fsa pour le D-xylulose et le D-Ru5P n'a pas été rapportée dans la littérature à notre connaissance. Il sera intéressant de caractériser l'activité enzymatique de Fsa sur ces substrats et s'assurer que l'efficacité catalytique ne soit pas plus élevée que sur D-A5P. Le D-xylulose serait clivé en DHA et GO, le glycoaldéhyde est valorisé en AG mais le DHA ne peut pas être pris en charge par le métabolisme carboné central pour la production d'AG.

Preuve de concept in vivo - La souche MG1655 *ΔtktA ΔtktB ΔglcD* a été utilisée pour le criblage de différentes constructions plasmidiques portant les gènes kdsD-Fsa et/ou AldA. Parmi les 23 systèmes plasmidiques testés, deux ont été sélectionnés. Ces constructions (n°09 et n°23) ont permis la croissance et la production significative d'AG des souches Screen09 et Screen23 à partir de xylose et d'arabinose. La production d'AG par la voie KdsD-Fsa-AldA à partir de pentoses et d'hexose a également été démontrée *in vivo* dans la souche W3CG *ΔgapA ΔmgsA ΔarcA ΔglcD ΔfucA ΔpfkA* galP ^{proD} exprimant le système constitutif n°23. Cette souche a été ingéniérié afin de diriger le carbone vers l'AG. Ainsi (i) le GAP et le F6P sont assimilés par la voie des PP et non pas par la voie basse de la glycolyse, (suppression de *gapA* et *pkfA*) (ii) la consommation du précurseur du GO est limité (suppression de *fucA*), (ii) l'oxydation du glycolate en glyoxylate est abolit (délétion de *glcD*), tandis que l'assimilation des sucres est favorisée par la surexpression de la perméase GalP. La production d'AG de la souche ingéniérée GA23 est de 0,34 g/L à partir de glucose et 0,4 g/L à partir de xylose.

Les suivis de croissance, l'analyse de l'endo- et de l'exométabolome et la réalisation d'un bilan carbone ont permis d'identifier les éléments limitants la production.
En parallèle nous avons démontré que la surexpression couplée de *kdsD* et *fsa* (avec *aldA* exprimé à son niveau naturel) induit une augmentation significative d'AG à partir de xylose comparé à la surexpression de *fsa* seule. Bien que l'analyse *in vitro* mette en évidence des activités secondaires de Fsa, il semblerait que l'efficacité catalytique de Fsa sur l'A5P soit prédominante *in vivo*. La formation d'arabinose-5P par KdsD est nécessaire à la production d'AG.

Eléments limitants la production d'AG - Les productions sont suffisantes pour démontrer la faisabilité de la production d'AG par la voie non-naturelle à partir d'hexose et pentose mais restent trop faibles pour un procédé économiquement viable.
L'assimilation des sucres par la souche GA23 est limitée, voir absente à partir l'arabinose. Le métabolisme des sucres dans *Escherichia coli* est dépendant de la souche utilisée. La souche W3110 est la seule souche K12 capable d'assimiler rapidement le glucose, le xylose ou l'arabinose (Jarmander, Hallström et Larsson, 2014). L'incapacité à absorber l'arabinose de la souche GA23, dérivée de W3110, serait en grande partie dû aux multiples délétions apportées. L'arabinose est transporté dans *Escherichia coli* par un système de symport xylose/H+ (araE) à faible affinité ou par un transporteur ATP dépendant (araFGH) à haute affinité (Kolodrubetz et Schleif, 1981; Horazdovsky et Hogg, 1989). Les modifications du métabolisme carboné central peuvent engendrer une déplétion d'ATP intracellulaire et impacter la capacité d'absorption d'arabinose. Le xylose utilise des systèmes de transport actif, symport xylose/H+ et essentiellement un transporteur ATP dépendant, tout comme l'arabinose, mais il a pu être en partie assimilé. La surexpression de GalP, galactose perméase d'*Escherichia coli*, permettrait de prendre en charge le xylose sans nécessité d'ATP. En effet, Alkim *et al.* (2016) ont montré que la surexpression de GalP améliorait l'assimilation du xylose en phase stationnaire. D'après nos résultats GalP n'aurait pas d'affinité pour l'arabinose. L'impact sur la production d'AG à partir de xylose et arabinose de chacune des modifications apportées à la souche W3110 devra être étudié pour conclure. Par ailleurs, il est prévu de mesurer le niveau d'ADP et d'ATP de la souche GA23 avec une méthode d'extraction adaptée afin de vérifier si le niveau d'ATP limite l'importation des pentoses comme nous le supposons.
L'assimilation du glucose dépend du système PTS qui a pour le moment été conservé dans la souche GA23, la disponibilité de PEP peut être limitée dans cette souche dépourvue de glycéraldéhyde-3P déshydrogénase. La surexpression de GalP permettrait l'assimilation complète du glucose.

D'autres facteurs tels que l'instabilité du système d'expression à deux plasmides et la possible consommation d'AG par la souche limitent la production d'AG. Le système d'expression était stable dans le triple mutant Δ*tktA* Δ*tktB* Δ*glcD*, ce qui nous a permis de réaliser des cultures de 200h. Dans ce mutant la viabilité cellulaire dépendait de l'expression la voie portée par les plasmides, imposant une pression de sélection suffisante pour le maintien des plasmides dans la souche. Dans la souche GA23, la voie n'est pas nécessaire à la croissance et un système représentant une charge métabolique trop importante pourrait rapidement être éliminé en cours de culture. L'AG produit les 40h premières heures de cultures, phase où la majorité des cellules exprimaient la voie, a pu être importé par les cellules. Nous avons recherché les réactions pouvant consommer le glycolate. Seule l'oxydation en glyoxylate est rapportée. Or la délétion de *glcD* prévient la formation de glyoxylate. L'accumulation intracellulaire d'AG observé pourrait être en partie dûe à la synthèse intracellulaire d'AG mais également à l'import d'AG qui ne pourrait être oxydé.

Les faibles rendements obtenus s'expliquent principalement par la perte de carbone sous forme de CO_2 provenant du cycle de Krebs quelle que soit la source de carbone utilisée. En cas de croissance sur glucose la décarboxylation du pyruvate généré par le système PTS et la décarboxylation du gluconate-6P dans la voie d'Entner Doudoroff entrainerait une légère augmentation de la libération de CO_2.

Les informations obtenues permettront de poursuivre de façon pertinente le développement d'un procédé de valorisation des sucres lignocellulosiques, en améliorant l'efficacité de la voie non naturelle KdsD-Fsa-AldA, et en poursuivant l'ingénierie de la souche W3CG *ΔgapA ΔmgsA ΔarcA ΔglcD ΔfucA ΔpfkA* galP ^{proD}. L'optimisation du procédé de fermentation n'a pas été traité dans cette étude, cette étape sera essentielle dans un second temps pour améliorer la capacité de production d'AG.

CHAPITRE V

CONCLUSIONS ET PERSPECTIVES

Notre propos a été de développer une voie de valorisation des pentoses et des hexoses renouvelables, en AG avec un rendement théorique augmenté comparé à l'existant. Pour cela deux approches ont été évaluées au cours de cette thèse avec un même objectif : créer une voie cyclique de production d'AG avec une conservation optimale du carbone. (**Figure 5.1**)

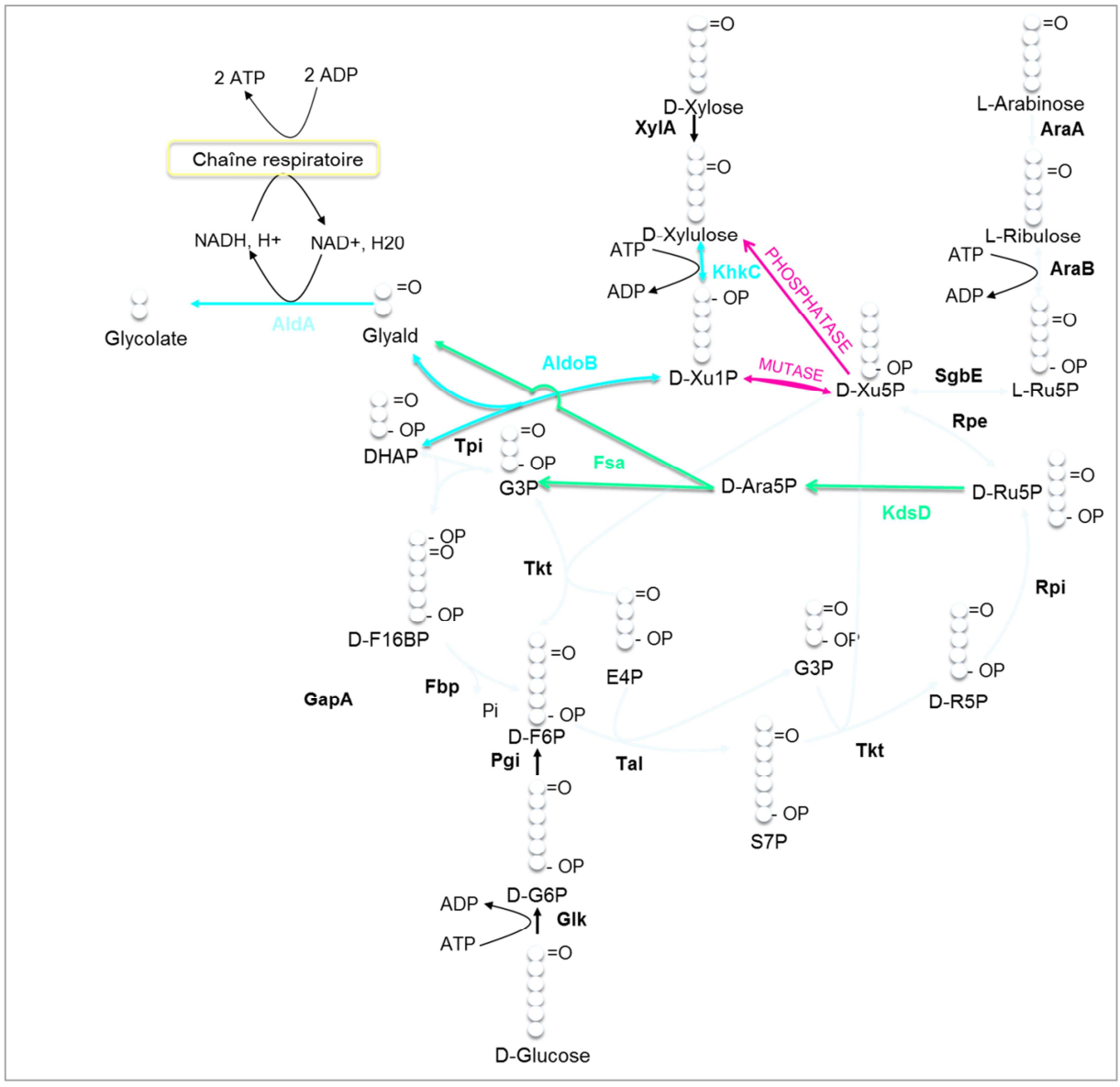


Figure 5.1 : Synthèse des approches évaluées pour la production d'acide glycolique à partir de glucose, de xylose et d'arabinose. *En noir : le métabolisme naturel, en rose : l'activité D-xylulose-5P mutase et phosphatas (approche développée dans le chapitre III), En Bleu : la voie xylulose-1P), en vert : la voie KdsD-Fsa-AldA (approche développée dans le chapitre IV)*

Dans **la première approche** conçue (**Figure 5.1**), les phosphatases et mutases candidates se sont révélées trop peu actives sur le D-xylulose-5P et spécifiques pour ouvrir la voie X1P au L-arabinose et D-glucose (**Chapitre III**). Plusieurs questions ont été soulevées pour la poursuite de ces travaux.

A) Phosphatases

Comment les rechercher ? – Les HADs représentent une des plus grandes superfamilles enzymatiques. Elles sont présentes dans les trois domaines du vivant : bactéries, achées et eucaryotes. Les motifs partagés au sein des membres de cette superfamille ont servi à leur identification à partir des bases de données de séquences protéiques, ainsi la majorité d'entre elles ont été identifié comme HAD à partir d'analyse *in silico*. Au moment de l'écriture de ce manuscrit, plus de 120 000 000 de séquences sont accessibles sur la base de données UniProtKB/TrEMBL, ce chiffre est en constante augmentation (**Figure 5.2).**

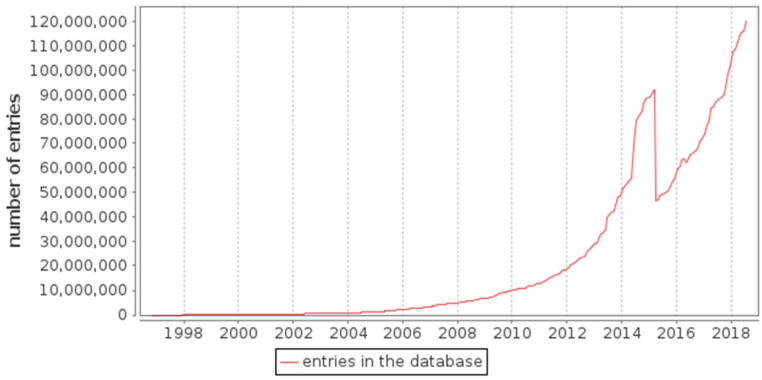


Figure 5.2 : Nombre d'entrée de séquences protéiques dans la base UniProtKB/TrEMBL de 1998 à juillet 2018 *(source : unitpro.org)*

152 306 HADs sont actuellement référencées, seules 341 HADS ont leur structure connue et peu ont été caractérisées. Le mécanisme catalytique est largement documenté tandis que les acides aminés impliqués dans la spécificité de substrat restent inconnus. Les travaux de recherches actuels portent sur (i) le développement de crible haut débit pour améliorer nos connaissances sur les activités biologiques des différentes HADs et sur (ii) l'identification de motifs ou structures déterminant la spécificité de substrat. Les cribles utilisent tous le vert de malachite. Ce test enzymatique permet le dosage du phosphate libre, il est adapté à la caractérisation de phosphatase quel que soit le substrat mais n'est pas adapté à la caractérisation des mutases. Parmi les études importantes du domaine : Kuznetsova a testé 23 enzymes HADs d'*Escherichia coli* sur 80 substrats, 19 sur 23 ont montré une activité sur une large gamme de substrats (Kuznetsova *et al.*, 2006). Huang *et al.* ont ainsi testés 167 métabolites phosphorylés sur plus de 200 enzymes d'espèces procaryotes diverses. Cet échantillonnage représentatif parmi les procaryotes a démontré une promiscuité de substrat des HADs commune à tous les organismes. Parmi les HADs caractérisées plus de 75% d'entre elles ont une affinité pour plus de 5 substrats (Huang *et al.*, 2015). Cela conduit à se demander comment la présence

d'enzymes à large spectre de substrats pourrait contribuer positivement au métabolisme microbien ? Plusieurs études ont porté sur le rôle de la promiscuité au sein des superfamilles enzymatiques (O'Brien and Herschlag, 1999; Allen et Dunaway-Mariano, 2004; Aharoni *et al.*, 2005; Khersonsky, Roodveldt et Tawfik, 2006; Jonas et Hollfelder, 2009; Fisher *et al.*, 2014; Copley, 2015; Pabis, Duarte et Kamerlin, 2016). L'accumulation de métabolites phosphorylés peut perturber le flux métabolique, ou représenter une toxicité l'organisme. Un grand nombre de phosphatase avec de faibles efficacité catalytique assurerait une détoxification cellulaire lorsque les niveaux de métabolites deviennent trop élevés (Huang *et al.*, 2015). Enfin les enzymes promiscuitaires peuvent représenter un avantage au cours de l'évolution en tant que 'réservoir de spécificité'.

Est-il possible de les rendre spécifiques de leur substrat ? Dans les premières études d'évolution datant des années 70's, les évolutionnistes étaient animés par le désir de comprendre comment des mutations pouvaient donner naissance à de nouvelles capacités métabolites (Brown *et al.,* 2018). Clarke et Hall, pionniers de l'évolution dirigée, ont utilisé un organisme déficient pour lequel l'activité recherchée permettait la croissance ce qui a permis l'évolution de nouvelles fonctions. Cela a confirmé la première règle d'évolution dirigée '*on obtient ce que l'on sélectionne'*, la deuxième étant qu'*'on ne peut pas faire évoluer ce qui n'existe pas déjà'*. Une expérience a peu de chance d'être couronnée de succès si l'enzyme de départ ne possède pas déjà l'activité recherchée ; même à un niveau très bas.

A notre connaissance aucun essai d'évolution dirigée n'a été mené pour améliorer la spécificité des HAD phosphatase bien que cette méthode ait été démontrée sur d'autres enzymes telles que la dialkylglycine decarboxylase (Yano, Oue et Kagamiyama, 1998), la biotine ligase (Lu *et al.*, 2014). Nous proposons de développer cette approche pour accroitre la spécificité et l'activité catalytique de YKR70W sur le Xu5P. Nous allons pour cela utiliser le mutant trancétolase (ScreenX1P) construit dans cette étude pour fournir le contexte biologique idéal à l'émergence de l'activité recherchée. En effet, des mutants pourraient acquérir une spécificité pour le D-xylulose-5P dans la mesure où cette activité est avantageuse tel que dans la souche ScreenX1P. A l'inverse, le métabolisme naturel des microorganismes ne favorise pas l'émergence d'enzyme déphosphorylant le Xu5P. Le Xu5P est un métabolite central de la voie des pentoses phosphate, une phosphatase avec une efficacité catalytique sur le D-xylulose-5P ne serait pas maintenue au cours de l'évolution. Cela expliquerait pourquoi Huang *et al.* n'a pas rapporté d'activité sur le Xu5P parmi les 200 phosphatases procaryotiques testées (Huang *et al.*, 2015). La première étape de l'étude consistera à transformer la souche ScreenX1P avec une banque du gène codant YR070W muté aléatoirement selon le protocole décrit par (You *et al.* 2012) et à cribler les transformants pour leur croissance sur L-arabinose. Les clones positifs seront isolés et le gène muté sera séquencé pour identifier les mutations et les variants protéiques correspondants seront caractérisés enzymatiquement. Si la spécificité n'est pas suffisante, nous pourront avoir recours à la technique de DNA shuffling qui permet l'évolution accélérée des protéines *in vitro* (Stemmer, 1994).

Comment valoriser l'activité de Pho13 sur les pentoses-P ? - Pho13 a montré une activité relativement spécifique des pentoses phosphate. Sa capacité à déphosphoryler le D-xylulose-5P, D-ribulose-5P et D-ribose-5P peut être exploité pour générer du xylulose, ribulose et ribose *in vivo*. Ceux-ci pourraient alors être pris en charge par les voies synthétiques existantes de production d'AG à partir de pentoses : (i) la voie X1P composée de KhkC, AldoB et AldA, (ii) la voie ribulose 1P P composée de FucK, FucA et AldA **(Figure 5.3)**

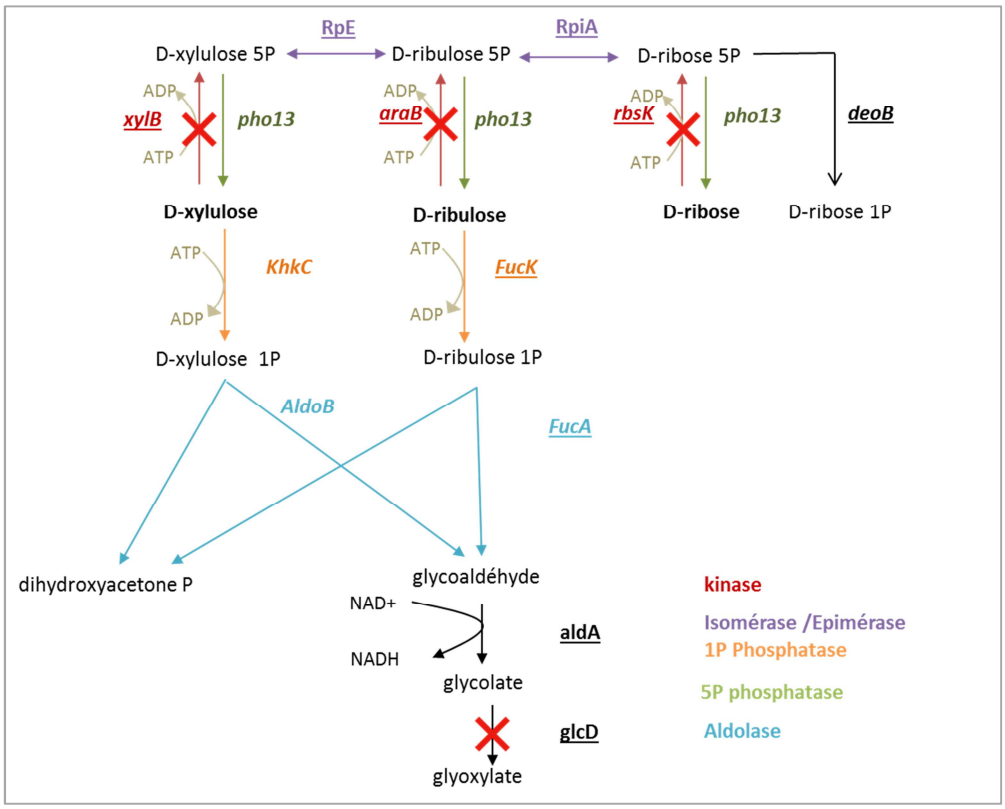


Figure 5.3 : Schématisation d'une voie de production d'AG tirant partie de la capacité de Pho13 à déphosphoryler le xylulose-5P, ribulose-5P et Ribose-5P.

B) Mutases

L'approche de mutagénèse/évolution dirigée proposée pour améliorer la spécificité de YKR070W est aussi adaptée pour l'ingénierie d'une mutase active sur Xu5P. Il faut pour cela identifier une mutase avec une activité secondaire sur ce métabolite. De nouvelles phosphopentomutases candidates pourront être testées, telle que DeoB d'*Escherichia coli*, ou bien PPM de *Bacillus cereus* qui catalysent l'interconversion du D-ribose-5P en D-ribose-1P.

La deuxième voie proposée repose sur la surexpression de la D-arabinose-5P isomérase (KdsD), la fructose-6P aldolase (Fsa), et l'aldéhyde déshydrogénase (AldA) d'*Escherichia coli (***Figure 5.1**) . La souche GA23 exprimant la voie KFSA produit 0,34 g/L d'AG à partir de glucose et 0,41 g/L à partir de xylose.

C) Amélioration de l'efficacité de la voie KdsD-Fsa-AldA

Comment optimiser cette voie ? - L'efficacité catalytique était la plus faible de la voie, elle possède une très bonne affinité pour l'arabinose-5P (0,5 mM) et une constante catalytique très faible (Kcat : 0,6 s-1) qui limite son efficacité. L'ingénierie d'un mutant de Fsa avec un kcat augmenté améliorerait la performance de la voie. La connaissance détaillée de la structure enzymatique de Fsa (PDB ID : 1L6W) (Thorell *et al.*, 2002) et des résidus impliqués dans la liaison avec le substrat, a permis l'ingénie de cette protéine afin d'étendre d'avantage sa gamme de substrat. Comme les enzymes n'ont pas évolué pour posséder l'activité catalytique idéale pour des applications biotechnologiques, plusieurs variants avec des préférences de substrats divergeant de FSA sauvage ont été développés avec succès (i.e FSA-A19S, FSA-A129S/A165G, FSA-L107Y/A129G ou le triple variant FSA-L107Y/A129G/A165G). Une étude a montré a l'inverse une ingénierie mineur avait engendré une promiscuité de substrat (Güclü *et al.*, 2016). Ce qui, pour faire le lien avec les phosphatases, indiquent qu'il est possible au cours de l'évolution de passer de non spécifique à spécifique, et inversement.

Nous avons pu constater que Fsa de type sauvage avait une activité *in vitro* sur les intermédaires métabolique D-xylulose, Xu5P et R5P, en plus de l'A5P. Il sera essentiel dans un premier temps de caractériser Fsa sur les intermédaires métaboliques D-xylulose-5P, D-ribose-5P et D-xylulose. L'activité de Fsa sur xylulose *in vivo* serait à l'origine d'une production d'AG, le DHA produit n'est pas valorisé ce qui représente une perte de carbone. La faible efficacité catalytique de Fsa sur A5P, du fait de la compétition avec d'autres substrats tels que le xylulose, est problématique. Les variants de Fsa publié n'ont pas été optimisé pour le clivage de l'A5P. Un travail de mutagénèse semi-dirigée sera menée afin d'ingénierier un variant au Kcat augmenté pour l'A5P et 'améliorer la performance de la voie KFA.

L'activité de Fsa *in vivo* sur des métabolites autres que l'A5P sera évaluée en étudiant la production d'AG en l'absence d'expression de KdsD. L'isomérase KdsD est impliquée dans la formation d'A5P, si une production significative est observée cela indiquera une activité de Fsa sur d'autres métabolites, l'ingéniérie de Fsa devra alors être poursuivie pour améliorer sa spécificité.

Le système d'expression de la voie fera également l'objet d'optimisation. Tout d'abord pour éliminer le 'metabolic burden' inhérent à l'utilisation de système d'expression à deux vecteurs et assurer la stabilité de l'expression nous envisageons d'intégrer l'opéron *kdsd-fsa-aldA* en plusieurs copies dans le génome d'*Escherichia coli.* L'opéron pourra être optimisé, tester une banque de RBS permettra

d'obtenir la meilleur niveau d'expression possible. Faire varier les RBS a l'avantage de permettre une optimisation de l'expression des gènes de l'opéron de façon individuelle. Le criblage fonctionnet sera utilisé pour identifier la meilleure combinaison sur la base de la lecture de la $DO_{600nm.}$

Enfin, le criblage fonctionnel MG1655 $\Delta tktA\ \Delta tktA\ \Delta glcD$, pourra être amélioré en ajoutant les délétions des gènes *kdsD, fsa, aldA.* De cette manière l'expression naturelle des gènes n'interfèrera pas avec l'évaluation de l'efficacité des systèmes d'expression candidats.

D) Questionnement sur la souche de production d'AG à partir de sucres lignocellulosiques

Comment expliquer les difficultés d'assimilation de l'arabinose et du xylose ?- La souche GA23 produit de l'acide glycolique à partir de xylose et de glucose mais ne consomme pas l'arabinose. La voie KFA a permis la synthèse d'AG dans la souche $\Delta tktA\ \Delta tktA\ \Delta glcD$ (Screen23) à partir d'arabinose. La production est donc thermodynamiquement favorable à partir de ce substrat, seul l'import d'arabinose est limitant.

Le transport d'arabinose est majoritairement ATP dépendant chez *Escherichia coli*. L'absence d'activité GAPDH dans la souche GA23 a pu perturber le métabolisme énergétique. La voie basse de la glycolyse qui fournit habituellement 2 ATP n'est plus alimentée en GAP, cela a pu contribuer à une déplétion en ATP. Le symporteur AraE/H^+ est le deuxième transporteur d'arabinose, il ne nécessite pas d'ATP. La surexpression d'AraE pourrait être une solution mais c'est un transporteur de faible affinité, l'efficacité dépendrait du pH employé pour la culture et l'import massif d'arabinose pourrait acidifier le milieu intracellulaire.

Le transport du xylose sollicite d'autres transporteurs spécifiques, un de haute affinité ATP dépendant, et un autre de faible affinité de type symporteur H^+ tout comme l'arabinose. A première vue un niveau bas d'ATP devrait limiter voir inhiber l'import de xylose. Or la galactose perméase GalP qui a été surexprimée dans GA23 améliore le transport du xylose, comme l'a montré Alkim *et al.,* (2016). L'amélioration du transport du L-arabinose n'a pas été décrit à notre connaissance bien que GalP ait 64% d'homologie avec le transporteur de l'arabinose AraE. Si GalP transporte de façon plus efficace le xylose que l'arabinose, cela pourrait expliquer l'absence d'assimilation observée.

Pour évaluer cette hypothèse il nous faudra dans un premier temps déterminer le niveau intracellulaire d'ATP. Si le niveau d'ATP est réduit comparé à une souche sauvage, la surexpression d'une réaction régénérant l'ATP participera à l'amélioration de l'absorption des sucres. Une fois assimilés par le voie KFA, le NADH formé par l'oxydation du GO est pris en charge par la chaine respiratoire qui forme 2 ATP par NADH, participant au maintien du niveau de consigne de la concentration d'ATP.

Comment s'approcher du rendement théorique ? -De faibles rendements mettent en évidence une perte de carbone considérable. La libération de CO_2 représente la principale perte de carbone à partir des sucres lignocellulosiques malgré les modifications métaboliques apportées à la souche GA23. Les autres pertes de carbone proviennent de l'assimilation de malate qui engendre la production de fumarate et d'acétate.

A court terme plusieurs réactions concourant à la production de CO2 seront éliminées :

- le gène *gapB*, codant pour une glycéraldéhyde-3P déshydrogénase mineure, sera délété pour assurer de la conservation du glycéraldéhyde-3P dans la voie des PP
- Le gène *zwf,* codant pour la glucose-6P déshydrogénase qui catalyse l'entrée du D-glucose dans la voie des pentoses phospate, sera délété. Le D-6P-gluconate est décarboxylé pour sa conversion en D-ribulose-5P (C5) ou bien clivé en pyruvate. Quel que soit le devenir du 6P-gluconate, celui-ci participe à la perte de carbone observée.
- Le système phosphotransferase phosphoenolpyruvate dépendant (PTS) sera supprimé. Dans une souche PTS^-, la surexpression de la galactose permease (GalP) pour le transport du glucose et de la glucokinase (Glk) pour la phosphorylation est nécessaire pour restaurer le phénotype PTS^+(Gosset, G. *et al.* 2005).
- La surexpression de glucokinase Glk et de GalP (Gosset, G. *et al.* 2005).

Une telle souche présente deux avantages significatifs, elle assimile efficacement les sources de carbone renouvelable et rend possible le découplage de la croissance et de la production dans le temps.

Un procédé bi-phasique sera développé pour améliorer le rendement. La première phase sera dédiée à la production de biomasse à partir d'une source de carbone abordable en C2, C3 ou C4 (ex : malate) alimentant le cycle de Krebs. Nous avons pu observer que l'apport immédiat de 0,4g/L de malate conduisait à l'accumulation des intermédiaires du cycle de Krebs et nuisait aux capacités respiratoires de la souche qui produit alors de l'acétate. Il sera important de déterminer la concentration optimale en malate à apporter à la souche. La deuxième phase de la fermentation sera dédiée à la production d'acide glycolique par la biomasse à partir de sucres lignocellulosiques, alors entièrement convertis par la voie cyclique en acide glycolique. De cette manière le rendement approchera le rendement théorique maximal de 3 mol d'AG par mol de C6 et 2,5 mol d'AG par mol de C5.

En conclusion l'atteinte d'un rendement optimal requiert :

- La réduction des activités secondaires de Fsa qui entrainent une perte de carbone (DHA, glycéraldéhyde-3P)
- L'ingénierie de la souche afin de permettre l'isolement de la voie d'assimilation des pentoses et hexoses, et de production de l'AG du reste du métabolisme carboné central
- La mise au point d'un procédé de production d'AG en phase stationnaire à partir de sucres lignocellulosiques, découplé de la croissance sur malate

La production d'AG quant à elle pourra être optimisée par :

- L'augmentation de l'activité de Fsa sur l'A5P
- L'optimisation de l'expression de l'opéron KdsD-Fsa-AldA
- L'intégration en multi-copie de l'opéron au génome d'*E.coli*
- L'amélioration de l'assimilation des sucres lignocellulosiques

Ces points résument les perspectives des travaux de thèse présentés, pour le développement d'un procédé microbien de production d'acide glycolique à partir des pentoses et hexoses lignocellulosiques, avec une conservation optimale du carbone.

RÉFÉRENCES BIBLIOGRAPHIQUES

Articles

Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., and Tawfik, D. S. (2005). The "evolvability" of promiscuous protein functions. Nat. Genet. 37, 73–76. doi:10.1038/ng1482.

Alfke, G., Irion, W. W., and Neuwirth, O. S. (2007). "Oil Refining," in Ullmann's Encyclopedia of Industrial Chemistry

Alkim, C., Cam, Y., Trichez, D., Auriol, C., Spina, L., Vax, A., et al. (2015). Optimization of ethylene glycol production from (d)-xylose via a synthetic pathway implemented in Escherichia coli. Microb. Cell Fact.

Alkim, C., Trichez, D., Cam, Y., Spina, L., François, J. M., and Walther, T. (2016). The synthetic xylulose-1 phosphate pathway increases production of glycolic acid from xylose-rich sugar mixtures. Biotechnol Biofuels 9.

Allen, K. N., and Dunaway-Mariano, D. (2004). Phosphoryl group transfer: Evolution of a catalytic scaffold. Trends Biochem. Sci. doi:10.1016/j.tibs.2004.07.008.

Alonso, S., Rendueles, M., and Díaz, M. (2015). Microbial production of specialty organic acids from renewable and waste materials. Crit. Rev. Biotechnol. 35, 497–513.

Alper, H., Miyaoku, K., and Stephanopoulos, G. (2005). Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. Nat. Biotechnol. 23, 612–616.

Amann, E., Brosius, J., and Ptashne, M. (1983). Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli. Gene 25, 167–178.

Atkinson, D. E., and Walton, G. M. (1965). Kinetics of regulatory enzymes Escherichia coli phosphofructokinase. J. Biol. Chem. 240, 757–764.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. Mol. Syst. Biol. 2.

Bachmann, B. J. (1996). Derivations and Genotypes of Some Mutant Derivatives of Escherichia coli K-12. Escherichia coli Salmonella typhimurium Cell. Mol. Biol., 2460–2488.

Baev, M. V., Baev, D., Jansco Radek, A., and Campbell, J. W. (2006). Growth of Escherichia coli MG1655 on LB medium: Monitoring utilization of amino acids, peptides, and nucleotides with transcriptional microarrays. Appl. Microbiol. Biotechnol. 71, 317–322.

Baldomà, L., and Aguilar, J. (1987). Involvement of lactaldehyde dehydrogenase in several metabolic pathways of Escherichia coli K12. J. Biol. Chem. 262, 13991–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3308886.

Baritugo, K.-A., Kim, H. T., David, Y., Choi, J.-I., Hong, S. H., Jeong, K. J., et al. (2018). Metabolic engineering of Corynebacterium glutamicum for fermentative production of chemicals in biorefinery. Appl. Microbiol. Biotechnol. 102.

Batt, C. A., Jamieson, A. C., and Vandeyar, M. A. (1990). Identification of essential histidine residues in the active site of Escherichia coli xylose (glucose) isomerase. Proc. Natl. Acad. Sci. U. S. A. 87, 618–22. doi:10.1073/pnas.87.2.618.

Baykov, A. A., Evtushenko, O. A., and Avaeva, S. M. (1988). A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. Anal. Biochem. 171, 266–270. doi:10.1016/0003-2697(88)90484-8.

Becker, J., and Wittmann, C. (2016). Systems metabolic engineering of Escherichia coli for the heterologous production of high value molecules. Curr. Opin. Biotechnol. 42, 178–188.

Bertani, G., Krenek, P., Samajova, O., Luptovciak, I., Doskocilova, A., Komis, G., et al. (2006). LB (Luria-Bertani) liquid medium. Cold Spring Harb. Protoc. 2006 ,8141.

Blattner, F. R. (1997). The Complete Genome Sequence of Escherichia coli K-12. Science 80. 277, 1453–1462.

Bogorad, I. W., Lin, T.-S., and Liao, J. C. (2013). Synthetic non-oxidative glycolysis enables complete carbon conservation. Nature.

Bozell, J. J., and Petersen, G. R. (2010). Cutting-edge research for a greener sustainable future Technology development for the production of biobased products from biorefinery carbohydrates — the US Department of Energy ' s " Top 10 " revisited.

Bremer, E., Silhavy, T. J., Weisemann, J. M., and Weinstock, G. M. (1984). λ placMu: A transposable derivative of bacteriophage lambda for creating lacZ protein fusions in a single step. J. Bacteriol. 158, 1084–1093.

Buschke, N., Schäfer, R., Becker, J., and Wittmann, C. (2013). Metabolic engineering of industrial platform microorganisms for biorefinery applications - Optimization of substrate spectrum and process robustness by rational and evolutive strategies. Bioresour. Technol. 135, 544–554.

Cabulong, R. B., Lee, W., Bañares, A. B., Ramos, K. R. M., Nisola, G. M., Valdehuesa, K. N. G., et al. (2018). Engineering Escherichia coli for glycolic acid production from D-xylose through the Dahms pathway and glyoxylate bypass.

Cabulong, R. B., Valdehuesa, K. N. G., Ramos, K. R. M., Nisola, G. M., Lee, W. K., Lee, C. R., et al. (2017). Enhanced yield of ethylene glycol production from D-xylose by pathway optimization in Escherichia coli. Enzyme Microb. Technol. 97, 11–20.

Cam, Y., Alkim, C., Trichez, D., Trebosc, V., Vax, A., Bartolo, F., et al. (2016). Engineering of a Synthetic Metabolic Pathway for the Assimilation of (d)-Xylose into Value-Added Chemicals. ACS Synth. Biol.

Castaño-Cerezo, S., Pastor, J. M., Renilla, S., Bernal, V., Iborra, J. L., and Cánovas, M. (2009). An insight into the role of phosphotransacetylase (pta) and the acetate/acetyl-CoA node in Escherichia coli. Microb. Cell Fact.

Chae, T. U., Ko, Y.-S., Kyu-Sang, H., and Lee, S. Y. (2017). Metabolic engineering of Escherichia coli for the production of lactam. Metab. Eng. 41, 82–91.

Cheng, K. K., Wang, G. Y., Zeng, J., and Zhang, J. A. (2013). Improved succinate production by metabolic engineering. Biomed Res. Int. 2013.

Choi, S. Y., Kim, W. J., Yu, S. J., Park, S. J., Im, S. G., and Lee, S. Y. (2017). Engineering the xylose-catabolizing Dahms pathway for production of poly (D -lactate- co -glycolate) and poly (D -lactate- co -glycolate- co - D -2-hydroxybutyrate) in Escherichia coli.

Chung, C. T., Niemela, S. L., and Miller, R. H. (1989). One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. 86, 2172–2175. doi:10.1073/pnas.86.7.2172.

Chung, T., Klumpp, D. J., and LaPorte, D. C. (1988). Glyoxylate bypass operon of Escherichia coli: cloning and determination of the functional map. J. Bacteriol. 170, 386–392.

Clark, D. P. (1989). The fermentation pathways of Escherichia coli. FEMS Microbiol. Lett. 63, 223–234.

Colanduoni, J. A., and Villafranca, J. J. (1985). Labeling of a specific arginine residue at the active site of glutamine synthetase (E. coli). Biochem. Biophys. Res. Commun. 126, 412–418. doi:10.1016/0006-291X(85)90621-7.

Copley, S. D. (2015). An evolutionary biochemist's perspective on promiscuity. Trends Biochem. Sci. 40, 72–78. doi:10.1016/j.tibs.2014.12.004.

Cox, R. S., Surette, M. G., and Elowitz, M. B. (2007). Programming gene expression with combinatorial promoters. Mol. Syst. Biol. 3.

Dai, Z., and Nielsen, J. (2015). Advancing metabolic engineering through systems biology of industrial microorganisms. Curr. Opin. Biotechnol. 36, 8–15.

Datsenko, K., and Wanner, B. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. … 2000, 6640–6645.

Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. Nucleic Acids Res. 39, 1131–1141. doi:10.1093/nar/gkq810.

de Boer, H. A., Comstock, L. J., and Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc. Natl. Acad. Sci. 80, 21–25.

Demain, A. L., and Adrio, J. L. (2008). Contributions of microorganisms to industrial biology. Mol. Biotechnol. 38, 41–55.

Deng, Y., Ma, N., Zhu, K., Mao, Y., Wei, X., and Zhao, Y. (2018). Balancing the carbon flux distributions between the TCA cycle and glyoxylate shunt to produce glycolate at high yield and titer in Escherichia coli. Metab. Eng. 46, 28–34.

Deng, Y., Mao, Y., and Zhang, X. (2015). Metabolic engineering of E . coli for efficient production of glycolic acid from glucose. Biochem. Eng. J. 103, 256–262.

Deutscher, J., Francke, C., Postma, P. W., Deutscher, J., Francke, C., and Postma, P. W. (2006). How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria How Phosphotransferase System-Related Protein Phosphorylation Regulates

Di Luccio, E., Petschacher, B., Voegtli, J., Chou, H. T., Stahlberg, H., Nidetzky, B., et al. (2007). Structural and Kinetic Studies of Induced Fit in Xylulose Kinase from Escherichia coli. J. Mol. Biol. 365, 783–798. doi:10.1016/j.jmb.2006.10.068.

Dong, J. M., Taylor, J. S., Latour, D. J., Iuchi, S., and Lin, E. C. C. (1993). Three overlapping lct genes involved in L-lactate utilization by Escherichia coli. J. Bacteriol. 175, 6671–6678. doi:10.1128/jb.175.20.6671-6678.1993.

Dykxhoorn, D. M., Pierre, R. S., and Linn, T. (1996). A set of compatible tac promoter expression vectors. Gene 177, 133–136.

EC 2010a, Communication from the Commission: Europe 2020, A strategy for smart, sustainable and inclusive growth, COM(2010) 2020.

EC 2012b, Commission Staff Working Document, accompanying the document Communication on Innovating for Sustainable Growth: A Bioeconomy for Europe.

EC 2012c, Communication from the Commission to the European Parliament and the Council: The EU Approach to Resilience: Learning from Food Security Crisis, COM(2012)

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the human intestinal microbial flora. Science 308, 1635–8.

Eiteman, M. A., and Altman, E. (2006). Overcoming acetate in Escherichia coli recombinant protein fermentations. Trends Biotechnol. 24, 530–536. doi:10.1016/j.tibtech.2006.09.001.

Ferguson, G. P., Battista, J. R., Annettee, T. L., and Booth, I. R. (2000). Protection of the DNA during the exposure of Escherichia coli cells to a toxic metabolite: The role of the KefB and KefC potassium channels. Mol. Microbiol. 35, 113–122. doi:10.1046/j.1365-2958.2000.01682.x.

Fischer, C. R., Klein-Marcuschamer, D., and Stephanopoulos, G. (2008). Selection and optimization of microbial hosts for biofuels production. Metab. Eng. 10, 295–304.

Fisher, A. K., Freedman, B. G., Bevan, D. R., and Senger, R. S. (2014). A review of metabolic and enzymatic engineering strategies for designing and optimizing performance of microbial cell factories. Comput. Struct. Biotechnol. J. 11, 91–99. doi:10.1016/j.csbj.2014.08.010.

Ganter, C., and Pliickthun, A. (1990). Glycine to Alanine Substitutions in Helices of Glyceraldehyde-3-phosphate Dehydrogenase : Effects on Stabilityt.

Garrabou, X., Castillo, J. A., Guérard-Hélaine, C., Parella, T., Joglar, J., Lemaire, M., et al. (2009). Asymmetric self- and cross-aldol reactions of glycolaldehyde catalyzed by d-fructose-6-phosphate aldolase. Angew. Chemie - Int. Ed. doi:10.1002/anie.200902065.

Georgellis, D., Kwon, O., and Lin, E. C. C. (1999). Amplification of Signaling Activity of the Arc Two-component System of Escherichia coli by Anaerobic Metabolites. J Biol Chem 274, 35950–35954.

Georgescu-Roegen, N. (1975). Energy and Economic Myths. South. Econ. J. 41, 347.

Ghalambor, M. A., and Heath, E. C. (1962). The metabolism of L-fucose. II. The enzymatic cleavage of L-fuculose 1-phosphate. J. Biol. Chem. 237, 2427–2433.

Gimenez, R., Nuñez, M. F., Badia, J., Aguilar, J., and Baldoma, L. (2003). The Gene yjcG, Cotranscribed with the Gene acs, Encodes an Acetate Permease in Escherichia coli. J. Bacteriol. 185, 6448–6455. doi:10.1128/JB.185.21.6448-6455.2003.

Goeddel, D. V., Heyneker, H. L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D. G., et al. (1979). Direct expression in Escherichia coli of a DNA sequence coding for human growth hormone. Nature 281, 544–548.

Goffeau, A., Barrell, G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., et al. (1996). Life with 6000 genes. Science (80-.). 274, 546–567.

Güclü, D., Szekrenyi, A., Garrabou, X., Kickstein, M., Junker, S., Clapés, P., et al. (2016). Minimalist Protein Engineering of an Aldolase Provokes Unprecedented Substrate Promiscuity. ACS Catal. 6, 1848–1852. doi:10.1021/acscatal.5b02805.

Guérard-Hélaine, C., Debacker, M., Clapés, P., Szekrenyi, A., Hélaine, V., and Lemaire, M. (2014). Efficient biocatalytic processes for highly valuable terminally phosphorylated C5 to C9 <scp>d</scp> -ketoses. Green Chem. 16, 1109–1113. doi:10.1039/C3GC42140F.

He, Y. C., Xu, J. H., Su, J. H., and Zhou, L. (2010). Bioproduction of Glycolic acid from glycolonitrile with a new bacterial isolate of alcaligenes sp. ECU0401. Appl. Biochem. Biotechnol. 160, 1428–1440.

Heermann, R., Zeppenfeld, T., and Jung, K. (2008). Simple generation of site-directed point mutations in the Escherichia coli chromosome using Red®/ET® Recombination. Microb. Cell Fact. 7.

Hellinga, H. W., and Evans, P. R. (1985). Nucleotide sequence and high-level expression of the major Escherichia coli phosphofructokinase. Eur. J. Biochem. 149, 363–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3158524.

Hong, K. K., and Nielsen, J. (2012). Metabolic engineering of Saccharomyces cerevisiae: A key cell factory platform for future biorefineries. Cell. Mol. Life Sci. 69, 2671–2690.

Horazdovsky, B. F., and Hogg, R. W. (1989). Genetic reconstitution of the high-affinity L-arabinose transport system. J. Bacteriol. 171, 3053–3059. doi:10.1128/jb.171.6.3053-3059.1989.

Huang, H., Pandya, C., Liu, C., Al-Obaidi, N. F., Wang, M., Zheng, L., et al. (2015). Panoramic view of a superfamily of phosphatases through substrate profiling. Proc. Natl. Acad. Sci. 112, E1974–E1983. doi:10.1073/pnas.1423570112.

Hunter, I. S., and Kornberg, H. L. (1979). Glucose Transport of Escherichia coli Growing in Glucose-Limited Continuous Culture. Biochem. J 178, 97–101.

International Energy Agency (2017). Renewables 2017: Analysis and Forecasts to 2022 - Executive Summary. J. Qual. Particip. doi:10.1073

Jarmander, J., Hallström, B. M., and Larsson, G. (2014). Simultaneous uptake of lignocellulose-based monosaccharides by Escherichia Coli. Biotechnol. Bioeng. 111, 1108–1115. doi:10.1002/bit.25182.

Jeschek, M., Gerngross, D., and Panke, S. (2016). Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. Nat. Commun. 7.

Jeschek, M., Gerngross, D., and Panke, S. (2017). Combinatorial pathway optimization for streamlined metabolic engineering. Curr. Opin. Biotechnol. 47, 142–151.

Jonas, S., and Hollfelder, F. (2009). Mapping catalytic promiscuity in the alkaline phosphatase superfamily. Pure Appl. Chem. 81, 731–742. doi:10.1351/PAC-CON-08-10-20.

Jørgensen, H., Kristensen, J. B., and Felby, C. (2007). Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities. Biofuels, Bioprod. Biorefining 1, 119–134.

Kataoka, M., Sasaki, M., Hidalgo, A.-R. G. D., Nakano, M., and Shimizu, S. (2001). Glycolic Acid Production Using Ethylene Glycol-Oxidizing Microorganisms. Biosci. Biotechnol. Biochem. 65, 2265–2270.

Kawe, M., Horn, U., and Plückthun, A. (2009). Facile promoter deletion in Escherichia coli in response to leaky expression of very robust and benign proteins from common expression vectors. Microb. Cell Fact. 8. doi:10.1186/1475-2859-8-8.

Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006). Enzyme promiscuity: evolutionary and mechanistic aspects. Curr. Opin. Chem. Biol. 10, 498–508. doi:10.1016/j.cbpa.2006.08.011.

Koivistoinen, O. M., Kuivanen, J., Barth, D., Turkia, H., Pitkänen, J.-P., Penttilä, M., et al. (2013). Glycolic acid production in the engineered yeasts Saccharomyces cerevisiae and Kluyveromyces lactis. Microb. Cell Fact.

Kolodrubetz, D., and Schleif, R. (1981). L-Arabinose transport systems in Escherichia coli K-12. J. Bacteriol. 148, 472–479.

Kolpak, F. J., and Blackwell, J. (1976). Determination of the Structure of Cellulose II. Macromolecules 9, 273–278.

Kornberg, H. L. (1957). The metabolism of C2 compounds in micro-organisms. Biochem. J. 68, 535–542.

Kornberg, H. L. (1966). The role and control of the glyoxylate cycle in Escherichia coli. Biochem. J. 99, 1–11.

Krymkiewicz, N. (1973). Reactions of methylglyoxal with nucleic acids. FEBS Lett. 29, 51–54. doi:10.1016/0014-5793(73)80013-4.

Kuhlman, T. E., and Cox, E. C. (2010). Site-specific chromosomal integration of large synthetic constructs. Nucleic Acids Res. 38.

Kundig, W., Ghosh, S., and Roseman, S. (1964). Phosphate Bound To Histidine in a Protein As an Intermediate in a Novel Phospho-Transferase System*. Pnas 52, 1067–1074.

Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. V., Borozan, I., et al. (2006). Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. J. Biol. Chem. 281, 36149–36161. doi:10.1074/jbc.M605449200.

Lam, V. M. S., Daruwalla, K. R., Henderson, P. J. F., and Jones-Mortimer, M. C. (1980). Proton-linked D-xylose transport in Escherichia coli. J. Bacteriol. 143, 396–402. doi:10.1016/0306-4492(80)90056-8.

Laporte, D. C., and Koshland, D. E. (1982). A protein with kinase and phosphatase activities involved in regulation of tricarboxylic acid cycle. Nature 300, 458–460.

Laurent, P., Roiz, J., Wertz, J. L., Richel, A., and Paquot, M. (2011). Le bioraffinage, une alternative prometteuse à la pétrochimie. Biotechnol. Agron. Soc. Environ. 15, 597–610.

Lee, L. V., Gerratana, B., and Cleland, W. W. (2001). Substrate specificity and kinetic mechanism of Escherichia coli ribulokinase. Arch. Biochem. Biophys. 396, 219–224. doi:10.1006/abbi.2001.2613.

Lee, L. V., Poyner, R. R., Vu, M. V., and Cleland, W. W. (2000). Role of metal ions in the reaction catalyzed by L-ribulose-5-phosphate 4-epimerase. Biochemistry 39, 4821–4830. doi:10.1021/bi9928952.

Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J., and Dueber, J. E. (2013). Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Res. 41, 10668–10678.

Liao, J. C., Mi, L., Pontrelli, S., and Luo, S. (2016). Fuelling the future: Microbial engineering for the production of sustainable biofuels. Nat. Rev. Microbiol. 14, 288–304.

Liu, X., and De Wulf, P. (2004). Probing the ArcA-P Modulon of Escherichia coli by Whole Genome Transcriptional Analysis and Sequence Recognition Profiling. J. Biol. Chem. 279, 12588–12597.

Lorca, G. L., Ezersky, A., Lunin, V. V., Walker, J. R., Altamentova, S., Evdokimova, E., et al. (2007). Glyoxylate and pyruvate are antagonistic effectors of the Escherichia coli IcIR transcriptional regulator. J. Biol. Chem. 282, 16476–16491.

Lord, J. M. (1972). Glycolate oxidoreductase in Escherichia coli. BBA - Bioenerg. 267, 227–237. doi:10.1016/0005-2728(72)90111-9.

Lorenz, M. C., and Fink, G. R. (2002). Life and death in a macrophage: Role of the glyoxylate cycle in virulence. Eukaryot. Cell 1, 657–662.

Lu, W. C., Levy, M., Kincaid, R., and Ellington, A. D. (2014). Directed evolution of the substrate specificity of biotin ligase. Biotechnol. Bioeng. 111, 1071–1081. doi:10.1002/bit.25176.

Mainguet, S. E., Gronenberg, L. S., Wong, S. S., and Liao, J. C. (2013). A reverse glyoxylate shunt to build a non-native route from C4to C2in Escherichia coli. Metab. Eng. 19, 116–127.

Majewski R A, D., and Majewski R A, Domach, M. M. (1990). Simple constrained optimization view of acetate overflow in E.coli. Biotechnol bioeng 35, 731–738.

Maloy, S. R., and Nunn, W. D. (1982). Genetic regulation of the glyoxylate shunt in Escherichia coli K-12. J. Bacteriol. 149, 173–180.

Martínez-Núñez, M., Rodríguez-Escamilla, Z., Rodríguez-Vázquez, K., and Pérez-Rueda, E. (2017). Tracing the Repertoire of Promiscuous Enzymes along the Metabolic Pathways in Archaeal Organisms. Life 7, 30.

McCormick, K., and Kautto, N. (2013). The Bioeconomy in Europe: An Overview. Sustain. 5, 2589–2608.

McDonald, T. P., Walmsley, A. R., and Henderson, P. J. F. (1997). Asparagine 394 in putative helix 11 of the galactose-H+symport protein (GalP) from Escherichia coli is associated with the internal binding site for cytochalasin B and sugar. J. Biol. Chem. 272, 15189–15199.

Meredith, T. C., and Woodard, R. W. (2003). Escherichia coli YrbH is a D-arabinose 5-phosphate isomerase. J. Biol. Chem. doi:10.1074/jbc.M303661200.

Meredith, T. C., and Woodard, R. W. (2006). Characterization of Escherichia coli D-arabinose 5-phosphate isomerase encoded by kpsF : implications for group 2 capsule biosynthesis. Biochem. J. doi:10.1042/BJ20051828.

Meyer, D., Schneider-Fresenius, C., Horlacher, R., Peist, R., and Boos, W. (1997). Molecular characterization of glucokinase from Escherichia coli K-12. J. Bacteriol. 179, 1298–306. doi:10.1128/jb.179.4.1298-1306.1997.

Ministère de la Transition écologique et solidaire (2017). https://www.ecologique-solidaire.gouv.fr/biocarburants.

Monk, J. M., Koza, A., Campodonico, M. A., Machado, D., Seoane, J. M., Palsson, B. O., et al. (2016). Multi-omics Quantification of Species Variation of Escherichia coli Links Molecular Features with Strain Phenotypes. Cell Syst. 3, 238–251.e12.

Moussatova, A., Kandt, C., O'Mara, M. L., and Tieleman, D. P. (2008). ATP-binding cassette transporters in Escherichia coli. Biochim. Biophys. Acta - Biomembr. 1778, 1757–1771. doi:10.1016/j.bbamem.2008.06.009.

Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., et al. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. Nat. Methods 10, 354–360. doi:10.1038/nmeth.2404.

Naraian, R., and Kumari, S. (2017). "Microbial Production of Organic Acids," in Microbial Functional Foods and Nutraceuticals, 93–121. doi:10.1002/9781119048961.ch5.

Núñez, M. F., Aguilar, J., Baldoma, L., Kwon, O., Wilson, T. H., and Lin, E. C. C. (2002). Transport of L-lactate, D-lactate, and glycolate by the LldP and GlcA membrane carriers of Escherichia coli. Biochem. Biophys. Res. Commun. 290, 824–829. doi:10.1006/bbrc.2001.6255.

Nuñez, M. F., Pellicer, M. T., Badia, J., Aguilar, J., and Baldoma, L. (2001). Biochemical characterization of the 2-ketoacid reductases encoded by ycdW and yiaE genes in Escherichia coli. Biochem. J. 354, 707–15. doi:10.1042/0264-6021:3540707.

O'Brien, P. J., and Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. Chem. Biol. 6. doi:10.1016/S1074-5521(99)80033-7.

Objectifs du Millénaire pour le développement Rapport 2008 (2008).

Oecd (2009). The Bioeconomy of 2030. doi:10.1787/9789264056886-en.

Ornston, L. N., and Ornston, M. K. (1969). Regulation of glyoxylate metabolism in Escherichia coli K-12. J. Bacteriol. 98, 1098–1108.

Pabis, A., Duarte, F., and Kamerlin, S. C. L. (2016). Promiscuity in the Enzymatic Catalysis of Phosphate and Sulfate Transfer. Biochemistry 55, 3061–3081. doi:10.1021/acs.biochem.6b00297.

Pacala, S., and Socolow, R. (2004). Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. Science (80-.). 305, 968–972. doi:10.1126/science.1100103.

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496, 528–532. doi:10.1038/nature12051.

Paes, B. G., and Almeida, J. R. (2014). Genetic improvement of microorganisms for applications in biorefineries. Chem. Biol. Technol. Agric. 1. doi:10.1186/s40538-014-0021-1.

P**anova, A., Mersinger, L. J., Liu, Q., Foo, T., Roe, D. C., Spillan, W. L., et al. (2007).** Chemoenzymatic Synthesis of Glycolic Acid. Adv. Synth. Catal. 349, 1462–1474. doi:10.1002/adsc.200700061.

Pellicer, M. T., Badía, J., Aguilar, J., and Baldomà, L. (1996). glc Locus of Escherichia coli: Characterization of genes encoding the subunits of glycolate oxidase and the glc regulator protein. J. Bacteriol. 178, 2051–2059. doi:10.1128/jb.178.7.2051-2059.1996.

Pellicer, M. T., Lynch, A. S., De Wulf, P., Boyd, D., Aguilar, J., and Lin, E. C. C. (1999). A mutational study of the ArcA-P binding sequences in the aldA promoter of Escherichia coli. Mol. Gen. Genet. 261, 170–176. doi:10.1007/s004380050954.

Pereira, B., Li, Z. J., De Mey, M., Lim, C. G., Zhang, H., Hoeltgen, C., et al. (2016). Efficient utilization of pentoses for bioproduction of the renewable two-carbon compounds ethylene glycol and glycolate. Metab. Eng. 34, 80–87. doi:10.1016/j.ymben.2015.12.004.

Perrenoud, A., and Sauer, U. (2005). Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in Escherichia coli. J. Bacteriol. 187, 3171–3179. doi:10.1128/JB.187.9.3171-3179.2005.

Pitera, D. J., Paddon, C. J., Newman, J. D., and Keasling, J. D. (2007). Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. Metab. Eng. 9, 193–207. doi:10.1016/j.ymben.2006.11.002.

Polisky, B., Bishop, R. J., and Gelfand, D. H. (1976). A plasmid cloning vehicle allowing regulated expression of eukaryotic DNA in bacteria. Proc. Natl. Acad. Sci. U. S. A. 73, 3900–4. doi:10.1073/pnas.73.11.3900.

Pontrelli, S., Chiu, T. Y., Lan, E. I., Chen, F. Y. H., Chang, P., and Liao, J. C. (2018). Escherichia coli as a host for metabolic engineering. Metab. Eng. doi:10.1016/j.ymben.2018.04.008.

Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993). Phosphoenolpyruvate: Carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. 57, 543–594. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=372926&tool=pmcentrez&rendertype=abstract.

Redden, H., and Alper, H. S. (2015). The development and characterization of synthetic minimal yeast promoters. Nat. Commun. 6. doi:10.1038/ncomms8810.

Rodríguez-Zavala, J. S., Allali-Hassani, A., and Weiner, H. (2006). Characterization of E. coli tetrameric aldehyde dehydrogenases with atypical properties compared to other aldehyde dehydrogenases. Protein Sci. 15, 1387–1396. doi:10.1110/ps.052039606.

Saier, M. H. (2000). Families of transmembrane sugar transport proteins. Mol. Microbiol. 35, 699–710. doi:10.1046/j.1365-2958.2000.01759.x.

Sanofi (2012). Prix Pierre Potier 2012 Sanofi récompensé par le prix Pierre Potier 2012 pour son innovation industrielle française contre le paludisme. 1–8.

Sathitsuksanoh, N., Zhu, Z., and Zhang, Y. H. P. (2012). Cellulose solvent- and organic solvent-based lignocellulose fractionation enabled efficient sugar release from a variety of lignocellulosic feedstocks. Bioresour. Technol. 117, 228–233. doi:10.1016/j.biortech.2012.04.088.

Sauer, M., Porro, D., Mattanovich, D., and Branduardi, P. (2008). Microbial production of organic acids: expanding the markets. Trends Biotechnol. 26, 100–108. doi:10.1016/j.tibtech.2007.11.006.

Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. Annu. Rev. Plant Biol. 61, 263–289. doi:10.1146/annurev-arplant-042809-112315.

Schmid, O., Padel, S., and Levidow, L. (2012). The bio-economy concept and knowledge base in a public goods and farmer perspective. Bio-based Appl. Econ. 1, 47–63. doi:10.13128/BAE-10770.

Schuchmann, K., and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. Nat. Rev. Microbiol. 12, 809–821. doi:10.1038/nrmicro3365.

Schürmann, M., and Sprenger, G. A. (2001). Fructose-6-phosphate Aldolase is a Novel Class I Aldolase from Escherichia coli and is Related to a Novel Group of Bacterial Transaldolases. J. Biol. Chem. doi:10.1074/jbc.M008061200.

Schwander, T., Von Borzyskowski, L. S., Burgener, S., Cortina, N. S., and Erb, T. J. (2016). A synthetic pathway for the fixation of carbon dioxide in vitro. Science. 354, 900–904. doi:10.1126/science.aah5237.

Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., and Liao, J. C. (2011). Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli. Society 77, 2905–15. doi:10.1128/AEM.03034-10.

Shen, X., Wang, J., Wang, J., Chen, Z., Yuan, Q., and Yan, Y. (2017). High-level De novo biosynthesis of arbutin in engineered Escherichia coli. Metab. Eng. 42, 52–58. doi:10.1016/j.ymben.2017.06.001.

Siegel, J. B., Smith, A. L., Poust, S., Wargacki, A. J., Bar-Even, A., Louw, C., et al. (2015). Computational protein design enables a novel one-carbon assimilation pathway. Proc. Natl. Acad. Sci., 201500545. doi:10.1073/pnas.1500545112.

Smanski, M. J., Bhatia, S., Zhao, D., Park, Y. J., Woodruff, L. B. A., Giannoukos, G., et al. (2014). Functional optimization of gene clusters by combinatorial design and assembly. Nat. Biotechnol. 32, 1241–1249. doi:10.1038/nbt.3063.

Sobota, J. M., and Imlay, J. A. (2011). Iron enzyme ribulose-5-phosphate 3-epimerase in Escherichia coli is rapidly damaged by hydrogen peroxide but can be protected by manganese. Proc. Natl. Acad. Sci. 108, 5402–5407. doi:10.1073/pnas.1100410108.

Sprenger, G. a, Schörken, U., Sprenger, G., and Sahm, H. (1995). Transketolase A of Escherichia coli K12. Purification and properties of the enzyme from recombinant strains. Eur J Biochem 230, 525–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7607225.

Staffas, L., Gustavsson, M., and McCormick, K. (2013). Strategies and policies for the bioeconomy and bio-based economy: An analysis of official national approaches. Sustain. 5, 2751–2769. doi:10.3390/su5062751.

Stemmer, W. P. (1994). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc. Natl. Acad. Sci. 91, 10747–10751. doi:10.1073/pnas.91.22.10747.

Stephen Dahms, A. (1974). 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochem. Biophys. Res. Commun. 60, 1433–1439. doi:10.1016/0006-291X(74)90358-1.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., et al. (1966). Frameshift Mutations and the Genetic Code. Cold Spring Harb. Symp. Quant. Biol. 31, 77–84. doi:10.1101/SQB.1966.031.01.014.

Szekrenyi, A., Soler, A., Garrabou, X., Guérard-Hélaine, C., Parella, T., Joglar, J., et al. (2014). Engineering the Donor Selectivity of D-Fructose-6-Phosphate Aldolase for Biocatalytic Asymmetric Cross-Aldol Additions of Glycolaldehyde. Chem. - A Eur. J. 20, 12572–12583. doi:10.1002/chem.201403281.

Teleman, A., Richard, P., Toivari, M., and Penttilä, M. (1999). Identification and quantitation of phosphorus metabolites in yeast neutral pH extracts by nuclear magnetic resonance spectroscopy. Anal. Biochem. 272, 71–79. doi:10.1006/abio.1999.4165.

Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010). The population genetics of commensal Escherichia coli. Nat.Rev.Microbiol. 8, 207–217. doi:10.1038/nrmicro2298.

Thakker, C., Martínez, I., San, K.-Y., and Bennett, G. N. (2012). Succinate production in Escherichia coli. Biotechnol. J. 7, 213–224. doi:10.1002/biot.201100061.

Thorell, S., Schürmann, M., Sprenger, G. A., and Schneider, G. (2002). Crystal structure of decameric fructose-6-phosphate aldolase from Escherichia coli reveals inter-subunit helix swapping as a structural basis for assembly differences in the transaldolase family. J. Mol. Biol. 319, 161–171. doi:10.1016/S0022-2836(02)00258-9.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin Biosynthesis and Structure. PLANT Physiol. 153, 895–905. doi:10.1104/pp.110.155119.

Vemuri, G. N., Altman, E., Sangurdekar, D. P., Khodursky, A. B., and Eiteman, M. A. (2006). Overflow metabolism in Escherichia coli during steady-state growth: Transcriptional regulation and effect of the redox ratio. Appl. Environ. Microbiol. 72, 3653–3661. doi:10.1128/AEM.72.5.3653-3661.2006.

Verduyn, C., Zomerdijk, T. P. L., van Dijken, J. P., and Scheffers, W. A. (1984). Continuous measurement of ethanol production by aerobic yeast suspensions with an enzyme electrode. Appl. Microbiol. Biotechnol. 19, 181–185. doi:10.1007/BF00256451.

Volkmer, B., and Heinemann, M. (2011). Condition-Dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling. PLoS One 6. doi:10.1371/journal.pone.0023126.

Walther, T., Baylac, A., Alkim, C., Vax, A., Cordier, H., and François, J. M. (2012). The PGM3 gene encodes the major phosphoribomutase in the yeast Saccharomyces cerevisiae. FEBS Lett. 586, 4114–4118. doi:10.1016/j.febslet.2012.09.042.

Wei, G., Yang, Æ. X., Gan, Æ. T., Zhou, W., Lin, Æ. J., Wei, Æ. D., et al. (2009). High cell density fermentation of Gluconobacter oxydans DSM 2003 for glycolic acid production. 1029–1034. doi:10.1007/s10295-009-0584-1.

Werpy, T., and Petersen, G. (2004). Top Value Added Chemicals from Biomass Volume I — Results of Screening for Potential Candidates from Sugars and Synthesis Gas Top Value Added Chemicals From Biomass Volume I : Results of Screening for Potential Candidates. Other Inf. PBD 1 Aug 2004, Medium: ED; Size: 76 pp. pages. doi:10.2172/15008859.

Wertz, J.-L., Richel, A., and Gérin, P. (2015). Molécules issues de la valorisation de la lignine. ValBiom Mars, 1–37. Available at: http://hdl.handle.net/2268/182162.

Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A. G. W., Collins, C. H., et al. (2013). Modular optimization of multi-gene pathways for fatty acids production in E. coli. Nat. Commun. 4. doi:10.1038/ncomms2425.

Xu, P., Rizzoni, E. A., Sul, S. Y., and Stephanopoulos, G. (2017). Improving metabolic pathway efficiency by statistical model-based multivariate regulatory metabolic engineering. ACS Synth. Biol. 6, 148–158. doi:10.1021/acssynbio.6b00187.

Yano, T., Oue, S., and Kagamiyama, H. (1998). Directed evolution of an aspartate aminotransferase with new substrate specificities. Biochemistry 95, 5511–5515.

Yoon, S. H., Han, M. J., Jeong, H., Lee, C. H., Xia, X. X., Lee, D. H., et al. (2012). Comparative multi-omics systems analysis of Escherichia coli strains B and K-12. Genome Biol. 13. doi:10.1186/gb-2012-13-5-r37.

You, C., and Percival Zhang, Y. H. (2012). Easy preparation of a large-size random gene mutagenesis library in Escherichia coli. Anal. Biochem. 428, 7–12. doi:10.1016/j.ab.2012.05.022.

Yu, J. H., Zhu, L. W., Xia, S. T., Li, H. M., Tang, Y. L., Liang, X. H., et al. (2016). Combinatorial optimization of CO2transport and fixation to improve succinate production by promoter engineering. Biotechnol. Bioeng. 113, 1531–1541. doi:10.1002/bit.25927.

Yuan, Y., Du, J., and Zhao, H. (2013). Customized optimization of metabolic pathways by combinatorial transcriptional engineering. Methods Mol. Biol. 985, 177–209. doi:10.1007/978-1-62703-299-5_10.

Zahoor, A., Otten, A., and Wendisch, V. F. (2014). Metabolic engineering of Corynebacterium glutamicum for glycolate production. J. Biotechnol. 192, 366–375. doi:10.1016/j.jbiotec.2013.12.020.

Zeng, A. P., and Sabra, W. (2011). Microbial production of diols as platform chemicals: Recent progresses. Curr. Opin. Biotechnol. 22, 749–757. doi:10.1016/j.copbio.2011.05.005.

Zhang, X., Chen, S., and Xu, F. (2017). Combining Raman Imaging and Multivariate Analysis to Visualize Lignin, Cellulose, and Hemicellulose in the Plant Cell Wall. J. Vis. Exp. doi:10.3791/55910.

Zhang, X., Jantama, K., Shanmugam, K. T., and Ingram, L. O. (2009). Reengineering Escherichia coli for succinate production in mineral salts medium. Appl. Environ. Microbiol. 75, 7807–7813. doi:10.1128/AEM.01758-09.

Zheng, J., Yates, S. P., and Jia, Z. (2012). Structural and mechanistic insights into the bifunctional enzyme isocitrate dehydrogenase kinase/phosphatase AceK. Philos. Trans. R. Soc. B Biol. Sci. 367, 2656–2668. doi:10.1098/rstb.2011.0426.

Zhu, L.-W., and Tang, Y.-J. (2017). Current advances of succinate biosynthesis in metabolically engineered Escherichia coli. Biotechnol. Adv. doi:10.1016/j.biotechadv.2017.09.007.

Brevets

U.S. 7,445,917 (2006) Process for producing glycolic acid from formaldehyde and hydrogen cyanide. Robert DiCosimo, Anna Panova, Jeffery Scott Thompson, Robert D. Fallon, F. Glenn Gallagher, Thomas Foo, Xu Li, George C. Fox, Joseph J. Zaher, Mark S. Payne, Daniel P. O'Keefe

US2012/0315682(2012) Method for producing high amount of glycolic acid by fermentation. Wanda Discher, Philippe Soucaille, Metabolic Explorer

US2152852A (1939) Process for manufacture of glycolic acid. Donald John Loder, Wilmington, Del., assignor to E. I. du Pont deNemours & Company, Wilmington, Del., a corporation of Delaware No Drawing.

US3141033A (1960) Production of glycolic acid nitrile. Koening Franz, Evonik Degissa GmbH

WO2006/069110 (2006) Production enzymatique d'acide glycolique **.** Robert Dicosimo, Mark S. Payne, Anna Panova, Daniel P. O'keefe, Jeffrey Scott Thompson

WO2007/141316 (2007) Production d'acide glycolique par fermentation à partir de ressources renouvelables. Philippe Soucaille, Metabolic Explorer

WO2010004032 (2010) Procédé pour la polymérisation d'acide glycolique avec des micro-organismes Philippe Soucaille, Wanda Dischert, Metabolic Explorer

WO2011036213A2 (2011) Processus de fermentation pour la production d'acide glycolique. Wanda Dischert, Cédric Collomb, Philippe Soucaille, Metabolic Explorer

WO2016079440 (2016) Procédé de production d'au moins un métabolite d'intérêt par transformation d'un pentose dans un microorganisme **.** Thomas Walther, Yvan Cam, Jean-Marie Francois, Ceren Alkim, Institut National De La Recherche Agronomique – Inra, Institut National Des Sciences Appliquees – Insa, Centre National De La Recherche Scientifique.

WO2017059236 (2017) Production microbienne de glycolate renouvelable. Gregory Stephanopoulos, Zheng-jun LI, Brian Pereira, Massachusetts Institute Of Technology