



AVERTISSEMENT

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

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

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

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

http://www.cfcopies.com/V2/leg/leg_droi.php

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

UNIVERSITE HENRI POINCARÉ – NANCY I

2000

FACULTE DE PHARMACIE



THESE

**MODELISATION DE L'INSTABILITE PLASMIDIQUE
SEGREGATIONNELLE DE SOUCHES RECOMBINEES.
APPLICATION A *ESCHERICHIA COLI***

Présentée et soutenue publiquement

le 12 décembre 2000

pour obtenir

Le Diplôme d'Etat de Docteur en Pharmacie

par Bao-Linh LE

Membres du Jury

Président : M. Alain BAGREL, Professeur des Universités

Juges : M. Joseph BOUDRANT, Directeur de Recherche au C.N.R.S.
M. Christian FONTEIX, Maître de Conférences à l'E.N.S.A.I.A.
M. Frantz FOURNIER, Maître de Conférences à l'E.N.S.A.I.A.

BU PHARM. ODONTOL.



D 104 054726 0

UNIVERSITE HENRI POINCARÉ – NANCY I

2000

DA 24023

FACULTE DE PHARMACIE



THESE

**MODELISATION DE L'INSTABILITE PLASMIDIQUE
SEGREGATIONNELLE DE SOUCHES RECOMBINEES.
APPLICATION A *ESCHERICHIA COLI***

Présentée et soutenue publiquement

le 12 décembre 2000

pour obtenir

Le Diplôme d'Etat de Docteur en Pharmacie

par Bao-Linh LE

Membres du Jury

Président : M. Alain BAGREL, Professeur des Universités

Juges : M. Joseph BOUDRANT, Directeur de Recherche au C.N.R.S.
M. Christian FONTEIX, Maître de Conférences à l'E.N.S.A.I.A.
M. Frantz FOURNIER, Maître de Conférences à l'E.N.S.A.I.A.

FACULTE DE PHARMACIE

UNIVERSITE Henri Poincaré - NANCY 1

Membres du personnel enseignant

Doyen : Mme Chantal FINANCE

Vice Doyen : Mme Anne ROVEL

DOYENS HONORAIRES

M. BERNANOSE André

M. VIGNERON Claude

PROFESSEURS HONORAIRES

Mle BESSON Suzanne

Mle GIRARD Thérèse

M. LECTARD Pierre

M. MIRJOLET Marcel

M. PIERFITTE Maurice

PROFESSEUR EMERITE

M. LOPPINET Vincent

PROFESSEURS

M.	ASTIER Alain	Pharmacie Clinique
M.	ATKINSON Jeffrey	Pharmacologie
M.	BAGREL Alain	Biochimie fondamentale et clinique, Biotechnologies
Mle	BATT Anne Marie	Toxicologie
M.	BLOCK Jean Claude	Santé et Environnement
M.	BONALY Roger	Biochimie microbienne
Mme	CAPDEVILLE-ATKINSON	Pharmacologie Cardiovasculaire
Mme	FINANCE Chantal	Microbiologie moléculaire
Mme	FRIANT-MICHEL Pascale	Biomathématiques, Biophysique et Audioprothèse
Mle	GALTEAU Marie Madeleine	Biochimie
M.	HENRY Max	Biologie végétale
M.	HOFFMAN Maurice	Pharmacie clinique
M.	JACQUE Michel	Pharmacodynamie
M.	LABRUDE Pierre	Physiologie
M.	LALLOZ Lucien	Chimie organique
M.	LEROY Pierre	Physico-chimie appliquée à la formulation pharmaceutique
M.	MAINCENT Philippe	Pharmacie galénique
M.	MARSURA Alain	Chimie thérapeutique
M.	MARTIN Jean Armand	Chimie minérale et Minéralogie
M.	MORTIER François	Pharmacognosie
M.	NICOLAS Alain	Chimie analytique et Bromatologie
M.	REGNOUF DE VAINS Jean Bernard	Chimie Thérapeutique
Mme	SCHWARTZBROD Janine	Bactériologie - Parasitologie
M.	SCHWARTZBROD Louis	Virologie - Immunologie
M.	SIEST Gérard	Chimie Biologique
M.	SIMON Jean Michel	Droit et Economie de la Santé
M.	VIGNERON Claude	Hématologie

MAITRES DE CONFERENCES

Mme	ALBERT Monique	Bactériologie - Virologie
M.	BONNEAUX François	Chimie Thérapeutique
M.	CATAU Gérard	Pharmacodynamie
M.	CHEVIN Jean Claude	Chimie minérale
M.	CHILLON Jean Marc	Pharmacologie
M.	COLLIN Jean François	Pôle européen
Mme	COLLOMB Jocelyne	Parasitologie
M.	COULON Joël	Biochimie
M.	DECOLIN Dominique	Chimie analytique
M.	DUCOURNEAU Joël	Biophysique, Audioprothèse, Acoustique
Mme	FAIVRE-FIORINA Béatrice	GBM - Hématologie
M.	FERRARI Luc	Biochimie
Mle	FONS Françoise	Biologie Végétale et Mycologie
Mme	FUZELLIER Marie Claude	Pharmacognosie
M.	GANTZER Christophe	Virologie
M.	GIBAUD Stéphane	Pharmacie Clinique
Mme	HASENFRATZ-SAUDER Marie Paule	Biologie Végétale
Mle	HINZELIN Françoise	Biologie végétale et Pharmacognosie
M.	HUMBERT Thierry	Interactions moléculaires
Mle	IMBS Marie Andrée	Bactériologie - Virologie et Parasitologie
M.	JORAND Frédéric	Santé et Environnement
Mme	KEDZIEREWICZ Francine	Pharmacie Galénique
Mme	LARTAUD-IDJOUADIENE Isabelle	Pharmacologie
Mme	LEININGER-MULLER Brigitte	Biochimie
Mme	LETOT Michèle	Bactériologie - Virologie et Parasitologie
Mme	LIVERTOUX Marie Hélène	Toxicologie
Mme	MARCHAL-HEUSSLER Emmanuelle	Chimie Analytique
Mme	MARCHAND-ARVIER Monique	Immunologie - Hématologie
M.	MENU Patrick	Physiologie
M.	MONAL Jean Louis	Chimie Thérapeutique
M.	NOTTER Dominique	Biologie cellulaire
Mme	PAULUS Francine	Informatique
Mme	PERDIAKIS Christine	Chimie organique
Mme	PICHON Virginie	Biophysique
Mme	POCHON Marie France	Chimie analytique
Mme	ROVEL Anne	Immunologie - Hématologie
M.	VISVIKIS Athanase	Toxicologie
Mme	WELLMAN-ROUSSEAU Maria Monika	Biochimie
Mme	ZINUTTI Colette	Pharmacie galénique

ASSISTANTS

Mme	BEAUD Mariette	Biologie Cellulaire
Mme	BERTHE Marie-Catherine	Biochimie
M.	DANGIEN Bernard	Botanique
Mme	MOREAU Blandine	Pharmacognosie
Mme	PAVIS Annie	Parasitologie
M.	TROCKLE Gabriel	Pharmacodynamie

PROFESSEUR ASSOCIE

Mme	GRISON Geneviève	Pratiques officinales
-----	------------------	-----------------------

PROFESSEUR AGREGE

M.	COCHAUD Christophe	Anglais
----	--------------------	---------

SERMENT DES APOTHICAIRES



Je jure, en présence des maîtres de la Faculté, des conseillers de l'ordre des pharmaciens et de mes condisciples :

D'honorer ceux qui m'ont instruit dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle à leur enseignement.

D'exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de l'honneur, de la probité et du désintéressement.

De ne jamais oublier ma responsabilité et mes devoirs envers le malade et sa dignité humaine ; en aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.

Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.



« LA FACULTE N'ENTEND DONNER AUCUNE APPROBATION,
NI IMPROBATION AUX OPINIONS EMISES DANS LES
THESES, CES OPINIONS DOIVENT ETRE CONSIDEREES
COMME PROPRES A LEUR AUTEUR ».

Avant-propos

Je remercie Monsieur Hans-Günther LINTZ, Directeur du Laboratoire des Sciences du Génie Chimique (L.S.G.C.) de Nancy de m'avoir accueillie au sein du laboratoire.

Ce travail a été réalisé sous la direction de Monsieur Joseph BOUDRANT, Directeur de Recherche au C.N.R.S. Je tiens à lui adresser mes vifs remerciements pour son soutien irremplaçable et ses conseils avisés tout au long de ce travail.

J'exprime ma sincère reconnaissance à Monsieur Christian FONTEIX, Maître de Conférences à l'E.N.S.A.I.A., de m'avoir fait profiter de l'étendue de ses connaissances, de la profondeur de ses idées et sans l'aide duquel ce travail n'aurait pu se faire.

Je remercie Monsieur Alain BAGREL, Professeur à la Faculté des Sciences Pharmaceutiques et Biologiques de Nancy, d'avoir accepté la présidence de ce jury et pour les conseils qu'il m'a prodigué.

Un grand merci à toute l'équipe du L.S.G.C. en particulier à Evelyne RENARD, à Marie-Thérèse BARTHE, à Bernard PARISSET pour la qualité de leur accueil et à Frantz FOURNIER pour son dynamisme et son aide efficace.

Introduction

L'évolution des techniques du génie génétique ont permis un essor considérable dans la production de métabolites et de protéines pouvant être à haute valeur ajoutée. Une des avancées majeures a été l'intervention des organismes recombinés dans la production de ces protéines. Les procédés mis en jeu impliquent plusieurs étapes dont :

- l'isolation du gène codant pour un produit recherché,
- l'insertion de ce gène dans un vecteur d'expression contenant des séquences de contrôle, le but étant de maximiser l'expression de ce gène. On parle alors de gène recombiné. On fait souvent appel aux plasmides en tant que vecteurs d'expression.
- l'introduction du vecteur d'expression dans une cellule hôte (opération de transformation) qui va prendre en charge la transcription et la traduction du gène. La cellule est alors dite recombinée. On utilise habituellement des bactéries, des levures ou des cellules animales dans la production de protéines recombinantes.

Les cultures de masse en bioréacteur permettent de produire de grandes quantités de protéines recombinantes, or un problème souvent rencontré dans les fermentations de souches recombinées est le phénomène d'instabilité du plasmide. En effet, pendant la phase de croissance et la phase stationnaire, les cellules recombinées se divisent et peuvent donner des cellules ayant perdu une partie ou la totalité de leurs plasmides ou encore des cellules contenant des plasmides modifiés. Ainsi ces cellules ne peuvent plus exprimer le gène étranger. De plus, les souches sans plasmides ont souvent une cinétique de croissance plus importante que celle des souches recombinées de départ.

On comprend alors que l'instabilité plasmidique soit une contrainte majeure dans les fermentations industrielles. D'une part, les pertes de plasmides affectent sévèrement les productivités et les rendements des cultures, ceci est d'autant plus vrai que les durées des fermentations sont souvent importantes. D'autre part, les modifications structurales du plasmide peuvent entraîner des changements dans la séquence des acides aminés de la protéine produite. Ainsi, la qualité et la reproductibilité du produit synthétisé peuvent être affectées par ces phénomènes d'instabilité.

Il existe deux types d'instabilité plasmidique. La première est de type ségrégationnel. Elle correspond à la perte du ou des plasmides du fait d'une partition défectueuse lors de la division cellulaire. Il en résulte une diminution voire une annulation de la productivité des fermentations. On parle également d'instabilité structurale lorsqu'intervient un changement dans la structure du plasmide dû à des insertions, délétions ou réarrangements de l'ADN. La cellule recombinée ne peut plus alors exprimer le gène d'intérêt. Ces instabilités génèrent donc des populations mixtes au sein desquelles les souches ségrégantes se développent plus rapidement. On peut ainsi aboutir à des cultures ne permettant plus de synthétiser le produit d'intérêt.

Un des principaux objectifs du génie biochimique est la mise en œuvre à l'échelle industrielle des souches recombinées. L'optimisation de la production passe par une maîtrise des procédures de conduite afin d'atteindre les meilleurs rendements de conversion.

Afin de réaliser ces objectifs, il est nécessaire de comprendre et de prévoir le comportement de ces souches bactériennes, afin de pouvoir optimiser, extrapoler et automatiser au mieux les procédés biotechnologiques. Pour cela, on fait de souvent appel à la modélisation. La formulation de modèles mathématiques permet de décrire le comportement et la croissance de souches recombinées et de comprendre les processus qui régissent ces systèmes complexes. Ces modèles sont basés sur des hypothèses, lesquelles sont vérifiées en comparant les résultats des simulations aux données expérimentales.

Nous avons dans cette étude cherché à modéliser le comportement de la souche recombinée *Escherichia coli* HB101 (GADPH) responsable l'expression de la glycéraldéhyde-3-phosphate déshydrogénase. Ce système modèle a été sélectionné du fait que la GADPH est directement proportionnelle au nombre de plasmides, sans présenter de phénomène de saturation. Cette protéine recombinée est soluble dans le cytoplasme bactérien, non toxique pour la cellule et non dénaturée. Enfin, le nombre de plasmides peut être facilement évalué par la technique des CMI.

Pour étudier le comportement de cette souche, nous avons mis en place un modèle mathématique explicitant le bilan des fractions de biomasse contenant de 0 à 30 plasmides et pour lequel la cinétique de croissance des souches bactériennes est bien différenciée de celle de répllication plasmidique. De plus, nous avons défini pour ce modèle deux expressions de la probabilité de perte de plasmides, l'une basée sur une distribution de répartition uniforme,

l'autre sur une distribution de type binomial. Les paramètres du modèle ont été déterminés par rapport à une culture en continu sans pression de sélection et avec une population recombinée de départ homogène où toutes les bactéries possèdent 30 plasmides. Les paramètres déterminés ont été ensuite appliqués à une culture en discontinu sans pression de sélection et avec une population initiale hétérogène.

Nous avons d'abord réalisé une identification paramétrique sur un modèle simple à deux populations : l'une contenant des plasmides, l'autre n'en contenant pas. Les paramètres obtenus ainsi que l'évolution des concentrations sont satisfaisants.

Etant donné que dans l'un de nos modèles ségrégués, une simple sommation des équations de bilan de populations permet de retrouver, à un terme près, le modèle non ségrégué à deux équations, nous avons appliqué ces paramètres à notre modèle. Les simulations obtenues alors ne montrent pas une adéquation suffisante avec les résultats expérimentaux, c'est pourquoi une identification paramétrique sur les modèles ségrégués s'est avérée nécessaire. Les paramètres déterminés suite à cette identification ainsi que l'évolution des simulations nous ont permis de conclure que :

- le modèle de probabilité de répartition uniforme, même s'il permet de simuler de façon satisfaisante les données expérimentales, ne constitue pas un modèle suffisant du fait de la valeur improbable d'un de ses paramètres. Etant donné que ce modèle ne différait de celui d'Imanaka et Aiba que par un terme, nous avons pu montrer l'importance de ce dernier,
- le modèle de probabilité de distribution binomiale permet, lui, d'avoir une approche des phénomènes d'instabilité, d'autres facteurs restant à prendre en compte pour une modélisation optimale.

Les paramètres appliqués aux données expérimentales en culture discontinue nous ont amené aux mêmes conclusions.

Chapitre 1 : Revue bibliographique

Sommaire

Avant propos.....	2
Introduction.....	3
Sommaire.....	6
Chapitre 1 :Revue bibliographique	

1. Production de protéines par des organismes recombinés.....	7
1.1. Généralités sur les systèmes recombinés	7
1.2. cellules hôtes.	8
1.2.1. Escherichia coli	10
1.2.2. Saccharomyces cerevisiae	11
1.2.3. Cellules de mammifères	11
1.3. Les vecteurs d'expression	12
2. Les facteurs influençant la production de protéines recombinantes	15
2.1. Instabilité plasmidique	15
2.1.1. Instabilité structurale	17
2.1.2. Instabilité ségrégationnelle.....	18
2.1.3. Méthodes d'évaluation de la stabilité plasmidique	24
2.1.4. Techniques de stabilisation	26
2.2. ARNm : Acide ribonucléique messager.....	30
2.2.1. Nombre de copies de plasmides	32
2.2.2. Optimisation de la transcription	34
2.2.3. Stabilité de l'ARNm.....	37
2.3. Traduction	37
2.4. Devenir post-traductionnel des protéines.....	38
2.4.1. Protéines cytoplasmiques	38
2.4.2. Dégradation protéolytique.....	40
2.4.3. Sécrétion de protéines chez E. coli.....	41
2.4.4. Excrétion des protéines recombinantes dans le milieu.....	41
2.5. Métabolisme de l'acétate chez E. coli.....	43
2.5.1. Mécanismes de production en aérobiose.....	43
2.5.2. Effet inhibiteur de l'acétate	44
3. Mise en œuvre en réacteur	46
4. Modèles mathématiques.....	48
4.1. Modèles non ségrégués	49
4.2. Modèles ségrégués	57

Chapitre 1 : Revue bibliographique

1. Production de protéines par des organismes recombinés

1.1. Généralités sur les systèmes recombinés

Les avancées du génie génétique ont permis de grands progrès dans la production de protéines et de métabolites actifs. Cette production fait généralement intervenir des micro-organismes recombinés.

Un système recombiné résulte de l'association entre un hôte, qui peut être un micro-organisme ou un organisme multicellulaire, et un vecteur d'expression (plasmide par exemple) contenant un gène codant pour une protéine d'intérêt.

La construction d'un tel système repose sur plusieurs étapes :

- isolement du gène codant pour la protéine recherchée,
- introduction de ce gène dans un vecteur d'expression, en l'occurrence un plasmide. Ce vecteur doit contenir des séquences de contrôle afin de permettre une expression optimale du gène,
- introduction du vecteur à l'intérieur de la cellule hôte (bactérie), c'est ce qu'on appelle une transformation bactérienne.

Le but de ces systèmes recombinés est d'assurer la synthèse par la cellule hôte d'une protéine bien déterminée codée par le gène introduit par le vecteur d'expression. Les protéines obtenues sont dites recombinantes. La cellule productrice résultant d'une combinaison entre le vecteur d'expression et la cellule hôte est dite recombinée.

1.2. cellules hôtes.

De nombreux micro-organismes, organismes ou cellules peuvent être utilisés comme hôtes. On peut citer les bactéries comme *Escherichia coli*, les levures comme *Saccharomyces cerevisiae*, les cellules CHO, les cellules végétales ou encore les organismes entiers dits transgéniques.

L'hôte idéal devrait :

- exécuter correctement les instructions fournies par le gène étranger et réaliser les modifications post-traductionnelles permettant d'obtenir une protéine identique à la protéine naturelle,
- se prêter à la culture de masse en fermenteur, être résistant et se multiplier à densités élevées,
- produire en grande quantité une protéine identique à la protéine naturelle,
- ne pas produire d'élément toxique,
- et être peu exigeant d'un point de vue nutritionnel.

Chaque système de production résulte d'une combinaison d'avantages et inconvénients. Ceux-ci sont regroupés dans le **Tableau I**. Le système hôte sera sélectionné en fonction de la protéine à produire et de son utilisation. Le choix peut être évident dans certains cas, par exemple, si la protéine nécessite des glycosylations, (les systèmes bactériens, ne permettant pas cette modification post-traductionnelle, seront éliminés d'office).

En règle générale, *E. coli*, *Saccharomyces cerevisiae* et les cellules CHO, sont les hôtes principalement utilisés pour la production de protéines recombinées, c'est pourquoi les paragraphes suivants leur seront consacrés. Toutefois la bactérie *Bacillus*, ainsi que les levures comme *Pichia pastoris*, *Hansenula* et *Yarrowia lipolitica* sont également utilisées du fait qu'elles sécrètent facilement les protéines produites dans le milieu de culture.

Pour déterminer les données expérimentales de notre étude, nous avons utilisé comme micro-organisme hôte *Escherichia coli* et plus particulièrement la souche HB101. Cette souche a été isolée par Boyer et Roulland-Dussoix en 1969 lors d'une étude du système restriction-modification d'*E. coli*. Au cours de cette étude, toute une série de mutants avaient été générés dont *E. coli* HB101. Cette dernière dérivait de la souche *E. coli* HB100 (r+m+), elle-même issue d'*E. coli* HB16 (*E. coli* K12 avec le système de restriction modification d'*E. coli* B). *E. coli* HB101 a été l'une de premières souches capable d'accepter un ADN étranger, à être isolée.

Tableau I. : Avantages et inconvénients de différents systèmes de production de protéines recombinées. (A. Gschaedler, 1994)

	Hôte	Avantages	Inconvénients
BACTERIES	<i>E. coli</i>	<ul style="list-style-type: none"> • beaucoup d'informations sur la physiologie et la biologie moléculaire de cette souche • nombreux vecteurs construits • taux d'expressions élevés 	<ul style="list-style-type: none"> • pas de sécrétion • pas de modifications post traductionnelles • protéines souvent insolubles • présence d'endotoxines
	<i>Bacillus</i>	<ul style="list-style-type: none"> • bonne sécrétion • non pathogène • utilisé industriellement 	<ul style="list-style-type: none"> • produit des protéases • modifications post-traductionnelles difficiles
	<i>Streptomyces</i>	<ul style="list-style-type: none"> • fortes densités cellulaires • utilisé industriellement • bonne sécrétion 	<ul style="list-style-type: none"> • souche moins bien connue
LEVURES ET CHAMPIGNONS	<i>Saccharomyces cerevisiae</i>	<ul style="list-style-type: none"> • beaucoup d'informations sur la physiologie et la biologie moléculaire de cette souche • utilisé industriellement • modifications post traductionnelles 	<ul style="list-style-type: none"> • faible taux d'expression • cellules difficiles à casser • glycosylation imparfaite
	Champignons filamenteux	<ul style="list-style-type: none"> • bonne sécrétion • bonne glycosylation • non pathogène 	<ul style="list-style-type: none"> • variabilité de l'expression • produit des protéases • moins bonnes connaissances sur l'expression
CELLULES EUCARYOTES	Cellules animales	<ul style="list-style-type: none"> • production de molécules de grande taille • bonne glycosylation 	<ul style="list-style-type: none"> • croissance lente, taux d'expression faible • milieux coûteux • problèmes de cultures en grandes quantités
	Cellules végétales	<ul style="list-style-type: none"> • protéines actives 	<ul style="list-style-type: none"> • cellules fragiles, culture délicate • croissance lente • milieux coûteux
	Cellules d'insectes	<ul style="list-style-type: none"> • bonne sécrétion • maturation correcte • milieux aprotéiques 	<ul style="list-style-type: none"> • cellules fragiles • glycosylation parfois imparfaite

1.2.1. *Escherichia coli*

La bactérie *E. coli* constitue le système hôte le mieux connu, c'est le premier qui a été utilisé pour l'obtention de protéines recombinantes (Georgiou, 1988). Cette bactérie dispose en effet de nombreux avantages, elle peut accepter un grand nombre de vecteurs plasmidiques, se prête bien à la culture de masse en fermenteur et permet d'obtenir des taux d'expression élevés de l'ordre de plusieurs grammes par litre.

Parmi ses inconvénients, ce système ne permet pas de maturations post-traductionnelles, notamment les glycosylations, les carboxylations et phosphorylations. Ceci empêche l'utilisation d'*E. coli* lorsque l'activité biologique de la molécule est soumise à ces modifications.

De plus, la bactérie sécrète mal les protéines, celles-ci sont alors soit rejetées dans le milieu extérieur, soit plus fréquemment accumulées entre la membrane cytoplasmique et la coque qui entoure la bactérie. Les protéines peuvent alors être récupérées par choc osmotique ce qui évite la destruction de la bactérie (Heppel, 1971).

Toutefois, la protéine reste dans le cytoplasme dans la majorité des cas, il faut alors pour l'extraire, détruire la bactérie. Ceci implique des phases de purification plus poussées afin d'éliminer les débris cellulaires et les protéines de la cellule hôte (Tsuchido *et al.*, 1985).

Dans certains cas, la protéine recombinante s'accumule dans le cytoplasme sous forme d'agrégats ou corps d'inclusion (Williams *et al.*, 1982). Ces derniers peuvent être aisément extraits mais doivent ensuite être solubilisés et renaturés.

Enfin *E. coli* possède des endotoxines, les protéines recombinantes obtenues doivent donc faire l'objet de purifications très poussées et de nombreux contrôles.

1.2.2. *Saccharomyces cerevisiae*

S. cerevisiae est une cellule eucaryote capable d'assembler correctement les polypeptides synthétisés, de réaliser certaines modifications post-traductionnelles comme les protéolyses ménagées, les carboxylations et les acylations. Elle effectue des glycosylations simples mais les glycosylations spécifiques sont différentes de celles observées dans les cellules de mammifères. Enfin, *S. cerevisiae* ajoute des résidus mannose en quantités importantes, ce qui peut conférer aux protéines recombinantes des propriétés immunogènes gênantes.

Les protéines synthétisées sont souvent intracytoplasmiques, elles doivent être récupérées après destruction de la levure. Notons que les protéines peuvent être sécrétées lorsqu'on utilise des vecteurs contenant des signaux de sécrétion de la levure ou des signaux de cellules de mammifères.

1.2.3. Cellules de mammifères

Le principal avantage des cellules de mammifères est leur capacité à produire des molécules complexes, de grande taille, correctement repliées dans l'espace, correctement maturées et généralement sécrétées vers le milieu extérieur. Ce système s'impose pour la production de protéines dont l'activité exige des transformations post-traductionnelles correctes. Toutefois les glycosylations peuvent être légèrement différentes de celles qui sont réalisées normalement dans le tissu où la protéine naturelle est produite.

On utilise surtout des cellules d'ovaires de hamster (cellules CHO). Ces cellules se prêtent à la culture de masse dans les bioréacteurs mais ces cellules sont fragiles, les milieux de cultures plus coûteux et les taux d'expression plus faibles. Toutes ces raisons font que ces cellules sont moins utilisées qu'*E. coli* ou *S. cerevisiae*.

1.3. Les vecteurs d'expression

Le vecteur d'expression a pour rôle d'introduire un gène étranger dans les cellules hôtes, c'est donc une construction qui comprend le gène d'intérêt sous le contrôle d'un promoteur reconnu par la cellule hôte, des signaux de fin de copie ou de traduction et éventuellement des signaux de sécrétion.

Les vecteurs d'expression appartiennent à quatre grandes familles :

- les plasmides utilisés pour transformer les bactéries et les levures. Les plasmides sont des portions d'ADN circulaire extrachromosomique, de petite dimension, dans lesquelles on insère le gène à exprimer. Ces plasmides recombinés sont introduits dans la cellule hôte (transformation), s'y maintiennent par réplication et assurent le bon fonctionnement du système recombiné.
- Les phages (virus) qui facilitent l'introduction du vecteur dans la cellule hôte,
- les cosmides (regroupant les caractéristiques des plasmides et des phages),
- les vecteurs dérivés de virus ou de rétrovirus utilisés pour l'infection des cellules animales.

La plupart des vecteurs actuellement utilisés proviennent d'éléments extra-chromosomiques naturels comme les plasmides ou les phages. Les premiers plasmides utilisés comme vecteurs chez les bactéries étaient des plasmides « sauvages » comme Col E1 (Clewell, 1972) ou pSC101 (Cohen et Chang, 1973). Mais ces plasmides sauvages ne possédaient pas toutes les caractéristiques nécessaires à l'expression d'un gène étranger dans les meilleures conditions. Des plasmides artificiels ont alors été construits.

D'après Balbas *et al.* (1988), les critères à respecter lors de la construction d'un plasmide sont les suivants :

- le plasmide doit être de petite taille car l'efficacité de la transformation diminue quand la taille du plasmide augmente,
- il doit pouvoir exister en de nombreux exemplaires dans la cellule hôte, afin que le gène étranger soit présent sous de multiples copies susceptibles d'être traduites en protéine,
- il doit se propager facilement dans la cellule hôte,
- il doit être stable et persister dans la cellule,

- il doit être parfaitement caractérisé (connaissance des gènes, de leur localisation, des sites reconnus par les enzymes de restrictions, de la séquence),
- il doit posséder un gène marqueur permettant de différencier les cellules recombinées non transformées,
- il peut également posséder un deuxième gène marqueur où l'ADN étranger peut-être inséré. Ceci permet de différencier les bactéries recombinées possédant le gène de celles qui ne le possèdent pas.
- Enfin, le plasmide doit posséder un nombre important de sites uniques reconnus par les enzymes de restriction.

Le plasmide pBR Eco gap (Branlant *et al.*, 1983) a été utilisé afin d'obtenir les données expérimentales de notre étude. C'est un plasmide artificiel, dérivé du plasmide pBR322, lequel a été construit par Bolivar *et al.* (1977). Les étapes de sa construction sont indiquées **figure 1.1**, et la **figure 1.2** montre sa structure simplifiée.

Le plasmide d'origine est le plasmide PMB1 codant pour la colicine. Différents fragments ont été ajoutés, en particulier le gène codant pour la résistance à l'ampicilline, ainsi que celui codant pour la résistance à la tétracycline.

Ce plasmide pBR322 possède une origine de réplication permettant une réplication en mode relaxé. En effet, chez les bactéries Gram-, la réplication de l'ADN se fait selon deux modes : stringent ou relaxé. Le mode stringent nécessite de la synthèse protéique, les plasmides présentant ce mode de réplication se trouvent en petites quantités dans les cellules (1 à 5 copies). La réplication en mode relaxé ne nécessite pas de synthèse protéique, le nombre de copies de plasmides par cellule est alors plus élevé (10 à 50 copies). De plus, lorsque la synthèse protéique est bloquée, le nombre de copies de plasmides par cellule peut être fortement augmenté (20 à 100) (Covarrubias *et al.*, 1981).

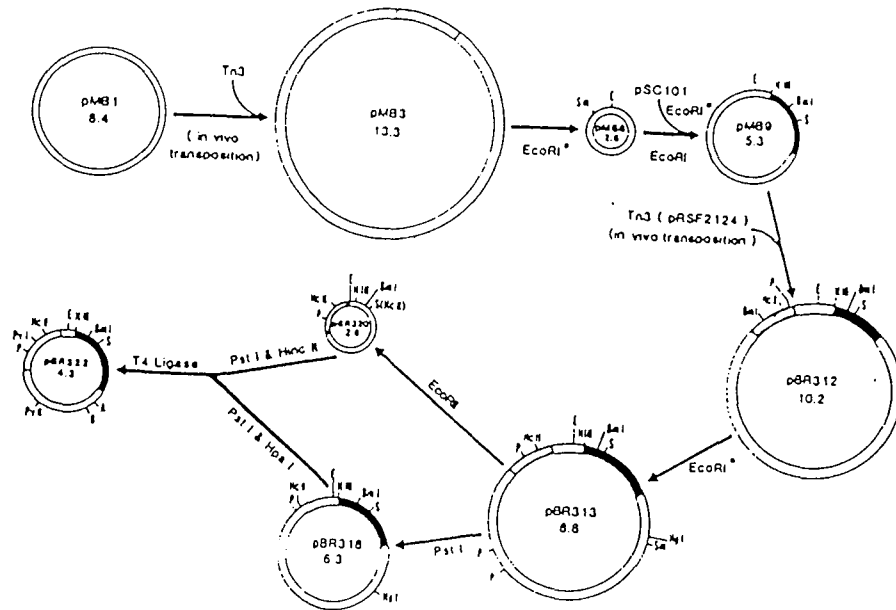


Figure 1.1. : Les différentes étapes de la construction du plasmide pBR322 (Balbas *et al.*, 1988)

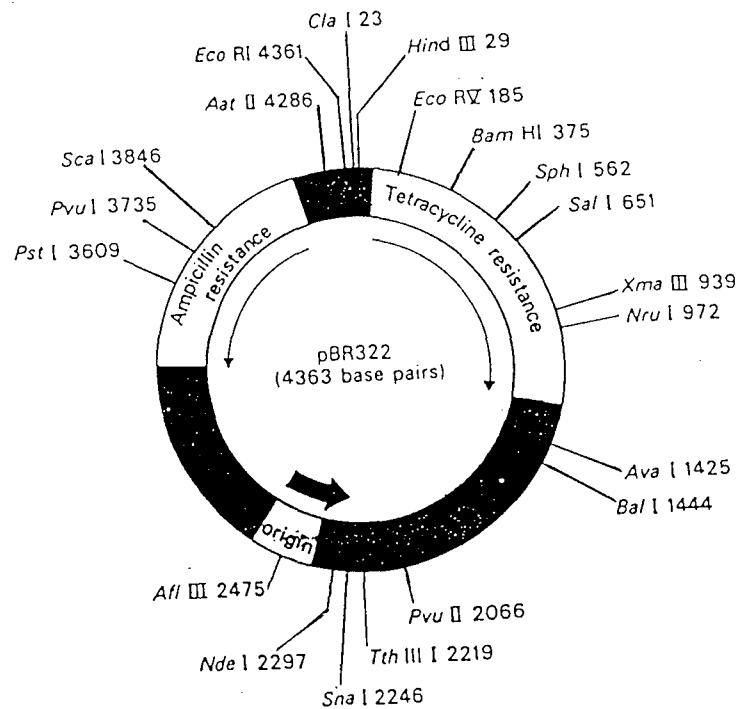


Figure 1.2. : Structure du plasmide pBR322 (Balbas *et al.*, 1988)

2. Les facteurs influençant la production de protéines recombinantes

2.1. Instabilité plasmidique

L'instabilité plasmidique peut se définir comme la variation du nombre de plasmides dans les cellules ou leur modification au cours des fermentations. Comme la productivité est directement liée au nombre de plasmides (cf. 2.2.1. Nombre de copies de plasmides), les formes d'instabilité plasmidique font partie des facteurs majeurs de contraintes dans les cultures au cours de fermentations (Karbasi et Keshavarz, 1997)

La distribution du nombre de copies de plasmides lors de la division cellulaire dépend de deux phénomènes totalement indépendants : la réplication et la partition (Nordström *et al.*, 1981). Ces deux phénomènes sont représentés dans la **figure 2.1**.

La répartition peut s'effectuer de façon aléatoire ou par un phénomène d'équipartition où chacune des cellules-filles reçoit la moitié des plasmides de la cellule mère.

L'instabilité plasmidique peut être, dans certains cas, directement reliée au phénomène de partition décrit précédemment. Il s'agit de l'instabilité plasmidique ségrégationnelle (Nugent *et al.*, 1983) qui résulte d'une partition défectueuse lors des divisions cellulaires, certaines cellules se trouvant dépourvues de plasmides et par conséquent incapables de produire des protéines. De plus, il a été montré que les cellules ne contenant pas de plasmides se développent plus rapidement que celles recombinées, ainsi, au fur et à mesure de la croissance, on observe une distribution du nombre de plasmides parmi les bactéries filles et une augmentation de cellules sans plasmides, ce qui entraîne une perte de produit (Wittrup et Bailey, 1988). Ce phénomène est pour une large part responsable des pertes de rendement en fermentation de protéines recombinées. Bentley et Quiroga, (1993) décrivent ce phénomène comme étant aléatoire.

L'instabilité plasmidique peut également être le fait de changements physiques dans la structure de l'ADN plasmidique (insertions, délétions, réarrangements). Il s'agit de l'instabilité structurale (Nugent *et al.*, 1983).

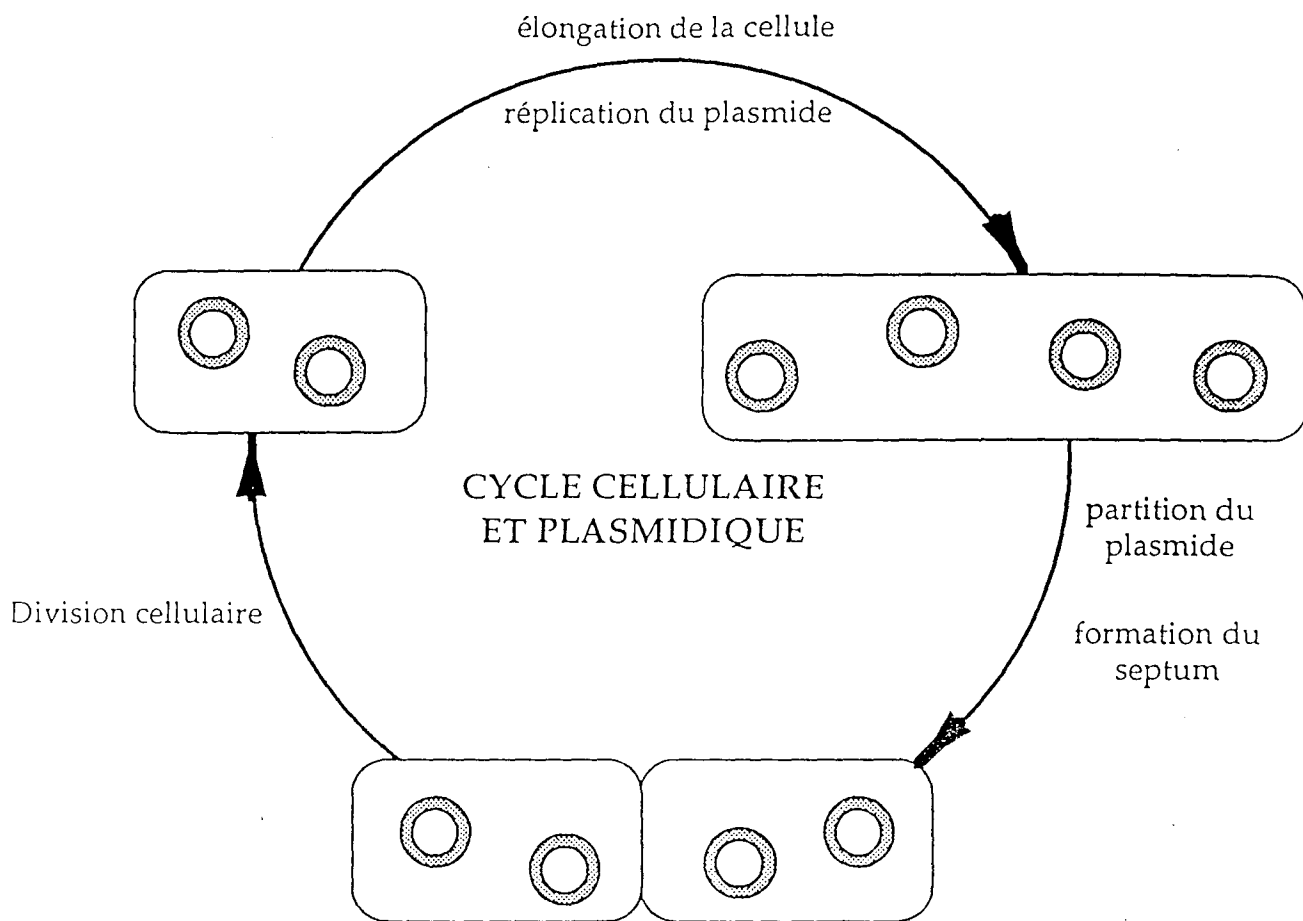


Figure 2.1. : Cycle cellulaire des bactéries recombinées (Nordström et Helle, 1984)

2.1.1. Instabilité structurale

Godwin et Slater (1979) ont étudié également l'influence de l'environnement de croissance sur la stabilité d'un plasmide codant pour plusieurs résistances aux antibiotiques en culture continue. Le plasmide utilisé codait pour la résistance à l'ampicilline, la streptomycine, le sulphonamide et la tétracycline. Quand le glucose est l'élément limitant, la population perd sa résistance à la tétracycline, mais conserve celle aux trois autres antibiotiques. Quand le phosphate est l'élément limitant, la résistance à l'un ou plusieurs des antibiotiques est perdue et les souches présentent différents modèles de résistance à l'antibiotique. Toutes les souches qui ont perdu le plasmide montrent un taux de croissance spécifique maximum supérieur à celui de la souche originale contenant le plasmide. Ainsi elles deviennent dominantes dans le milieu de fermentation. La perte de quelques marqueurs de résistance est un indicateur d'un problème d'instabilité structurale. Le fait que la perte du marqueur varie avec des limitations nutritives différentes indique l'influence de l'environnement sur la stabilité du plasmide.

La stabilité plasmidique a été également étudiée chez des procaryotes autres qu'*E. coli* : Pinches *et al.* (1985) se sont penchés sur la stabilité du plasmide codant pour l' α -amylase chez *Bacillus subtilis*, en discontinu, sur milieu complexe supplémenté par du glucose ou de l'amidon en présence ou en l'absence de pression de sélection. Avec du glucose ou de l'amidon et en l'absence de pression de sélection, la perte du plasmide durant la phase exponentielle reste négligeable mais elle atteint 70 % durant la phase stationnaire. Le taux de perte de plasmide dépend de la source de carbone utilisée. En effet, cette perte est plus rapide dans un milieu glucosé que dans un milieu à base d'amidon. En présence de pression de sélection, l'activité de l' α -amylase est toujours perdue. Les études avec électrophorèse sur gel d'agarose indiquent que cette perte est corrélée avec des délétions du plasmide. Ces résultats indiquent la présence d'instabilités structurales.

Kapralk et Jecmen (1993) ont étudié la stabilité d'*E. coli* HB101 transformée par un vecteur dérivé du plasmide de type pUC. Leurs travaux ont montré que l'expression du gène porté par le plasmide (gène codant pour la prochymosine de veau) conduit à une perte rapide de l'expression de ce dernier, liée à une instabilité structurale. A l'inverse, la même souche cultivée dans des conditions de non-expression du gène cloné présente un plasmide structurellement stable pendant 200 générations.

2.1.2. Instabilité ségrégationnelle

2.1.2.1. Influence des conditions de culture

L'effet de la température a été étudié par Wouters *et al.* (1980) sur la souche *E. coli* PC221 (pBR322). Ils ont montré que l'apparition de cellules dépourvues de plasmide survenait plus tôt lorsque la température de culture était plus élevée (42°C au lieu de 37°C). Cortinho et Hjortso (1988) ont obtenu des résultats identiques avec *E. coli* B (pECA306). Ces auteurs ont de plus démontré que l'augmentation de l'acidité du milieu de culture contribuait à augmenter la probabilité de perte de plasmide.

Les travaux de Karbasi et Keshavarz (1997) ont révélé l'influence des conditions opératoires telles que le pH et la température sur la stabilité du plasmide TOL chez *Pseudomonas putida* PPK1 en culture continue. De plus, sur milieux limités en benzoate ou en succinate, ils ont remarqué que la stabilité du plasmide diminuait parallèlement à une augmentation du taux de dilution.

Hopkins *et al.* (1987) ont montré que l'instabilité plasmidique pouvait être induite par un apport massif d'oxygène dissous. Par contre en 1997, Brigidi et Gonzalez ont conclu que le taux d'oxygène dissous n'avait pas d'influence sur la stabilité plasmidique en culture continue sur *Bacillus stearothermophilus* NUB3621 (BGSC) transformée par les vecteurs pPAM2, pPCH2 et pPLY2 obtenus par combinaison du plasmide pRP9 avec respectivement les gènes de l' α -amylase de *Bacillus licheniformis*, la cholestérol oxydase de *Streptomyces sp* et la lipase chez *Pseudomonas fluorescens*.

L'influence de la vitesse d'agitation sur la stabilité plasmidique de cultures d'*E. coli* recombinées, immobilisées ou non, a été étudiée par Huang *et al.* (1990). Une agitation douce serait favorable à la stabilité des plasmides dans les deux types de cultures.

Sur *Bacillus stearothermophilus* NUB3621 (BGSC), Brigidi et Gonzalez (1997) ont montré que les conditions optimales en culture continue en terme de concentration de biomasse étaient obtenues avec une température de 65 °C, un pH variant de 6.8 à 7.2 et une pression partielle en oxygène supérieure à 10 %. De plus, cette souche recombinée avec

différents plasmides a montré une instabilité ségrégationnelle croissant avec la température de fermentation.

2.1.2.2. Influence du taux de dilution

L'influence du taux de dilution sur la stabilité plasmidique en culture continue a été étudiée pour de nombreuses souches d'*E. coli* transformées par différents plasmides. Deux comportements opposés ont été mis en évidence. En effet, on observe une stabilité plus grande aux forts taux de dilution pour certains systèmes hôte-vecteur alors que pour d'autres, la stabilité est accrue aux faibles taux de dilution.

2.1.2.2.1. Stabilité accrue pour de forts taux de dilution

Une étude de Sayadi *et al.* (1989) concernant *E. coli* W3101 (pTG201) en culture continue a montré que 75 % des cellules étaient pourvues de plasmides après 20 générations pour un taux de dilution de 0.06 h^{-1} . Pour des taux de dilution de 0.13 h^{-1} et 0.25 h^{-1} , le même pourcentage était atteint après respectivement 35 et 60 générations.

Chew *et al.* (1988) ont réalisé le même type d'étude avec la souche *E. coli* HB101 (pAT153). Ce plasmide dérivait de pBR322 mais présentait un nombre de copies beaucoup plus élevé que ce dernier (plus de 100 copies de plasmide par cellule). Ces auteurs ont mis en évidence une augmentation de la stabilité plasmidique avec le taux de dilution en culture continue et ceci dans différentes conditions de limitation nutritionnelle (carbone, azote, phosphate ou sulfate).

Le même comportement a été observé par Caulcott *et al.* (1987) pour la souche *E. coli* RV308 (pHSG415). Ce plasmide est présent dans les cellules à faible nombre de copies de plasmide (5 copies par cellule).

Sayadi *et al.* (1989) expliquent ce phénomène par le fait qu'aux faibles taux de dilution, les cellules sont porteuses d'un plus grand nombre de plasmides. Or Helling *et al.*, (1981), Seo et Bailey (1985), Ryan et Parulekar (1991) ont montré que le taux de croissance de la cellule est d'autant plus élevé qu'elle comporte peu de plasmides. Ainsi en culture continue à faible taux de dilution, les cellules pourvues d'un petit nombre de plasmides prennent le pas

sur celles qui n'en possèdent plus. Di Biasio et Sardonini (1985) ont modélisé cette affirmation.

Brigidi et Gonzalez (1997) ont étudié en culture continue *Bacillus stearothermophilus* NUB3621 (BGSC), transformé par les vecteurs pPAM2, pPCH2, et pPLY2 obtenus par combinaison du plasmide pRP9 avec respectivement les gènes de l' α -amylase de *Bacillus licheniformis*, la cholesterol-oxydase de *Streptomyces sp.* et la lipase chez *Pseudomonas fluorescens*. Ces souches recombinées ont montré une instabilité ségrégationnelle décroissante avec une augmentation de la température de fermentation.

2.1.2.2.2. Stabilité accrue pour de faibles taux de dilution

Weber et San (1987) ont montré avec *E. coli* RR1 (pBR322), une stabilité plus importante des cultures à faible taux de dilution. De plus, ils ont mis en évidence que la production de β -lactamase (codée par le plasmide) ne dépend pas de ce taux.

Noack *et al.* (1981) ainsi que Nancib *et al.* (1992) ont obtenu des cultures plus stables à de faibles taux de dilutions avec des plasmides dérivés du pBR322. Ces derniers auteurs ont réalisé des cultures continues d'*E. coli* C600 galK (GADPH) à différents taux de dilution sur milieu complexe sans pression de sélection. Ils ont montré que la fraction de cellules recombinées décroissait avec l'augmentation du taux de dilution. Contrairement à d'autres études (notamment celle de Weber et San, 1987), ils ont montré que l'activité enzymatique de la protéine recombinée produite : la GADPH dépend du taux de dilution D (elle diminue lorsque D augmente). Il existerait une plage de dilution comprise entre 0.1 et 0.2 h⁻¹ qui n'entraînerait pas de perte de plasmide et assurerait donc une production optimale.

Karbasi et Keshavarz (1997), ont étudié l'influence du taux de dilution sur la stabilité du plasmide TOL pTKO chez *Pseudomonas putida* en culture continue sur milieu limité en succinate. A des taux de dilution de 0.115 h⁻¹ et 0.23 h⁻¹, les souches restaient stables pendant respectivement 400 et 540 h. Par contre à un taux de dilution plus élevé : 0.46 h⁻¹, ces auteurs ont remarqué une perte plasmidique de 25.7 % au bout de 290 h de fermentation. En présence d'une limitation en benzoate, on a observé des instabilités plasmidiques au bout de 67 h, 45.5 h et 28 h pour des taux de dilution respectifs de 0.115 h⁻¹, 0.172 h⁻¹, et 0.46 h⁻¹.

Une explication à ce phénomène peut être donnée en se basant sur un résultat obtenu par Engberg et Nordström (1975) qui ont montré que les mécanismes de réplication du plasmide sont totalement indépendants de ceux régissant la réplication du chromosome. Ainsi, à de forts taux de dilution, donc à des taux de croissance élevés, une dilution du nombre de plasmides peut survenir et ceci d'autant plus rapidement que sa réplication est lente et son nombre initial de copies de plasmide est faible. Si on ajoute à cela, des processus de partition défectueux et le fait que les cellules dépourvues de plasmides prennent le dessus en chemostat sur les cellules porteuses, la culture se verra de plus en plus constituée de cellules ségréguées lors de cultures à taux de dilution élevé.

2.1.2.3. Influence du taux de croissance

La relation entre instabilité plasmidique et taux de croissance a été mise en évidence par les travaux de Lee et Bailey (2000) sur les plasmides pBR322, NR1, R1 et col E1. Ils ont observé expérimentalement une diminution du nombre de plasmides lors de l'augmentation du taux de croissance.

De même Lamotte, Gschaedler et Boudrant (1994) ont mené des investigations sur *E. coli* HB101 (GADPH) dont les plasmides codent pour la GADPH. Les auteurs ont montré une diminution transitoire de la GADPH (directement liée au nombre de plasmides) par rapport à la biomasse pendant la phase exponentielle de croissance. En effet pendant cette phase, les cellules se multiplient sans doute plus vite que les plasmides, entraînant une instabilité ségrégationnelle.

2.1.2.4. Influence d'une limitation nutritionnelle en culture continue.

Melling *et al.* (1977) ont employé la technique des répliques pour étudier la maintenance de gènes de résistance en culture continue chez *E. coli*. La résistance phénotypique est maintenue stable pendant 144 h en chemostat limité en carbone, phosphate ou magnésium à des taux de dilution de 0.05 à 1 h⁻¹, en l'absence de sélection par antibiotique.

Wouters *et al.* (1979) ont montré que cette résistance peut être maintenue stable en l'absence de pression de sélection à des taux de dilution de 0.05 à 0.5 h⁻¹ sous une limitation en carbone, azote ou phosphate et avec des conditions aérobies ou anaérobies. Ces résultats indiquent que la stabilité ségrégationnelle des gènes de résistance est élevée.

L'effet de trois limitations nutritionnelles (glucose, phosphate et magnésium) a été étudié par Jones *et al.* (1984) pour un même hôte *E. coli* HB 101 transformé par cinq plasmides différents dérivés du pBR322 (pBR322, pAT153, pBR325, pBR327, pBR328). Plusieurs comportements ont été mis en évidence. La souche HB101 (pAT153) est stable durant 100 générations, quelle que soit la limitation. Par contre, *E. coli* HB101 (pBR327) et HB101 (pBR328) perdent leurs plasmides après 50 générations dans les trois types de limitation nutritionnelle. Le système contenant le plasmide pBR322 est resté quasiment stable durant les 100 générations de culture sauf en cas de limitation en phosphate. Cette étude a permis de mettre en évidence l'importance du facteur nutritionnel et de la structure du vecteur.

La stabilité du plasmide a été étudiée chez *Pseudomonas putida* en culture continue (Keshavarz *et al.*, 1985). Le plasmide (TOL) code pour l'utilisation du toluène et du xylène via la voie du clivage du meta. La stabilité a été suivie par la technique des répliques et par électrophorèse sur gel. Les cellules ont poussé jusqu'à l'état stationnaire sur m-toluate et ont été transférées sur milieu succinate ou benzoate. Le transfert sur succinate ne provoque aucune perte pendant 600 h et les auteurs suggèrent que la stabilité peut être le résultat de la répression catabolique favorisée par le succinate. Par contre le transfert sur benzoate donne 98 % de perte de plasmides après 100 h. Les données d'électrophorèse sur gel indiquent que la perte de plasmide est due à une instabilité ségrégationnelle. Des résultats similaires ont été obtenus avec *E. coli* (Jones *et al.*, 1980).

Caulcott *et al.* (1987) ont démontré l'inverse avec la souche *E. coli* RV308 (pHSG415). En effet, la culture la plus instable est alors obtenue lorsque le sulfate est l'élément limitant. Les cultures sont en revanche plus stables avec des limitations en glucose, carbone ou phosphate.

Chew *et al.* (1988) ont testé différents éléments limitants lors d'une culture continue réalisée à un taux de dilution donné avec la souche *E. coli* HB101 (pAT153). Leurs résultats montrent qu'une culture limitée en sulfate est plus stable qu'une culture où l'élément limitant serait le glucose, l'azote ou le phosphate.

Sur *Bacillus stearothermophilus* NUB3621, Brigidi et Gonzalez (1997) ont montré qu'une stabilité plasmidique supérieure était obtenue avec des cultures sur milieu complexe par rapport aux cultures sur milieu simple.

2.1.2.5. Instabilité chez les levures

Futcher et Cox (1984) ont étudié la stabilité des plasmides 2μ chez plusieurs souches de *Saccharomyces cerevisiae*. La stabilité plasmidique a été quantifiée par la méthode des répliques. D'après ces auteurs, la stabilité varie d'un plasmide à un autre et les systèmes à nombre de copies élevé sont plus stables que les systèmes à faible nombre de copies.

Walmsley *et al.* (1983) ont étudié la persistance de plusieurs plasmides chez les levures en culture continue. Les levures avec ou sans les plasmides 2μ ont été utilisées comme souches hôtes. Toutes les souches utilisées étaient auxotrophes à la leucine et tous les plasmides présentaient une prototrophie à la leucine. Les résultats indiquent que la stabilité augmente avec le nombre de copies de plasmide. En général, l'augmentation du nombre de copies diminue le taux de croissance spécifique maximal de la culture. De plus, les plasmides 2μ confèrent aux levures une augmentation du taux de croissance et contribuent à la stabilité de la souche désirée en culture.

En conclusion, on peut donc dire que la stabilité du plasmide est fonction des caractéristiques génétiques des plasmides et qu'elle est influencée par la génétique des souches hôtes. En effet, différents plasmides peuvent présenter des degrés de stabilité différents dans une même souche hôte. A l'inverse, le même plasmide peut présenter différents degrés de stabilité dans des souches différentes. En culture continue, le taux de croissance cellulaire et la nature de l'élément nutritif limitant interviennent sur les phénomènes de stabilité, et on peut dire qu'en général, les plasmides sont moins stables dans des cultures limitées en phosphate. Enfin, le fait d'utiliser une pression de sélection durant la phase de croissance permet de compenser quelque peu l'instabilité ségrégationnelle. Elle n'est toutefois pas applicable à grande échelle et ne résout pas les problèmes liés à l'instabilité structurale.

2.1.3. Méthodes d'évaluation de la stabilité plasmidique

De nombreuses méthodes ont été mises au point pour déterminer le nombre de copies de plasmide présent dans les cellules ou pour différencier les cellules comportant des plasmides de celles qui n'en contiennent plus. Ces méthodes peuvent être classées en deux groupes : celles utilisant des techniques de mesure directes, et celles se basant sur des méthodes indirectes.

2.1.3.1. Les méthodes directes

La technique d'hybridation sandwich sur filtre (Korpela *et al.*, 1987) permet de déterminer le nombre moyen de copies de plasmides. A cet effet, on peut également utiliser la chromatographie HPLC (Coppela *et al.*, 1987) qui toutefois nécessite un investissement plus important.

La technique la plus utilisée est basée sur une extraction puis une séparation par électrophorèse des ADN plasmidiques et chromosomiques (Projan *et al.*, 1983). Après migration, le gel est coloré au bromure d'éthidium puis photographié sous U.V.. On mesure alors par densitométrie, la quantité des deux ADN à partir du négatif obtenu. A partir de ces mesures et des masses moléculaires du plasmide et du chromosome, on peut déterminer le nombre moyen de copies de plasmide par cellule.

2.1.3.2. Les méthodes indirectes

De nombreux plasmides utilisés comme vecteur d'expression, codent pour une β -lactamase, enzyme qui dégrade l'ampicilline. La synthèse de β -lactamase permet donc aux bactéries transformées de résister à l'action antibiotique de l'ampicilline.

Deux méthodes d'évaluation du nombre de copies de plasmide sont basées sur cette caractéristique.

La première méthode consiste à doser la quantité de β -lactamase produite par les cellules. Cette quantité peut être reliée directement au nombre de copies de plasmide (Seo et Bailey, 1985).

La seconde méthode consiste à déterminer la concentration minimum inhibitrice (CMI) définie comme la dose minimale d'antibiotique capable d'empêcher le développement d'une bactérie. Plus le nombre de copies de plasmide est élevé, plus la concentration minimum inhibitrice sera forte (Dennis *et al.*, 1985). Grâce à cette technique, on ne considère plus une population bactérienne mais des cellules de façon individuelle. En effet, les CMI sont déterminées par étalements individualisés de bactéries d'une culture, sur boîtes de Pétri contenant du milieu gélosé à différentes concentrations en antibiotique. Le nombre de copies de plasmide est relié linéairement à la CMI. C'est cette méthode qui a permis de déterminer les données expérimentales de notre étude.

Le nombre de copies de plasmide peut également être évalué par mesure de la taille des colonies se développant sur milieu gélosé (Park et Ryu, 1992). En effet, la présence du plasmide ralentissant la croissance, plus le nombre de copies de plasmide est élevé, plus la croissance sera ralentie et plus la colonie sera de petite taille par rapport aux autres à un temps donné.

La cytométrie de flux a été utilisée par Seo et Bailey (1987) pour étudier l'influence du nombre de copies de plasmide sur le cycle cellulaire d'*E. coli*. Cette technique consiste à mesurer la fluorescence des cellules après coloration spécifique des ADN par la mithramycine.

En règle générale, dans de nombreuses études de la stabilité plasmidique, seule la présence ou l'absence de plasmide est déterminée (Dinçbas *et al.*, 1993). Ceci est basé, comme précédemment, sur la capacité des cellules transformées à résister à un antibiotique. Des étalements de bactéries d'une culture sont réalisés sur milieu gélosé contenant ou ne contenant pas un antibiotique (présent en faible quantité). Le rapport entre les colonies développées sur milieu avec et sans antibiotique permet de déterminer la fraction de bactéries transformées.

2.1.4. Techniques de stabilisation

2.1.4.1. Stratégies moléculaires et cellulaires

Ces différentes techniques de stabilisation ont été classées par Ensley (1985) en deux catégories :

- les méthodes sélectives qui visent à détruire les cellules dépourvues de plasmides,
- les méthodes non sélectives visant à éviter l'apparition de ces cellules ségréguées.

2.1.4.1.1. Méthodes sélectives

La première méthode sélective repose sur l'utilisation de la résistance à certains antibiotiques conférée par le plasmide, cette méthode est dite de pression de sélection. On dit alors que les plasmides contiennent des gènes marqueurs de sélection. Ces gènes codent pour des protéines qui sont nécessaires à la survie de l'organisme dans son milieu de culture. Les cellules contenant le plasmide peuvent se développer dans des milieux contenant l'antibiotique car elles produisent des protéines inactivant ce dernier. L'ajout d'antibiotique dans un milieu de culture permet d'éliminer les cellules ségréguantes. Cette technique est largement utilisée à l'échelle du laboratoire mais est difficilement applicable au niveau industriel.

En effet :

- l'utilisation d'antibiotiques implique des coûts supplémentaires non négligeables,
- l'élimination de cet antibiotique pose des problèmes lors des étapes de séparation et de purification,
- les souches peuvent développer des résistances aux antibiotiques.

De plus, tout l'antibiotique utilisé dans une culture peut être inactivé par une petite fraction de cellules contenant le plasmide qui coexiste avec une proportion importante de cellules ayant perdu leur plasmide. Les cellules ségréguantes sont alors protégées. (Pierce et Gutteridge, 1985).

De plus, il faut noter que la pression de sélection diminue au cours de la fermentation, notamment lorsque l'antibiotique utilisé est l'ampicilline. En effet, l'ampicilline est dégradée par la β -lactamase synthétisée par les cellules recombinées. Ainsi, la concentration en antibiotique diminue au cours de la fermentation et des ajouts sont nécessaires. (Zabriske et

Arcuri, 1986). La résistance à la tétracycline, elle, repose sur l'action des peptides codés par le plasmide pSC101. Ces peptides bloquent le transport de la tétracycline dans la cellule et la protègent ainsi de l'action de l'antibiotique. Ainsi la concentration en antibiotique reste constante dans la cellule et la pression de sélection reste la même (Bentley et Quiroga, 1992).

Une deuxième technique consiste à incorporer dans le plasmide un gène palliant une auxotrophie de la souche hôte. Nilsson et Skogman (1986) ont inséré sur un plasmide donné, un gène codant pour la valyl tRNA synthétase (vals). L'utilisation d'une souche d'*E. coli* K12 possédant une vals mutée thermosensible permet de sélectionner une population uniquement composée de cellules recombinées lorsque la culture est réalisée à une température permettant la destruction de la vals codée par le chromosome.

Dans le système proposé par Gerdes et Molin (1986), la cellule produit une protéine « suicide » dont la synthèse est empêchée par la présence du plasmide. En utilisant cette combinaison hôte-plasmide, la fréquence expérimentale de la ségrégation du plasmide a pu être divisée par trois. Finalement, la stabilité de la culture peut être augmentée lorsque le plasmide codant pour la protéine est essentiel pour le métabolisme cellulaire et que la cellule hôte est incapable de synthétiser cette protéine. Ce système a été employé avec succès et a conféré 100 % du maintien d'un plasmide normalement instable, pendant plus de 150 générations (Nilson et Skogman, 1986).

Rosteck et Hershberger (1983) ont inséré sur le plasmide un fragment d'ADN codant pour un répresseur thermosensible du phage lambda. La culture de bactéries recombinantes est infectée par un phage lambda qui ne produit pas le répresseur. La perte du plasmide entraîne donc la disparition du répresseur et donc l'induction du phage qui va détruire la cellule.

2.1.4.1.2. Méthodes non sélectives

Les méthodes non sélectives sont moins nombreuses. La première a été découverte par Meacock et Cohen en 1980. Ces auteurs ont isolé un fragment d'ADN dans le plasmide pSC101 responsable d'une équiartition rigoureuse du nombre de copies de plasmides entre la cellule mère et les cellules filles lors de la division cellulaire. Il s'agit du locus *par* qui a ensuite été inséré dans de nombreux plasmides afin de les stabiliser notamment dans le plasmide pBR322 (Manen et Goebel, 1990, Weber *et al.*, 1991). De plus, ces derniers auteurs ont remarqué une augmentation du nombre de copies de plasmide par cellule après insertion de ce locus.

Une méthode proposée par Yukawa *et al.* (1985), consiste à insérer sur le plasmide un fragment du plasmide mini-F qui est stable et soumis à une réplication synchrone avec le chromosome.

Nudel *et al.*, (1993), ont montré qu'il était possible de stabiliser une culture d'*E. coli* HB 101 (pBR322) par insertion dans le plasmide d'un fragment d'ADN Tn1000. Ce fragment éviterait de plus la multimérisation des plasmides, laquelle favorise l'instabilité.

Il est à noter que des plasmides stables peuvent apparaître spontanément lors de cultures continues. Chew *et al.* (1986) ont isolé un plasmide pLCX1 apparu spontanément lors d'une culture continue d'*E. coli* HB101 (pAT153) limitée en source de carbone. Ce nouveau plasmide résultait de l'insertion d'un fragment d'ADN chromosomique dans le gène codant pour la résistance à la tétracycline du plasmide. Ce gène était un des facteurs influençant la stabilité plasmidique.

Enfin, le problème d'instabilité peut être entièrement résolu si au lieu d'utiliser des plasmides, le gène désiré est inséré directement dans le chromosome bactérien. Dans ce cas, en contre partie de la stabilité, on aura une synthèse protéique plus faible. En effet, les gènes insérés dans le chromosome peuvent être maintenus à un nombre de copies moindre que celui des gènes insérés dans des vecteurs d'expression.

2.1.4.2. Stratégies relatives au procédé

Stephens *et al.* (1988, 1992) ainsi que Weber et San (1988, 1989) ont démontré la possibilité de ralentir la perte plasmidique en faisant alterner de faibles et de forts taux de dilution en culture continue. Weber et San (1988) ont obtenu les résultats suivants : après 300 heures de culture d'*E. coli* (pBR322), le pourcentage de cellules porteuses de plasmides était de 20%, 40% et de 70% pour des taux de dilution respectifs de 0.45 h^{-1} , 0.31 h^{-1} et une alternance entre 0.45 h^{-1} et 0.31 h^{-1} toutes les deux heures.

Deux hypothèses ont été émises pour expliquer ce phénomène (Weber et San, 1989) : d'une part, les cellules soumises à des conditions transitoires s'adaptent plus facilement aux changements d'environnement et par conséquent la différence de taux de croissance entre les cellules recombinées et non recombinées est faible. D'autre part, dans de telles conditions d'alternance, le nombre de copies de plasmide augmenterait d'où une diminution des pertes.

Une deuxième technique consiste à immobiliser les cellules recombinées. De Taxis du Poët *et al.*, 1986, 1987, ainsi que Marin-Iniesta *et al.*, 1987, 1988 décrivent une stabilité plasmidique accrue observée au cours de ces cultures. Dinçbas *et al* (1993) ont étudié la stabilité de cultures mixtes d'*E. coli* HB101 recombinées ou non, avec ou sans immobilisation. Ils ont mis en évidence le fait que l'immobilisation des cellules sur billes d'alginate entraîne une disparition moins rapide des cellules recombinées. Le délai de disparition de ces dernières est proportionnel au nombre de cellules immobilisées. Ce phénomène s'expliquerait par la compartimentation des cellules qui retarde la compétition des bactéries recombinées ou non recombinées envers le substrat. Lors de cultures classiques sans immobilisation cette compétition est immédiate et favorise le développement de bactéries non recombinées.

2.2. ARNm : Acide ribonucléique messager

Les principales étapes de la transcription et de la traduction sont présentées dans la **figure 2.2**. Le gène codant pour la protéine désirée agit comme un modèle pour la synthèse de la séquence de ribonucléotides correspondante ou ARNm. Cet ARNm est ensuite traduit en une séquence d'acides aminés (polypeptide) qui constitue la forme naissante de la protéine produite. La séquence d'acides aminés peut être par la suite modifiée par l'élimination ou l'adjonction d'un ou de plusieurs acides aminés, de certains groupes chimiques, et finalement la protéine se replie sur elle-même pour atteindre sa conformation tridimensionnelle. Cette dernière forme correspond à la protéine désirée biologiquement active.

La synthèse de la protéine est cinétiquement contrôlée. Une augmentation de la synthèse de la protéine désirée peut être obtenue par augmentation du taux de formation des intermédiaires limitants et c'est la synthèse d'ARNm qui habituellement, est limitante.

La maximisation de la quantité de protéines recombinantes produites passe donc par une transcription optimale, pour ce faire, on va chercher à :

- accroître le taux des ARNm en augmentant le nombre de copies de gène dans la cellule.
On comprend alors aisément l'intérêt de maintenir un nombre de plasmides élevé dans la cellule hôte,
- augmenter la transcription du même gène,
- assurer la stabilité de l'ARNm synthétisé.

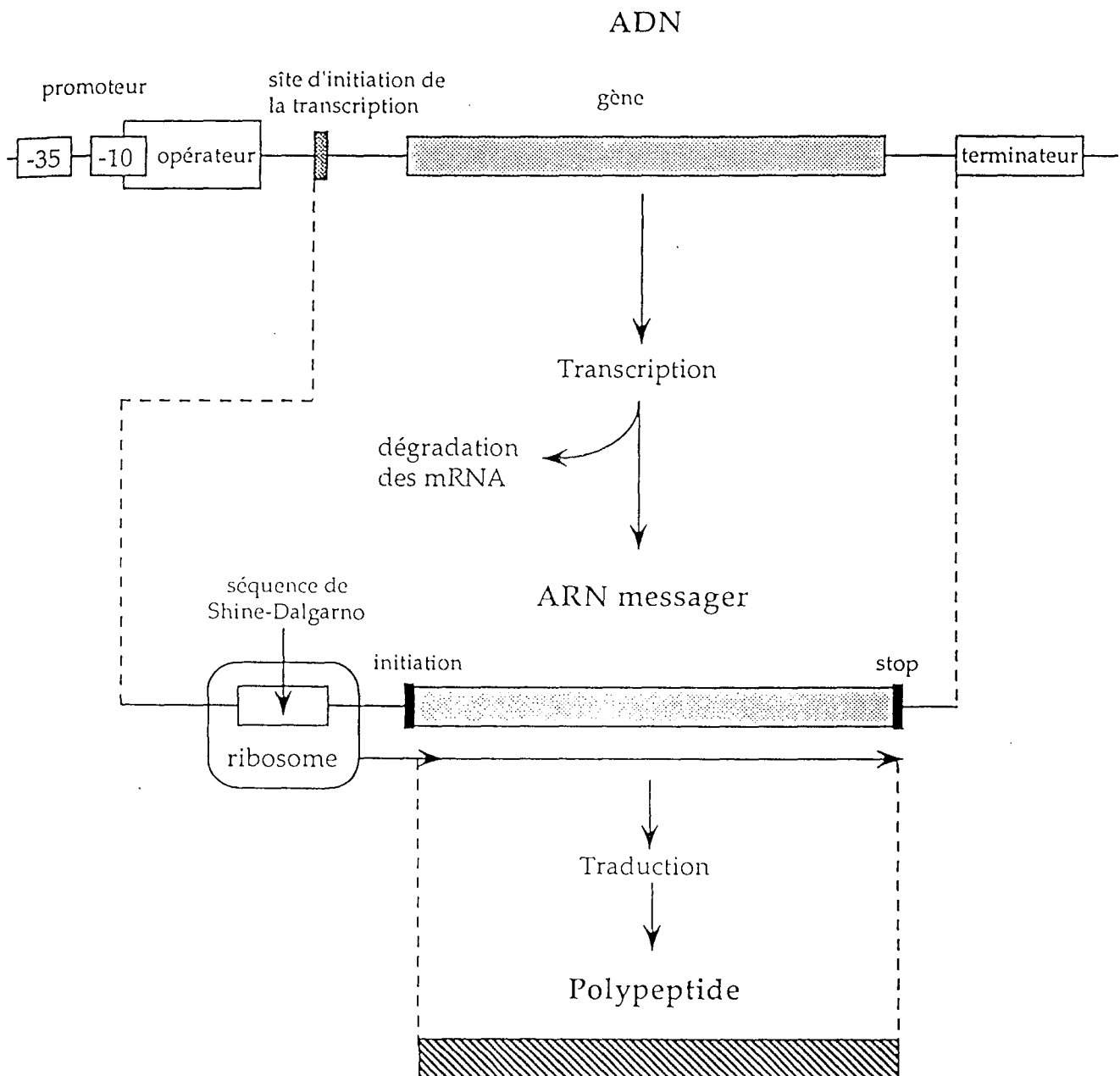


Figure 2.2. : Les principales étapes de transcription et de la traduction (Georgiou, 1988).

2.2.1. Nombre de copies de plasmides

Pour un gène exprimé de façon constitutive, le taux d'ARNm est proportionnel au nombre de copies du gène correspondant. Par conséquent, une façon pour augmenter le taux de formation de la protéine d'intérêt est de maintenir un nombre élevé de copies de gènes dans la cellule.

Pour cela le gène est inséré dans un plasmide qui est ensuite introduit dans la cellule hôte. Les plasmides se propagent d'une génération à l'autre dans les cellules par réplication et existent en plusieurs exemplaires dans les cellules. Chaque plasmide contient normalement une copie du gène désiré. Donc, plus le nombre de plasmides par cellule est élevé, plus le taux de synthèse de la protéine est important (Uhlen et Nordström, 1977).

Toutefois, il existe une valeur critique au delà de laquelle toute addition de copies supplémentaires est inutile. Seo et Bailey (1985) ont étudié la production de β -lactamase en fonction du nombre de copies du gène chez *E. coli* HB101. Pour cela, la souche hôte a été transformée par cinq plasmides, portant un gène codant pour la β -lactamase, structurellement proches mais présents à des nombres de copies différents dans la cellule hôte (de 12 à 408 copies). La production de β -lactamase augmente linéairement en fonction du nombre de copies de gènes et ce jusqu'à 60 copies. Entre 60 et 120 copies, la production augmente légèrement mais sans plus être proportionnelle au nombre de copies de gène. Enfin, au delà de 120 copies, la production reste quasiment constante.

Dans des conditions de culture constantes, le nombre moyen de plasmides par cellule (nombre de copies de plasmide) est déterminé par l'origine de la réplication. Celle-ci est un segment spécifique du plasmide contenant l'information essentielle qui régule le nombre de copies de plasmide dans la cellule. Beaucoup de vecteurs d'expression chez *E. coli* ayant différents nombre de copies existent. Pour un taux d'expression élevé de protéines étrangères chez *E. coli*, les plasmides les plus largement utilisés contiennent une origine de réplication bien caractérisée dénotée ColE1. Ces plasmides sont dérivés de celui dénoté pBR322. Dans des circonstances normales, ces plasmides sont présents à des taux d'environ 15-50 copies par cellule. Le mécanisme moléculaire qui contrôle le nombre de copies de plasmide avec l'origine de la réplication ColE1 est bien identifié (Thomas 1988).

L'influence de la physiologie cellulaire sur le nombre de copies de plasmide a fait l'objet de plusieurs études. Enberg et Nordström (1975) ont étudié l'effet du taux de croissance sur le nombre de copies du plasmide R1 chez *E. coli*. Les résultats indiquent que le nombre de copies décroît avec l'augmentation du taux de croissance.

Les systèmes à nombres de copies élevés sont en général plus productifs que ceux à faible nombre de plasmides, de plus Boudrant, Lamotte et Fonteix (1997) ont montré qu'ils ont une tendance à être plus stables. Cela pourrait s'expliquer par le fait que le plasmide exercerait une action comparable à celui d'un fardeau (Boudrant *et al.*, 1996) qui provoquerait une diminution du taux de croissance cellulaire. Comme les vitesses de croissance cellulaire sont indépendantes des vitesses de réplication plasmidique et leur sont souvent supérieures, leur diminution par un effet « fardeau » les rapproche des vitesses de réplication. Ainsi les pertes de plasmides au cours des générations sont diminuées.

Sachant que les taux de croissance peuvent être modifiés lors de ces cultures discontinues en employant différents milieux de culture, il n'est pas évident que le changement du nombre de copies soit seulement dû à la variation du taux de croissance. Les changements observés peuvent également être le résultat des effets des nutriments. Ceci a été observé par Seo et Bailey (1985) au cours d'expériences où le nombre de copies de plasmide a varié en employant différents milieux de culture. Dans la plupart des cas, la productivité spécifique augmente avec l'augmentation du nombre de copies. A des nombres de copies élevés, la productivité spécifique la plus grande est obtenue dans des expériences où le nombre de copies est contrôlé par les variations des milieux.

Koizumi *et al.* (1985) ont étudié les influences de la température et du taux de croissance sur le nombre de copies du plasmide pLP11, et la productivité du plasmide codant pour l'activité de la pénicillinase en culture continue chez *B. stearothermophilus*. A 44°C, la concentration cellulaire est à peu près moitié moindre que celles observées à 47.5 et 50°C. A 44°C, la teneur en plasmide (mg plasmide/g cellule) est approximativement le double de celle observée à 50°C et celle-ci reste constante avec l'augmentation du taux de dilution. Le nombre de copies observé à 44°C décroît avec l'augmentation du taux de dilution. A 47.5 et 50°C, le nombre de copies et la teneur en plasmides augmentent avec le taux de dilution. Ainsi l'influence du taux de croissance sur le nombre de copies de plasmides chez *B. stearothermophilus* est différent de celui rapporté chez *E. coli* (Engberg et Nordström,

1975). Koizumi *et al.* (1985) ont suggéré que la différence observée était due à la différence entre les mécanismes de réplication de ces deux souches bactériennes. De plus les procédures expérimentales employées étaient différentes.

En résumé on peut dire que le nombre de copies de plasmides est déterminé par la génétique du plasmide, les conditions de croissance et par la génétique et la physiologie de la souche hôte. Le nombre de copies de plasmide affecte le taux de production des protéines recombinantes. En général, les systèmes à fort nombre de copies sont plus productifs et plus stables que ceux à faible nombre de copies. Toutefois un nombre de copies de plasmide élevé peut entraîner une diminution du taux spécifique de croissance maximum de la souche recombinée. Cette instabilité ségrégationnelle conduira à un remplacement progressif de la souche recombinée par une souche sans plasmides.

2.2.2. Optimisation de la transcription

Le taux de transcription d'une protéine est proportionnel à la concentration de l'ARNm correspondant. Pour optimiser la formation de la protéine produite, le taux de transcription doit être élevé. La synthèse d'ARNm est catalysée par l'ARN polymérase. Il est possible d'accroître le taux d'ARNm en augmentant le taux de transcription par copie de gène. Celui-ci dépend principalement de la fréquence avec laquelle l'ARN polymérase initie la transcription. Deux facteurs interviennent dans ce phénomène :

- la quantité d'ARN polymérase présente dans la cellule, plus il y a d'enzyme et plus le taux de transcription sera élevé.
- L'efficacité du promoteur. Le promoteur est considéré comme fort ou faible en fonction du taux auquel il est transcrit.

2.2.2.1. ARN polymérase

Nomura *et al.* (1987) ont démontré la possibilité d'augmenter la quantité d'ARNm en augmentant la quantité d'ARN polymérase dans la cellule.

2.2.2.2. *Promoteurs forts*

La technique la plus employée pour augmenter la transcription est l'utilisation de promoteurs dits forts. D'après Peretti et Bailey (1987), une majeure partie de l'ARN polymérase présente dans la cellule est fixée de façon non spécifique à l'ADN. En augmentant la force du promoteur, l'ARN polymérase sera utilisée de façon plus efficace. La force relative des promoteurs chez *E. coli* varie d'un facteur quatre.

Les promoteurs utilisés pour la production de protéines recombinantes ont été obtenus à partir de séquences naturelles d'*E. coli*. Un promoteur standard d'*E. coli* est une portion d'ADN d'environ 40 paires de bases et se positionne en amont du site de démarrage de la transcription. On distingue deux types de promoteurs :

- les promoteurs constitutifs qui sont toujours opérationnels comme celui du gène de la β -lactamase. Le taux de synthèse d'ARNm n'y est pas modulé par la présence d'autres composés cellulaires.
- les promoteurs régulés permettant d'obtenir un taux de synthèse de l'ARNm variable. Les promoteurs régulés sont très importants pour une production efficace de protéines recombinantes. Chez les bactéries, les promoteurs régulés sont contrôlés principalement par des protéines répresseurs. La liaison de cette protéine inhibe l'interaction de l'ARN polymérase avec le promoteur et la transcription est diminuée. L'affinité de la protéine répresseur pour l'ADN est souvent modulée par la concentration de certaines petites molécules (nucléotides, acides aminés). On distingue d'une part les promoteurs inductibles, où une petite molécule (inducteur) se lie à la protéine répresseur pour diminuer son affinité pour l'ADN et augmenter la synthèse d'ARNm et d'autre part les promoteurs réprimés où une concentration élevée de petites molécules (corépresseur) diminue la transcription. Des principes similaires sont appliqués aux promoteurs induits/réprimés directement par les changements physiques (température, pression osmotique) des conditions de croissance.

D'après Georgiou (1988), un promoteur utilisable pour les productions à grande échelle doit posséder les qualités suivantes :

- avoir un niveau de transcription maximal,
- être fortement régulé : la synthèse de l'ARNm doit être totalement inhibée en dehors des périodes d'induction,
- l'induction ne doit pas avoir d'effet sur la physiologie de la cellule.

De plus un promoteur est considéré comme fort à partir d'une accumulation de la protéine recombinante représentant au moins 10 % des protéines cellulaires.

Des promoteurs de types *trp* ou *pho* ont été utilisés pour des productions à grande échelle, ils sont actifs dans des cultures limitées en tryptophane (*trp*) ou en phosphate (*pho*). L'utilisation de ces promoteurs pose des difficultés d'application. En effet, la synthèse de protéines à partir de promoteurs réprimés est initialisée quand la concentration du corépresseur diminue. Les cellules poussent donc en présence d'une concentration élevée de corépresseur (inhibant la synthèse prématurée de la protéine). La formation du produit est donc initiée par le transfert des cellules en milieu à faible concentration en corépresseur. Des cultures à deux étages sont souvent utilisées pour ce type de promoteur. Dans la première phase de culture la croissance est favorisée, dans la deuxième phase le répresseur est supprimé et la synthèse induite. Toutefois ceci n'est pas facilement réalisable à grande échelle c'est pourquoi, l'utilisation de promoteurs induits est préférable dans ces cas de figure.

Les principaux promoteurs inductibles sont *lac* UV5, *tac* et λp_L . Ils permettraient une accumulation de la protéine à un taux de 30 % de la protéine totale cellulaire. Les deux premiers promoteurs sont induits par de petites quantités d'IPTG (isopropyl-B-D-thiogalactoside). Le promoteur le plus utilisé : λp_L est thermoinductible. La synthèse de l'ARNm est initiée par une élévation de température (à 42°C), ceci pose quelques problèmes :

- les bactéries doivent croître à des températures faibles (inférieures à 37 °C) afin de réprimer la production de la protéine. Or la température optimale pour *E. coli* est 37°C, par conséquent, les temps de fermentation seront allongés,
- lors de la fermentation à 42°C, on peut observer l'apparition de protéines non spécifiques.

Celles-ci peuvent être des enzymes protéolytiques et accélérer la dégradation du produit.

Ceci a conduit au développement de promoteurs non thermoinductibles.

Le promoteur fort de l' α -amylase de *Bacillus subtilis* a été caractérisé et utilisé de façon intensive pour la production de protéines recombinées (Palva, 1982).

L'utilisation de promoteurs doit être couplée avec l'utilisation de séquences de terminaison de transcription de même force. En effet, en l'absence d'un terminateur fort, l'ARN polymérase continue au delà de l'extrémité du gène et forme une molécule plus longue

laquelle interfère avec l'origine de réplication entraînant une instabilité du plasmide (Stueber et Bujard, 1982).

2.2.3. Stabilité de l'ARNm

La stabilité de l'ARNm diminue avec la température (Kennell, 1986). Il est donc possible de stabiliser les ARNm en diminuant la température de fermentation.

2.3. Traduction

Dans la synthèse de protéines, l'étape suivante est la traduction de la séquence ribonucléotidique de l'ARNm en une séquence d'acides aminés correspondant à la protéine désirée. Le ribosome reconnaît le site ribosomique, séquence spécifique de ribonucléotides à l'intérieur de la molécule d'ARNm et s'y lie. La structure du site ribosomique est très importante car elle détermine l'efficacité de synthèse des protéines par molécule d'ARNm (Stormo, 1986). La séquence nucléotidique entourant le site ribosomique affecte aussi l'interaction du ribosome avec l'ARNm et exerce un effet important sur le taux de synthèse des protéines.

Affectent également le taux de traduction :

- la séquence nucléotidique précédant et suivant le premier triplet de nucléotides qui est traduit en un acide aminé. Souvent il est possible de modifier certaines séquences nucléotidiques critiques du début du gène et augmenter le taux de synthèse de la protéine sans changer le taux d'ARNm (Hui *et al.*, 1984),
- la séquence nucléotidique du gène désiré ou plutôt l'abondance relative des différents triplets nucléotidiques à l'intérieur du gène (codon usage) (De Boer et Kastelein, 1986). Ce dernier détermine probablement la vitesse avec laquelle le ribosome complète la synthèse de la molécule de protéine.

2.4. Devenir post-traductionnel des protéines.

La protéine synthétisée peut être modifiée suite à la formation de liaisons entre groupes chimiques, dégradée ou précipitée dans la cellule pour former des agrégats. Si elle reste dans la cellule, elle est dite cytoplasmique, si elle traverse la membrane interne bactérienne, elle est sécrétée, enfin si elle traverse les deux membranes pour se retrouver dans le milieu extérieur, elle est dite excrétée. Dans ce dernier cas seulement elle évite la dégradation par les enzymes bactériens.

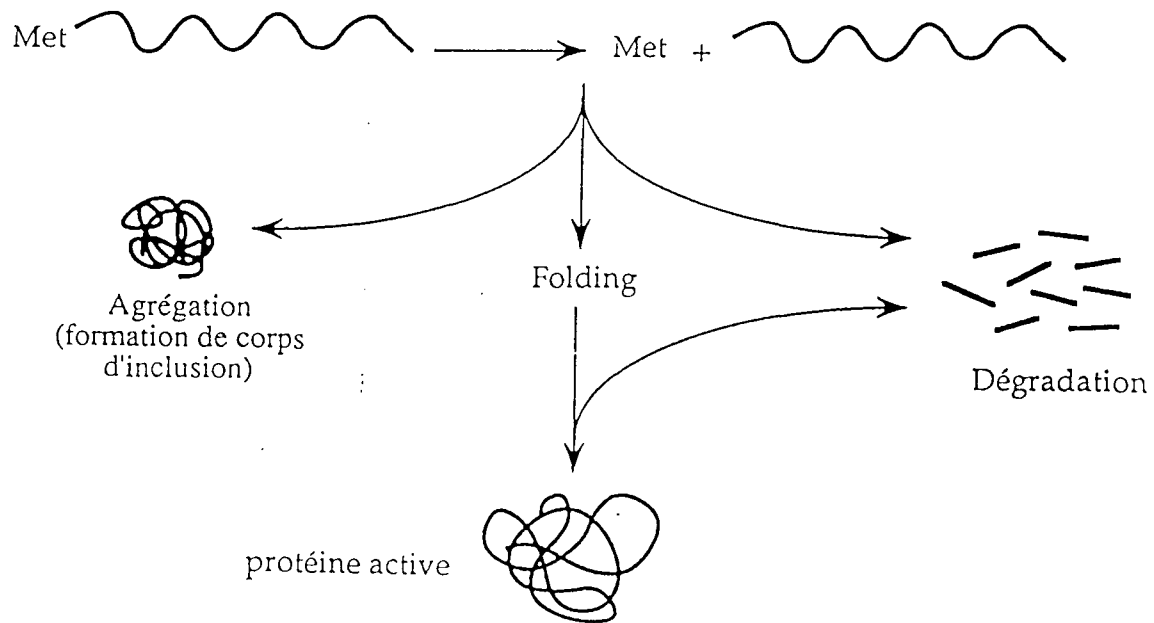
Le devenir des protéines dépend de leur localisation. L'importance de la modification chimique de la protéine varie d'un organisme à un autre. Il est préférable que la protéine recombinée soit identique à la protéine naturelle. La localisation intracellulaire et la forme de la protéine (agrégée ou soluble) permettent de déterminer les techniques de purification. Ainsi les étapes post-traductionnelles de la synthèse de la protéine déterminent souvent le rendement total et la rentabilité de la production de protéines recombinées.

2.4.1. Protéines cytoplasmiques

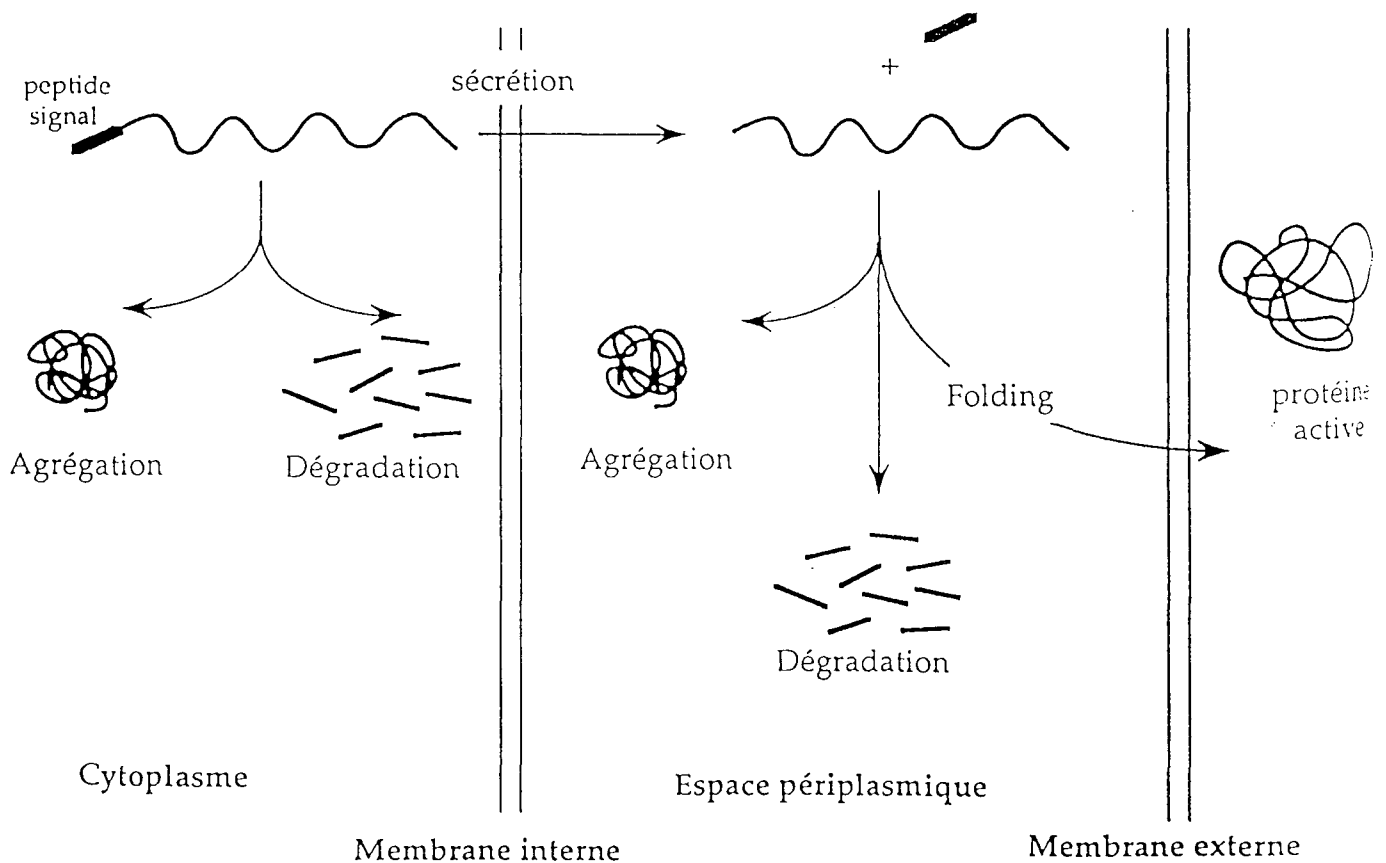
Les protéines cytoplasmiques restent dans le cytoplasme des cellules.

Toutes les protéines (sauf les mitochondriales et les chloroplastiques) sont synthétisées dans le cytoplasme. Le devenir des protéines étrangères dans le cytoplasme chez *E. coli* est indiqué **figure 2.3**. La synthèse de la protéine est initiée avec la méthionine ou son dérivé la N-formylméthionine. Chez la plupart des protéines cytoplasmiques, la méthionine initiatrice est éliminée post-traductionnellement par l'enzyme méthionine amino-peptidase. Les protéines cytoplasmiques peuvent être dégradées par l'action des enzymes protéolytiques, agrégées pour former de grandes particules : les corps d'inclusion où elles peuvent acquérir leur structure active ou rester solubles à l'intérieur de la cellule. Ces dernières correspondent aux protéines les plus recherchées.

Protéines cytoplasmiques



Protéines sécrétées

Figure 2.3. : Devenir post-traductionnel des protéines (Georgiou *et al.*, 1986)

2.4.2. Dégradation protéolytique

La dégradation protéolytique du produit est un problème important dans la production des protéines recombinantes. La cellule identifierait certaines protéines recombinées comme différentes ou « anormales » par rapport aux protéines natives. De telles protéines anormales sont rapidement exposées à la destruction.

La séquence des premiers acides aminés de la protéine peut être un signal important dans la dégradation. Ainsi, *E. coli* aurait huit enzymes qui dégradent les protéines en attaquant les liaisons peptidiques internes (Ces enzymes ont pour noms : *Do*, *Re*, *Mi*, *Fa*, *So*, *La*, *Ci* et *Pi* Goldberg et Goff, 1986). La protéase *La* semble jouer un rôle prédominant dans ces phénomènes, en effet dans les cellules qui ne produisent pas cette enzyme, la dégradation des protéines anormales est plus faible (Maurizi *et al.*, 1985). De plus, une surproduction de protéines étrangères cause une augmentation du taux de *La*, ce qui accélère la dégradation du produit (Goff et Goldberg, 1985).

La protéolyse peut être réduite en jouant sur les conditions de croissance. Par exemple, la limitation en source carbonée ou en acides aminés se traduit par une augmentation de la protéolyse (Tsai *et al.*, 1987). Ceci est important car des limitations en glucose et oxygène peuvent se manifester lors de fermentations à haute densité et à grande échelle. Une possibilité pour réduire la dégradation est de cultiver les cellules à des températures suboptimales, par exemple à 30 °C (Oka *et al.*, 1985).

On peut aussi fusionner un gène codant pour la protéine instable derrière un autre gène codant pour une protéine native d'*E. coli* (Germino et Bastia, 1984). La protéine résultante est un polypeptide hybride. Il comprend une séquence d'acides aminés précédée par une séquence d'acides aminés correspondant à un autre gène. La présence de la séquence d'acides aminés ajoutés permet de stabiliser la protéine hybride. Ce dernier peut ainsi être produit en grande quantité et sera purifié par les techniques conventionnelles. Ensuite, la partie correspondant au polypeptide désiré est clivée par des moyens enzymatiques ou chimiques. De plus, le produit peut être protégé de la dégradation protéolytique s'il forme des agrégats ou corps d'inclusion à l'intérieur de la cellule. En effet, ces derniers ne sont pas accessibles à la protéolyse enzymatique.

Toutefois, l'utilisation de protéines hybrides n'est pas toujours possible car un clivage correct du produit à partir de l'hybride est difficile.

2.4.3. Sécrétion de protéines chez *E. coli*

Le cytoplasme d'*E. coli* est protégé de l'environnement externe par la présence de deux couches membranaires constituées de lipides hydrophobes : les membranes interne et externe. La plupart des protéines ne peuvent pas traverser normalement la membrane cellulaire. Chez *E. coli*, la membrane interne retient les protéines à l'intérieur du cytoplasme. Cependant, il y a un mécanisme élaboré qui permet à certaines protéines d'être sécrétées à travers cette membrane. Les protéines sont alors synthétisées sous forme de précurseurs, elles sont complétées par une extension d'environ 20 à 30 acides aminés (Oliver, 1985). Grâce à cette séquence signal, la protéine peut traverser la membrane interne. Une fois de l'autre côté, la séquence signal est clivée et la forme mature de la protéine se plie dans sa structure native. Dans ce cas, la forme précurseur de la protéine peut s'agréger et former des corps d'inclusion cytoplasmiques (Georgiou *et al.*, 1986). Ou bien elle peut être dégradée par la protéolyse. Une partie de la quantité totale de la protéine précurseur peut être sécrétée au travers de la membrane interne, auquel cas, la séquence signal est éliminée. La forme mature prend alors sa forme tridimensionnelle. Si les taux de sécrétion et de synthèse sont élevés, la concentration de la protéine dans l'espace périplasmique peut former des corps d'inclusion périplasmiques.

La sécrétion dans l'espace périplasmique est en général recherchée car elle évite une dégradation du produit (Talmadge et Gilbert, 1982) et entraîne des rendements élevés, des protéines correctement repliées et une terminaison précise du dernier acide aminé de la protéine. Apparemment, la structure d'un grand nombre de protéines cytoplasmiques semble incompatible avec la sécrétion, toutefois, des protéines à intérêt thérapeutique telles l'insuline ou l'hormone de croissance sont naturellement sécrétées. Quand ces protéines sont synthétisées par *E. coli*, elles sont transportées à travers la membrane interne et finissent dans l'espace périplasmique (Barbero *et al.*, 1986).

2.4.4. Excrétion des protéines recombinantes dans le milieu

L'excrétion des protéines recombinées dans le milieu de culture est souhaitable. Notons que la sécrétion est définie comme la translocation de la protéine à travers la seule membrane lipidique, tandis que l'excrétion est la seule libération complète de la protéine dans le milieu de fermentation. Une protéine excrétée doit passer à travers au moins une membrane cellulaire.

Tout d'abord, la récupération et la purification de la protéine est simplifiée. L'excrétion évite de casser les cellules pour récolter le produit. La masse cellulaire est enlevée par centrifugation ou filtration, et le produit est purifié à partir de la suspension diluée. L'excrétion permet aussi d'éviter la dégradation protéolytique du produit, ce qui montre que la cellule hôte ne peut pas libérer les enzymes protéolytiques dans le milieu de culture. Le cas de *Bacillus subtilis* est toutefois à part, cette bactérie excrète normalement ses protéases qui peuvent ainsi dégrader le produit recombiné. Enfin, l'excrétion permet d'utiliser des procédés de fermentation en continu.

La présence de la membrane externe chez *E. coli* empêche la libération de la forme mature de la protéine dans le milieu extracellulaire. Il n'y a pas de phénomène d'excrétion. Il faut afin de libérer les protéines, altérer la perméabilité de la membrane externe. Chez *E. coli*, la protéine sécrétée se retrouve soit au niveau de la membrane externe soit dans l'espace périplasmique. Ces dernières peuvent alors diffuser à travers la membrane externe modifiée. Ainsi, on a utilisé des méthodes physiques pour réaliser la libération des protéines dans l'espace périplasmique telles le choc osmotique (Heppel, 1971) ou le traitement thermique (Tsuchido *et al.*, 1985). Toutefois ces méthodes sont difficiles à reproduire à grandes échelles, de plus elles s'accompagnent souvent de produits indésirables tels que les endotoxines qui doivent être ultérieurement éliminées.

Le transport de protéines à travers la ou les membranes d'autres organismes suit les mêmes principes que ceux appliqués à la sécrétion des protéines à travers la membrane interne d'*E. coli*. Toutefois beaucoup de micro-organismes excrètent la protéine dans le milieu de culture comme *Bacillus subtilis*, qui n'a pas de membrane externe. Toutefois un inconvénient à l'utilisation de cette cellule hôte est la faible stabilité des protéines obtenues après excrétion par *Bacillus*. En effet, ce dernier libère également beaucoup d'enzymes protéolytiques dans le milieu ce qui cause une dégradation rapide.

Les mécanismes d'excrétion des protéines par les levures sont plus complexes que chez les bactéries. Les protéines sont insérées dans des vésicules intracellulaires, lesquelles sont transportées vers la surface cellulaire où elles libèrent leur contenus. Seules de petites protéines recombinées sont facilement excrétées par les levures, toutefois, un des avantages important de leur utilisation réside dans le fait que les protéines excrétées subissent certaines modifications compliquées qui ne peuvent pas être réalisées chez les bactéries (glycosylation et isomérisation des ponts disulfures).

2.5. Métabolisme de l'acétate chez *E. coli*

2.5.1. Mécanismes de production en aérobiose

Les cultures d'*E. coli*, réalisées en aérobiose en présence de fortes concentrations de glucose et à des taux de croissance élevés, produisent des acides organiques et particulièrement de l'acétate (Pan *et al.*, 1987), c'est l'effet Crabtree bactérien. Plusieurs hypothèses ont été avancées afin d'expliquer ce phénomène :

Selon Majewski et Domach (1990), la production d'acétate serait liée à la fois à une limitation du cycle de Krebs et de la chaîne de transport des électrons. La limitation du cycle de Krebs serait responsable de l'initialisation de la production d'acétate, la limitation au niveau du transport des électrons intervenant à des niveaux de croissance supérieurs.

D'après Doelle *et al.* (1982), l'accumulation de NADH₂ dans la cellule est responsable de la déviation du flux de carbone vers l'acétate. Cette accumulation d'acétate serait liée à une répression des enzymes du cycle de Krebs ainsi que celles de la chaîne respiratoire.

Une autre hypothèse attribue la formation d'acétate à la capacité respiratoire d'*E. coli*. Reiling *et al.* (1985) ont montré que la vitesse spécifique de consommation de l'oxygène atteint une valeur maximale à laquelle elle se stabilise même si le taux de croissance continue à augmenter. Cette limitation de la capacité respiratoire entraînerait un passage en métabolisme fermentaire et donc conduirait à une production d'acétate.

Han *et al.* (1992) ont étudié la production d'acétate en culture continue à des taux de dilution différents. A faible taux de dilution, les demandes en énergie du métabolisme sont satisfaites par la voie oxydative. Par contre, lorsque le taux de dilution (et donc le taux de croissance en chemostat) augmente la demande énergétique des voies anabolique et catabolique dépasse les capacités de la bactérie. Celle-ci doit alors réorganiser son métabolisme oxydatif et produire de l'acétate, pour fournir de l'énergie à l'anabolisme. Les voies cataboliques utilisent l'énergie du métabolisme oxydatif (non mobilisée par les voies anaboliques) ainsi que celle de l'acétate.

2.5.2. Effet inhibiteur de l'acétate

La production d'acétate ainsi que son accumulation lors de cultures avec *E. coli* entraîne une inhibition de la croissance (Pan *et al.*, 1987), et peut inhiber la production de protéines recombinantes (Brown *et al.*, 1985).

Selon Smirnova et Ooktyabr'skii (1985, 1989). La croissance d'*E. coli* nécessite une valeur élevée de $\Delta\mu\text{H}^+$ de la membrane ainsi que le maintien d'une valeur de pH optimum dans la cellule (entre 7.4 et 7.8). Or l'acétate est capable de pénétrer dans la cellule et de perturber le pH en acidifiant le contenu intracellulaire. En effet, la forme protonique de l'acétate est capable de traverser la membrane de la cellule, l'acétate est alors dissocié en CH_3COO^- et H^+ au pH interne de 7.5, ce qui entraîne une diminution du pH intracellulaire. Du fait d'un grand volume de milieu tamponné, le pH externe total ne varie que très peu. Ainsi cette diminution du pH intracellulaire devrait causer un effet de découplage. Les mécanismes homéostatiques de la cellule sont capables de rétablir le pH moyennant une dépense énergétique importante. Toutefois ce mécanisme de régulation a une limite au delà de laquelle la croissance cellulaire est ralentie ou inhibée.

La production d'acétate est donc un élément essentiel à suivre et à contrôler lors des cultures d'*E. coli*. Des études ont été menées mettant en évidence des variations de production suivant les souches utilisées et les conditions de culture appliquées.

Des cultures discontinues réalisées avec onze souches d'*E. coli* dont *E. coli* HB101, B, JM 105, et DH1 par Luli et Strohl (1990), ont permis de comparer la croissance, l'inhibition de la croissance et la production d'acétate pour ces différentes souches. En culture en fiole d'Erlenmeyer, *E. coli* HB101 a le taux de croissance spécifique maximum le plus faible parmi les souches testées mais présente un rendement de biomasse sur glucose le plus élevé ainsi que la production d'acétate la plus faible.

Sun *et al.* (1993) ont testé des souches produisant une protéine de fusion lors de différentes cultures en fioles d'Erlenmeyer en présence d'acétate. Ils ont mis en évidence une variation du niveau d'inhibition suivant la souche utilisée.

Selon Koh *et al.* (1992), les souches recombinées d'*E. coli* HB101 seraient plus inhibées par l'acétate que les souches hôtes non recombinées.

De plus le milieu influe sur les phénomènes d'inhibition, ceux-ci sont plus importants en milieu défini qu'en milieu complexe. Meyer *et al.* (1984) ont réalisé des cultures continues à des taux de dilution croissants en milieu complexe et en milieu synthétique. En milieu synthétique, la production d'acétate commence à des taux de dilution plus élevés qu'en milieu complexe (0.35 h^{-1} contre 0.2 h^{-1}). Par contre, la vitesse spécifique de production d'acétate est plus élevée en milieu complexe (0.7 g/g.h contre 0.3 g/g.h en milieu défini).

3. Mise en œuvre en réacteur

Les objectifs à atteindre lors de cultures de micro-organismes recombinés sont définis par Georgiou (1988) comme étant :

- taux de croissance et vitesse de formation de produit élevés,
- obtention de hautes densités cellulaires,
- augmentation de la stabilité du produit.

En culture discontinue, tous les substrats sont présents dès le début de la fermentation. Au départ, la croissance de la culture n'est pas limitée et le taux de formation de biomasse est élevé. Toutefois, des concentrations importantes en glucose ainsi qu'une croissance non limitée (en oxygène ou autre élément nutritif) saturent certaines voies métaboliques cellulaires (Holms, 1986). Par conséquent la croissance diminue et des inhibitions métaboliques par des produits accumulés dans le milieu de culture se manifestent. De ce fait, la densité cellulaire finale et le rendement global de la masse cellulaire par gramme de carbone sont diminués. Ce phénomène appelé « effet Crabtree » est un exemple d'inhibition par le substrat et il est observé chez différents organismes.

Chez *E. coli*, on observe également une inhibition par un produit qui est l'acétate. Les levures, elles, accumulent l'éthanol qui est aussi inhibiteur de la croissance. Pour éviter l'accumulation de substrat et obtenir de fortes densités cellulaires, on cherche à maintenir de faibles concentrations en glucose durant toute la fermentation. En culture continue ou semi-continue, on utilise de faibles débits de glucose. De plus, dans de nombreux procédés, la concentration en glucose est maintenue à de faibles valeurs pour éviter la formation d'acétate inhibiteur de la croissance et de la production de protéines recombinantes.

La concentration en oxygène dissous est également un facteur à prendre en compte. En effet, lorsque la vitesse de consommation en oxygène dépasse la vitesse à laquelle il est ajouté, la culture devient limitée en oxygène, la bactérie passe en métabolisme fermentaire et produit de l'acétate, le taux de croissance et le rendement en carbone en sont affectés. Ce problème est d'autant plus important que les densités cellulaires sont plus élevées. En effet, la consommation de l'oxygène est proportionnelle à la concentration de biomasse et au taux de croissance. Pour remédier à cet inconvénient, les cultures sont alimentées en air enrichi en

oxygène, voire en oxygène pur. De plus, la viscosité du milieu augmente avec la concentration cellulaire, ce qui entraîne une diminution du coefficient de transfert de masse. La vitesse de transfert de l'oxygène peut être augmentée soit par augmentation de l'agitation et/ou en dispersant de l'oxygène pur.

Lee et Chang (1988) ont réalisé des cultures continues à hautes densités cellulaires d'*E. coli* HB101 (pAKAS2) productrice de pénicilline G acylase. L'utilisation d'oxygène pur a permis d'obtenir de hautes densités cellulaires sans mettre en évidence une inhibition par l'acétate.

Enfin, l'utilisation de cultures continues pour la production de protéines recombinantes présente deux inconvénients : le risque de contamination lié à des cultures plus longues dans le temps et l'instabilité des cellules contenant des plasmides (Imanaka et Aiba, 1981, Siegel et Ryu, 1985). Comme elles croissent moins rapidement que les cellules ayant perdu leurs plasmides, elles peuvent être lessivées hors du fermenteur, diminuant voire annulant le rendement de la production. La stabilité plasmidique en culture continue peut être augmentée par l'immobilisation des cellules recombinantes (De Taxis du Poët *et al.*, 1987). Dans les réacteurs utilisant des cellules immobilisées, les cellules sont enfermées soit dans une membrane poreuse, soit à travers un piège à l'intérieur d'une phase solide. Le liquide circule à travers le réacteur pour fournir le substrat aux cellules et enlever le produit, les cellules restant à l'intérieur du réacteur.

Ariga *et al.*, (1991) ont cultivé *E. coli* HB101 (pHI301) immobilisée sur billes de carraghénane pour synthétiser une α -amylase de *Bacillus stearothermophilus*. Cette technique a permis d'augmenter la stabilité plasmidique. Dinçbas *et al.* (1993) sont arrivés aux mêmes conclusions avec des cultures d'*E. coli* HB101 immobilisées sur billes d'alginate.

En 1990, Lee et Chang ont réalisé des cultures continues avec recyclage cellulaire. Comme pour les cultures en semi-continu, la production d'acétate doit être minimale pour obtenir de hautes densités cellulaires, c'est pourquoi la possibilité d'éliminer l'acétate produit par recyclage du milieu constitue un avantage prépondérant de ce type de culture par rapport aux cultures continues et semi-continues.

Park *et al.* (1990) ont utilisé des cultures bi-étagées d'*E. coli* M72 recombinées productrices de la tryptophane synthétase. L'intérêt de ce type de culture est de séparer la phase de croissance de la phase de production des protéine recombinantes. Les paramètres importants de ce type de culture sont la température et les taux de croissance dans les deux compartiments de la culture. Ce type de technique permet d'obtenir des cultures stables pendant des temps plus longs.

4. Modèles mathématiques

Un des principaux objectifs du génie biochimique est la mise en œuvre à l'échelle industrielle des souches recombinées. L'optimisation de la production passe par une maîtrise des procédures de conduite afin d'atteindre les meilleurs rendements de conversion.

Afin de réaliser ces objectifs, il est nécessaire de comprendre et de prévoir le comportement des souches bactériennes, afin de pouvoir optimiser, extrapoler et automatiser au mieux les procédés biotechnologiques.

Pour cela, on fait souvent appel à la modélisation. La formulation de modèles mathématiques permet de décrire le comportement et la croissance de souches recombinées. Ces modèles sont basés sur des hypothèses, lesquelles sont vérifiées en comparant les résultats des simulations aux données expérimentales.

Les phénomènes d'instabilités plasmidiques sont pour une grande part responsables de pertes de rendements observées au cours des fermentations de cellules recombinées. C'est pourquoi la modélisation de ces phénomènes s'avère essentielle afin de mieux comprendre les processus qui les régissent. On pourra ainsi prévoir l'évolution des cellules recombinées et améliorer en conséquent les processus à mettre en œuvre pour optimiser les cultures.

Les modèles développés jusqu'ici ont surtout étudié les instabilités plasmidiques de type ségrégationnel.

On distingue :

- les modèles structurés prenant en compte les phénomènes moléculaires, des modèles non structurés ne considérant que les phénomènes se déroulant au niveau cellulaire,
- les modèles dits non ségrégués lorsque toutes les cellules sont considérées comme identiques, des modèles dits ségrégués considérant les variations d'une cellule à l'autre.

Un des principes communs aux modèles mathématiques utilisés pour décrire les cultures microbiennes repose sur une considération du caractère hétérogène de la population (Shuler , 1985). En effet les souches de bactéries recombinées se caractérisent par une plus ou moins grande instabilité plasmidique. Celle-ci s'explique principalement par un phénomène de ségrégation plasmidique au moment de la division cellulaire d'une part, et par une différence entre les cinétiques de croissance des sous populations d'autre part, les populations ségrégantes se développant sans doute plus rapidement que les recombinées.

4.1. Modèles non ségrégués

Les premiers modèles considéraient deux sous-populations : l'une comportant des plasmides, l'autre n'en comportant pas. Dans ces modèles dits non ségrégués, les cellules au sein de chaque sous-population interagissent de façon identique avec l'environnement. Cela reviendrait à considérer deux « espèces » en culture dans la fermentation.

Le premier modèle de croissance non ségrégué a été proposé par Imanaka et Aiba (1981). Il est constitué d'un système de deux équations différentielles du premier ordre et décrit la croissance de chaque population :

$$\left\{ \begin{array}{l} \frac{dX^+}{dt} = \mu^+ X^+ - p\mu^+ X^+ - DX^+ \\ \frac{dX^-}{dt} = \mu^- X^- + p\mu^+ X^+ - DX^- \end{array} \right. \quad \begin{array}{l} (1) \\ (2) \end{array}$$

X^+ et X^- représentent les concentrations en g/l des cellules avec et sans plasmides, t est le temps en h, μ^+ et μ^- sont les taux de croissance spécifiques en h^{-1} , p est un paramètre de ségrégation, il correspond à la probabilité de perte de plasmide par instabilité ségrégationnelle. D est le taux de dilution en h^{-1}

Luedeking et Piret (1959) ont décrit la variation de la concentration du produit P :

$$\frac{dP}{dt} = AX^+ + B \cdot (\mu^+ X^+ - p\mu^+ X^+) - Dp$$

A et B sont des constantes.

La variation de la concentration du substrat est décrite par Pirt (1975) :

$$\frac{dS}{Dt} = D(S_{ali} - S) - \frac{\mu^+ X^+}{y^+} - \frac{\mu^- X^-}{y^-}$$

y^+ et y^- sont les rendements de conversion du substrat en biomasse pour les deux populations, S_{ali} est la concentration en substrat à l'alimentation.

Dans ce modèle, la croissance globale repose sur deux paramètres clés : le taux de croissance différentiel entre les deux souches : $\delta\mu$, et le coefficient de ségrégation : $p\mu^+$. C'est pourquoi de nombreux auteurs se sont attachés à les étudier.

Selon Bhattacharya et Roy (1995), une intégration analytique des deux équations fondamentales de la cinétique des cellules recombinantes (1 et 2) est possible lorsque :

- la croissance est exponentielle c'est-à-dire que μ^+ et μ^- sont des constantes, Ceci est possible pendant la phase de croissance exponentielle notamment en batch. (Roy et Bhattacharya, 1998),
- le paramètre p est indépendant du temps et de μ et le taux de dilution D est nul.

Une intégration de l'équation (1) donne :

$$\ln[X^+ / X_0^+] = (1 - p)\mu^+ t$$

X_0^+ étant la valeur de X^+ au temps $t=0$. p est obtenu à partir de la pente de la droite $\ln[X^+ / X^-]$ en fonction du temps.

Toutefois, d'après les études expérimentales de Sardonini et Di Basio (1987) ainsi que de Ryan et Parulekar (1991), p , paramètre d'instabilité ségrégationnel serait dépendant du taux de croissance μ . Pour relier alors p à μ , Bhattacharya et Roy (1995) ont considéré la biomasse totale : $X^T = X^+ + X^-$. En différentiant cette équation et en y substituant les équations fondamentales (1) et (2), on obtient :

$$\frac{dX^T}{dt} = \mu^+ X^+ + \mu^- X^- - DX^T. \text{ En divisant cette équation par } X^T \text{ et en y substituant la}$$

fraction de cellules contenant des plasmides : $F = X^+ / X^T$; on a :

$$\frac{1}{X^T} \cdot \frac{dX^T}{dt} + D = F\mu^+ + (1 - F)\mu^- \quad (3)$$

De plus, si on considère que les taux de croissance spécifiques suivent une loi de Monod telle que :

$$\mu^+ = \frac{\mu_m^+ S}{K_s^+ + S}$$

$$\mu^- = \frac{\mu_m^- S}{K_s^- + S}$$

μ_m^+ et μ_m^- étant les taux de croissance maximum des populations et K_s le paramètre de demi-saturation.

Et que le K_s est constant pour les deux sous-populations, soit : $S \gg K_s$ auquel cas celui-ci peut

être négligé, alors : $\frac{\mu^-}{\mu^+} = \frac{\mu_m^-}{\mu_m^+} = \alpha$

Enfin, en considérant $g(X^T) = \frac{1}{X^T} \cdot \frac{dX^T}{DT}$, on obtient

$$\begin{cases} g(X^T) + D = \mu^+[F+(1-F)\alpha] & (4) \\ g(X^+) + D = \mu^+(1-p) \end{cases}$$

à partir des équations (1) et (3).

On peut alors estimer p en fonction de μ^+ au moment où X^T et F sont mesurés expérimentalement. Pour estimer α , on se met en conditions batch ($D = 0$) à $t = 0$, à ce moment, on a :

$\mu^+ = \mu_m^+$ et $\mu^- = \mu_m^-$. Et l'équation (4) est réduite à :

$$g(X^T) = \mu_m^+ F^0 + \mu_m^- (1-F^0).$$

Sachant que μ_m^- est déterminé expérimentalement à partir de souches complètement ségrégantes, on calcule alors μ_m^+ et on en déduit une valeur de α .

Selon Roy et Bhattacharya (1998), si on considère μ constant et si on prend en compte le fait que p paramètre de ségrégation peut être défini comme le rapport du taux de formation de cellules sans plasmides P^- sur le taux de croissance des cellules avec plasmides P^+ , il est ainsi hautement probable que p et μ soient fonction l'un de l'autre.

Ainsi si on considère μ constant, il est probable que p le soit aussi.

San et Weber (1989) se sont basés sur cette hypothèse pour résoudre de façon analytique le système d'Imanaka et Aiba.

En considérant la biomasse totale : $X^T = X^+ + X^-$ et la fraction de biomasse contenant des plasmides : $y^+ = X^+/X^T$, ces auteurs parviennent à la relation algébrique :

$$\frac{dy^+}{dt} = y^+[(\mu^+ - \mu^-)(1 - y^+) - p\mu^+] = y^+[(\delta\mu)(1 - y^+) - p\mu^+] \quad (5)$$

une intégration analytique de cette équation est possible comme $p\mu^+$ et $\delta\mu$ sont constants, on obtient :

$$\frac{y^+}{\delta\mu - p\mu^+ - \delta\mu y^+} = \frac{y_0^+}{\delta\mu - p\mu^+ - \delta\mu y_0^+} \cdot e^{(\delta\mu - p\mu^+)t}$$

Dans un système plasmidique relativement stable, on a : $|\delta\mu| \gg p\mu^+$, $y_0^+ = 1$ et $y^+ \ll 1$, on obtient alors :

$$\frac{y^+}{1 - y^+} = \frac{y^+}{y^-} = -[\delta\mu / p\mu^+] \cdot e^{(\delta\mu - p\mu^+)t}$$

Un tracé semilog du rapport y^+ / y^- du temps permet alors de déterminer $\delta\mu$ et $p\mu^+$ à partir de la pente et de l'ordonnée à l'origine de la droite de régression.

Toutefois, l'a priori selon lequel les paramètres clés seraient constants n'est pas toujours applicable (Nancib *et al.*, 1992). C'est pourquoi ces auteurs ont proposé une méthode de résolution des équations (1) et (2) en prenant en compte le cas plus général où l'équation (5) de San et Weber ne serait pas linéaire et où les paramètres pourraient être variables.

L'équation (5) divisée par y^+ donne :

$$\frac{d\ln y^+}{dt} = \delta\mu(1 - y^+) - p\mu^+$$

cette équation est de forme $Y = \delta\mu X - p\mu^+$

Une régression linéaire des valeurs de Y en fonction de X permet de déterminer si les valeurs de $\delta\mu$ et $p\mu^+$ sont constantes ou pas, il suffit pour cela d'étudier le coefficient de régression (qui est égal à 1 si on a une fonction parfaitement linéaire). Si $d\ln y^+ / dt$ en fonction de $1 - y^+$ donne une droite, alors sa pente et son ordonnée à l'origine seront respectivement $\delta\mu$ et $p\mu^+$.

Les données expérimentales pour y^+ sont d'abord filtrées numériquement pour éliminer le bruit de fond, puis un lissage polynomial est appliqué pour obtenir les valeurs de y^+ , les valeurs de dy^+ / dt puis de $d\ln y^+ / dt$ sont alors calculées. On réalise alors une régression linéaire de $d\ln y^+ / dt$ en fonction de $1 - y^+$ pour obtenir les valeurs des paramètres $\delta\mu$ et $p\mu^+$.

Les auteurs ont remarqué que les tracés obtenus n'étaient pas véritablement linéaires, ce qui suggère que les paramètres $\delta\mu$ et $p\mu^+$ ne sont pas constants et varient en fonction du taux de croissance. De plus, ces méthodes d'estimation n'ont donné que des valeurs moyennes.

Le paramètre p de ségrégation plasmidique a été étudié en détail par Mosrati *et al.*, (1993). Ils ont étudié d'abord de façon analytique la variation de p et ont montré que p variait de façon croissante avec μ^+ puis ont proposé une modélisation de ce paramètre.

Ils ont proposé la méthode suivante pour déterminer les valeurs des paramètres réels μ^+ , μ^- et p en fonction de conditions de culture données.

Le bilan total des équations de biomasse d'Imanaka et Aiba donne :

$$\frac{dX^T}{dt} = \mu^{app} X^T - DX^T \quad (6)$$

avec $\mu^{app} = (\mu^+ X^+ + \mu^- X^-) / X^T$ et la biomasse totale : $X^T = X^+ + X^-$

En considérant la fraction de biomasse contenant des plasmides : $y^+ = X^+ / X^T$ et la relation algébrique (5) donnée par San et Weber :

$$\frac{dy^+}{dt} = y^+[(\mu^+ - \mu^-)(1 - y^+) - p\mu^+] = y^+[(\delta\mu)(1 - y^+) - p\mu^+]$$

les auteurs obtiennent la relation suivante de μ^{app} :

$$\mu^{app} = \mu^+ y^+ + \mu^- (1 - y^+) \quad (7)$$

Les équations (6) et (7) conduisent à :

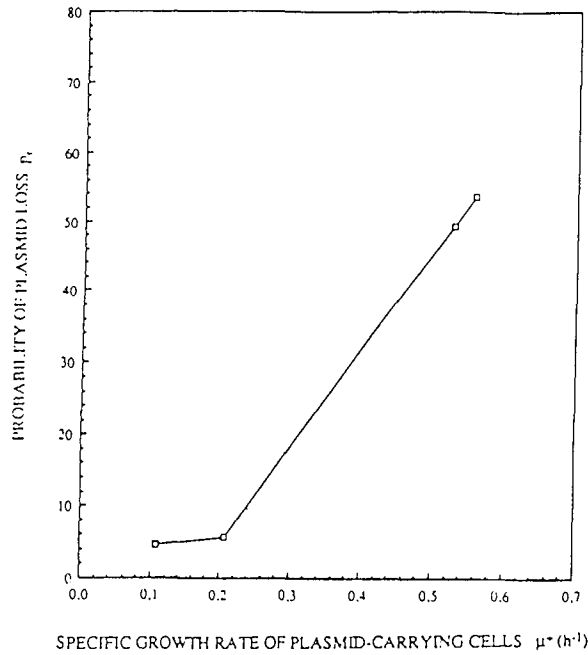
$$\frac{dLnX^T}{dt} = \delta\mu y^+ + \mu^- - D$$

A l'état stationnaire, le membre de gauche de l'équation peut être considéré comme nul, d'où :

$$\mu^- = D - \delta\mu y^+$$

Les valeurs de $\delta\mu$ et de $p\mu^+$ sont estimées selon la méthode proposée par Nancib *et al.*, (1992). A des valeurs données de D et y^+ , on détermine les valeurs de μ^- connaissant $\delta\mu$ et $p\mu^+$. On détermine alors μ^+ puis p grâce à $p\mu^+$.

En traçant les valeurs de p obtenues en fonction de μ^+ , ces auteurs ont observé que p était une fonction linéaire croissante de μ^+ en deux phases (2 pentes) séparées par une valeur critique $\mu_c^+ = 0.2 \text{ h}^{-1}$.



Ces auteurs ont ensuite réalisé une modélisation de la variation de ce paramètre p en fonction de μ^+ . Pour cela, il se sont basés sur le modèle de Satyagal et Agrawal (1989), lesquels ont proposé un modèle de réplication plasmidique selon lequel le taux net de variation de la concentration en plasmides dans une cellule est égal à la différence entre les taux de synthèse et de dilution due à la croissance cellulaire, ainsi :

$$\frac{dp}{dt} = V_p(p, h) - \mu^+ p$$

où p étant la concentration intracellulaire de plasmides et $V_p(p, h)$ est le taux intrinsèque de synthèse de plasmide.

Mosrati *et al.*, (1993) ont démontré que

$$p = \alpha \cdot \left[\mu^+ - V^0 \cdot \frac{\left(\frac{\mu^+}{K_h} \right)^n}{1 + \left(\frac{\mu^+}{K_h} \right)^n} \right] \quad (8)$$

où α et n sont des constantes, K_h est une constante de saturation (h^{-1}), et V^0 le taux maximal de réplication plasmidique (mg/g de cellule.h).

Pour déterminer les paramètres, les auteurs ont montré que lorsque $\mu^+ \gg K_h$, on a $(\mu^+/K_h^n) \gg 1$, soit l'équation (8) est approximée par :

$P = \alpha(\mu^+ - V^0)$ qui est une relation linéaire.

Le tracé des valeurs de p en fonction de μ^+ permet d'obtenir la pente α et de déterminer V^0 à partir de l'ordonnée à l'origine.

L'équation (8) peut être arrangée sous la forme suivante :

$$\left(\frac{\mu^+}{K_h}\right)^n \cdot \left[1 - \frac{\alpha\mu - p}{\alpha V^0}\right] = \frac{\alpha\mu - p}{\alpha V^0}$$

et si $Z = \frac{\alpha\mu^+ - p}{\alpha V^0}$, et $1 - Z = 1 - \frac{\alpha\mu^+ - p}{\alpha V^0}$, l'équation précédente devient :

$$\frac{Z}{1-Z} = \left(\frac{\mu^+}{K_h}\right)^n \text{ ou encore : } n \ln(\mu^+) - n \ln(K_h) = \ln\left(\frac{Z}{1-Z}\right)$$

En traçant $\ln [Z / (1-Z)]$ en fonction de $\ln \mu^+$, on obtient n en tant que pente de la droite et K_h à partir de l'ordonnée à l'origine.

La simulation obtenue avec les paramètres trouvés corrobore bien les résultats expérimentaux

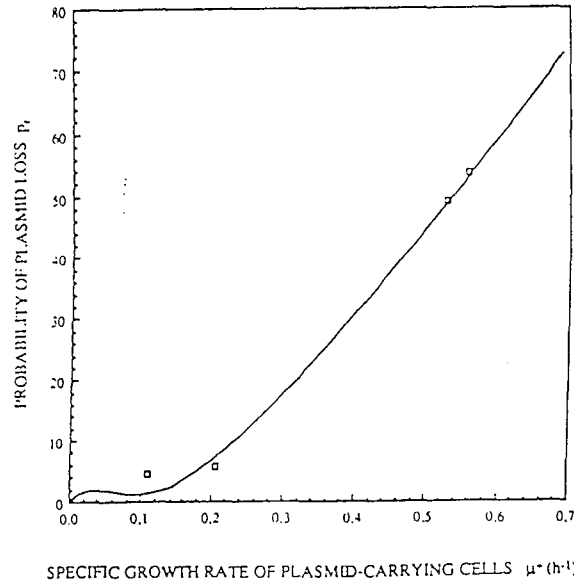


Figure 9. Result of the simulation of the evolution of the probability of plasmid loss as a function of the specific growth rate of the plasmid-carrying cells.

Patnaik (1993) a déterminé à partir des résultats de Mosrati *et al.*, (1993), une solution analytique permettant de relier directement la variation de p en fonction du taux de dilution. D'après les travaux de ces derniers auteurs, Patnaik (1993) a remarqué que μ^+ et μ^- étaient des fonctions linéaires du taux de dilution et que les tracés passaient par l'origine, soit :

$\mu^+ = \mu_0^+ D$ et $\mu^- = \mu_0^- D$. En incorporant ces équations dans l'équation (8), on obtient :

$$p = \alpha \cdot \left[\mu_0^+ D - \frac{V_0 \left(\frac{\mu_0^+ D}{K_h} \right)^n}{1 + \left(\frac{\mu_0^+ D}{K_h} \right)^n} \right] \text{ soit en dérivant cette expression par rapport au taux de}$$

dilution :

$$\frac{dp}{dD} = \alpha \mu_0^+ - \frac{n \alpha V_0 \left(\frac{\mu_0^+ D}{K_h} \right)^n}{D \left[1 + \left(\frac{\mu_0^+ D}{K_h} \right)^n \right]^2}$$

Les paramètres de cette équation sont repris dans la publication de Mosrati *et al.*, (1993) et les valeurs de μ_0^+ et μ_0^- sont obtenues à partir des régressions linéaires des valeurs de μ en fonction du taux de dilution. On s'aperçoit que p augmente avec le taux de dilution pour se stabiliser ensuite au delà de 0.7 h^{-1} . Ce phénomène serait dû au lessivage des cellules à de forts taux de dilution.

Koizumi et Aiba (1986) ont construit un modèle afin de rendre compte de l'évolution globale du nombre de copies de plasmide par cellule, en fonction du taux de croissance de la cellule. Ils ont considéré que la réplication du plasmide apparaissait au hasard, ou encore qu'un plasmide pouvait se répliquer quel que soit le temps qui le sépare de sa précédente division. Le temps nécessaire est très inférieur au temps de doublement de la population.

On obtient alors l'équation suivante :

$$m_p^{i+1} = E(N_p) = \frac{1}{2} [m_p^i - E(N_p)]$$

où m_p^{i+1} est le contenu plasmidique d'une cellule à la $i^{\text{ème}}$ génération, $E(N_p)$ représente le nombre de réplifications attendues au cours d'une génération (entre i et $i+1$).

Donc $[m_p^{i+1} - E(N_p)]$ suit une loi géométrique de raison 0.5. Lorsque le nombre de générations augmente, m_p tend vers $E(N_p)$, soit au bout d'un temps t fini, le nombre de copies de plasmides dans une cellule naissante tend vers le nombre de réplifications effectuées pendant une génération de la cellule hôte. Finalement, ils obtiennent une relation entre le taux de croissance (μ) et le nombre de copies de plasmide par cellule (N_p) :

$$N_p = \lambda / \mu$$

En utilisant différentes expressions pour λ , ce modèle permet d'envisager différentes relations existant entre μ et N_p :

- si λ est une constante, N_p augmente quand μ diminue,
- si λ est proportionnel à μ , on a alors N_p et μ indépendants,
- si $\lambda = f(\mu)$, on a N_p varie en fonction de μ .

4.2. Modèles ségrégués

Comme nous l'avons vu précédemment, de nombreux auteurs ont développé des modèles à partir des équations d'Imanaka et Aiba (1981) et ont cherché à estimer par des méthodes analytiques ou numériques les valeurs de deux paramètres clés de ce modèle : $\delta\mu$ (taux de croissance différentiel entre les deux sous-populations) et $p\mu^+$ coefficient de ségrégation. [San et Weber, (1989), Roy et Bhattacharya, (1995, 1998), Mosrati *et al.*, (1993), Nancib *et al.*, (1992)].

Bentley et Quiroga (1992) ont toutefois montré que l'exploitation de ces modèles donnaient des estimations du nombre de copies de plasmides par cellule inférieures à la réalité. Ainsi ces modèles à deux équations ne permettent pas de comprendre au mieux le comportement des plasmides même s'ils permettent de bien simuler les fractions de biomasse recombinée. Ils ont alors proposé un modèle ségrégué palliant à cet inconvénient. De plus, les modèles ségrégués présentent l'avantage de reconnaître explicitement la distribution des propriétés entre les individus d'une population mixte. Ils incorporent directement des informations sur les mécanismes de réplication plasmidique et de partition à l'échelle de la cellule (Dunn et Davidson, 1995)

Bentley et Quiroga (1992) ont exploité les données expérimentales de plusieurs études en continu et leur ont appliqué la méthode de San et Weber afin d'estimer les paramètres $p\mu^+$ et $\delta\mu$. Leurs résultats sont regroupés dans le tableau ci-après. Notons que ces valeurs permettent de bien simuler les fractions de cellules recombinées en fonction du temps.

En approximant ensuite la valeur du taux de dilution D à la valeur de μ^+ , et à partir des valeurs estimées du paramètre $p\mu^+$, ils ont ensuite calculé des valeurs de p . Enfin, ils ont déterminé les valeurs de N_{pd} (nombre de copies de plasmide par cellule au moment de la division) selon des relations suivantes données par Seo et Bailey (1985).

$$p = 1 - \frac{\ln(2 - \delta)}{\ln 2} \text{ et } \delta = 2^{(1 - N_{pd})}$$

Notons que le Npd (nombre de copies par cellule au moment de la division), doit être à peu près deux fois plus grand que le nombre moyen de copies de plasmides par cellule. Or les valeurs ainsi calculées de Npd sont de l'ordre de 10, ce qui donne des nombres de copies de plasmides moyens de 5, ce qui, au vu du taux de dilution appliqué, est très inférieur à la réalité. D'après ces auteurs, le Npd devrait s'approcher de 60 à ces taux de dilution. Ils en déduisent que le modèle à deux équations est inadéquat pour les plasmides pBR322 et leurs dérivés et que d'autres considérations doivent entrer en compte.

Table I. Model parameters and copy number calculated from two-population model.

Plasmid	Host	Replication Origin	Media	D (hr ⁻¹)	$\delta\mu$ (hr ⁻¹)	$p\mu^+$ (hr ⁻¹)	Average N_p measured	N_{pd} est ^a	Ref.
pBREcogap	C600 gal k (GAPDH)	pBR322	Complex	0.13	0.040	5.0×10^{-5}	N/A	11.8	35
				0.23	0.047	1.2×10^{-4}		11.4	
				0.57	0.083	2.6×10^{-3}		8.3	
				0.65	0.190	3.0×10^{-3}		8.3	
pBR322	K12	pBR322	Complex	0.31	0.023	3.0×10^{-5}	N/A	13.8	37
				0.46	0.023	1.0×10^{-4}		12.7	
pBR329	RR1	pBR322	M9CA	0.39	0.067	3.4×10^{-4}	50 ^d	10.7	3
			M9	0.11	0.085	1.0×10^{-6}	100	17.7	
			M9	0.20	0.075	2.8×10^{-7}	80	20.0	
pBR325	RR1	pBR322	M9	0.27	0.059	1.4×10^{-2}	N/A	8.3	
				0.40	0.119	1.4×10^{-5}		15.3	
pHSG415	K12(RV308)	pSC101	Defined	0.16	0.005 ^c	2.8×10^{-2c}	3.6 ^b	3.2	11
				0.30	0.004	3.3×10^{-2}	4.1	3.8	
				0.40	3.4×10^{-5}	8.8×10^{-3}	3.8	6.0	
				0.50	1.7×10^{-4}	9.8×10^{-3}	3.5	6.2	
pLc23-trpA1	M72	pKN402	Minimal	0.25	0.078	3.1×10^{-3}	60	6.9	42
				0.48	0.048	1.2×10^{-3}	44	9.2	
				0.79	0.054	1.9×10^{-3}	37	9.3	
				1.08	0.052	2.0×10^{-3}	31	9.5	
p17	B/r	pBR322	Minimal	0.27	0.068	3.5×10^{-4}	N/A	10.4	1
				0.7	0.20	8.6×10^{-5}		13.5	

^a Estimated by method of San and Weber (1990) and Seo and Bailey (1985).

^b Estimated by conversion of 5 plasmids/chromosome equivalent to per cell basis with Cooper-Helmstetter model.

^c As reported in original reference.

^d Estimated from Bentley and Kompala (1990).

Les références indiquées sur le tableau correspondent à :

35 : Nancib, N., Mosrati, R., Boudrant, J., 1992. *Biotechnol. Bioeng.* **39** : 398-407

37 : San, K. Y., Weber, A.E., 1989. *Biotechnol. Bioeng.* **33** : 451-459.

3 : Bentley, W. 1989. Effects of plasmid-mediated activity on bacterial metabolism and culture stability. PhD thesis, University of Colorado, Boulder, USA.

11 : Caulcott, C., A., Dunn, A., Robertson, H. A., Cooper, N. S., Brown, M., E., Rhodes, P., M., 1987. *J. Gen. Micro.* **133** : 1881-1889.

42 : Siegel, R., Ryu, D.D.Y. 1985. *Biotechnol. Bioeng.* **27** : 28-33.

1 : Ataai, M.M., Shuler, M.L. 1987. *Biotechnol. Bioeng.* **30** : 1389-1397.

Les auteurs ont proposé un modèle ségrégué où les phénomènes de ségrégation plasmidique et de réplication sont considérés explicitement. Pour cela, les auteurs ont établi un bilan pour un réacteur parfaitement mélangé en écoulement continu, à volume fini, entre les temps t et $t + \Delta t$. On considère une population où i est le nombre de plasmides par cellule. Les termes du bilan sont indiqués ci-dessous :

Accumu- -lation de cellules à i plasmides entre t et t + Δt	=	Nombre de cellules possédant au moins i plasmides et donnant par division des cellules à i plasmides	-	Nombre de cellules possédant i plasmides et donnant par division des cellules à moins de i plasmides	+	Nombre de cellules possédant moins de i plasmides, lesquels par réplication donnent des cellules à plus de i plasmides	-	Nombre de cellules à i plasmides, lesquels par réplication donnent des cellules à plus de i plasmides	-	Nombre de cellules à i plasmides qui quittent le réacteur
--	---	--	---	--	---	--	---	---	---	---

De plus, les auteurs considèrent comme étant aléatoire, la probabilité qu'une cellule mère à j plasmides, donne naissance à deux cellules ayant i plasmides et $j-i$ plasmides. Cette probabilité peut être estimée au moyen d'une distribution binomiale telle que :

$$\delta_{ji} = 2 \cdot \left(\frac{j!}{(j-i)!i!} \right) \cdot \left(\frac{1}{2} \right)^j$$

Le bilan de biomasse exprimé en nombre de cellules à i plasmides est :

$$\frac{dN(i)}{dt} = \sum_{j=i+1}^{Np \max} [\delta_{ji} \cdot \mu(j, S) \cdot N(j)] - (1.0 - 2 \cdot (0.5)^i) \cdot \mu(i, S) N(i) + R(i-1) - R(i) - DN(i)$$

Avec Np_{\max} le nombre maximal de copies de plasmides par cellule, $\delta_{ii} = 2 \cdot (0.5)^i$, D le taux de dilution.

Le bilan pour $i=0$ est :

$$\frac{dN(0)}{dt} = \sum_{i=1}^{Np \max} [2(0.5)^i \mu(i, S) N(i)] + \mu(0, S) N(0) - DN(0)$$

Le bilan en substrat est :

$$\frac{dS}{dt} = \frac{1}{Y_{N/S}} \sum_{i=0}^{Np \max} \left[\frac{dN(i)}{dt} + DN(i) \right] + (S_f - S)D$$

avec $Y_{N/S}$ le rendement biomasse sur substrat (nombre de cellule/g de substrat), S_f le substrat à l'alimentation.

De plus les termes μ et $R(i)$ sont explicités comme suit :

$$\mu(i, S) = \mu_{\max} \left(\frac{K}{K + N_p^n} \right) \left(\frac{S}{S + K_s} \right)$$

$R(i)$ est calculé à partir d'un modèle de réplication plasmidique où :

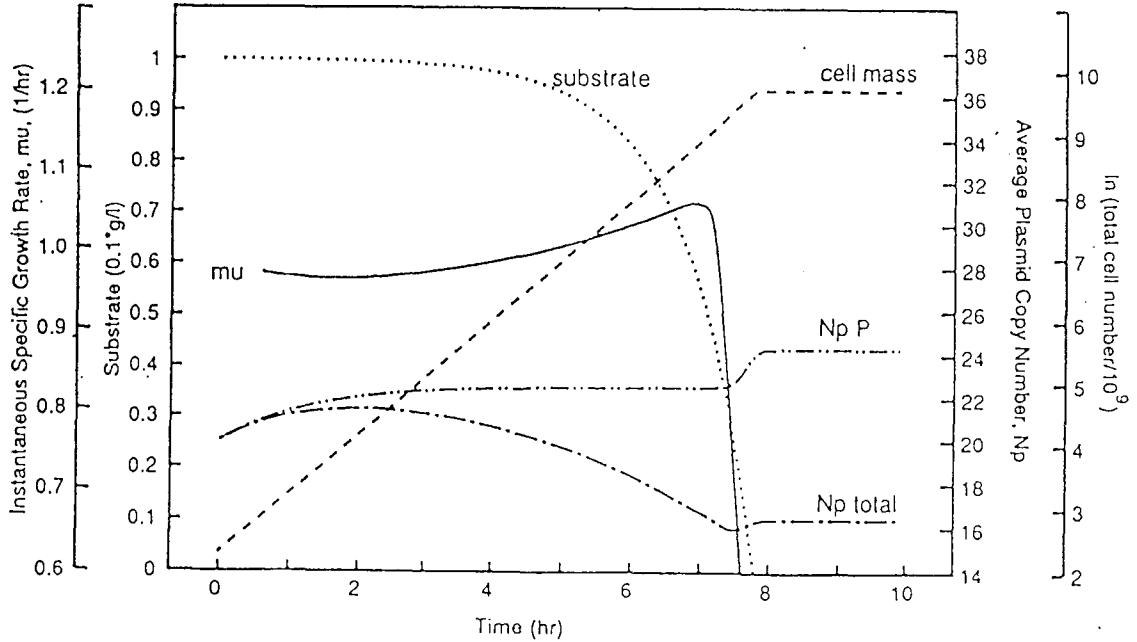
$$R(i) = Vm \left[\frac{i}{i + Km} \right] - i * C_i$$

où Vm est le taux de réplication maximum (nombre de plasmides/cellule.h),

C_i est la concentration en plasmides (nombre de plasmides par cellule),

Km la constante de demi-saturation liée au plasmide

Les auteurs ont comparé les résultats des simulations avec les résultats expérimentaux obtenus en culture batch ($D=0$) avec *E. coli* RR1(pBR329). Ce modèle permet une bonne adéquation aux résultats expérimentaux en ce qui concerne la croissance cellulaire et le nombre de copies de plasmide. Une partie des résultats modélisés est présentée dans la figure suivante :



Après 4 heures, le nombre de copies de plasmides dans les cellules contenant des plasmides reste à peu près constant, à une valeur d'environ 23 plasmides par cellule. Toutefois, le nombre de copies de plasmide dans la population totale diminue du fait de l'apparition de cellules sans plasmide. De plus, ce nombre de plasmides augmente légèrement pendant la phase stationnaire, ce qui avait déjà été observé, toutefois du fait que peu de données expérimentales existent dans ces zones autour de $\mu=0$, les auteurs n'ont pas fait d'études plus approfondies.

De plus, on remarque que la masse cellulaire augmente de façon exponentielle (linéaire en échelle logarithmique) avec un taux de croissance spécifique global constant (la pente de la droite) étant donné que : $\ln \frac{N_{tot}}{N_{tot0}} = \mu_{global} \Delta t$ avec N_{tot} = le nombre total de cellules.

Toutefois, Le taux de croissance instantané (μ sur la figure) lui, augmente en fonction du temps, μ est calculé selon la relation suivante :

$$\mu = \frac{1}{N_{tot}} \times \frac{dN_{tot}}{dt}$$

μ correspond donc à la valeur de la pente de la courbe N_{tot} en fonction du temps à un moment t donné, divisée par N_{tot} à ce moment précis. μ permet donc d'évaluer l'état métabolique d'une culture, sa valeur augmente du fait que la population évolue vers une prépondérance de cellules à croissance rapide. En effet, si on considère à un temps t donné, deux cellules à croissance rapide (à faible nombre de copies de plasmide) et deux à croissance faible (nombre de plasmides élevé), après un incrément de temps donné, les cellules à croissance rapide se seront déjà divisées, donc la relation entre les deux populations (rapide et lente) est passée de 2/2 à 4/2. Il est évident qu'une mesure de μ aura montré une augmentation. Le taux de croissance global, lui, mesuré comme étant la pente de la courbe exponentielle de croissance néglige cet aspect et représente une moyenne entre les taux de croissance des populations. Ainsi le taux de croissance instantané qui a été observé expérimentalement a pu être quantifié par cette approche ségréguée, alors qu'un modèle basé sur deux populations ne l'aurait pas permis.

Boudrant *et al.*, (1997) ont réalisé une modélisation de l'instabilité plasmidique d'une souche recombinée d'*E. coli* HB101 (GADPH) (transformée par le plasmide pBR 322 par l'insertion du gène de la GAPDH (glyceraldéhyde 3-phosphate déshydrogénase) dans le gène de la tétracycline).

Le modèle est basé sur le fait qu'une culture de cellules recombinantes peut être considérée comme une culture mixte où peuvent coexister plusieurs sortes de souches ayant de 0 à 30 plasmides. Ce modèle explicite le bilan de chacune des fractions de biomasse en différenciant la cinétique de croissance de la cinétique de réplication des plasmides. L'identification a été réalisée en culture batch au moyen d'un algorithme de type Simplex. Le modèle ainsi obtenu a aussi été confronté aux résultats en culture continue.

L'évolution de la biomasse s'exprime par l'équation suivante :

$$\frac{dX_i}{dt} = \frac{(a - ib)X_i S}{K + S} - (c + id)X_i + [c + (i - 1)d]X_{i-1} - DX_i$$

L'évolution du substrat s'exprime par :

$$\frac{dS}{dt} = -\frac{1}{Y} \sum_{i=0}^{30} \frac{(a - ib)X_i S}{K + S} + D(S_{ali} - S)$$

Y étant le rendement de conversion du substrat en biomasse, K la constante de demi-saturation et les coefficients a, b, c, d des constantes.

L'allure des courbes modélisées permet de corroborer les résultats expérimentaux.

Chapitre 2 : Matériels et Méthodes

Sommaire

1	Matériels utilisés par Lamotte (1992) pour déterminer les résultats expérimentaux	64
1.1	Présentation des souches et du vecteur	64
1.1.1	Souche hôte	64
1.1.2	Vecteur (Branlant et al, 1983)	64
1.1.3	Souche transformée	65
1.2	Milieux utilisés	65
1.3	Méthodes d'analyse	66
1.3.1	Evaluation de la biomasse	66
1.3.2	Dosage du glucose	66
1.4	Mesure de l'instabilité plasmidique par la technique des CMI (concentration minimum inhibitrice)	67
2	Modélisation	68
2.1	Démarche utilisée	68
2.2	Méthodes numériques	69
2.2.1	Intégration numérique	69
2.2.2	Identification paramétrique	70

Chapitre 2 : Matériels et méthodes

Nous avons utilisé les résultats de Lamotte (1992) comme données expérimentales de notre étude de modélisation. Les matériels utilisés par ces auteurs sont présentés ci-après.

1 Matériels utilisés par Lamotte (1992) pour déterminer les résultats expérimentaux

1.1 Présentation des souches et du vecteur

1.1.1 Souche hôte

Escherichia coli HB101 a été utilisée comme souche hôte. Cette souche a été isolée par Boyer et Roulland-Dussoix en 1969 lors d'une étude du système restriction-modification d'*E. coli*. Au cours de cette étude, une série de mutants avaient été générés dont *E. coli* HB101. Cette dernière dérivait de la souche *E. coli* HB 100 (r+m+), elle-même issue d'*E. coli* HB16 (*E. coli* K12 avec le système de restriction-modification d'*E. coli* B). *E. coli* HB101 a été l'une des premières souches capable d'accepter un ADN étranger, à être isolée.

1.1.2 Vecteur (Branlant et al, 1983)

On a utilisé un dérivé du plasmide pBR322, celui-ci possède de nombreux sites uniques reconnus par des enzymes de restriction ainsi que deux gènes permettant à la bactérie de se développer dans des milieux contenant de la tétracycline et/ou de l'ampicilline. Le plasmide utilisé : pBR Eco gap dérive de pBR322 par l'insertion du gène de la glyceraldéhyde-3-phosphate déshydrogénase dans le gène de la tétracycline.

1.1.3 Souche transformée

On a utilisé la souche *E. coli* HB101 (GADPH) transformée par le plasmide pBR Eco gap. La souche n'est plus capable de se développer dans un milieu contenant de la tétracycline, le gène codant pour la GADPH étant cloné dans le gène de résistance à la tétracycline. La résistance à l'ampicilline est toutefois conservée.

1.2 Milieux utilisés

Les milieux utilisés n'exercent aucune pression de sélection, on distingue :

- Le milieu de Luria Bertani (Milieu LB)

Ce milieu est utilisé pour mettre en culture les souches après décongélation, il est composé de :

Tryptone	10 g/l
Extrait de levure	5 g/l
NaCl	5 g/l
Glucose	2 g/l

- Le milieu complexe M2

Ce milieu a été mis au point et optimisé en 1990 par Nguyen Thi Le afin d'obtenir une quantité optimale de GADPH. Il est composé de :

Tryptone	10 g/l
Extrait de levure	10 g/l
Casaminoacides	4 g/l
NaCl	10 g/l
Glucose	4 g/l
K ₂ HPO ₄	4 g/l
KH ₂ PO ₄	1 g/l

1.3 Méthodes d'analyse

1.3.1 Evaluation de la biomasse

La biomasse est évaluée par spectrophotométrie à 660 nm. Une corrélation est établie entre les unités DO mesurées et les concentrations de bactéries en g de poids sec :

$$\text{Biomasse (g/l)} = 0.52 \cdot \text{DO}$$

Lorsque cela est nécessaire, les échantillons sont dilués de façon à obtenir des densités optiques dans la zone linéaire du spectrophotomètre (inférieures à 0.5).

1.3.2 Dosage du glucose

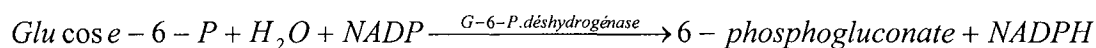
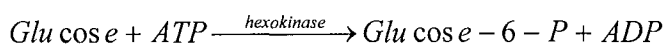
Le glucose est dosé par deux techniques différentes :

1.3.2.1 La méthode de Miller

On l'utilise pour doser des concentrations inférieures à 1 g/l. On mélange pour cela 200 µl de l'échantillon avec 300 µl de réactif de Miller. L'absorbance du milieu réactionnel est mesurée à 640 nm après chauffage à 100°C pendant 15 min et refroidissement. (Une droite d'étalonnage est réalisée avec des solutions diluées de glucose).

1.3.2.2 Dosage enzymatique sur Autoanalyseur Technicon II

Le réactif utilisé contient de l'hexokinase, de l'ATP, du NADP et de la glucose-6-phosphate déshydrogénase. Les réactions du dosage sont les suivantes :



L'apparition de NADPH est suivie en mesurant l'absorbance à 340 nm, la quantité de NADPH produite est proportionnelle à la quantité de glucose initiale pour des concentrations comprises entre 0.1 et 5 g/l. Lorsque la concentration en glucose est supérieure à cette limite, des dilutions sont nécessaires

1.4 Mesure de l'instabilité plasmidique par la technique des CMI (concentration minimum inhibitrice)

On définit la concentration inhibitrice minimum comme étant la dose minimum d'antibiotique capable d'empêcher le développement de bactéries. Cette technique est employée pour évaluer le nombre de copies de plasmide présent dans une cellule. En effet, la bactérie possède sur son plasmide un gène codant pour la β -lactamase, l'expression de ce gène est constitutive et proportionnelle au nombre de copies de plasmide. Plus le nombre de copies de plasmide est élevé et plus la bactérie sera capable de résister à des concentrations élevées en ampicilline.

Pour cela, les cellules prélevées sont diluées dans de l'eau physiologique stérile (eau distillée à 0.9 % de NaCl) de façon à obtenir 500 à 1000 bactéries par ml. Des étalements de la dilution précédente (200 μ l) sont réalisés sur milieu LB gélosé contenant différentes concentrations d'ampicilline. Les concentrations d'ampicilline choisies sont : 0.1, 0.5, 1.0, 2.0, 4.0 et 6 g/l. Des étalement sont réalisés sur milieu LB gélosé sans ampicilline. Après 12 h d'incubation à 37°C, les bactéries se développant sur les différents milieux sont comptées.

0.1 g/l est la concentration généralement utilisée pour différencier les cellules contenant des plasmides de celles qui n'en contiennent plus. 6 g/l est la concentration au-dessus de laquelle *E. coli* HB101 (GADPH) ne peut plus croître. La production de la β -lactamase est proportionnelle au nombre de plasmides jusqu'à 60 (Seo et Bailey, 1985, Dennis *et al.*, 1985). Le nombre de plasmides a été estimé à l'aide de la relation linéaire liant ce nombre à la concentration en ampicilline sachant que 6 g/l d'ampicilline correspond à 30 copies par cellule, le maximum déterminé expérimentalement dans notre étude (Balbas *et al.*, 1988).

Pour chaque concentration d'ampicilline, le pourcentage de cellules résistantes peut être calculé de la manière suivante :

$$\frac{\text{nombre} \cdot \text{de} \cdot \text{colonies} \cdot \text{sur} \cdot \text{LB} + \text{ampicilline}}{\text{nombre} \cdot \text{de} \cdot \text{colonies} \cdot \text{sur} \cdot \text{LB}} \times 100$$

Le pourcentage de cellules recombinées est obtenu en considérant l'étalement réalisé avec la concentration en ampicilline la plus faible.

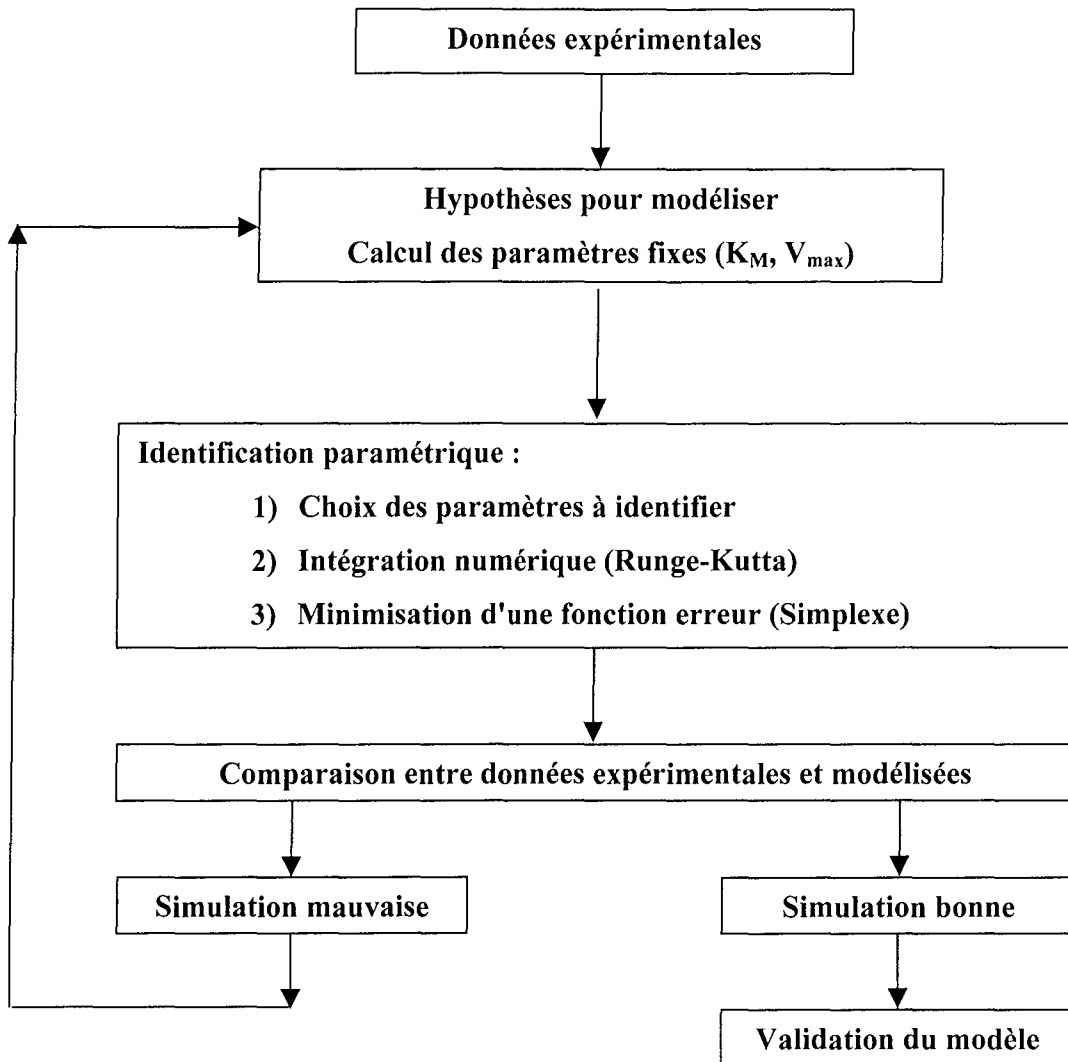
2 Modélisation

2.1 Démarche utilisée

La démarche mise en œuvre pour l'élaboration des modèles se déroule selon les étapes suivantes :

- A partir de l'étude de Lamotte (1992), nous avons émis des hypothèses sur la cinétique de croissance et de réplication bactérienne. A partir de ces dernières, nous avons établi des bilans de matière concernant notamment la biomasse et le substrat. Ces bilans sont ensuite présentés sous forme de systèmes d'équations différentielles. Les paramètres considérés comme constants dans ces équations (μ , K_s etc.) sont ensuite calculés par identification paramétrique.
- La modélisation a été réalisée sous Matlab 5.3.0. pour l'identification paramétrique, les systèmes d'équations différentielles sont intégrés par la méthode de Runge et Kutta, tandis que les paramètres à identifier sont déterminés par une fonction objectif par la méthode du Simplex.
- Les résultats de la simulation sont comparés aux données expérimentales de l'étude de Lamotte (1992).

Le schéma suivant résume les étapes de la modélisation :



2.2 Méthodes numériques

2.2.1 Intégration numérique

Les hypothèses émises pour la modélisation ont impliqué la mise en place de bilans matière sous la forme d'équations différentielles couplées. Nous avons utilisé la méthode numérique de Runge et Kutta pour la résolution de ces systèmes. Cette méthode est disponible sous Matlab avec la fonction « ode 23 » pour la méthode d'ordre 2 et « ode 45 » pour la méthode d'ordre 4.

Les méthodes de Runge et Kutta ainsi que les méthodes développées antérieurement : celles d'Euler et de Taylor figurent en **Annexe 1**.

Les avantages de la méthode de Runge et Kutta sont leur stabilité, le fait que le pas d'intégration peut être modifié sans difficulté et qu'elle « démarrent » toutes seules, la connaissance de $y(0)$ suffit à intégrer l'équation différentielle. Ces méthodes présentent l'inconvénient de générer des erreurs locales difficiles à estimer. Nous avons utilisé la méthode de Runge et Kutta d'ordre 2 dont la précision est suffisante pour notre étude et dont le temps de calcul est moins important que celle d'ordre 4.

2.2.2 Identification paramétrique

L'identification paramétrique a pour but de déterminer les valeurs de paramètres pour lesquelles le modèle simule au mieux les données expérimentales, ceci en minimisant une fonction objectif. La fonction J à minimiser dans notre calcul est :

$$j = \sum_{i=1}^n \left[m_i \times \ln \sum_{j=1}^{m_i} \left[X_i \text{ mod}(t_{ij}) - X_i \exp(t_{ij}) \right]^2 \right]$$

où :

- n est le nombre de variables mesurées, dans notre étude, il s'agira de la biomasse X , du substrat S et des fractions cumulées de biomasse contenant au moins 1, 3, 5, 10, 20 et 30 plasmides,
- m_i est le nombre d'instants de mesure, dans notre étude, nous en avons 14 pour la biomasse ainsi que le substrat et 7 pour chacune des fractions cumulées,
- $X_i \text{ mod}(t_{ij})$ est la valeur simulée de la variable X_i à l'instant t_{ij}
- $X_i \exp(t_{ij})$ est la valeur mesurée de la variable X_i à l'instant t_{ij} .

L'algorithme permettant de minimiser J est de type Simplex. Cette méthode est disponible sous Matlab avec la fonction « fminsearch ». C'est une méthode de recherche directe n'utilisant pas de gradient numérique ou analytique.

Les conditions initiales pour la recherche des valeurs des paramètres ont été évaluées à partir des travaux de Lamotte (1992).

Chapitre 3 : Résultats et Discussions

Sommaire

1	Structure des modèles.....	72
1.1	Hypothèses de base	73
1.1.1	Hypothèses de génie chimique	73
1.1.2	Hypothèses de cinétique physique	74
1.2	Développement des termes du modèle général	75
1.2.1	Bilan biomasse	75
1.2.2	Bilan en substrat	79
1.3	Modèle de probabilité binomiale.....	80
1.3.1	Définition de la probabilité P_{ji}	80
1.3.2	Bilans de biomasse détaillés pour chaque fraction de population.....	81
1.3.3	Bilan total de biomasse	82
1.4	Modèle de probabilité de répartition uniforme	82
1.4.1	Définition de la probabilité P_{ji}	82
1.4.2	Bilans de biomasse détaillés pour chaque fraction de population.....	83
1.4.3	Bilan total de biomasse	84
2	Résultats et discussions	85
2.1	Etude préliminaire	85
2.1.1	Modèle non ségrégué d'Imanaka et Aiba (1981)	85
2.1.2	Application au modèle de probabilité de répartition uniforme	88
2.2	Identification paramétrique : Modèles ségrégués.....	91
2.2.1	Paramètres obtenus.....	91
2.2.2	Culture continue	93
2.2.3	Culture discontinue	99

Chapitre 3 : Résultats et discussions

1 Structure des modèles

Dans cette partie, nous allons présenter un modèle ségrégué général. En effet, les modèles ségrégués présentent l'avantage de reconnaître explicitement la distribution des propriétés entre les individus d'une population mixte. Ils incorporent directement des informations sur les mécanismes de réplication plasmidique et de partition à l'échelle de la cellule. (Dunn et Davidson, 1995).

Notre modèle est basé sur le fait qu'une culture de cellules recombinantes peut être considérée comme une culture mixte où peuvent coexister plusieurs sortes de souches ayant de 0 à N copies de plasmides. Ce modèle explicite le bilan de chacune des fractions de biomasse et considère séparément la cinétique de croissance bactérienne et la cinétique de réplication plasmidique.

A partir de ce modèle général, nous allons expliciter deux modèles qui se distinguent par des distributions de probabilité différentes concernant la perte de plasmides. Ces modèles seront ensuite confrontés aux résultats expérimentaux.

1.1 Hypothèses de base

Nous nous sommes basés sur les conditions expérimentales de Lamotte (1992) pour réaliser notre modèle et pouvoir le confronter à leurs résultats d'expérience.

1.1.1 Hypothèses de génie chimique

Les cultures ont été réalisées en fermenteur de 2 l admettant un flux d'entrée (alimentation) et un flux de sortie. La première hypothèse revient donc à assimiler ce réacteur à un **système ouvert** c'est-à-dire susceptible d'échanger de la matière avec le milieu extérieur. La conservation de la matière exige qu'à chaque instant, s'applique pour un constituant chimique donné, la relation suivante :

$$\text{Entrée} + \text{Production} = \text{Sortie} + \text{Consommation} + \text{Accumulation}$$

NB : en réacteur fermé, les termes d'entrée et de sortie sont donc nuls.

De plus, le fermenteur est muni d'un système d'agitation. Si celui-ci est suffisamment performant, on peut supposer que la composition et l'état instantané dans le réacteur sont parfaitement uniformes. Dans ce cas, on assimile le fermenteur à un réacteur ouvert uniforme ou encore **réacteur parfaitement agité**. Dans ce cas, on considère que :

- les réactifs ajoutés sont mélangés dans un temps infiniment court,
- la composition du mélange dans le courant de sortie est à chaque instant identique à celle dans le réacteur, ce qui nous permet d'écrire pour un bilan en g/l.h :

$$\frac{Q_E C_E}{V} + r_s^+ = \frac{Q_S C_S}{V} + r_s^- + \frac{1}{V} \frac{d(V C_S)}{dt}$$

où Q est le débit volumique, C la concentration massique, V le volume réactionnel, r^+ et r^- les vitesses de consommation et de production (g/l.h), ces vitesses sont les mêmes en tout point du réacteur. Les indices S et E correspondent à l'entrée et à la sortie.

Développons les termes de sorties et d'accumulation, ils s'expriment comme il suit :

$$\frac{1}{V} \frac{d(C_S V)}{dt} + \frac{Q_S C_S}{V} = \frac{C_S}{V} \cdot \frac{dV}{dt} + \frac{dC_S}{dt} + \frac{Q_S C_S}{V}$$

De plus, on a $\frac{dV}{dt} = Q_E - Q_S$, on obtient alors :

$$\frac{(Q_E - Q_S) \cdot C_S}{V} + \frac{dC_S}{dt} + \frac{Q_S C_S}{V} = \frac{Q_E C_S}{V} + \frac{dC_S}{dt} = DC_S + \frac{dC_S}{dt}$$

On définit ainsi, le taux de dilution D (h^{-1}), en fonction de Q_E débit à l'entrée du réacteur quelle que soit la valeur de Q_S . Ce taux de dilution devient nul en réacteur fermé uniforme (batch).

Enfin nous nous intéresserons particulièrement à la marche en **régime transitoire**, c'est à dire au cas où le terme d'accumulation est non nul du fait de la variation du mélange réactionnel au cours du temps. Bien entendu, le régime pseudo-stationnaire est atteint au bout d'un certain temps de fermentation en réacteur ouvert.

1.1.2 Hypothèses de cinétique physique

Dans ce modèle ségrégué, nous considérons chaque fraction de biomasse, caractérisée par un nombre de copies de plasmides donné, comme un constituant chimique propre. Ainsi, la fermentation est le siège d'un système de réactions chimiques, basées sur la transformation de 32 constituants : les fractions de biomasse ayant 0, 1, 2... à 30 plasmides ainsi que le substrat.

De plus, chaque constituant est caractérisé par une vitesse de transformation qui est ici fonction du taux spécifique de croissance : $\mu_i X_i$ (g/l.h) pour les fractions de biomasse.

Par ailleurs, lors du développement de nos modèles, nous noterons : j le nombre de copies de plasmide contenues dans la cellule mère et i le nombre de copies de plasmide contenues dans la cellule fille. i est également le rang considéré par bilan de population en biomasse, sachant que l'on considère les populations ayant 0, 1, 2 à N plasmides. 30 est le nombre maximum de plasmides déterminés expérimentalement, nous fixerons donc ce nombre N à 30.

On considère une population de départ homogène où toutes les cellules possèdent 30 plasmides.

Enfin la cinétique de croissance bactérienne et la cinétique de réplication plasmidique seront dissociées.

1.2 Développement des termes du modèle général

Ce modèle général regroupe les deux modèles que nous allons développer par la suite, les expressions des bilans biomasse et substrat sont applicables aux deux modèles. Ces derniers se distinguent uniquement par leur probabilité de perte de plasmide. Les bilans sont exprimés en g/l.h.

1.2.1 Bilan biomasse

Le bilan de biomasse s'exprime de la façon suivante :

$$\text{Entrée} + \text{Production} = \text{Sortie} + \text{Consommation} + \text{Accumulation}$$

Comme la croissance est dissociée de la réplication, la production et la consommation sont constituées chacune de deux termes. De façon plus explicite, on a, si i est le nombre de plasmides par cellule, dans le bilan de population considéré :

0	+	Nombre de cellules possédant au moins i plasmides et donnant par division des cellules à i plasmides	+	Nombre de cellules possédant moins de i plasmides, lesquels par réplication donnent des cellules à i plasmides	=	Nombre de cellules à i plasmides qui quittent le réacteur	+	Nombre de cellules possédant i plasmides et donnant par division des cellules à moins de i plasmides	+	Nombre de cellules à i plasmides, lesquels par réplication donnent des cellules à plus de i plasmides	+	Accumu- -lation
---	---	--	---	---	---	---	---	--	---	---	---	--------------------

Notons que dans notre étude, l'ensemble de l'inoculum est introduit dès le départ, le terme d'entrée est donc nul.

1.2.1.1 Termes explicitant la division cellulaire

Dans notre modèle, nous considérons que la croissance bactérienne suit une loi de Monod, ce qui jusqu'à présent, a été classiquement admis.

Notre modèle prend en compte le phénomène d'instabilité plasmidique de type ségrégationnel. Pour cela, nous allons noter P_{ji} la probabilité de perte de plasmides où une cellule mère à j plasmides, donne naissance par division à deux cellules filles ayant respectivement i plasmides et $j-i$ plasmides, le phénomène de répllication des plasmides étant considéré de façon séparé.

Ainsi toutes les cellules à nombre de copies de plasmide au moins égal à i peuvent donner par division des cellules filles à i plasmides.

- soit $j=2$

les possibilités de répartition des plasmides dans les cellules filles sont :

- 1) 0 - 2
- 2) 1 - 1
- 3) 2 - 0

Il y a donc deux possibilités de répartition, les cas 1) et 3) étant équivalents comme les cellules ne sont pas différenciées.

- Soit $j=3$

les possibilités de répartition sont alors :

- 1) 0 - 3
- 2) 1 - 2
- 3) 2 - 1
- 4) 3 - 0

Comme les cas 1) et 4) d'une part, 2) et 3) d'autre part sont équivalents, on peut compter 2 possibilités de répartition différentes.

Ainsi, on se rend compte que pour deux entiers j pair et impair qui se suivent, le nombre de possibilités de répartition est le même.

Si on généralise le cas à j plasmides, on obtient :

$$\begin{array}{c}
 (0) - (j) \\
 (1) - (j-1) \\
 (2) - (j-2) \\
 \vdots \\
 \downarrow \\
 (j/2) - (j/2) \\
 \vdots \\
 \downarrow \\
 (j-2) - (2) \\
 (j-1) - (1) \\
 (j) - (0)
 \end{array}$$

Les différentes possibilités de répartition des plasmides dans les cellules filles allant de $(0) - (j)$ à $(j/2) - (j/2)$, les cas recensés après $(j/2) - (j/2)$ étant équivalents à ceux recensés avant.

L'examen de ces exemples simples permet de conclure que :

- le nombre total de possibilités de répartition de plasmides est $Ent\left(\frac{j+2}{2}\right)$,
- i peut prendre des valeurs allant de 0 à $j/2$, de telle façon que $\sum_{i=0}^{Ent(\frac{j}{2})} P_{ji} = 1$,
- l'utilisation de la partie entière permet de décrire correctement la probabilité quel que soit j pair ou impair.

Enfin, si on considère que chaque bilan de biomasse (de rang i , possédant i plasmides) a un taux de croissance spécifique bien déterminé, le terme de production dans le bilan est donc :

$$\sum_{j=i}^N \frac{\mu_{mj} S}{K_j + S} \cdot X_j \cdot P_{ji}$$

μ_{mj} étant le taux de croissance spécifique maximum de la souche mère à j plasmides,

N le nombre maximum de copies de plasmides par cellules que nous fixerons à 30 (chiffre déterminé expérimentalement),

K_j la constante de demi-saturation relative aux cellules mères à j plasmides,

S étant la concentration de substrat en g/l,

X_j la concentration en biomasse (cellules mères), à j copies de plasmide (g/l),

Et avec $\sum_{i=0}^{Ent(\frac{j}{2})} P_{ji} = 1$, P_{ji} étant la probabilité de perte de plasmides.

Remarque : On doit considérer le cas particulier où une cellule mère possède j plasmides tels que $j=2i$. Ainsi cette cellule donnera forcément par division deux cellules à i plasmides. Les cellules filles ont alors la même probabilité de se former et le terme de production exprimé précédemment doit être compté deux fois. Ce cas est possible lorsque j est le double de i, soit quand $j \leq N/2$. Nous tiendrons compte de ce phénomène dans le bilan global de biomasse.

De la même façon, le terme de consommation dans le bilan est :

$$\frac{\mu_{mi} S}{K_i + S} \cdot X_i$$

où l'indice i correspond au rang considéré de bilan de population.

1.2.1.2 Termes explicitant la réplication plasmidique

Soit v le taux de réplication des plasmides, dans nos modèles, nous considérerons ce taux constant. On considère par ailleurs, que chaque plasmide de la cellule a la même probabilité de se répliquer, le terme de réplication est donc directement fonction du nombre de plasmides.

Le terme de production du bilan est proportionnel à la fraction de biomasse avant réplication donnant des cellules à i plasmides, soit possédant i-1 plasmides. Ce terme est donc :

$$(i-1) \cdot v \cdot X_{i-1}$$

Le terme de consommation du bilan est proportionnel à la fraction de biomasse avant répllication, possédant déjà i plasmides. Ce terme de consommation est donc :

$$i \cdot \nu \cdot X_i$$

1.2.1.3 Bilan global de biomasse

$$\frac{dX_i}{dt} = \underbrace{\sum_{j=i}^N \frac{\mu_{mj} S}{K_j + S} \cdot X_j \cdot P_{ji}}_{\text{Production par croissance}} + \underbrace{\frac{\mu_{m2i} S}{K_{2i} + S} \cdot X_{2i} \cdot P_{(2i)i}}_{\text{Production supplémentaire quand } j=2i} - \underbrace{\frac{\mu_{mi} S}{K_i + S} \cdot X_i}_{\text{Consommation par croissance}} + \underbrace{(i-1) \cdot \nu \cdot X_{i-1}}_{\text{Production par répllication}} - \underbrace{i \cdot \nu \cdot X_i}_{\text{Consommation par répllication}} - \underbrace{DX_i}_{\text{Terme de sortie}}$$

Le terme de rang $j=2i$ n'est comptabilisé que lorsque $i \leq N/2$.

Ce bilan de biomasse sera appliqué pour nos deux modèles ségrégués, la différence entre ces modèles portant essentiellement sur la définition de la probabilité de perte de plasmides P_{ji} .

1.2.2 Bilan en substrat

Le bilan en substrat est le même que pour le modèle d'Imanaka et Aiba (1981) à deux équations différentielles :

$$\frac{dS}{Dt} = D(S_{ali} - S) - \frac{\mu_p X_p}{y_p} - \frac{\mu_0 X_0}{y_0}$$

Où y_p et y_0 sont les rendements de conversion du substrat en biomasse respectivement pour les populations avec et sans plasmide et S_{ali} la concentration en substrat à l'alimentation (g/l).

1.3 Modèle de probabilité binomiale

Dans ce modèle, la probabilité P_{ji} correspond à un cas de figure où chacune des cellules filles issues de la division, tirerait au sort les plasmides contenus dans la cellule mère, et ceci à tour de rôle jusqu'à épuisement du patrimoine plasmidique.

1.3.1 Définition de la probabilité P_{ji}

Nous allons définir un modèle ségrégué, caractérisé par une probabilité de perte de plasmide suivant une loi de type binomial. Cette probabilité a été définie par Bentley et Quiroga en 1992.

D'après ces auteurs, P_{ji} est la probabilité qu'une cellule mère à j plasmides, donne naissance à deux cellules ayant i plasmides et $j-i$ plasmides. Cette probabilité peut être estimée au moyen d'une distribution binomiale telle que :

$$P_{ji} = 2 \cdot \left(\frac{j!}{(j-i)! \times i!} \right) \cdot \left(\frac{1}{2} \right)^j \text{ Cette probabilité tient compte du cas où } j=2i.$$

Pour la démontrer, notons P_{ji}^* la probabilité pour qu'une cellule à j plasmides donne naissance à une cellule à i plasmides. Quels que soient i et j , et si la distribution est binomiale, on a :

$$P_{ji}^* = \frac{j!}{(j-i)! \times i!} \cdot \left(\frac{1}{2} \right)^j$$

En reprenant l'expression du bilan global de biomasse décrit au 1.2.1.4. on a :

$$\frac{dX_i}{dt} = \underbrace{\sum_{j=i}^N \frac{\mu_{mj} S}{K_j + S} \cdot X_j \cdot \left[\frac{j!}{(j-i)! \times i!} \cdot \left(\frac{1}{2} \right)^j \right]}_{\text{Production}} + \underbrace{\frac{\mu_{m2i} S}{K_{2i} + S} \cdot X_{2i} \cdot \left[\frac{j!}{i! \times i!} \cdot \left(\frac{1}{2} \right)^j \right]}_{\text{Production supplémentaire quand } j=2i} - \frac{\mu_{mi} S}{K_i + S} \cdot X_i + (i-1) \cdot \nu \cdot X_{i-1} - i \cdot \nu \cdot X_i - DX_i$$

En sommant les deux termes de production, on obtient :

$$\frac{dX_i}{dt} = \sum_{j=i}^N \frac{\mu_{mj} S}{K_j + S} \cdot X_j \cdot \left[2 \cdot \frac{j!}{(j-i)! i!} \cdot \left(\frac{1}{2} \right)^j \right] - \frac{\mu_{mi} S}{K_i + S} \cdot X_i + (i-1) \cdot \nu \cdot X_{i-1} - i \cdot \nu \cdot X_i - DX_i$$

On retrouve donc la probabilité P_{ji} décrite par Bentley et Quiroga.

A partir de cette expression générale, détaillons les bilans de biomasse en fonction de i.

1.3.2 Bilans de biomasse détaillés pour chaque fraction de population

1.3.2.1 Pour i = 0

Pour i=0, il n'y a pas de terme de réplication plasmidique.

$$\frac{dX_0}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} X_0 + \frac{\mu_p \cdot S}{K_p + S} \sum_{j=1}^N 2 \times 0.5^j X_j - DX_0$$

1.3.2.2 Pour i = 1

$$\frac{dX_1}{dt} = \frac{\mu_p \cdot S}{K_p + S} \cdot \left[\sum_{j=2}^N (2 \cdot j \cdot 0.5^j \cdot X_j) \right] - \nu \cdot X_1 - DX_1$$

1.3.2.3 Pour 1 < i < N

$$\frac{dX_i}{dt} = \frac{\mu_p \cdot S}{K_p + S} \cdot \left[-X_i + \sum_{j=i}^N 2 \cdot \left(\frac{j!}{(j-i)! i!} \right) \cdot 0.5^j X_j \right] + (i-1) \cdot \nu \cdot X_{i-1} - i \cdot \nu \cdot X_i - DX_i$$

1.3.2.4 Pour i = N

Quand i = N, il n'y pas de terme de production par réplication plasmidique.

$$\frac{dX_N}{dt} = \frac{\mu_p \cdot S}{K_p + S} \cdot \left[(2 \times 0.5^N - 1) \cdot X_N \right] + (N-1) \cdot \nu \cdot X_{N-1} - DX_N$$

1.3.3 Bilan total de biomasse

Le bilan de biomasse total correspond à :

$$\frac{d(X_0 + X_p)}{dt} = \underbrace{\sum_{j=0}^N 2 \cdot \frac{\mu_{mj} \cdot S}{K_j + S} \cdot X_j}_{\text{Apparition par division des cellules mères à j plasmides donnant 2 cellules filles}} - \underbrace{\sum_{i=0}^N \frac{\mu_{mi} \cdot S}{K_i + S} \cdot X_i}_{\text{Disparition par division des cellules à i plasmides}} - \underbrace{D(X_0 + X_p)}_{\text{Terme de sortie du réacteur}}$$

La somme de tous les bilans donne le bilan de biomasse total suivant :

$$\frac{d(X_0 + X_p)}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} \cdot X_0 + \frac{\mu_p \cdot S}{K_p + S} \cdot X_p - D(X_0 + X_p)$$

1.4 Modèle de probabilité de répartition uniforme

Dans ce modèle, la probabilité P_{ji} correspond à un cas de figure où la cellule en division verrait son patrimoine plasmidique partagé « d'un coup » en deux blocs, chacun des blocs allant aux cellules filles. Le nombre de plasmides contenus dans chacun des blocs est aléatoire et varie donc de 0 à j .

1.4.1 Définition de la probabilité P_{ji}

D'après le modèle d'Imanaka et Aiba (1981), on note P , la probabilité d'avoir une cellule fille à 0 plasmide à partir d'une cellule à j plasmides.

Ainsi $P_{j0} = P_{jj} = P$

Le nombre total de possibilités de répartition de j plasmides dans les cellules filles est :

$$1 + Ent\left(\frac{j}{2}\right) = Ent\left(\frac{j+2}{2}\right)$$

L'ensemble des probabilités de répartition des plasmides différentes de $P_{j0} = P_{jj} = P$ sera donc $(1-P)$. Si toutes les répartitions de plasmides autres que $P_{j0} = P_{jj} = P$ sont des probabilités égales,

alors la probabilité P_{ji} sera : $\frac{1-P}{Ent(\frac{j+2}{2}-1)} = \frac{1-P}{Ent(\frac{j}{2})}$

En résumé :

Si $j \neq 0$, on a : $P_{j0} = P_{jj} = P$

Notons que $P_{00}=P_{10}=P_{11}=1$

Et pour tout i différent de 0 et j , on a : $P_{ji} = \frac{1-P}{Ent(\frac{j}{2})}$

On retrouve bien $\sum_{i=0}^{Ent(\frac{j}{2})} P_{ji} = 1$ tel que défini dans le modèle global.

1.4.2 Bilans de biomasse détaillés pour chaque fraction de population

Comme dans le modèle de Imanaka et Aiba (1981), on considère pour les paramètres μ_m et K_s , deux valeurs en fonction du nombre de plasmides :

- $\mu_{mj} = \mu_0$ pour $j=0$ et $\mu_{mj} = \mu_p$ pour tout $j>0$
- $K_j = K_0$ pour $j=0$ et $K_j = K_p$ pour tout $j>0$
- $X_p = \sum_{j=1}^N X_j$

1.4.2.1 Pour $i=0$

Pour $i=0$, il n'y a pas de terme de répllication plasmidique.

$$\frac{dX_0}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} X_0 + (1-P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_1 + P \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_p - DX_0$$

1.4.2.2 pour $i=1$

$$\frac{dX_1}{dt} = (1-P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot \left[X_2 + \sum_{j=2}^N \frac{X_j}{Ent(\frac{j}{2})} \right] - \nu \cdot X_1 - DX_1$$

1.4.2.3 Pour $1 < i \leq N/2$

$$\frac{dX_i}{dt} = (1-P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot \left[-X_i + \frac{X_{2i}}{i} + \sum_{j=i+1}^N \frac{X_j}{Ent(\frac{j}{2})} \right] + (i-1) \cdot \nu \cdot X_{i-1} - i \cdot \nu \cdot X_i - DX_i$$

Nb : le terme de production par division cellulaire où $j=2i$ est bien compté deux fois. (Cf. 1.2.1.1. termes explicitant la division cellulaire).

1.4.2.4 Pour $N/2 < i < N$

$$\frac{dX_i}{dt} = (1-P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot \left[-X_i + \sum_{j=i+1}^N \frac{X_j}{Ent(\frac{j}{2})} \right] + (i-1) \cdot \nu \cdot X_{i-1} - i \cdot \nu \cdot X_i - DX_i$$

1.4.2.5 Pour $i = N$

Quand $i=N$, il n'y a pas de terme de production par réplication plasmidique.

$$\frac{dX_N}{dt} = (1-P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot (-X_N) + (N-1) \cdot \nu \cdot X_{N-1} - DX_N$$

1.4.3 Bilan total de biomasse

La somme de tous les bilans redonne le bilan de biomasse total trouvé également avec le modèle binomial:

$$\frac{d(X_0 + X_p)}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} \cdot X_0 + \frac{\mu_p \cdot S}{K_p + S} \cdot X_p - D(X_0 + X_p)$$

2 Résultats et discussions

2.1 Etude préliminaire

2.1.1 Modèle non ségrégué d'Imanaka et Aiba (1981)

Nous avons réalisé une identification paramétrique en utilisant le modèle simple d'Imanaka et Aiba afin de déterminer son niveau d'adéquation par rapport aux données expérimentales.

L'identification paramétrique s'est faite à partir d'une culture continue avec un taux de dilution de $D=0.55 \text{ h}^{-1}$ et une concentration de glucose à l'alimentation de 9.5 g/l . La mise en régime continu a été précédée d'une étape discontinue qui a duré $8\text{h}15$.

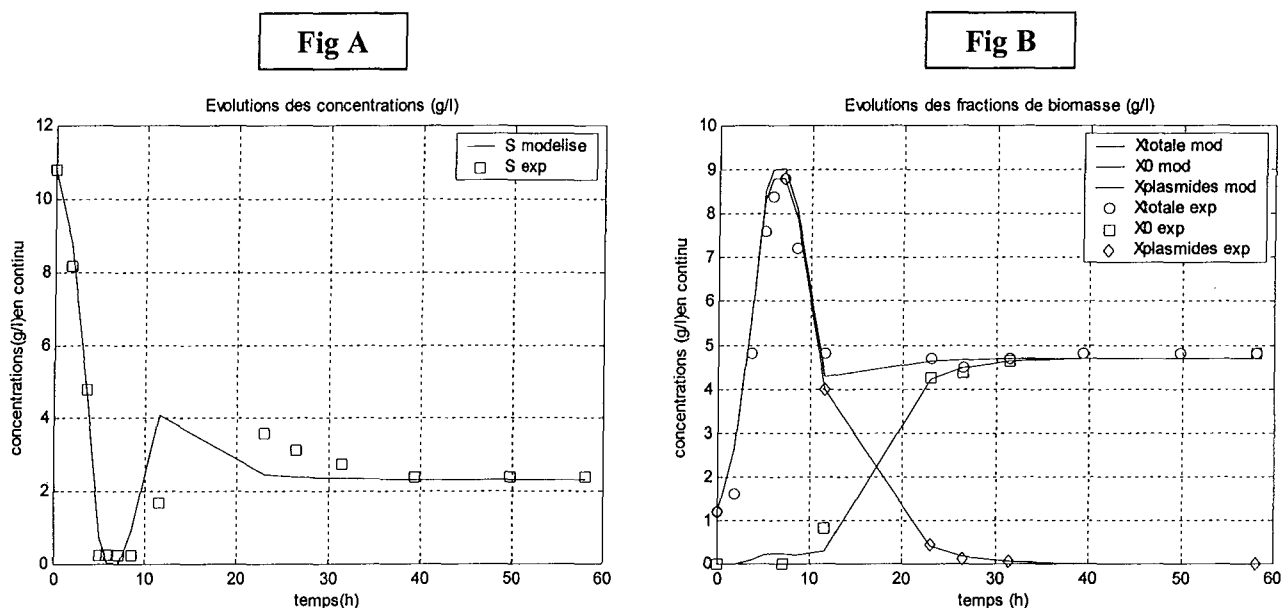
On a obtenu une valeur très faible de la fonction de minimisation $J = -15.6616$. Le modèle permet donc de bien corroborer les résultats d'expérience.

Les paramètres obtenus sont les suivants :

Paramètre	$\mu_0 \text{ (h}^{-1}\text{)}$	$\mu_p \text{ (h}^{-1}\text{)}$	$K_0 \text{ (g/l)}$	$K_p \text{ (g/l)}$	p	$y_0 \text{ (g/g)}$	$y_p \text{ (g/g)}$
Signification	Taux spécifique de croissance des cellules sans plasmides	Taux spécifique de croissance des cellules avec plasmides	Constante de demi-saturation des cellules sans plasmides	Constante de demi-saturation des cellules avec plasmides	Probabilité	Rendement de conversion du substrat en biomasse des cellules sans plasmides	Rendement de conversion du substrat en biomasse des cellules avec plasmides
valeur	3.528	0.503	12.648	1.468	0.0014	0.728	0.654

Les paramètres déterminés sont proches de ceux donnés dans la littérature.

Les graphiques suivants présentent les évolutions du modèle et des résultats d'expérience concernant le substrat ainsi que des fractions de biomasse avec et sans plasmides



Evolutions des concentrations expérimentales et simulées avec le modèle d'Imanaka et Aiba (1981) en culture continue. **Fig. A** : Evolution du substrat, **Fig. B** : Evolution de la biomasse totale X_{totale} , de la biomasse sans plasmides X_0 et de la biomasse contenant des plasmides X_p .

Les tableaux de valeurs des concentrations sont regroupés en **Annexe 2**.

Les valeurs des paramètres obtenus en simulation sont proches de celles données par la littérature. Notons que le taux de croissance spécifique des cellules sans plasmides $\mu_0 = 3.528 \text{ h}^{-1}$ est plus élevé que celui des cellules contenant des plasmides $\mu_p = 0.503 \text{ h}^{-1}$. Ceci avait également été remarqué par Boudrant *et al.* (1996) pour lesquels les plasmides agiraient comme des « fardeaux », diminuant le taux de croissance cellulaire.

La concentration expérimentale en substrat diminue pendant la phase batch jusqu'à une valeur de 0.24 g/l à partir de la 5^{ème} heure de fermentation. Elle augmente après l'introduction de l'alimentation pour atteindre un pic de 3.6 g/l à la 23^{ème} heure pour ensuite se stabiliser à 2.4 g/l.

La simulation permet de bien rendre compte de ces variations, on observe toutefois que le pic de substrat est décalé par rapport aux résultats d'expérience : la concentration maximale est de 4.07 g/l après 11h30 de fermentation.

Pendant la phase batch, la biomasse expérimentale augmente fortement (phase de croissance exponentielle) puis chute du fait de l'appauvrissement puis de la disparition du substrat dans le milieu. La mise en place de l'alimentation en phase continue (après 8h15) permet alors de stabiliser la biomasse à une valeur de 4.8 g/l. La fraction de biomasse contenant des plasmides évolue pendant la phase de croissance exponentielle comme la biomasse totale, en effet, toutes les cellules possèdent des plasmides au début de la culture. On remarque que cette fraction diminue de la même façon que la biomasse totale lors de l'épuisement du substrat mais continue au-delà de 11h30 à diminuer jusqu'à s'annuler. Parallèlement à ce phénomène, la biomasse sans plasmide, nulle au départ, augmente pour aller rejoindre la courbe de biomasse totale. Cela montre que l'ensemble de la biomasse contenant des plasmides a été entièrement remplacée à partir de la 30^{ième} heure de fermentation par de la biomasse sans plasmide. Ceci rend compte du phénomène d'instabilité plasmidique.

Les simulations décrivent bien ces évolutions.

De ces résultats, nous pouvons conclure que le modèle non ségrégué d'Imanaka et Aiba permet de corroborer de façon satisfaisante les données expérimentales.

2.1.2 Application au modèle de probabilité de répartition uniforme

Le modèle ségrégué de probabilité de répartition uniforme peut être ramené au modèle d'Imanaka et Aiba.

En effet, si on additionne toutes les équations correspondant aux fractions de biomasse de 1 à 30, on obtient l'équivalent de la fraction contenant des plasmides X^+ . Notons X_p cette fraction résultant d'une somme. On a pour X_0 et X_p :

$$\begin{cases} \frac{dX_0}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} X_0 + (1 - P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_1 + P \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_p - DX_0 \\ \frac{dX_p}{dt} = -(1 - p) \frac{\mu_p S}{K_p + S} X(1) + (1 - p) \frac{\mu_p S}{K_p + S} X_p - DX_p \end{cases} \quad (1)$$

Dans le modèle d'Imanaka et Aiba, on a :

$$\begin{cases} \frac{dX^-}{dt} = \mu^- X^- + p\mu^+ X^+ - DX^- \\ \frac{dX^+}{dt} = \mu^+ X^+ - p\mu^+ X^+ - DX^+ \end{cases}$$

Les indices correspondants entre ces deux systèmes d'équations étant :

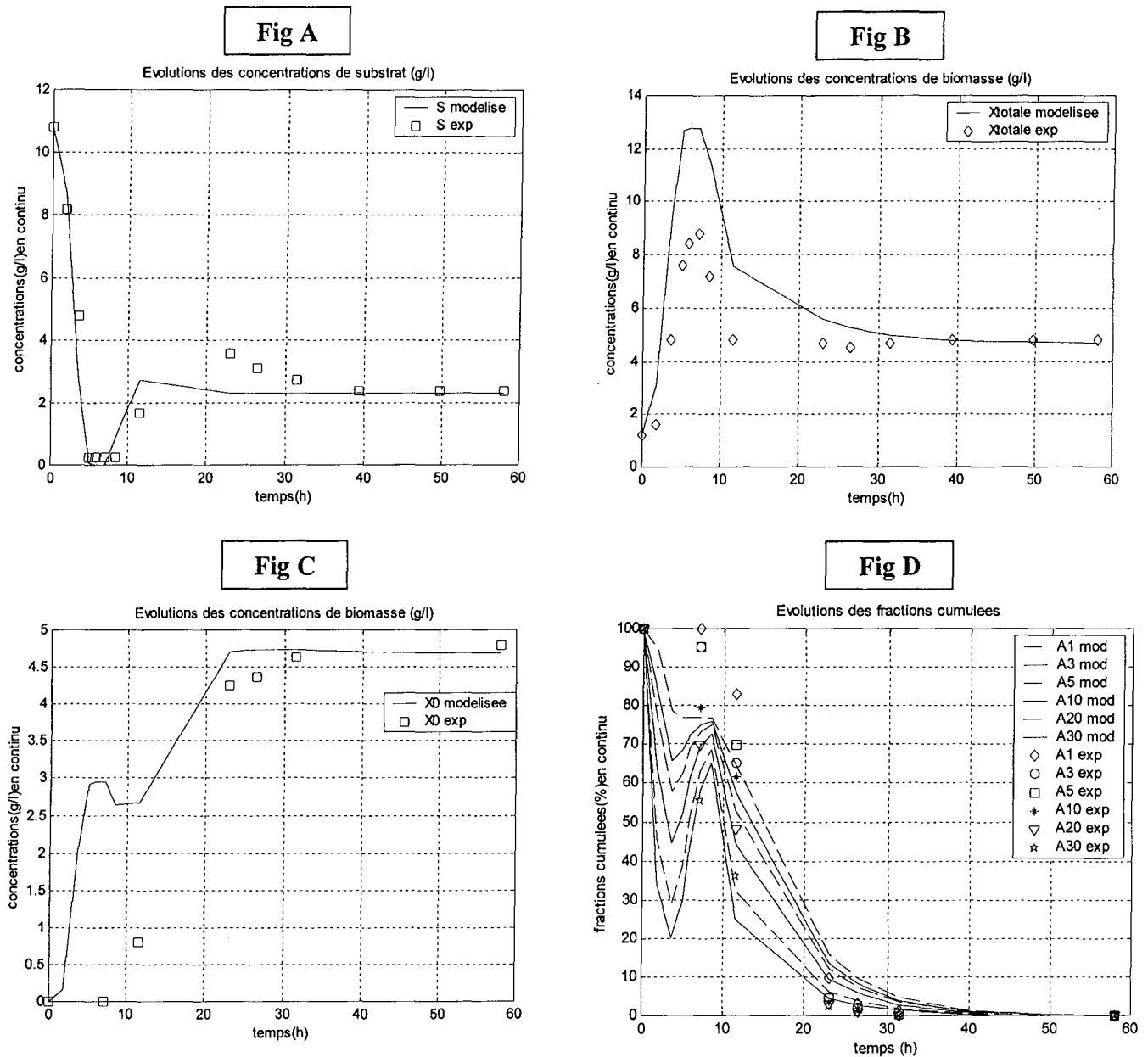
- 0 et –
- p et +

On se rend donc compte qu'à un terme fonction de $X(1)$ près, notre modèle permet de retrouver le modèle non ségrégué d'Imanaka et Aiba. Nous allons donc appliquer au modèle de probabilité de répartition uniforme, les paramètres déterminés au moyen du modèle non ségrégué.

Ceci permettra de déterminer l'importance du rôle du terme en $X(1)$ sur les évolutions de concentrations.

Par ailleurs, notons que la probabilité de perte de plasmides est très différente dans le modèle de probabilité binomiale, son expression ne permet pas de retrouver celle du modèle d'Imanaka et Aiba par réarrangements. C'est pourquoi nous ne chercherons pas à appliquer les paramètres déjà déterminés sur le modèle de probabilité de distribution binomiale.

Les graphiques suivants présentent les évolutions du modèle et des résultats d'expérience concernant les concentrations et les fractions cumulées.



Evolutions des valeurs expérimentales et simulées avec le modèle de répartition uniforme en culture continue. Les paramètres ont été obtenus par identification paramétrique à partir du modèle d'Imanaka et Aiba (1981). Fig A : Evolution du substrat S, Fig B : Evolution de la biomasse totale X_{totale} , Fig C : Evolution de la biomasse sans plasmides X_0 , Fig D : Evolution des fractions cumulées A_i .

Les valeurs correspondant aux graphiques ci-dessus figurent en **Annexe 3**.

Le taux de réplication ν n'est pas pris en compte dans le modèle d'Imanaka et Aiba, pour réaliser la simulation ci-dessus nous avons pris fixé sa valeur à 0.75 h^{-1} (données de la littérature).

Les valeurs expérimentales du substrat (Fig. A) sont bien simulées. En revanche, on remarque que la biomasse sans plasmide modélisée (Fig. C) : X_0 présente un pic à 2.94 g/l à 6-7 h contrairement aux valeurs expérimentales. Or, dans notre modèle, l'expression de X_0 est :

$$\frac{dX_0}{dt} = \frac{\mu_0 \cdot S}{K_0 + S} X_0 + (1 - P) \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_1 + P \cdot \frac{\mu_p \cdot S}{K_p + S} \cdot X_p - DX_0$$

alors qu'elle est égale dans le modèle d'Imanaka et Aiba à :

$$\frac{dX^-}{dt} = \mu^- X^- + p\mu^+ X^+ - DX^-$$

Seul le terme en $X(1)$ différentie ces deux équations. Ainsi, du fait de ce terme, on observe une différence importante entre les simulations des deux modèles. Ce terme est donc loin d'être négligeable. De même, on remarque que la biomasse totale modélisée (Fig. B) présente un pic vers 6-7 h que l'on ne retrouve pas avec les valeurs expérimentales, ce pic est aussi dû au terme en $X(1)$.

Enfin, la distribution des fractions cumulées (Fig. D) est beaucoup plus resserrée pour les simulations que pour les données expérimentales. Si les fractions à fort nombre de plasmides (A30 et A20) sont relativement bien modélisées, on observe en revanche un décalage vers le bas des fractions A1, A3, A5 et A10 par rapport aux valeurs d'expérience. En effet, les fractions cumulées sont définies par : $A_j = \sum_{i=j}^{30} \alpha_i$, α_i étant la fraction de biomasse à i plasmides telle que :

$$\alpha_i = 100 \cdot \frac{X_i}{X_T}$$

Ainsi comme la valeur de la fraction de biomasse sans plasmide X_0 est plus importante dans notre modèle (du fait du terme en $X(1)$), elle entraîne une valeur de biomasse totale plus importante et donc une diminution des valeurs des fractions cumulées. Plus on considère de fractions de biomasse et plus ce phénomène est accentué, c'est pourquoi les fractions cumulées A30 et A20 (qui ne prennent en compte que les fractions de biomasse à 30 et 20 plasmides) sont peu affectées par ce phénomène.

Ainsi, les paramètres déterminés à partir du modèle d'Imanaka et Aiba ne permettent pas de simuler un modèle ségrégué plus complet. Il est donc nécessaire de procéder à des identifications paramétriques sur ce modèle ségrégué. Ceci est décrit dans les paragraphes suivants.

2.2 Identification paramétrique : Modèles ségrégués

2.2.1 Paramètres obtenus

L'identification paramétrique a été réalisée à partir d'une culture continue avec un taux de dilution de $D=0.55 \text{ h}^{-1}$ et une concentration de glucose à l'alimentation de 9.5 g/l . La mise en régime continu a été précédée d'une étape discontinue qui a duré 8h15.

Les paramètres obtenus pour les deux modèles sont présentés dans le tableau suivant :

Paramètres	Signification	Modèle de probabilité de répartition uniforme	Modèle de probabilité binomiale
$\mu_0 \text{ (h}^{-1}\text{)}$	Taux spécifique de croissance des cellules sans plasmides	5.813	1.778
$\mu_p \text{ (h}^{-1}\text{)}$	Taux spécifique de croissance des cellules avec plasmides	0.370	0.231
$K_0 \text{ (g/l)}$	Constante de demi-saturation des cellules sans plasmides	28.697	7.255
$K_p \text{ (g/l)}$	Constante de demi-saturation des cellules avec plasmides	1.686	1.033
p	Probabilité	-0.016	NA
$v \text{ (h}^{-1}\text{)}$	Taux de réplication plasmidique	0.789	0.266
$y_0 \text{ (g/g)}$	Rendement substrat / biomasse sans plasmides	0.455	0.292
$y_p \text{ (g/g)}$	Rendement substrat / biomasse avec plasmides	0.710	0.123

Nous avons obtenus les valeurs suivantes pour la fonction J de minimisation :

	Modèle de probabilité de répartition uniforme	Modèle de probabilité binomiale
Valeur de J	152.76	246.96

Tout d'abord, au vu des valeurs de J, nous pouvons conclure que le modèle de probabilité de répartition uniforme se rapproche davantage des valeurs expérimentales que le modèle de probabilité binomiale. Une valeur de p constante permettrait donc de mieux simuler le comportement des souches recombinées qu'une valeur de p variant avec le nombre de plasmides.

Toutefois, ce sont les valeurs des paramètres obtenus qui permettront de mieux apprécier la validité du modèle. On remarque d'abord que les paramètres des deux modèles présentent le même ordre de grandeur, à l'exception de la valeur obtenue pour μ_0 où le modèle de probabilité de répartition uniforme suggère un écart plus important entre les vitesses de croissance des souches avec et sans plasmide. Il est normal que les valeurs des paramètres soient proches du fait que seule l'expression de la probabilité p varie entre ces deux modèles.

Notons que nous avons obtenu une valeur négative de p pour le modèle de probabilité de répartition uniforme, ainsi, il ne nous sera pas possible de considérer ce modèle comme ayant un sens physique relativement aux phénomènes d'évolution des souches recombinantes. Dans le modèle de probabilité binomiale, la probabilité est calculée en fonction du nombre de plasmides et reste par conséquent strictement positive.

Les paragraphes suivants présentent les résultats expérimentaux ainsi que les résultats simulés obtenus avec les paramètres calculés.

2.2.2 Culture continue

2.2.2.1 Evolution de la biomasse totale et du substrat

Les concentrations expérimentales et les simulations de biomasse et de substrat en culture continues sont présentées sur les graphiques suivants. Les valeurs correspondantes sont regroupées en **Annexe 4**.

Fig A

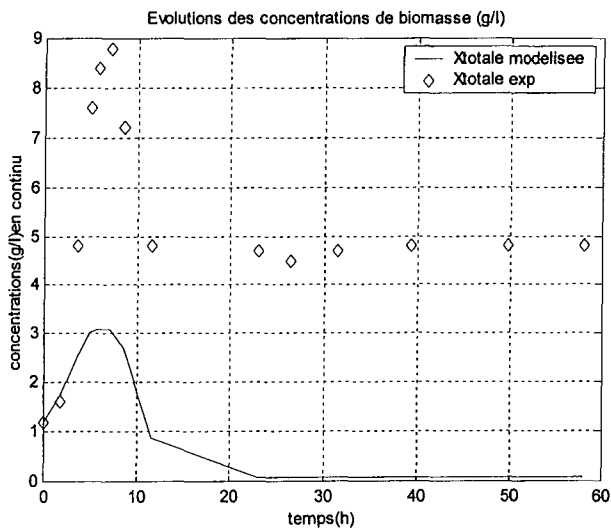


Fig C

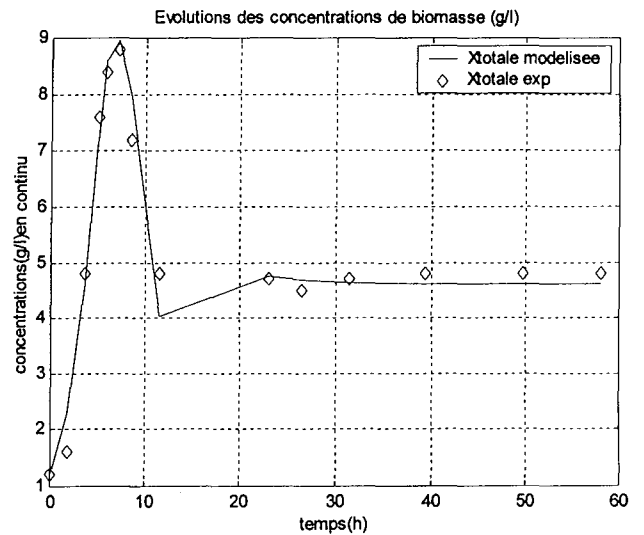


Fig B

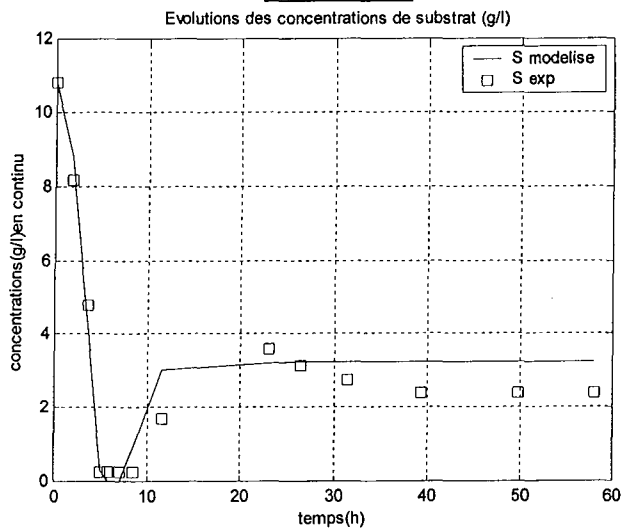
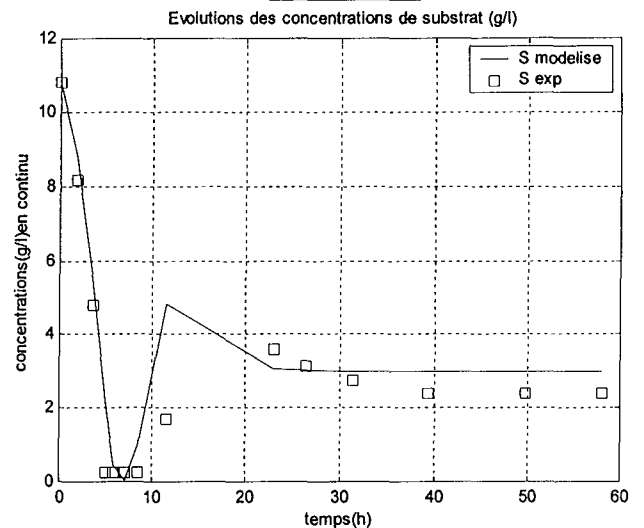


Fig D



Evolution des concentrations expérimentales et simulées avec le modèle de probabilité binomiale en culture continue. Fig A : Evolution de la biomasse X_{totale} , Fig B : Evolution du substrat S

Evolution des concentrations expérimentales et simulées avec le modèle de probabilité de répartition uniforme en culture continue. Fig C : Evolution de la biomasse X_{totale} , Fig D : Evolution du substrat S

Pendant la phase batch, la biomasse expérimentale augmente fortement (phase de croissance exponentielle) puis chute du fait de l'appauvrissement puis de la disparition du substrat dans le milieu. La mise en place de l'alimentation en phase continue (après 8h15) permet alors de stabiliser la biomasse à une valeur de 4.8 g/l.

Le modèle de probabilité binomiale ne permet pas de rendre compte de l'évolution de la biomasse, la simulation présente la même allure que la courbe expérimentale mais les concentrations simulées sont trop faibles.

Par contre, le modèle de probabilité de répartition uniforme la simule correctement avec toutefois, un léger décrochement avant l'établissement du régime stationnaire (après 11 h 30).

La concentration expérimentale en substrat diminue pendant la phase batch jusqu'à une valeur de 0.24 g/l à partir de la 5^{ème} heure de fermentation. Elle augmente après l'introduction de l'alimentation pour atteindre un pic de 3.6 g/l à la 23^{ème} heure pour ensuite se stabiliser à 2.4 g/l.

Le modèle de probabilité de répartition uniforme simule correctement les données expérimentales, toutefois, le pic de substrat semble trop important et être décalé à 11 h 30. De plus la valeur en régime stationnaire est légèrement trop élevée (2.99 au lieu de 2.4 g/l).

Le modèle de probabilité binomiale ne présente pas de pic ainsi le régime stationnaire est établi plus tôt à 11h30. On remarque que la concentration en substrat est beaucoup plus proche des données expérimentales que l'autre modèle.

2.2.2.2 Evolution des fractions cumulées de biomasse

Les valeurs expérimentales des fractions cumulées ayant au moins 1, 3, 5, 10, 20 et 30 plasmides ont été mesurées.

Si on note X_T la biomasse totale, on a $X_T = \sum_{i=0}^{i=30} X_i$

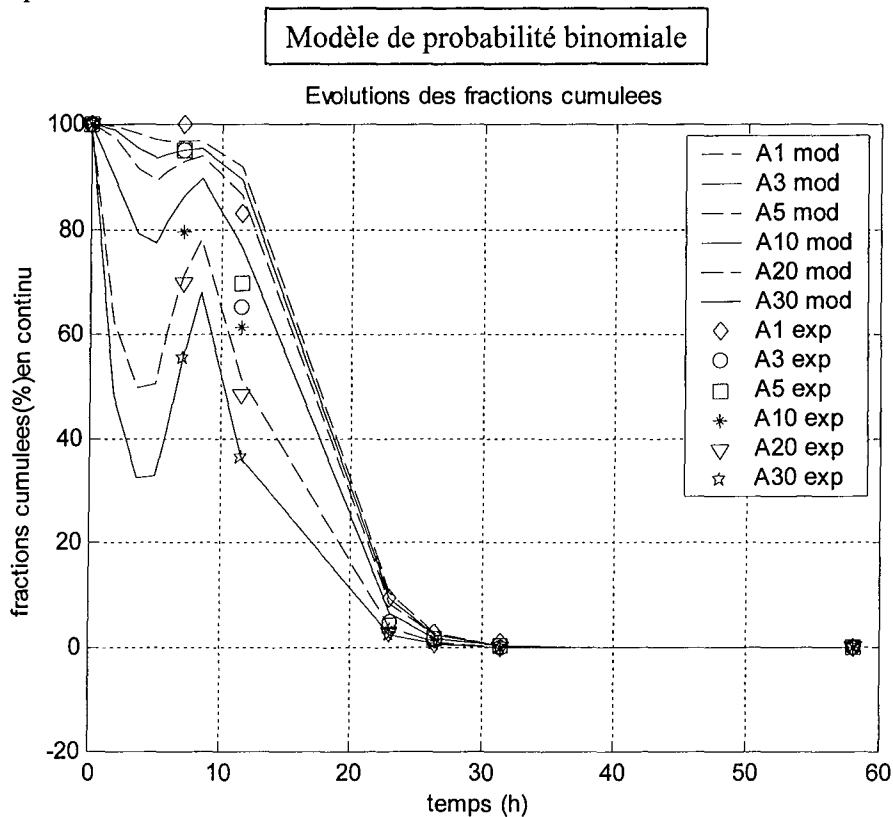
La fraction de biomasse à i plasmides est : $\alpha_i = 100 \cdot \frac{X_i}{X_T}$

La fraction cumulée de biomasse contenant au moins j plasmides est définie par :

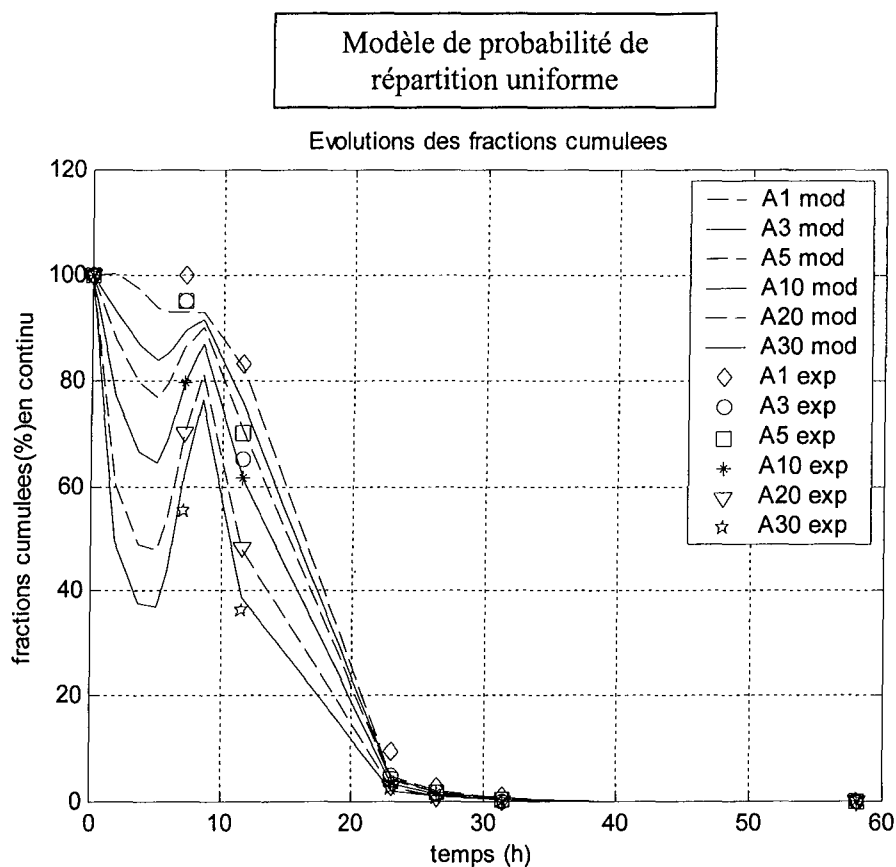
$$A_j = \sum_{i=j}^{30} \alpha_i$$

Ainsi la fraction cumulée à 30 plasmides est égale à la fraction de biomasse à 30 plasmides

Les fractions cumulées expérimentales et simulées pour les deux modèles sont présentées sur les graphiques suivants :



Evolutions des valeurs expérimentales et simulées avec le modèle de probabilité binomiale en culture continue. Evolution des fractions cumulées A_i .



Les valeurs correspondant à ces graphiques sont regroupées en **Annexe 5**.

En ce qui concerne le modèle de probabilité binomiale, les simulations des fractions cumulees A30 et A20 corroborent parfaitement les résultats expérimentaux. Par contre, la fraction A10 leur est légèrement supérieure. La fraction expérimentale A5 se situe de part et d'autre des courbes simulées. Enfin la fraction A1 simulée chute légèrement plus rapidement que les valeurs expérimentales alors que la fraction A3 présente des simulations légèrement supérieures aux valeurs expérimentales aux environs de 11h30.

Les fractions cumulees A30, A20 et A10 du modèle de probabilité de répartition uniforme corroborent parfaitement les résultats expérimentaux. La fraction cumulée expérimentale A3 se situe de part et d'autre de la courbe simulée. Les simulations des fractions cumulees A1 et A5

sont légèrement trop basses par rapport aux résultats d'expérience et ce jusqu'à 11h30 de fermentation.

Les deux modèles sont proches des valeurs expérimentales, le modèle de probabilité de répartition uniforme s'en rapprochant tout de même davantage. C'est pourquoi les explications suivantes concernant l'évolution des fractions cumulées s'appliquera aussi bien aux simulations qu'aux données expérimentales.

On constate qu'à partir de la population homogène initiale, les fractions cumulées (populations possédant des plasmides) chutent très fortement au départ, ceci correspond à la phase exponentielle de croissance bactérienne. Les bactéries se multiplient alors beaucoup plus rapidement que les plasmides ne se répliquent. Une ségrégation plasmidique apparaît alors, expliquant la chute de toutes les fractions plasmidiques. Ceci a également observé cette fois-ci uniquement expérimentalement par Lamotte, Gschaedler et Boudrant (1994). Ces auteurs ont en effet montré une diminution transitoire de la protéine recombinée (directement liée au nombre de plasmides) par rapport à la biomasse pendant la phase exponentielle de croissance.

Les fractions cumulées subissent dans un deuxième temps une augmentation qui va de pair avec une diminution de la biomasse totale, laquelle est causée par l'appauvrissement en substrat dans le milieu. Cette diminution de la vitesse de croissance bactérienne permet de réduire l'écart entre la vitesse de croissance des bactéries et celle de la répllication des plasmides. Ainsi le phénomène de ségrégation plasmidique diminue, ce qui explique la remontée des courbes de fractions cumulées.

Enfin, peu après la mise en régime continu (aux environs de 9 h 00) toutes les fractions cumulées chutent jusqu'à s'annuler traduisant ainsi le phénomène de perte de plasmides au niveau de toutes les fractions de biomasse.

Tout d'abord, cette instabilité plasmidique s'explique par le fait qu'un taux de dilution élevé provoque un lessivage des cellules contenant des plasmides, leur taux de croissance n'étant pas suffisant pour assurer leur stabilité, contrairement aux cellules ne contenant pas de plasmides.

Ensuite, ce phénomène est fidèlement traduit au niveau des modèles par le fait que le taux de répllication plasmidique v est toujours très inférieur au taux de croissance spécifique des bactéries sans plasmides μ_0 et ce, quel que soit le modèle (cf. 2.1.1. paramètres obtenus), les bactéries

ségréguées croissent plus vite que les plasmides ne se répliquent, entraînant une diminution permanente de la population recombinée. Ainsi ces deux valeurs de paramètres permettent aux deux modèles de suivre correctement les évolutions des fractions cumulées.

Enfin, on remarque que le taux spécifique de croissance des cellules possédant des plasmides : μ est très inférieur à μ_0 pour les deux modèles, ce qui permet de rejoindre les travaux de Boudrant *et al.* (1996) pour lesquels les plasmides agiraient comme des « fardeaux », diminuant le taux de croissance cellulaire.

2.2.2.3 Discussions

En culture continue, le modèle de probabilité de répartition uniforme permet de bien simuler le comportement des souches recombinées, toutefois, la valeur négative de p déterminée par identification paramétrique, ne nous permet pas de valider ce modèle comme étant représentatif des phénomènes physiques se déroulant dans les fermentations.

Le modèle de probabilité binomiale présente quant à lui des valeurs de paramètres tous positifs, ce qui permet de ne pas l'exclure comme étant un modèle sans signification physique, toutefois, il simule de façon très insuffisante la biomasse totale.

De plus, il semblerait plus logique qu'une probabilité de répartition uniforme s'applique plutôt qu'une probabilité binomiale concernant la répartition des plasmides lors de la division cellulaire. Une probabilité binomiale signifiant que chaque cellule fille « piocherait au hasard » et à tour de rôle dans le patrimoine plasmidique de la cellule mère. La probabilité de répartition uniforme implique simplement que le patrimoine plasmidique est séparé en deux blocs de taille aléatoire, chacun des blocs allant à une cellule fille.

En conclusion, on peut dire que le modèle de probabilité binomiale permet d'approcher les évolutions des concentrations en biomasse et substrat mais ne permet pas d'expliquer l'ensemble des phénomènes régissant l'instabilité plasmidique.

2.2.3 Culture discontinue

Les modèles obtenus à partir de la culture continue ont été confrontés aux résultats expérimentaux d'une fermentation discontinue à population initiale mixte. Les résultats expérimentaux et simulés sont présentés sur les graphiques suivants.

2.2.3.1 Evolution de la biomasse totale et du substrat

Fig A

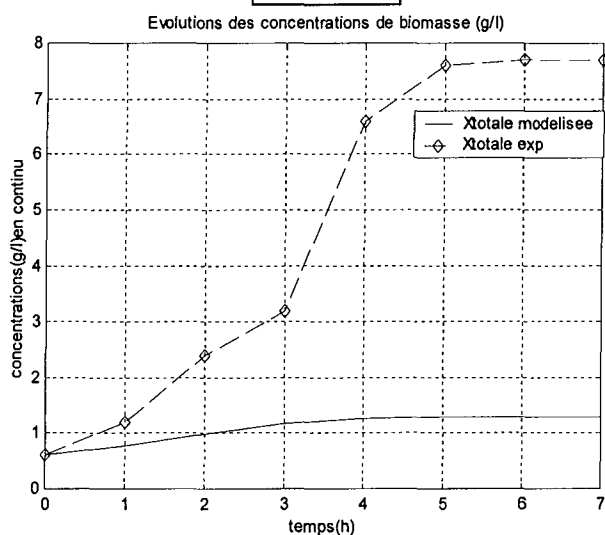


Fig C

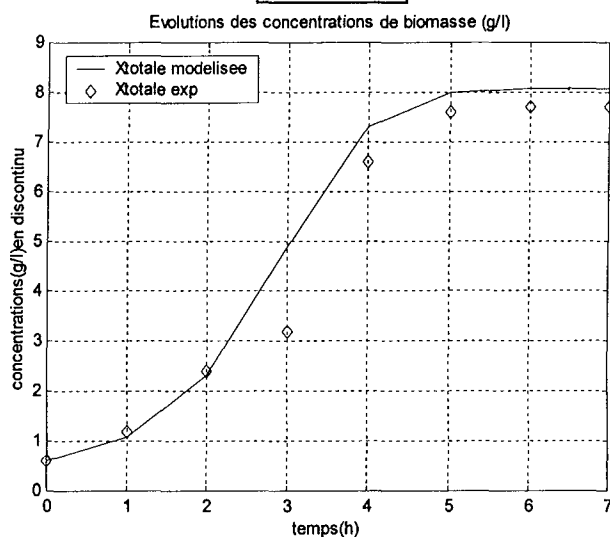
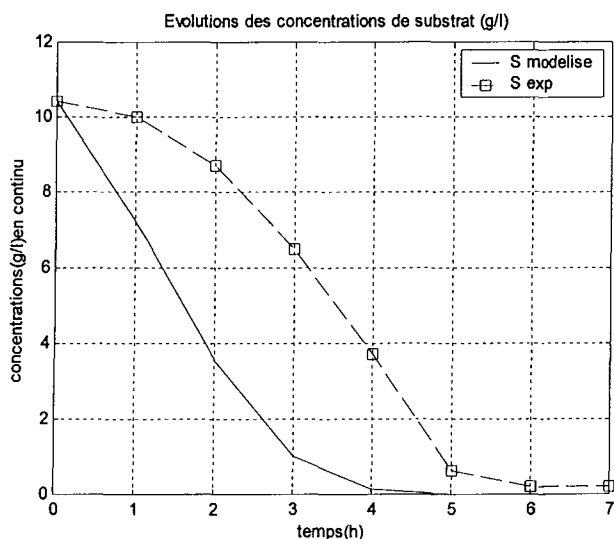
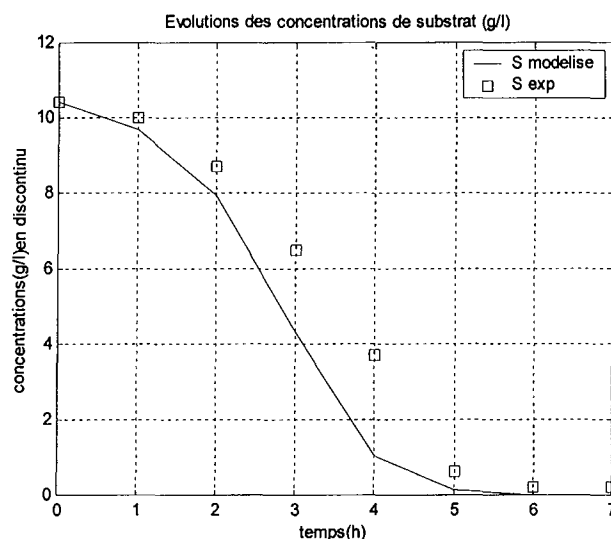


Fig B



Evolution des concentrations expérimentales et simulées avec le modèle de probabilité binomiale en culture discontinue. Fig A : Evolution de la biomasse X_{totale} , Fig B : Evolution du substrat S

Fig D



Evolution des concentrations expérimentales et simulées avec le modèle de répartition uniforme en culture discontinue. Fig C : Evolution de la biomasse X_{totale} , Fig D : Evolution du substrat S

Les valeurs correspondant à ces graphiques sont regroupées en **Annexe 6**.

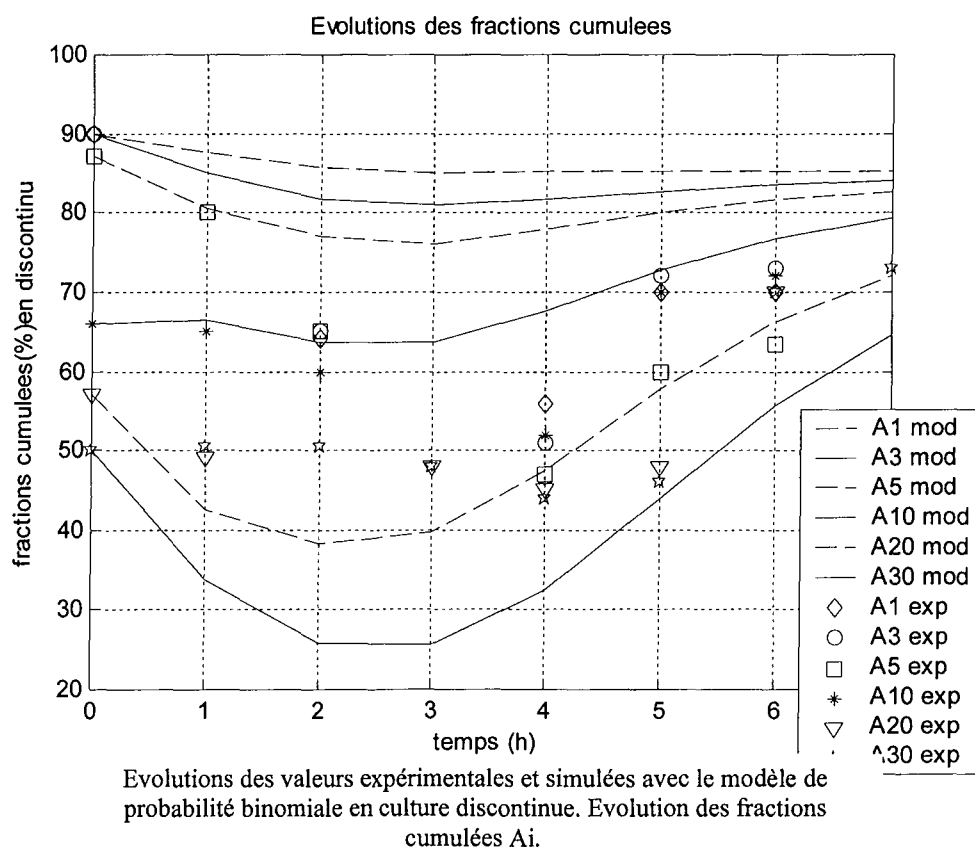
On observe une distorsion dans l'évolution du substrat (Fig. B) et de la biomasse (Fig. A) concernant le modèle de probabilité binomiale. Il ne permet donc pas une simulation satisfaisante.

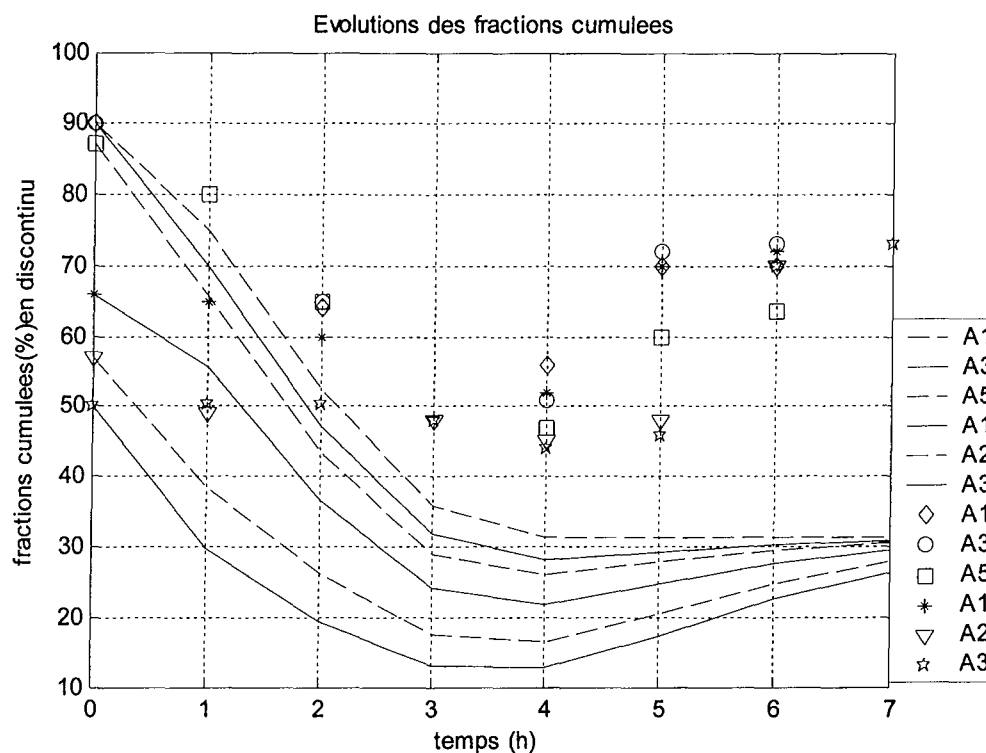
Le modèle de probabilité de répartition uniforme s'approche bien mieux des résultats d'expérience que le modèle de probabilité binomiale bien qu'il n'ait pas de signification physique (la probabilité p déterminée par identification étant négative), le substrat (Fig. D) chute légèrement trop rapidement tandis que la biomasse reste très proche des valeurs expérimentales (Fig. C).

Les courbes modélisées ne permettent donc pas de corroborer de façon suffisante les résultats expérimentaux obtenus en culture discontinue. Notons toutefois que l'allure générale des courbes expérimentales est respectée.

2.2.3.2 Evolution des fractions cumulées de biomasse

Les graphiques suivants montrent l'évolution des fractions cumulées en culture discontinue.





Evolutions des valeurs expérimentales et simulées avec le modèle de probabilité de répartition uniforme en culture discontinue. Evolution des fractions cumulees A_i .

Les valeurs correspondant à ces graphiques figurent en **Annexe 7**.

En ce qui concerne les fractions cumulees en culture discontinue, le modèle de probabilité binomiale offre une meilleure simulation des valeurs expérimentales. En effet, les simulations montrent un domaine de variation plus large en début de fermentation et qui se restreint à partir de la troisième heure de fermentation. Toutefois, les simulations sont décalées vers les faibles valeurs de pourcentage par rapport aux valeurs expérimentales.

Le modèle de probabilité à distribution uniforme ne permet pas de simuler correctement les fractions cumulees en fermentation discontinue. En effet, les simulations sont fortement décalées vers les bas et se restreignent à un domaine de variation très faible autour de 20 à 30 % alors que les valeurs expérimentales se situent entre 50 et 90 %.

2.2.3.3 Discussion

Au vu des graphiques obtenus en fermentation discontinue, on peut conclure que l'application des paramètres identifiés à partir d'une culture continue, ne permet pas de simuler de façon suffisamment satisfaisante l'évolution des souches recombinées. Notons toutefois que l'allure générale des courbes est à chaque fois respectée même si les courbes modélisées présentent des décalages par rapport aux courbes expérimentales. Le modèle de probabilité binomiale (valable d'un point de vue signification physique du fait de la valeur de ses paramètres) permet donc de se rapprocher des phénomènes se déroulant en fermentation de cellules recombinées sans pour autant en apprécier toutes les composantes. Des études plus approfondies restent donc à faire.

Ainsi, bien que le modèle simpliste à deux équations différentielles ne permette pas d'accéder aux fractions de populations, il présente néanmoins l'avantage de simuler de façon satisfaisante l'évolution de la biomasse totale.

Le modèle de répartition uniforme diffère de celui d'Imanaka et Aiba uniquement par un terme supplémentaire en $X(1)$. La présence de ce terme est donc responsable des résultats incohérents obtenus en identification paramétrique. On retrouve ce terme en $X(1)$ notamment dans l'expression de la fraction de biomasse sans plasmide $X(0)$. D'après le modèle de répartition uniforme, on a :

$$\frac{dX_0}{dt} = \sum_{j=0}^N \frac{\mu_j S}{K_j + S} \cdot X_1 \cdot P_{j0} - DX_0$$

Afin de comprendre l'impact de ce terme en $X(1)$, considérons μ_1 , le taux de croissance spécifique de la population à 1 plasmide (dans notre modèle, nous avons considéré μ_p taux de croissance spécifique constant pour toute fraction de biomasse ayant au moins 1 plasmide), on a alors :

$$\frac{dX_1}{dt} = \frac{\mu_0 S}{K_0 + S} X_0 + \frac{\mu_1 S}{K_p + S} X_1 - p \frac{\mu_p S}{K_p + S} X_1 + p \frac{\mu_p S}{K_p + S} X_p - DX_0$$

Pour que le terme en $X(1)$ n'influe pas, il faudrait qu'il soit nul. Ainsi, il faudrait que $\mu_1 = p\mu_p$. Or la probabilité p est très faible (de l'ordre de 10^{-3} à 10^{-4}). Le taux spécifique de croissance μ_1 de la population à un plasmide devrait alors être beaucoup plus faible que celui des populations à nombre de copies de plasmides supérieurs, ce qui est fort peu probable étant donné que le taux de croissance diminue avec le nombre de plasmides. Ainsi, le terme en $X(1)$ est loin d'être négligeable et influe de façon prépondérante sur l'évolution de la biomasse.

Boudrant, Lamotte et Fonteix (1997), ont proposé un modèle ségrégué dans lequel ils ont inclus un terme permettant de diminuer celui en $X(1)$. Les simulations corroboraient alors correctement les résultats expérimentaux. Dans ce modèle, l'équation de la biomasse se définit comme il suit :

$$\frac{dX_i}{dt} = \frac{(a - ib)X_i S}{K + S} - (c + id)X_i + [c + (i - 1)d]X_{i-1} - DX_i$$

Pour $i = 1$, on a :

$$\frac{dX_1}{dt} = \frac{(a - b) \cdot X_1 \cdot S}{K + S} - (c + d)X_1 + cX_0 - DX_0$$

La présence du terme $(c+d)X_1$ diminue l'importance de la fraction X_1 et permet d'obtenir des simulations satisfaisantes, toutefois, ce modèle inclut une apparition spontanée de plasmides (cX_0), ce qui est plus difficile à concevoir. Ainsi l'adéquation entre les simulations et les valeurs d'expérience peut être obtenue par un moyen mathématique mais qui reste toutefois artificiel.

Du fait de la complexité des phénomènes d'instabilité plasmidique, certains aspects nous échappent pour l'instant. Pour améliorer le modèle, il faudrait avoir accès à davantage de phénomènes régissant les évolutions de biomasse. Par exemple, nous avons considéré deux taux

de croissance distincts, μ_0 et μ_P , que nous avons déterminés par identification paramétrique. Sachant que le taux spécifique de croissance diminue avec le nombre de plasmides, il serait plausible de considérer non pas deux valeurs de μ mais autant de valeurs qu'il y a de nombre de copies de plasmide. Pour cela, il faudrait avoir accès aux lois de variation du taux de croissance en fonction du nombre de plasmide. Il serait nécessaire de réaliser des études expérimentales puis de simuler plusieurs comportements. De plus, rien n'indique que l'évolution de μ en fonction du nombre de plasmides soit monotone, ainsi une loi mathématique nous permettra de nous approcher davantage des phénomènes d'instabilité sans obligatoirement en apprécier toutes les composantes.

*Chapitre 4 : Conclusions et
Perspectives*

Chapitre 4 : Conclusions et Perspectives

Ce travail a permis d'étudier la modélisation et une partie du comportement de la souche recombinée *Escherichia coli* HB101 (GADPH) capable d'exprimer la glycéraldéhyde-3-phosphate déshydrogénase. Ce système modèle a été sélectionné du fait que la quantité de GADPH produite est pratiquement proportionnelle au nombre de copies de plasmide, sans présenter de phénomène de saturation. Cette protéine recombinée est également soluble dans le cytoplasme bactérien, non toxique pour la cellule et non dénaturée. Enfin, le nombre de plasmides peut être facilement évalué par la technique des CMI.

Pour étudier le comportement de cette souche, nous avons mis en place un modèle mathématique ségrégué explicitant le bilan des fractions de biomasse contenant de 0 à 30 plasmides et pour lequel la cinétique de croissance des souches bactériennes est bien différenciée de celle de la réplication plasmidique.

De plus, nous avons défini pour ce modèle deux expressions de la probabilité de perte de plasmides, l'une basée sur une distribution de répartition uniforme, l'autre sur une distribution de type binomial.

Les paramètres des modèles ont été déterminés par rapport à une culture en continu sans pression de sélection et avec une population recombinée initiale homogène où toutes les bactéries possèdent 30 plasmides. Les paramètres déterminés ont été ensuite appliqués à une culture en discontinu sans pression de sélection et avec une population initiale hétérogène.

Nous avons d'abord réalisé une identification paramétrique sur un modèle simple à deux populations : l'une contenant des plasmides, l'autre n'en contenant pas. Les paramètres obtenus ainsi que l'évolution des concentrations sont satisfaisants.

Etant donné que dans l'un des modèles ségrégués, une simple sommation des équations de bilan de populations permet de retrouver, à un terme près, le modèle non ségrégué à deux équations, nous avons appliqué ces paramètres au modèle. Les simulations obtenues alors ne montrent pas une adéquation suffisante avec les résultats expérimentaux, c'est pourquoi une identification paramétrique sur les modèles ségrégués s'est avérée nécessaire.

La valeur des paramètres obtenus autorise à conclure que le modèle de probabilité de distribution binomiale permet d'avoir une approche des phénomènes d'instabilité, d'autres facteurs restant sans doute à prendre en compte pour une modélisation plus complète. En effet les paramètres sont plausibles mais les simulations restent insatisfaisantes. L'identification paramétrique sur le modèle de probabilité de répartition uniforme a donné une valeur improbable d'un des paramètres. Ainsi, il ne nous sera pas possible de le considérer comme représentatif des phénomènes d'instabilité.

Au vu de l'adéquation des simulations par rapport aux données expérimentales, et en faisant totale abstraction de la valeur des paramètres, on remarque que le modèle de probabilité de répartition uniforme semble meilleur que celui de probabilité binomiale. Le concept même d'une probabilité de répartition uniforme semblerait d'ailleurs plus plausible que celui d'une distribution de type binomial. Ceci tout en notant que la valeur négative du paramètre p du modèle de probabilité de répartition uniforme n'autorise pas à le valider.

Etant donné que le modèle simple non ségrégué d'Imanaka et Aiba permet de bien simuler les données expérimentales et qu'il ne diffère du modèle de probabilité de répartition uniforme que par l'absence d'un terme en $X(1)$ (fraction de biomasse à 1 plasmide), nous avons pu montrer l'importance de ce terme. Boudrant, Lamotte et Fonteix (1997) ont proposé un modèle ségrégué permettant de minimiser ce terme par l'addition d'un terme en $-X(1)$. Bien que ce moyen permettait d'obtenir des simulations satisfaisantes, il restait néanmoins artificiel parce que le modèle impliquait une création spontanée de plasmides. Le modèle proposé dans cette étude, bien qu'imparfait présente l'avantage d'être complet et de comprendre pourquoi celui de Boudrant, Lamotte et Fonteix (1997) donnait des simulations satisfaisantes.

En réalité, du fait de la complexité des phénomènes d'instabilité plasmidique, il est évident que certains aspects nous échappent encore complètement pour l'instant. Pour améliorer le modèle, il faudrait avoir accès à davantage de phénomènes régissant les évolutions de la biomasse et de sa physiologie. Par exemple, nous n'avons considéré que deux taux de croissance distincts, μ_0 (correspondant à la biomasse sans plasmide) et μ_p (correspondant à la biomasse contenant des plasmides). Ces paramètres ont été déterminés par identification paramétrique. Sachant que le taux spécifique de croissance diminue avec le nombre de plasmides, il serait plausible de considérer non pas deux valeurs de μ mais autant de valeurs qu'il y a de nombre de copies de plasmide. Mais pour cela, il faudrait avoir accès

aux lois réelles de variation du taux de croissance en fonction du nombre de plasmides. Il faudrait pour cela réaliser des études expérimentales puis simuler plusieurs comportements. Notons aussi que rien n'indique que l'évolution du taux de croissance spécifique en fonction du nombre de plasmides est monotone. Ainsi une loi mathématique permettrait de nous approcher davantage des phénomènes d'instabilité sans obligatoirement pouvoir en apprécier toutes les composantes.

Annexes

Annexe 1

Méthode d'Euler

Soit l'équation à intégrer : $\frac{dy}{dx} = f(x, y) = y'$ (1)

La méthode d'Euler fournit une approximation de $y(x)$ en des points d'abscisses donnés. Ayant ainsi obtenu des points (x_i, y_i) , une solution approchée en d'autres points peut être obtenue par interpolation. Ici les nœuds d'intégration sont choisis équidistants c'est-à-dire que $x_i = x_0 + i \cdot h$ où le pas d'intégration h est constant. Soient $y(x_i)$, la valeur exacte et y_i la valeur approchée.

On remplace dans (1) y' par la formule de dérivation $\frac{y(x_i) - y(x)}{h}$. Puisqu'à la $i^{\text{ème}}$ étape, on approxime $y(x_i) \approx y_i$, on obtient : $y(x_i + h) \approx y_i + hf(x_i, y_i)$

Ce qui suggère de nouveau l'équation :

$$y_{i+1} = y_i + hf(x_i, y_i) \quad (2) \text{ appelée équation aux différences pour la méthode d'Euler.}$$

On peut montrer que dans cette méthode l'erreur $|y(x_i) - y_i|$ est fonction de h . Les méthodes ultérieures auront une erreur qui diminuera comme h^p avec $p > 1$.

Méthode de Taylor

On améliore la méthode d'Euler en utilisant un développement de Taylor jusqu'à l'ordre

2. Si $y(x_i)$ est donné, on a :

$$y(x_{i+1}) = y(x_i) + (x_{i+1} - x_i)y'(x_i) + \frac{1}{2}(x_{i+1} - x_i)^2 y''(x_i) + \frac{1}{6}(x_{i+1} - x_i)^3 y'''(\xi_i)$$

où ξ_i appartient à $[x_{i+1}, x_i]$

Si $h = x_{i+1} - x_i$ est assez petit, on a l'égalité approchée :

$$y(x_{i+1}) \approx y(x_i) + hy'(x_i) + \frac{(x_{i+1} - x_i)^2}{2} y''(x_i)$$

Puisque $y'(x) = f(x, y(x))$, on a :

$$y''(x) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'(x)$$

$$y''(x) = \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x)) \cdot f(x, y(x))$$

Tout ceci suggère la méthode aux différences suivante :

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2} \left[\frac{\partial f}{\partial x}(x_i, y_i) + \frac{\partial f}{\partial y}(x_i, y_i) \cdot f(x_i, y_i) \right] \quad (3) \quad \text{appelée méthode de}$$

Taylor d'ordre 2

Remarque :

- On peut montrer que pour chaque x_i fixé, l'erreur varie avec h^2 ,
- La nécessité de calculer $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ constitue l'inconvénient de cette méthode,
- cette méthode se généralise facilement, cependant, la méthode de Taylor d'ordre plus élevé nécessitera des calculs de dérivées partielles d'ordre important, ce qui peut être long.

La méthode de Runge et Kutta

Cette méthode permet d'avoir une erreur qui se comporte localement comme h^p avec $p > 1$ et de corriger les inconvénients dus au calcul des dérivées partielles de f . Cette méthode est basée sur la formule de Taylor à plusieurs variables :

$$f(x, y) = f(x_0, y_0) + \left[(x - x_0) \frac{\partial f}{\partial x} + (y - y_0) \frac{\partial f}{\partial y} \right] +$$

$$\frac{1}{2!} \left[(x - x_0)^2 \frac{\partial^2 f}{\partial x^2} + 2(x - x_0)(y - y_0) \frac{\partial^2 f}{\partial x \partial y} + (y - y_0)^2 \frac{\partial^2 f}{\partial y^2} \right] + \dots \quad (4)$$

$$\dots + \frac{1}{(n+1)!} \sum_{j=0}^{n+1} \left(\frac{n+1}{j} \right) \left[(x - x_0)^{n+1-j} (y - y_0)^j \frac{\partial^{n+1} f}{\partial x^{n+1} \partial y^j} (\xi, n) \right]$$

où $\xi \in (x, x_0)$ et $n \in (y, y_0)$

Dans cette méthode, l'idée est de remplacer le terme :

$$f(xi, yi) + \frac{h}{2} \left[\frac{\partial f}{\partial x} (xi, yi) f(xi, yi) \right] \text{ dans la formule (3) par une expression de type :}$$

$$af(xi + \alpha, yi + \beta)$$

En utilisant l'équation (4), on obtient :

$$af(xi + \alpha, yi + \beta) = af(xi, yi) + a\alpha \frac{\partial f}{\partial x} (xi, yi) + a\beta \frac{\partial f}{\partial y} (xi, yi) + aR_2(\xi i, ni)$$

où le reste R_2 est tel que :

$$R_2(\xi i, ni) = \frac{1}{2} \left[\alpha^2 \frac{\partial^2 f}{\partial x^2} (\xi i, ni) + 2\alpha\beta \frac{\partial^2 f}{\partial x \partial y} (\xi i, ni) + \beta^2 \frac{\partial^2 f}{\partial y^2} (\xi i, ni) \right]$$

avec $\xi i \in (xi, xi + \alpha)$ et $ni \in (yi, yi + \beta)$

En identifiant les coefficients de $f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$, on voit que l'on doit choisir :

$$a = 1 \quad \alpha = \frac{h}{2} \quad \beta = \frac{h}{2} f(xi, yi)$$

En substituant dans l'équation (3), on obtient la formule aux différences :

$$y_{i+1} = yi + hf \left(xi + \frac{h}{2}, yi + \frac{h}{2} f(xi, yi) \right) \quad (5) \text{ Cette équation est utilisée dans les algorithmes.}$$

Méthode de Runge et Kutta d'ordre 2. Algorithme.

1. poser $i=0$
2. soit $k = hf(xi, yi)$
3. $y_{i+1} = yi + hf\left(xi + \frac{h}{2}, yi + \frac{k}{2}\right)$
4. $x_{i+1} = xi + h$
5. si $i=N-1$, allez à l'étape 7
6. $i=i+1$, allez à l'étape 2
7. fin

N étant le nœud terminal d'intégration, l'intégration se termine donc à x_N

Le fait que l'erreur soit d'ordre 2 peut être constaté en examinant la forme du reste

$R_2(\xi i, ni)$, si $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial x \partial y}$ sont bornées, puisque $\alpha = \frac{h}{2}$ et $\beta = \frac{h}{2} f(xi, yi)$, il suffit que f soit bornée pour que $|R_2(\xi i, ni)| < Ch^2$. C étant une constante.

Méthode de Runge et Kutta d'ordre 4. Algorithme.

1. Poser $N = \frac{x_N - x_0}{h}$, le nombre de pas nécessaires
2. Poser $i=0$
3. $k1 = hf(xi, yi)$
4. $k2 = hf\left(xi + \frac{h}{2}, yi + \frac{1}{2}k1\right)$
5. $k3 = hf\left(xi + \frac{h}{2}, yi + \frac{1}{2}k2\right)$
6. $k4 = hf(xi + h, yi + k3)$
7. $y_{i+1} = yi + \frac{1}{6}(k1 + 2k2 + 2k3 + k4)$
8. $x_{i+1} = xi + h$
9. si $i=N-1$, allez à l'étape 11
10. $i=i+1$, allez à l'étape 3
11. fin

Annexe 2

**Valeurs des concentrations : Résultats d'expérience et du modèle non ségrégué
d'Imanaka et Aiba (1981).**

Temps (h)	S mod. (g/l)	S exp (g/l)	X ₀ mod. (g/l)	X ₀ exp. (g/l)	X _p mod. (g/l)	X _p exp (g/l)	Xtotale mod. (g/l)	Xtotale exp (g/l)
0.00	10.80	10.80	0.00	0.00	1.20	1.2	1.20	1.20
1.80	8.81	8.16	0.01		2.61		2.66	1.60
3.60	4.67	4.80	0.10		5.55		5.65	4.80
4.92	0.72	0.24	0.22		8.29		8.51	7.60
5.82	0.03	0.24	0.23		8.79		9.02	8.40
7.00	0.00	0.24	0.23	0.00	8.80	8.8	9.04	8.80
8.51	0.91	0.24	0.21		7.89		8.09	7.20
11.50	4.07	1.68	0.28	0.81	4.00	3.99	4.28	4.80
23.00	2.47	3.60	4.22	4.25	0.42	0.45	4.64	4.70
26.50	2.39	3.12	4.49	4.37	0.18	0.13	4.67	4.50
31.50	2.35	2.76	4.63	4.64	0.06	0.06	4.68	4.70
39.40	2.34	2.40	4.68		0.01		4.69	4.80
49.70	2.34	2.40	4.69		0.00		4.69	4.80
58.00	2.34	2.40	4.69	4.8	0.00	0.0	4.69	4.80

Annexe 3

Valeurs des concentrations : Résultats d'expérience et du modèle de répartition de probabilité uniforme. Application des paramètres du modèle d'Imanaka et Aiba (1981).

Temps (h)	Substrat mod.(g/l)	Substrat exp. (g/l)	X ₀ mod (g/l)	X ₀ exp (g/l)	Xtotale mod. (g/l)	Xtotale exp. (g/l)
0.00	10.80	10.80	0.00	0.00	1.20	1.20
1.80	8.68	8.16	0.16		3.16	1.60
3.60	2.75	4.80	2.02		9.46	4.80
4.92	0.06	0.24	2.92		12.69	7.60
5.82	0.00	0.24	2.94		12.77	8.40
7.00	0.00	0.24	2.94	0.00	12.77	8.80
8.51	0.90	0.24	2.64		11.45	7.20
11.50	2.75	1.68	2.67	0.81	7.54	4.80
23.00	2.33	3.60	4.71	4.25	5.59	4.70
26.50	2.32	3.12	4.74	4.37	5.25	4.50
31.50	2.33	2.76	4.73	4.64	4.96	4.70
39.40	2.33	2.40	4.70		4.77	4.80
49.70	2.34	2.40	4.69		4.70	4.80
58.00	2.34	2.40	4.69	4.8	4.69	4.80

Valeurs des fractions cumulées : Résultats d'expérience et du modèle de répartition de probabilité uniforme. Application des paramètres du modèle d'Imanaka et Aiba (1981).

Temps (h)	A1 simulé	A3 simulé	A5 simulé	A10 simulé	A20 simulé	A30 simulé	A1 exp	A3 exp	A5 exp	A10 exp	A20 exp	A30 exp
0.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1.80	94.78	83.84	76.60	63.41	45.45	34.10						
3.60	78.61	65.77	57.82	44.60	28.89	20.13						
4.92	76.95	68.74	62.80	52.24	38.72	29.89						
5.82	76.97	72.51	68.78	61.43	51.06	43.55						
7.00	76.97	75.08	73.32	69.42	63.11	58.06	100.0	95.18	95.18	79.52	69.88	55.42
8.51	76.93	75.93	74.99	72.79	68.67	65.11						
11.50	64.58	57.55	52.80	44.14	32.39	24.79	83.13	65.06	69.88	61.45	48.19	36.14
23.00	15.74	13.42	11.92	9.38	6.29	4.51	9.64	4.82	4.34	3.61	2.89	2.41
26.50	9.63	8.22	7.31	5.76	3.87	2.78	2.89	1.93	1.69	1.69	0.84	0.72
31.00	4.61	3.93	3.50	2.76	1.85	1.33	1.20	0.00	0.24	0.00	0.00	0.00
39.40	1.37	1.17	1.04	0.82	0.55	0.40						
49.70	0.27	0.23	0.21	0.16	0.11	0.08						
58.00	0.07	0.06	0.06	0.04	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00

Annexe 4

**Valeurs des concentrations : Résultats d'expérience et du modèle de probabilité
binomiale.**

Temps (h)	Xtotale mod. (g/l)	Xtotale exp. (g/l)	Substrat mod. (g/l)	Substrat exp. (g/l)
0.000	1.200	1.200	10.800	10.800
1.800	1.750	1.600	8.815	8.160
3.600	2.532	4.800	4.167	4.800
4.920	3.038	7.600	0.273	0.240
5.820	3.081	8.400	0.008	0.240
7.000	3.083	8.800	0.000	0.240
8.510	2.723	7.200	0.907	0.240
11.500	0.882	4.800	3.031	1.680
23.000	0.091	4.700	3.219	3.600
26.500	0.083	4.500	3.238	3.120
31.500	0.081	4.700	3.247	2.760
39.400	0.081	4.800	3.249	2.400
49.700	0.081	4.800	3.249	2.400
58.000	0.081	4.800	3.249	2.400

Valeurs des concentrations : Résultats d'expérience et du modèle de répartition de probabilité uniforme.

Temps (h)	Xtotale mod. (g/l)	Xtotale exp. (g/l)	Substrat mod. (g/l)	Substrat exp. (g/l)
0.000	1.200	1.200	10.800	10.800
1.800	2.292	1.600	8.800	8.160
3.600	4.617	4.800	5.424	4.800
4.920	7.260	7.600	2.068	0.240
5.820	8.596	8.400	0.456	0.240
7.000	8.952	8.800	0.020	0.240
8.510	7.963	7.200	1.002	0.240
11.500	4.032	4.800	4.824	1.680
23.000	4.750	4.700	3.049	3.600
26.500	4.687	4.500	3.012	3.120
31.500	4.641	4.700	3.001	2.760
39.400	4.621	4.800	2.999	2.400
49.700	4.617	4.800	2.999	2.400
58.000	4.617	4.800	2.999	2.400

Annexe 5

Valeurs des fractions cumulées : Résultats d'expérience et du modèle de probabilité binomiale.

Temps (h)	A1 simulé	A3 simulé	A5 simulé	A10 simulé	A20 simulé	A30 simulé	A1 exp	A3 exp	A5 exp	A10 exp	A20 exp	A30 exp
0.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1.80	99.81	99.02	97.44	89.94	62.50	48.34						
3.60	98.21	95.50	91.65	79.18	49.68	32.59						
4.92	97.01	93.72	89.58	77.37	50.51	32.83						
5.82	96.96	94.30	91.08	81.58	60.07	41.28						
7.00	96.95	95.04	92.81	86.30	71.27	55.42	100.0	95.18	95.18	79.52	69.88	55.42
8.51	96.92	95.58	94.09	89.80	78.10	67.97						
11.50	91.96	89.57	86.57	76.92	51.54	36.15	83.13	65.06	69.88	61.45	48.19	36.14
23.00	10.49	9.54	8.67	6.82	3.90	2.40	9.64	4.82	4.34	3.61	2.89	2.41
26.50	2.91	2.62	2.37	1.84	1.04	0.64	2.89	1.93	1.69	1.69	0.84	0.72
31.50	0.42	0.38	0.34	0.26	0.15	0.09	1.20	0.00	0.24	0.00	0.00	0.00
39.40	0.02	0.02	0.02	0.01	0.01	0.00						
49.70	0.00	0.00	0.00	0.00	0.00	0.00						
58.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Valeurs des fractions cumulées : Résultats d'expérience et du modèle de répartition de probabilité uniforme.

Temps (h)	A1 simulé	A3 simulé	A5 simulé	A10 simulé	A20 simulé	A30 simulé	A1 exp	A3 exp	A5 exp	A10 exp	A20 exp	A30 exp
0.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1.80	100.0	92.83	87.46	76.92	60.98	49.52						
3.60	97.00	86.53	79.39	66.45	49.00	37.70						
4.92	93.61	83.71	76.85	64.43	47.86	37.07						
5.82	92.92	85.25	79.57	68.93	54.18	43.93						
7.00	92.89	89.21	86.07	79.61	69.79	62.15	100.0	95.18	95.18	79.52	69.88	55.42
8.51	92.88	91.37	89.97	86.78	81.16	76.43						
11.50	82.24	75.49	70.66	61.43	48.03	38.65	83.13	65.06	69.88	61.45	48.19	36.14
23.00	5.09	4.58	4.22	3.57	2.68	2.10	9.64	4.82	4.34	3.61	2.89	2.41
26.50	2.12	1.91	1.77	1.50	1.13	0.89	2.89	1.93	1.69	1.69	0.84	0.72
31.00	0.60	0.54	0.50	0.42	0.32	0.25	1.20	0.00	0.24	0.00	0.00	0.00
39.40	0.08	0.072	0.07	0.057	0.04	0.03						
49.70	0.01	0.01	0.00	0.00	0.00	0.00						
58.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Annexe 6

Valeurs des concentrations : Résultats d'expérience et du modèle de probabilité de répartition uniforme.

Temps (h)	Xtotale mod. (g/l)	Xtotale exp. (g/l)	Substrat mod. (g/l)	Substrat exp. (g/l)
0.00	0.6000	0.600	10.407	10.407
1.00	0.7827	1.200	7.210	9.982
2.00	0.9934	2.400	3.605	8.708
3.00	1.1801	3.200	1.028	6.478
4.00	1.2664	6.600	0.147	3.717
5.00	1.2816	7.600	0.015	0.637
6.00	1.2832	7.700	0.001	0.212
7.00	1.2833	7.700	0.000	0.212

Valeurs des concentrations : Résultats d'expérience et du modèle de probabilité de répartition uniforme.

Temps (h)	Xtotale mod. (g/l)	Xtotale exp. (g/l)	Substrat mod. (g/l)	Substrat exp. (g/l)
0.00	0.600	0.600	10.407	10.407
1.00	1.090	1.200	9.667	9.982
2.00	2.315	2.400	7.925	8.708
3.00	4.893	3.200	4.346	6.478
4.00	7.301	6.600	1.059	3.717
5.00	7.998	7.600	0.124	0.637
6.00	8.081	7.700	0.011	0.212
7.00	8.089	7.700	0.001	0.212

Annexe 7

Valeurs des fractions cumulées : Résultats d'expérience et du modèle de probabilité binomiale.

Temps (h)	A1 simulé	A3 simulé	A5 simulé	A10 simulé	A20 simulé	A30 simulé	A1 exp	A3 exp	A5 exp	A10 exp	A20 exp	A30 exp
0.00	90.00	90.00	87.00	66.00	57.00	50.00	90.00	90.00	87.00	66.00	57.00	50.00
1.00	87.51	85.01	80.64	66.43	42.60	33.86			80.00	65.00	49.00	50.50
2.00	85.60	81.81	76.94	63.54	38.34	25.87	64.00	65.00	65.00	60.00		50.50
3.00	85.06	80.88	76.10	63.57	39.74	25.65					48.00	48.00
4.00	85.20	81.70	77.86	67.74	47.42	32.34	56.00	51.00	47.00	52.00	45.00	44.00
5.00	85.26	82.74	80.03	72.83	57.79	43.91	70.00	72.00	60.00	70.00	48.00	46.00
6.00	85.27	83.50	81.63	76.67	66.19	55.56	70.00	73.00	63.50	72.00	70.00	70.00
7.00	85.27	84.03	82.74	79.34	72.12	64.64						73.00

Valeurs des fractions cumulées : Résultats d'expérience et du modèle de probabilité de répartition uniforme.

Temps (h)	A1 simulé	A3 simulé	A5 simulé	A10 simulé	A20 simulé	A30 simulé	A1 exp	A3 exp	A5 exp	A10 exp	A20 exp	A30 exp
0.00	90.00	90.00	87.00	66.00	57.00	50.00	90.00	90.00	87.00	66.00	57.00	50.00
1.00	75.41	73.10	69.35	57.16	36.74	29.21			80.00	65.00	49.00	50.50
2.00	66.80	63.88	60.44	50.75	31.96	21.85	64.00	65.00	65.00	60.00		50.50
3.00	66.35	64.06	61.56	54.53	39.83	27.84					48.00	48.00
4.00	66.35	64.73	63.03	58.30	48.00	37.82	56.00	51.00	47.00	52.00	45.00	44.00
5.00	66.35	65.21	64.03	60.86	53.83	46.51	70.00	72.00	60.00	70.00	48.00	46.00
6.00	66.35	65.54	64.73	62.58	57.84	52.79	70.00	73.00	63.50	72.00	70.00	70.00
7.00	66.35	65.78	65.21	63.74	60.56	57.13						73.00

Bibliographie

Bibliographie

Ariga, O., Y. Andoh, Y. Fujishita, T. Watani, Y. Sano. Production of thermophilic alpha-amylase using immobilized transformed *E. coli* by addition of glycine. *J. Ferment. Bioeng.*, **6**, 397, (1991).

Balbas, P., X. Soberon, F. Bolivar et R. L. Rodriguez. The plasmid pBR322. Vectors. A survey of molecular cloning vectors and their uses. (R. L. Rodriguez, D. T. Denhardt eds), Butterworths, 5-41, (1988).

Barbero, J. L., J. M. Buesa, M. A. Penalva, A. Perez-Aranda et J. L. Garcia. Secretion of mature human interferon alpha 2 in the periplasmic space of *Escherichia coli*. *J. Biotechnol.*, **4**, 255, (1986).

Bentley, W. E. et O. E., Quiroga. Investigation of subpopulation heterogeneity and plasmid stability in recombinant *E. coli* via a simple segregated model. *Biotechnol. Bioeng.*, **42**, 224, (1993).

Bhattacharya, P. et D. Roy. Estimating segregational plasmid instability in recombinant cell cultures : a generalised approach. *J. Ferment. Bioeng.*, **5**, 520, (1995).

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker et H. W. Boyer. Construction and characterization of new cloning Vehicles II. A multiple cloning system. *Gene*, **2**, 95, (1977).

Boudrant, J., A. Gschaedler, N. Robas et C. Branlant, Environmental dependance and behavioral changes of a recombinant *E. coli*. Recombinant, DNA Biotechnology III, Engineering Foundation, Deauville, 16-21 oct. (1994). In Recombinant DNA Biotechnology III, The Integration of Biological and Engineering Sciences, J. A. Asenjo and B. A. Andrews editors, *Ann. N. Y. Acad. Sci.*, **782**, 159, (1996).

Boudrant, J., D. Lamotte et C. Fonteix. Modélisation de l'instabilité plasmidique de souches recombinantes, application à *Escherichia coli*. *Rec. Prog. Génie Procédés*, **57**, 1, (1997).

Boyer, H. W. et D. Roulland-Dussoix. A complementation analysis of the restriction and purification of DNA in *E. coli*. *J. Mol. Biol.*, **41**, 459, (1969).

Branlant, G., G. Flesch et C. Branlant. Molecular cloning of the glyceraldehyde-3-phosphate deshydrogenase genes of *Bacillus stearothermophilus* and *Escherichia coli*, and their expression in *Escherichia coli*. *Gene*, **25**, 1, (1983).

Brigidi, P., R. A. Gonzalez-vara y, M. Rossi et D. Matteuzzi. Study of recombinant plasmids during the continous culture of *Bacillus stearothermophilus* NUB3621 in selective medium. *Biotechnol. Bioeng.*, **5**, 507, (1997).

Brown, S. W., H.-P. Meyer et A. Fiechter. Continuous production of human leucocyte interferon with *Escherichia coli* and continuous cell lysis in a two stage chemostat. *Appl. Microbiol. Biotechnol.*, **23**, 5, (1985).

Caulcott, C. A., A. Dunn, H. A. Robertson, N. S. Cooper, M. E. Brown et P. M. Rhodes. Investigation on the stability of low-copy number plasmids in *E. coli*. *J. Gen. Microbiol.*, **133**, 1881, (1987).

Chew, L. C. K., W. C. A. Tacon et J. A. Cole. Increased stability of maintenance of pAT153 in *E. coli* HB101 due to the transposition of IS1 from the chromosome into the tetracycline resistance site of pAT153. *FEMS Microbiol. Lett.*, **36**, 275, (1986).

Chew, L. C. K., W. C. A. Tacon et J. A. Cole. Effects of growth conditions on the rate of loss of plasmid pAT153 from continuous cultures of *E. coli* HB101. *FEMS Microbiol. Lett.*, **56**, 101, (1988).

Clewell, D. B. Nature of ColE1 plasmid replication in *E. coli* in the presence of chloramphenicol. *J. Bacteriol.*, **110**, 667, (1972).

Cohen, S. N. et A. C. Y. Chang. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. USA*, **70**, 3240, (1973).

Coppela, S. J., C. M. Acheson et P. Dhurjati. Measurement of copy number using HPLC. *Biotechnol. Bioeng.*, **29**, 646, (1987).

Cortinho, T. M. et M. A. Hjortso. A new method for quantifying the probability of segregative plasmid loss in *E. coli* B/r. *Biotechnol. Tech.*, **2**, 141, (1988).

Covarrubias, L., L. Cervantes, A. Covarrubias, X. Soberon, I. Vichido, A. Blanco, J.-M. Kuperstoch-Portnoy et F. Bolivar. Construction and characterization of new cloning vehicles : Mobilization and coding properties of pBR322 and several deletion derivatives including pBR322an. *Gene*, **13**, 25, (1981).

De Boer, H. et R. A. Kastelein. Biased codon usage : an exploration of its role in optimisation of translation. *Maximising Gene expression*, W. Resnikoff and L. Gold eds., Butterworths (1986).

Dennis, K., F. Srienc, et J. E. Bailey. Ampicillin effects on five recombinant *Escherichia coli* strains : implications for selection pressure design. *Biotechnol. Bioeng.*, **27**, 1490, (1985).

De taxis du Poët, P., P. Dhulster, J.-N. Barbotin et D. Thomas. Plasmid instability and biomass production : comparison between free and immobilized cell cultures of *Escherichia coli* BZ18 (pTG201) without selection pressure. *J. Bacteriol.*, **165**, 871, (1986).

De Taxis du Poët, P., Y. Arcand, R. Bernier, J.-N. Barbotin et D. Thomas. Plasmid stability in immobilized and free recombinant *Escherichia coli* JM105 (pKK2236220) : importance of oxygen diffusion, growth rate and plasmid copy number. *Appl. Environ. Microbiol.*, **53**, 1548, (1987).

Di Biasio A. et C. Sardonini. Stability of continuous culture with recombinant organisms. *Ann. N. Y. Acad. Sci.*, **4**, 111, (1985).

Dinçbas, U., A. Hortaçsu et A. Camurdan. Plasmid stability in immobilized mixed cultures of recombinant *E. coli*. *Biotechnol. Prog.*, **9**, 218, (1993).

Doelle, H. W., K. N. Hollywood et A. W. Westwood. Regulation of glucose metabolism in bacterial systems. *Adv. Biochem. Eng.* **23**, 1, (1982).

Dunn, A., A. M. Davidson, M. J. Day, et P. F. Randerson. An analysis of the instability kinetics of plasmid pHSG415 during continuous culture of *Escherichia coli*. *J. Gen. Microbiology*. **141**, 63, (1995).

Engberg, B. et K. Nordström. Replication of the R-factor R1 in K-12 at different growth rates. *J. Bacteriol.*, **123**, 179, (1975).

Ensley, B. D. Stability of recombinant plasmids in industrial microorganisms. *CRC Critical Reviews in Biotechnology.*, **3**, 263, (1985).

Futcher, A. B. et B. S. Cox. Copy number and the stability of 2 uncircle-based artificial plasmids of *Saccharomyces cerevisiae*. *J. Bacteriol.*, **157**, 283, (1984).

Gschaedler, A. "Etudes physiologiques et moléculaires intracellulaires de la surexpression de la biosynthèse de protéines lors de la mise en œuvre de *Escherichia coli* recombinée. Application à la glycéraldéhyde-3-phosphate déshydrogénase". I.N.P.L., 14 avril 1994.

Georgiou, G. Optimizing the production of recombinant proteins in microorganisms. *AIChE Journal*, **8**, 1233, (1988).

Georgiou, G., T. N. Telford, M. L., Shuler et D. B. Wilson. Localization of inclusion bodies in *Escherichia coli* overproducing β -lactamase or alkaline phosphatase. *Appl. Environ. Microbiol.*, **52**, 1157, (1986).

Gerdes, K., P. B. Rasmussen et S. Molin. Unique type of plasmid maintenance function : post segregational killing of plasmid free cells. *Proc. Natl. Acad. Sci. USA.*, **83**, 3116, (1986).

Germino, J. et D. Bastia. Rapid purification of cloned gene product by genetic fusion and site specific proteolysis. *Proc. Natl. Acad. Sci. USA*, **83**, 4692, (1984).

Godwin D. et J. H. Slater. The influence of the growth environment on the stability of a drug resistance plasmid in *Escherichia coli* K12. *J. Gen. Micro.*, **111**, 201, (1979).

Goff, S. A. et A. L. Goldberg. Production of abnormal proteins in *E. coli* stimulates transcription of *Lon* and other heat shock genes. *Cell.*, **41**, 587, (1985).

Goldberg, A. L. et S. A. Goff. The selective degradation of abnormal proteins in bacteria. *Maximising gene expression*, W. Reznikoff and L. Gold eds., Butterworths, (1986).

Han, K., H. C. Lim et J. Hong. Acetic acid formation in *Escherichia coli* fermentation. *Biotechnol. Bioeng.*, **39**, 663, (1992).

Helling, R. B., T. Kinney et J. Adams. The maintenance of plasmid containing organisms in populations of *E. coli*. *J. Gen. Microbiol.*, **123**, 129, (1981).

Heppel, L. A. « The concept of periplasmic enzymes. » Structure and function of biological membranes, L. I. Rothfield eds., Academic Press, New York. (1971).

Holms, W. H. The central metabolic pathway of *Escherichia coli* : relationship between flux and control at a branch point, efficiency of conversion to biomass and excretion of acetate. *Topics in cellular regulation*, **28**, 69, (1986).

Hopkins, D. J., M. J. Betenbaugh et P. Dhurjati. Effects of dissolved oxygen shock on the stability of recombinant *E. coli* containing plasmid PKN401. *Biotechnol. Bioeng.*, **29**, 85, (1987).

Huang, J., P. Dhulster, D. Thomas et J. N. Barbotin. Agitation rate effect on plasmid stability in immobilized and free-cell continuous cultures of recombinant *E. coli*. *Enz. Microbiol. Technol.*, **12 dec**, 933, (1990).

Hui, A., J. Hayflick, K. Dinkelspiel et H. A. de Boer. Mutagenesis of the three bases preceding the start codon of the β -galactosidase mRNA and its effects on translation in *Escherichia coli*. *EMBO J.*, **3**, 623, (1984).

Imanaka, I. T. et S. Aiba, A perspective on the application of genetic engineering : stability of recombinant plasmid. *Ann. N. Y. Acad. Sci.*, **1**, 369, (1981).

Jones, I. M., S. B. Primrose, A. Robinson et D. C., Ellwood. Maintenance of some ColE1-type plasmids in chemostat culture. *Mol. Gen. Genet.*, **180**, 579, (1980).

Jones, S. A. et J. Melling. Persistence of pBR322-related plasmids in *Escherichia coli* grown in chemostat cultures. *FEMS Microbiol. Lett.*, **22**, 239, (1984).

Kapralk, F. et P. Jecmen. The structural stability of an expression plasmid bearing a heterologous cloned gene depends on whether this gene was expressed or not. *Biotechnol. Lett.*, **4**, 251, (1993).

Karbasi, M. et T. Keshavarz, Effect of dilution rate on the stability of TOL plasmid pTKO in *Pseudomonas putida* PPK1. *Biotechnol. Lett.*, **10**, 985, (1997).

Kennell, D. E. The instability of messenger RNA in bacteria. *Maximising Gene Expression*. W. Reznikoff and L. Gold eds., Butterworths, (1986).

Keshavarz, T., M. D. Lilly et P. H. Clarke. Stability of a catabolic plasmid in continuous culture. *J. Gen. Microbiol.*, **131**, 1193, (1985).

Koh, B. T., U. Nakashimida, M. Pfeiffer et M. G. S. Yap. Comparison of acetate inhibition on growth of host and recombinant *E. coli* K12 strains. *Biotechnol. Lett.*, **14**, 1115, (1992).

Koizumi, J. L. et S. Aiba. Some consideration of plasmid number in a proliferating cell. *Biotechnol. Bioeng.*, **28**, 311, (1986).

- Koizumi, J., Y. Monden et S. Aiba. Effect of temperature and dilution rate on the copy number of recombinant plasmids in continuous culture of *Bacillus Stearothermophilus* (pLP11). *Biotechnol. Bioeng.*, **27**, 721, (1985).
- Korpela, K., P. Buchert et H. Söderlund. Determination of plasmid copy number with nucleic acid hybridation. *J. Biotechnol.*, **5**, 267, (1987).
- Lamotte, D., "Etude de la stabilité plasmidique de bactéries recombinées productrices de glycéraldéhyde-3-phosphate déshydrogénase". Diplôme d'ingénieur chimiste de l'Université Libre de Bruxelles, septembre 1992.
- Lamotte, D., A., Gschaedler et J. Boudrant. Transitory plasmid instability during the exponential phase of growth of recombinant *E. coli*. *Biotechnol. Lett.*, **16**, 1113, (1994).
- Lee, S. B. et J. E. Bailey. Analysis of growth rate effects on productivity of recombinant *E. coli* populations using molecular mechanism models. *Biotechnol. Bioeng.*, **6**, 805, (2000).
- Lee, Y. L. et H. N. Chang. High cell density continuous culture of *E. coli* producing penicillin acylase. *Biotechnol. Lett.*, **11**, 787, (1988).
- Lee, Y. L. et H. N. Chang. High density culture of recombinant *E. coli* producing penicillin acylase in a membrane cell recycle fermentor. *Biotechnol. Bioeng.*, **36**, 330, (1990).
- Luedeking, R. et E. L. Piret. Transient and steady state in continuous fermentations. Theory and experiment. *J. Biochem. Microbiol. Technol. Eng.*, **1**, 431, (1959).
- Luli, G. W. et W. R. Strohl. Comparison of growth, acetate production and acetate inhibition of *E. coli* strains in batch and fed batch fermentations. *Appl. Environ. Microbiol.*, **apr.**, 1004, (1990).
- Majewski, R. A. et M. M. Domach. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnol. Bioeng.*, **35**, 732, (1990).

Manen, D., T. Goebel et L. Caro. The *par* region of pSC101 affects plasmid copy number as well as stability. *Mol. Microbiol.*, **4**, 1839, (1990).

Marin-Iniesta, F., P. De Taxis du Pöet, P Dhulster, D. Thomas et J. N. Barbotin. Immobilized bacteria and plasmid stability. *Ann. N. Y. Acad. Sci.*, **501**, 217, (1987).

Marin-Iniesta, F., M. Nasri, P Dhulster, J. N. Barbotin et D. Thomas. Influence of Oxygen supply on the stability of recombinant plasmid pTG201 in immobilized *E. coli* cells. *Appl. Microbiol. Biotechnol.*, **28**, 455, (1988).

Maurizi, M. R., P. Trisler et S. Gottesman. Insertional mutagenesis of the *Lon* gene in *Escherichia coli*, *Lon* is indispensable. *J. Bacteriol.*, **164**, 1124, (1985).

Meacock, P. et S. N. Cohen. Partitioning of plasmids during cell division : a cis-acting locus that accomplishes stable plasmid inheritance. *Cell.*, **20**, 529, (1980).

Melling, J., D. C. Ellwood et A. Robinson. Survival of R-factor carrying *E. coli* in mixed cultures in the chemostat. *FEMS Microbiol. Lett.*, **2**, 87, (1977).

Meyer, H.-P., H.-J. Kuhn, S. W. Brown et A. Fiechter. Production of human leukocyte interferon by *Escherichia coli*. In proceedings of the third european congress on biotechnology, verlag chemie gmbH, Weinheim, Federal Republic of Germany, **1**, 499, (1984).

Mosrati, R., N. Nancib et J. Boudrant. Variation and modelling of the probability of plasmid loss as a function of growth rate of plasmid-bearing cells of *Escherichia coli* during continuous cultures. *Biotechnol. Bioeng.*, **41**, 395, (1993).

Nancib, N., R. Mosrati et J. Boudrant. Study of population dynamics for a recombinant bacterium during continuous cultures : Application of data filtering and smoothing. *Biotechnol. Bioeng.*, **39**, 398, (1992).

- Nancib, N et J. Boudrant. Effect of growth rate on stability and gene expression of a recombinant plasmid during continuous culture of *Escherichia coli* in a non selective medium. *Biotechnol. Lett.*, **8**, 643, (1992).
- Nilsson J. et S. G. Skogman. Stabilization of *Escherichia coli* tryptophan-production vectors in continuous cultures : a comparison of three different systems. *Bio. Technol.*, **4**, 901, (1986).
- Noack, D., M. Toth, R. Geuther, G. Muller, K. Undisz, C. Hoeffmeier et S. Caspar. Maintenance and genetic stability of vector plasmids pBR322 and pBR325 in *E. coli* K12 strains grown in chemostat. *Mol. Gen. Genet.*, **184**, 121, (1981).
- Nomura, M., D. M., Bedwell, M. Yamagishi, J. R. Cole et J. M. Kolb. RNA polymerase and regulation of RNA synthesis in *E. coli* : RNA polymerase concentration stringent control, and ribosome feedback regulation. RNA polymerase and the regulation of transcription, Reznikoff W. S. (ed). Elsevier, New York., (1987).
- Nordström, K. et A. H. Helle. Maintenance of bacterial plasmids : comparison of theoretical calculations and experiments with plasmids R1 in *E. coli*. *Mol. Gen. Genet.*, **197**, 1, (1984).
- Nordström, K. S., S. Molin et H. Aagard-Hansen. Partitioning of plasmid R1 in *E. coli* : Kinetics of loss of plasmid derivatives deleted of the *par* region. *Plasmid*, **4**, 215, (1981).
- Nudel, B. C., B. Vosman, M. J. Teixeira de Mattos, K. J. Hellingwerf et O. N. Neijssel. Increased stability of recombinant plasmids Tn1000. Insertion in chemostat cultures of recombinant *E. coli* GT123. *Curr. Microbiol.*, **26**, 281, (1993).
- Nugent, M. E., S. B. Primrose et W. C. A. Tacon. The stability of recombinant DNA. *Dev. Ind. Microbiol.*, **24**, 271, (1983).
- Oka, T., S. Sakamoto, K. I. Moyoshi, T. Fuwa, K. Yoda, M. Yamasaki, G. Tamura et T. Miyake. Synthesis and secretion of human epidermal growth factor by *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **82**, 615, (1985).
- Oliver, D. Protein secretion in *Escherichia coli*. *Ann. Rev. Microbiol.* **39**, 615, (1985).

- Palva, I. Molecular cloning of alpha-amylase from *Bacillus amyloliquefaciens* and its expression in *Bacillus subtilis*. *Gene*, **19**, 81 (1982).
- Pan, J. G., J. S. Rhee et J. M., Lebeault. Physiological constraints in increasing biomass concentration of *Escherichia coli* β in fed batch culture. *Biotechnol. Lett.*, **9**, 89, (1987).
- Park, S., D. D. Y. Ryu et J. Y. Kim. Effect of growth rate on the performance of a two stage continuous culture system in a recombinant *E. coli*. fermentation. *Biotechnol. Bioeng.*, **36**, 493, (1990).
- Park, S. et D. D. Y. Ryu. Analysis of plasmid stability by the difference in colony size. *J. Ferment. Bioeng.*, **3**, 185, (1992).
- Patnaik, P. R. A simple analytical solution of a model for probabilistic plasmid loss and its implications in continuous cultures. *Biotechnol. Tech.*, **5**, 357, (1993).
- Peretti, S. W. et J. E. Bailey. Simulations of host-plasmid interactions in *Escherichia coli* : copy number, promoter strength and ribosome binding site strength effects on metabolic activity and plasmid gene expression. *Biotechnol. Bioeng.*, **29**, 316, (1987).
- Pierce J. et S. Gutteridge. Large scale preparation of ribulobiphosphate carboxylase from a recombinant system in *Escherichia coli* characterized by extreme plasmid instability. *Appl. Environ. Microbiol.*, **49**, 1094, (1985).
- Pinches, A., M. E. Louw, et T. G. Watson. Growth, plasmid stability and amylase production in batch fermentations using recombinant *Bacillus subtilis* strain. *Biotechnol. Lett.*, **7**, 621, (1985).
- Pirt, S. J., Principles of microbe and cell cultivation. Oxford : Blackwell Scientific Publication. (1975).
- Projan, S. J., S. Carleton et R. P. Novick. Determination of plasmid copy number by fluorescence densitometry. *Plasmid*, **9**, 182, (1983).

Reiling, H. A., H., Laurila et A. Fiechter. Mass culture of *Escherichia coli*: medium development for low and high density cultivation of *Escherichia coli* B/r in minimal and complex media. *J. Biotechnol.* **2**, 191, (1985).

Rosteck, P. R. et C. Hersberger. Selective retention of recombinant plasmids coding for human insulin. *Gene*, **25**, 29, (1983).

Roy, D. et P. Bhattacharya, A method for estimation of growth and instability parameters of batch recombinant cultures. *Biochem. Eng. J.*, **1**, 219, (1998).

Ryan, W. et S. J. Parulekar. Recombinant protein synthesis and plasmid instability in continuous cultures of *Escherichia coli* JM103 harboring a high plasmid copy number, *Biotechnol. Bioeng.*, **37**, 415, (1991).

San K. Y. et A. E. Weber. Data analysis of plasmid maintenance in a CSTR. *Biotechnol. Bioeng.*, **33**, 451, (1989).

Sardonini, C. A. et D. Di Basio. A model for growth of *Saccharomyces cerevisiae* containing a recombinant plasmid in selective media. *Biotechnol. Bioeng.*, **29**, 469, (1987).

Satyagal V. N. et P. Agrawal. A generalized model of plasmid replication. *Biotechnol. Bioeng.*, **33**, 1135, (1989).

Sayadi, S., N. Nasri, J. N. Barbotin et D. Thomas. Effect of environmental growth conditions on plasmid stability, plasmid copy number and catecho-2,3-dioxygenase activity in free and immobilized *E. coli* cells. *Biotechnol. Bioeng.*, **33**, 801, (1989).

Seo, J. H. et J. E. Bailey. Effect of recombinant plasmid content on growth properties on cloned gene product formation in *Escherichia coli*. *Biotechnol. Bioeng.*, **27**, 1668, (1985).

Seo J. H. et J. E. Bailey. Cell cycle analysis of plasmid-containing *E. coli* HB101 populations with flow cytometry. *Biotechnol. Bioeng.*, **30**, 297, (1987).

- Shuler M. L. On the use of chemically structured models for bioreactors. *Chem. Eng. Commun.*, **36**, 161, (1985).
- Siegel, R. et D. D. Y. Ryu. Kinetic study of instability of recombinant plasmid pPLc23trpA1 in *E. coli* using a two stage continuous culture system. *Biotechnol. Bioeng.*, **27**, 28, (1985).
- Smirnova G. V. et O. N. Oktyabr'skii. Effect of activity of primary proton pumps on growth of *Escherichia coli* in the presence of acetate. *Microbiology [USSR]*. **57**, 554, (1988).
- Stephens, M. L. et G. Lyberatos. Effect of cycling on the stability of plasmid bearing microorganisms in continuous culture. *Biotechnol. Bioeng.*, **31**, 464, (1988).
- Stephens, M. L., C. Christensen et G. Lyberatos. Plasmid stabilization of an *E. coli* culture through cycling. *Biotechnol. Bioeng.*, **8**, 1, (1992).
- Stormo, G. « Translation Initiation », *Maximising Gene Expression*, W. Reznikoff and L. Gold eds., Butterworths, (1986).
- Stueber, D. et H. Bujard. Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. *EMBO J.*, **1**, 1399, (1982).
- Sun, W.-J., C. Lee, H. A., George, A. L. Powell, M. E. Dahlgren, R. Greasham et C.-H. Park. Acetate inhibition on growth of recombinant *E. coli* and expression of fusion protein TGF α -PE40. *Biotechnol. Lett.*, **8**, 809, (1993).
- Talmadge, K. et W. Gilbert. Cellular location affects protein stability in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **81**, 1830, (1982).
- Thomas, C. M. Recent studies on the control of plasmid replication. *Biochem. Biophys. Acta.*, **949**, 253, (1988).

Tsai, L. B., M. Mann, F. Morris, C. Rotgers et D. Fenton. The effect of organic nitrogen and glucose on the production of recombinant human insulin-like growth factor in high cell density *Escherichia coli* fermentations. *J. Ind. Microbiol.*, **2**, 181, (1987).

Tsuchido, T., N. Katsui, A. Takeuchi, M., Takano et J. Shibasaka. Destruction of the outer membrane permeability barrier of *Escherichia coli* by heat treatment. *Appl. Environ. Microbiol.*, **50**, 298, (1985).

Uhlen, B. et K. Nordström. R plasmid gene dosage effects in *E. coli* K12 : Copy mutants of the R plasmids R1drd19. *Plasmid*, **1**, 1, (1977).

Walmsley, R. M., D. C. J. Gardner et S. G. Oliver. Stability of a cloned gene in yeast grown in chemostat culture. *J. Bacteriol.*, **162**, 176, (1983).

Weber, A. E., Y. Peng et K. Y. San. The effect of partition locus on plasmid stability and expression of a prolonged chemostat culture. *J. Biotechnol.*, **18**, 141, (1991).

Weber A. E. et San K. Y. Persistence and expression of the plasmid pBR322 in *Escherichia coli* K12 cultured in complex medium. *Biotechnol. Lett.* **9**, 757, (1987).

Weber, A. E. et K. Y. San. Enhanced plasmid maintenance in a CSTR upon square-wave oscillations in the dilution rate. *Biotechnol. Lett.*, **8**, 531, (1988).

Weber, A. E. et K. Y. San. Dynamics of plasmid maintenance in a CSTR upon square-wave perturbations in the dilution rate. *Biotechnol. Bioeng.*, **34**, 1104, (1989).

Williams, D. C., R. M. Van Franck, W. L. Muth et J. P. Burnett. Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Sci.*, **215**, 687, (1982).

Wittrup, K. D. et J. E. Bailey. A segregated model of recombinant multicopy plasmid propagation. *Biotechnol. Bioeng.*, **31**, 304, (1988).

Wouters, J. T. M. et J. G. Van Andel. R plasmid persistence in *E. coli* grown in chemostat cultures. *Antonie van leeuwenhoek*, **45**, 317, (1979).

Wouters, J. T., F. L. Driehuis, P. J. Polaczek, M.-L. H. A. Van Oppenraay et J. Van Andel. Persistence of the plasmide pBR322 in *E. coli* K12 grown in chemostat cultures. *Antonie van Leeuwenhoek.*, **46**, 353, (1980).

Yukawa, H., Y. Kurusu, M. Shimazu, H. Yamagata et M. Terasawa. Maintenance of a mini-F recombinant plasmid in *E. coli* and expression of its tryptophan synthase gene. *Process Biochem.*, **Dec**, 165, (1985).

Zabriske, D. W. et J. E. Arcuri. Factors influencing productivity of fermentations employing recombinant microorganisms. *Enzyme Microb. Technol.*, **8**, 706, (1986).



DEMANDE D'IMPRIMATUR**DIPLOME D'ETAT DE DOCTEUR EN
PHARMACIE**Présenté par **Bao-Linh LE**Sujet :**Modélisation de l'instabilité plasmidique
ségrégationnelle de souches recombinantes.
Application à *Escherichia coli*.**Jury :

Président : M. Alain BAGREL, Professeur

Juges : M. Joseph BOUDRANT, Directeur de recherche
M. Christian FONTEIX, Maître de ConférencesVu,
Nancy, le 20 septembre 2000

Le Président du jury

Le Directeur de Thèse

M. Alain BAGREL,
Professeur*M. Joseph BOUDRANT,*
Directeur de Recherche

Vu et approuvé

Nancy, le

07 novembre 2000

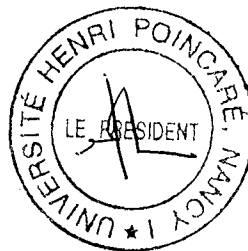
Doyen de Faculté de Pharmacie
de l'Université Henri Poincaré – Nancy I,

Vu,

Nancy, le

10 Nov. 2000 n° 929

Le président de l'Université Henri Poincaré – Nancy I,



N° d'identification :

PH Nancy 80 n°70

Modélisation de l'instabilité plasmidique ségrégationnelle de souches recombinantes. Application à *Escherichia coli*.

Thèse soutenue le 12 décembre 2000

Par Bao-Linh LE

RESUME :

Nous avons étudié en modélisation le phénomène d'instabilité plasmidique ségrégationnelle chez la souche recombinée *E.coli* HB101 (GADPH). Après avoir défini les modèles sur le plan théorique, nous avons réalisé les identifications paramétriques à partir des résultats expérimentaux d'une fermentation en continu, à l'aide d'un algorithme de type Simplex. Les modèles ainsi obtenus ont été confrontés à des résultats expérimentaux en culture batch. L'application dans un premier temps d'un modèle simple non ségrégué permet de simuler de façon satisfaisante l'évolution de la souche. Nous proposons ensuite un modèle ségrégué, basé sur le fait qu'une culture de cellules recombinantes est une culture mixte où coexistent plusieurs souches différenciées par leur nombre de copies de plasmides variant de 0 à N. Ce modèle explicite le bilan des fractions de biomasse et différencie la cinétique de croissance des souches bactériennes de celle de répllication plasmidique. De plus, nous avons défini pour ce modèle deux expressions de la probabilité de perte de plasmides, l'une basée sur une distribution de répartition uniforme, l'autre sur une distribution de type binomial. Le modèle de probabilité de répartition uniforme, même s'il permet de simuler de façon satisfaisante les données expérimentales, ne constitue pas un modèle suffisant du fait de la valeur improbable d'un de ses paramètres. Le modèle de probabilité de distribution binomiale permet lui d'avoir une approche des phénomènes d'instabilité, d'autres facteurs restant sans doute à prendre en compte pour une modélisation optimale.

MOTS CLES : Fermenteur – *Escherichia coli* – Organisme recombiné – Stabilité plasmidique – Modélisation mathématique

Directeur de thèse	Intitulé du laboratoire	Nature
Monsieur Joseph BOUDRANT Directeur de Recherche au C.N.R.S.	Laboratoire des Sciences du Génie Chimique (L.S.G.C.)	Expérimentale <input checked="" type="checkbox"/> Bibliographique <input type="checkbox"/> Thème 5

Thèmes :

1-Sciences fondamentales
3-Médicament
5-Biologie

2-Hygiène/Environnement
4-Alimentation-Nutrition
6-Pratique professionnelle