HAL
open science

Structure, biogenèse et expression de la protéine T du complexe de la Glycine décarboxylase des plantes supérieures (Pisum sativum)

Pierre Vauclare

To cite this version:

Pierre Vauclare. Structure, biogenèse et expression de la protéine T du complexe de la Glycine décarboxylase des plantes supérieures (Pisum sativum). Biologie cellulaire. Université Joseph-Fourier - Grenoble I, 1996. Français. NNT : . tel-01011285

HAL Id: tel-01011285

https://theses.hal.science/tel-01011285

Submitted on 23 Jun 2014

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

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

THESE

présentée par

Pierre VAUCLARE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER (GRENOBLE I)

(Arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Spécialité: BIOLOGIE

STRUCTURE, BIOGENESE ET EXPRESSION DE LA PROTEINE T DU COMPLEXE DE LA GLYCINE DECARBOXYLASE DES PLANTES SUPERIEURES

soutenue le 29 mars 1996, devant la commission d'examen

P. V. VIGNAIS	*Président*
J. M. GRIENENBERGER	*Rapporteur*
J. F. MOROT-GAUDRY	*Rapporteur*
R. DOUCE	
J. BOURGUIGNON	

Thèse préparée au sein du laboratoire de Physiologie Cellulaire Végétale
Commissariat à l'Energie Atomique de Grenoble

La terre nous en apprend plus long sur nous que tous les livres. Parce qu'elle nous résiste. L'homme se découvre quand il se mesure avec l'obstacle...

Saint-Exupéry
Terre des hommes

AVANT-PROPOS

J'exprime toute ma gratitude à Monsieur le Professeur Roland Douce, de l'Institut Universitaire de France pour m'avoir donné la chance de débuter au sein de son laboratoire de Physiologie Cellulaire Végétale du Commissariat à l'Energie Atomique de Grenoble. Ses vastes connaissances scientifiques, son enthousiasme communicatif et la confiance qu'il a pu me témoigner, m'ont permis d'effectuer mes travaux dans les meilleures conditions possibles.

J'adresse mes plus vifs remerciements à Messieurs les Professeurs P. V. Vignais, J. M. Grienenberger et J. F. Morot-Gaudry pour l'honneur qu'il me font en acceptant de juger cette thèse.

Je tiens à remercier tout particulièrement Jacques Bourguignon, Ingénieur CEA, pour m'avoir guidé durant ces années. Son expérience, ses conseils éclairés et nos nombreuses discussions souvent animées mais fructueuses m'ont permis une avancée rapide dans mes travaux.

Certaines études ont été réalisées en collaboration avec les Docteurs Eric Forest et Véronique Mérand au laboratoire de Spectrométrie de Masse des Protéines de L'IBS. Je tiens à les remercier sincèrement.

Je voudrais remercier chaleureusement David Macherel avec lequel j'ai eu une collaboration amicale. Ses compétences scientifiques, son intuition et ses connaissances anglicistes m'ont été précieuses.
Je ne peux oublier Nicolas Diallo qui contribua à cette étude. Je lui témoigne ici toute ma gratitude pour la compétence et la gentillesse dont il a fait preuve.

Je tiens à exprimer toute ma sympathie aux autres membres du laboratoire pour leurs conseils éclairés.

Enfin, que tous ces pionniers de la recherche qui ont su me communiquer leur enthousiasme et m'ont soutenu au cours de ces années, reçoivent toute mon amitié.

SOMMAIRE

SOMMAIRE

INTRODUCTION 1

DONNEES BIBLIOGRAPHIQUES 3

I - LA PHOTORESPIRATION CHEZ LES ORGANISMES CHLOROPHYLLIENS.....3

A - RAPPEL SUR LA PHOTORESPIRATION CHEZ LES VEGETAUX A METABOLISME DE TYPE C3..... 3

B - ACTIVITE DE LA PHOTORESPIRATION CHEZ LES ORGANISMES POSSEDANT UN SYSTEME DE CONCENTRATION DU CO_2 7

1 - Les plantes à métabolisme de type C4.....8
2 - Des espèces intermédiaires: les plantes à métabolisme de type C3-C4.....9
3 - Quelques aspects de la photorespiration chez les algues.....10

II - L'OXYDATION DE LA GLYCINE PAR UN COMPLEXE MULTI-ENZYMATIQUE: LA GLYCINE DECARBOXYLASE (GDC).....11

A - OXYDATION DE LA GLYCINE CHEZ LES VEGETAUX A METABOLISME DE TYPE C3....11

B - QUELQUES RAPPELS SUR L'OXYDATION DE LA GLYCINE CHEZ LES ORGANISMES NON CHLOROPHYLLIENS.....13

III - STRUCTURE MOLECULAIRE ET MECANISME REACTIONNEL DU COMPLEXE DE LA GLYCINE DECARBOXYLASE.....14

*A - STRUCTURE MOLECULAIRE DE LA **GDC**..... 14*

*B - MECANISME REACTIONNEL DU COMPLEXE DE LA **GDC**.....15*

1 - Description.....15
2 - Régulations métaboliques.....16
3 -Inhibiteurs spécifiques des protéines du complexe de la GDC.....17

IV - LA PROTEINE T: PROTEINE A TETRAHYDROFOLATE.....18

A - REACTION CATALYSEE PAR LA PROTEINE T.....18

B - LE TETRAHYDROFOLATE.....19

C - INTERACTION ENTRE LA PROTEINE T ET LE TETRAHYDROFOLATE..... 20

V - OBJECTIFS DU TRAVAIL.....21

MATERIEL ET METHODES 23

I - MATERIELEXPERIMENTAL.....23

A - MATERIEL VEGETAL.....23

B - CONDITIONS DE CULTURE.....24

II - ISOLEMENT DES ORGANITES.....24

A - EXTRACTION ET PURIFICATION DES MITOCHONDRIES DE FEUILLES DE POIS.....24

1 - Composition des milieux.....24
2 - Préparation des mitochondries brutes.....25
3 - Purification des mitochondries.....25

B - METHODES D'ETUDES DES MITOCHONDRIES ISOLEES.....27

1 - Mesure des activités oxydatives.....27
2 - Test d'intégrité des mitochondries.....27

C - ISOLEMENT DE CHLOROPLASTES DE FEUILLES DE POIS (METHODE DE MOURIOUX ET DOUCE 1981).....28

1 - Composition des milieux.....28
2 - Purification des chloroplastes.....28

D - ISOLEMENT DES NOYAUX.....29

1 - Méthode de Watson et Thompson (1986).....29
2 - Méthode de Green *et al.* (1989).....29
a - Composition du milieu.....29
b - Purification des noyaux.....29
c - Préparation de l'extrait nucléaire.....30

III - MESURES D'ACTIVITES ENZYMATIQUES.....30

A - MESURE D'ACTIVITE DE LA GLYCINE DECARBOXYLASE.....30

B - MESURE DE L'ACTIVITE DE LA RIBULOSE 1,5 BISPHOSPHATE *CARBOXYLASE/ OXYGENASE (SELON LORIMER 1982).....31*

IV - PURIFICATION DE LA PROTEINE T IMPLIQUEE DANS LE COMPLEXE DE LA GLYCINE DECARBOXYLASE.....31

A - PREPARATION DE L'EXTRAIT MATRICIEL.....31

B - TECHNIQUE DE PURIFICATION DE LA PROTEINE T.....32

1 - Chromatographie sur colonne Séphacryl S-300.....32
2 - Chromatographie sur colonne mono Q HR 5/5.....32

V - TECHNIQUES D'ETUDE DES PROTEINES.....33

A - EXTRACTION DES PROTEINES TOTALES DE FEUILLES.....33

B - DOSAGE DES PROTEINES.....33

C - ELECTROPHORESE DES PROTEINES SUR GEL DE POLYACRYLAMIDE EN CONDITIONS DENATURANTES ET NON DENATURANTES.....33

1 - Préparation des échantillons protéiques.....33
2 - Composition des gels.....34
3 - Coloration des gels.....34

D - IMMUNODETECTION DES PROTEINES PAR LA TECHNIQUE DU WESTERN-BLOT.....34

E - MICROSEQUENÇAGE DE LA PROTEINE T.....35

1 - Coupure au bromure de cyanogène.....36
2 - HPLC préparative.....37

VI - SPECTROMETRIE DE MASSE.....38

A - RAPPEL THEORIQUE.....38

B - DESCRIPITON TECHNIQUE.....39

1 - La spectrométrie de masse FAB (Fast Atom Bombardment).....39
a - Principe.....39
b - Domaine d'application.....40
2 - La spectrométrie de masse par électrospray (ESI).....41
a - Principe.....41
b - Domaine d'application.....42
3 - Principe d'un analyseur de masse.....43

C - PROTOCOLE OPERATOIRE POUR L'ETUDE DE LA PROTEINE T.....43

1 - FAB/MS.....43
2 - HPLC/ESI/MS.....44

VII - TECHNIQUES GENERALES D'ETUDE DES ACIDES NUCLEIQUES.....44

A - TECHNIQUES D'ANALYSE DES ACIDES DEOXYRIBONUCLEIQUES.....45

1 - Extraction de l'ADN nucléaire de pois.....45
2 - Techniques d'hybridation.....45
a - Marquage de l'ADN.....45
b - Transfert de l'ADN sur membrane de nylon (Thomas, 1980).....45
c - Préhybridation, hybridation et rinçage.....46
3- Amplification in vitro de fragments d'ADN.....46
4 - Technique de séquençage de l'ADN 48
a - Réaction de séquençage.....48
b - Electrophorèse sur gel de polyacrylamide.....49
5 - Synthèse des oligonucléotides.....49

B - TECHNIQUES D'ANALYSE DES ACIDES RIBONUCLEIQUES.....50

1 - Extraction des ARN totaux.....50
2 - Electrophorèse des ARN sur gel d'agarose et transfert sur membrane de nylon.....50
3 - Hybridation, quantification et coloration des ARN.....51
4 - Détermination du site d'initiation de la transcription par la méthode d'extension d'amorce.....52

C - ISOLEMENT ET CARACTERISATION D'UN ADNc ET D'UN GENE.....53

1 - Criblage, à l'aide d'anticorps polyclonaux et d'une sonde, d'une banque d'expression dans lambda gt 11.....53
a - Utilisation d'anticorps.....54
b - Utilisation d'une sonde.....54
2 - Criblage d'une banque génomique dans lambda EMBL 3 à l'aide d'une sonde.....55
3- Isolement de l'ADN des phages.....55

D - SOUS-CLONAGE DE L'ADNc.....56

1 - Préparation des bactéries compétentes.....56
a - Première méthode.....56
b - Deuxième méthode.....56
2 - Transformation des bactéries compétentes.....57
3 - Préparation de l'ADN des plasmides 58

VIII - TECHNIQUE D'EXPRESSION PROCARYOTIQUE.....58

A - CONSTRUCTION DU VECTEUR D'EXPRESSION PROCARYOTIQUE.....58

B - MUTAGENESE DIRIGEE.....59

C - CLONAGE DANS LES VECTEURS D'EXPRESSION DU GENE argU.....59

D - CONDITIONS D'EXPRESSION.....60

IX - EXPERIENCE DE GEL RETARD.....61

A - FABRICATION ET MARQUAGE DES SONDES.....62

1 - Hybridation des oligonucléotides monomériques complémentaires.....62
2 - Marquage et purification de l'oligonucléotide double brin.....62

B - PROTOCOLE EXPERIMENTAL D'UN GEL RETARD.....63

CHAPITRE I: DETERMINATION ET ANALYSE DE LA STRUCTURE PRIMAIRE DE LA PROTEINE T 65

I - PURIFICATION DE LA PROTEINE T.....65

A - ETAPE DE PURIFICATION.....65

B - PREPARATION D'ANTICORPS POLYCLONAUX DIRIGES CONTRE LA PROTEINE T ET ANALYSE IMMUNOLOGIQUE.....68

C - MICROSEQUENÇAGE DE LA PROTEINE T.....69

II - ISOLEMENT DE L'ADN COMPLEMENTAIRE (ADNc) CORRESPONDANT A LA PROTEINE T.....70

A - CRIBLAGE D'UNE BANQUE D'EXPRESSION DANS λGT 11 A L'AIDE DES ANTICORPS DIRIGES SPECIFIQUEMENT CONTRE LA PROTEINE T.....70

1 - Caractérisation d'un ADNc de taille incomplète correspondant à la protéine T (TD3).....70

2 - Validation de l'identité du clone TD3.....72

B - CRIBLAGE DE LA BANQUE D'EXPRESSION DANS λGT 11 A L'AIDE DE L'ADNc TD3'.....73

1 - Caractérisation d'un ADNc de taille complète correspondant à la protéine T.....73

2 - Validation et détermination de la séquence primaire de la protéine T déduite de l'ADNc.....74

a - Stratégie de séquençage de l'ADNc codant pour la protéine T.....74

b - Analyse de la séquence de l'ADNc.....75

c - Composition globale en acides aminés de la protéine T mature.....78

III - ETUDE DE LA PROTEINE T A L'AIDE DE LA SPECTROMETRIE DE MASSE.....79

A - DETERMINATION DE LA MASSE MOLECULAIRE DE LA PROTEINE T NATIVE.....80

B - ANALYSE DES PRODUITS DE LA DIGESTION AU BROMURE DE CYANOGENE DE LA PROTEINE T.....81

1 - Analyse des fragments de la protéine T par couplage de la chromatographie liquide avec la spectrométrie de masse par électrospray ou ionspray (HPLC/ESI-MS).....82

2 - Analyse par FAB (Fast Atom Bombardment) et dégradation d'Edman des fragments purifiés par HPLC.....85

IV - ANALYSE DE LA SEQUENCE PRIMAIRE DE LA PROTEINE T.....86

A - QUELQUES PARAMETRES BIOCHIMIQUES DE LA PROTEINE T, DEDUITS DE LA SEQUENCE EN ACIDES AMINES.....86

1 - Estimation du pH isoélectrique (pI) et de la charge globale de la protéine T mature au pH matriciel.....86

2 - Analyse du degré d'hydrophobicité de la protéine T.....87

3 - Caractérisation du peptide de transit.....88

B - RECHERCHE D'HOMOLOGIES DE SEQUENCES.....90

V - DISCUSSION ET CONCLUSION.....93

CHAPITRE II: ETUDE DE L'EXPRESSION DE LA PROTEINE T DE LA MISE EN PLACE DU COMPLEXE DE LA GLYCINE DECARBOXYLASE AU COURS DU DEVELOPPEMENT 99

I - DISTRIBUTION TISSULAIRE ET INFLUENCE DE LA LUMIERE SUR L'EXPRESSION DES DIFFERENTES PROTEINES DU COMPLEXE DE LA GLYCINE DECARBOXYLASE.....99

*A - ANALYSE PAR WESTERN-BLOT DE LA PRESENCE DE LA PROTEINE T ET DES AUTRES PROTEINES DE LA **GDC** DANS DES MITOCHONDRIES ISOLEES DE DIFFERENTS TISSUS.....99*

B - ANALYSE PAR NORTHERN-BLOT DU NIVEAU DES TRANSCRITS CORRESPONDANT A LA PROTEINE T.....100

C - DISCUSSION.....102

II - ETUDE DE LA MISE EN PLACE DU COMPLEXE DE LA GLYCINE DECARBOXYLASE AU COURS DU DEVELOPPEMENT DES FEUILLES DE POIS.....103

A - ETUDE DE LA MISE EN PLACE DES PROTEINES IMPLIQUEES DANS L'OXYDATION DE LA GLYCINE AU COURS DU DEVELOPPEMENT.....104

1 - Evolution de l'oxydation de la glycine par des mitochondries isolées aux différents stades de développement.....104

2 - Analyse électrophorétique et détection immunologique des protéines impliquées dans l'oxydation de la glycine.....106

3 - Modification de la densité des mitochondries au cours du développement du pois.....108

B - ETUDE DE LA MISE EN PLACE DE LA RIBULOSE 1,5 BISPHOSPHATE CARBOXYLASE/OXYGENASE DANS LES CHLOROPLASTES.....109

C - EVOLUTION AU COURS DU DEVELOPPEMENT DE L'EXPRESSION DES TRANSCRITS CORRESPONDANT AUX DIFFERENTES PROTEINES DU COMPLEXE MULTIENZYMATIQUE DE LA GDC, DE LA RUBISCO ET DE LA PROTEINE cab.....111

D - EVOLUTION DES TRANSCRITS AU COURS DE LA CROISSANCE DU POIS APRES EXPOSITION DES FEUILLES A L'OBSCURITE A DIFFERENTS STADES DU DEVELOPPEMENT.....114

E - DISCUSSION-CONCLUSION.....116

CHAPITRE III: CLONAGE ET CARACTERISATION DU GENE NUCLEAIRE CODANT POUR LA PROTEINE T 121

I - ISOLEMENT ET IDENTIFICATION DU GENE CODANT POUR LA PROTEINE T....121

A - DETECTION PAR SOUTHERN-BLOT DU GENE CODANT POUR LA PROTEINE T DANS LE GENOME DE POIS....121

B - CRIBLAGE DE LA BANQUE GENOMIQUE DANS λEMBL3 A L'AIDE DE L'ADNc TB2: PURIFICATION ET AMPLIFICATION....123

1 - Caractérisation du gène codant pour la protéine T (gdcT)....123
2 - Cartographie du clone....124
3 - Stratégie de séquençage du gène *gdcT*....126

II - ANALYSE DE LA STRUCTURE DU GENE CODANT POUR LA PROTEINE T...127

A - ORGANISATION DU GENE gdcT....127

1 - Détermination de la séquence nucléotidique....127
2 - Analyse de la structure primaire de la protéine T....130

B - DETERMINATION DU SITE D'INITIATION DE LA TRANSCRIPTION DU GENE CODANT POUR LA PROTEINE T....132

1 - Résultat....132
2 - Discussion....134

III - ETUDE DE LA REGION PROMOTRICE....137

A - IDENTIFICATION DE TROIS POTENTIELS SEQUENCES CONSENSUS DE REGULATION DANS LA REGION PROMOTRICE DE gdcT....137

1 - Caractérisation d'une région riche en nucléotides AT....137
2 - Présence d'une boîte GT-1 dans la région promotrice du gène *gdcT*....138
3 - La région 5' en amont de *gdcT* possède l'élément GATA....139

B - EXPERIENCE D'INTERACTION PROTEINES/ADN....141

1 - Relation entre la région riche en nucléotides AT du promoteur de *gdcT* et l'élément cis, riche en nucléotides AT, du promoteur de rbcS 3,6....142
2 - Relation entre la boîte II située dans la région promotrice de *gdcT* et celle située dans la région 5' du gène codant pour rbcS3A....144

IV - CONCLUSION....146

CHAPITRE IV: EXPRESSION DE LA PROTEINE T DU COMPLEXE DE LA GLYCINE DECARBOXYLASE CHEZ ESCHERICHIA COLI 149

I - ELABORATION D'UN GENE SYNTHETIQUE CORRESPONDANT A LA PROTEINE T MATURE EN VUE DE SON EXPRESSION DANS ESCHERICHIA COLI.....149

A - RAPPEL SUR LE MECANISME DE SUREXPRESSION DE PROTEINES DANS UN ORGANISME PROCARYOTE.....149

B - CONSTRUCTION ET CLONAGE DE LA FORME MATURE DE LA PROTEINE T.....150

II - UTILISATION DU SYSTEME D'EXPRESSION pET (STUDIER *et al.*, 1990).....152

A - PRINCIPE DE LA SUREXPRESSION DANS ESCHERICHIA COLI.....152

B - TENTATIVE D'EXPRESSION DE LA PROTEINE T MATURE DANS E. COLI EN UTILISANT LE VECTEUR D'EXPRESSION pET.....153

C - TENTATIVE DE COEXPRESSION DES PROTEINES T ET H MATURES.....155

1 - Construction du vecteur d'expression procaryote pET-HTM contenant en tandem les séquences codant pour les protéines T et H matures.....155

2 - Essai de coexpression de la protéine H et T dans Escherichia coli à l'aide des vecteurs pET-HTA et pET-HTU.....156

D - MODIFICATION DU SYSTEME D'EXPRESSION DANS E.COLI PAR INSERTION DANS LES VECTEURS pET-TM ET pET-HTA ET pET-HTU DU GENE argU CODANT POUR UN ARNt^{ARG} RARE.....159

1 - Insertion du gène argU dans les vecteurs pET-TM, pET-HTA et pET-HTU.....159

2 - Expression et coexpression des protéines H et T dans Escherichia coli à l'aide des vecteurs pETY-TM, HTA, HTU.....160

III - DISCUSSION - CONCLUSION.....161

CONCLUSION GENERALE ET PERSPECTIVES 163

REFERENCES BIBLIOGRAPHIQUES 167

PUBLICATIONS ET PARTICIPATION A DES CONGRES 187

I - PUBLICATIONS.....187

A - ARTICLE PUBLIE DANS UNE REVUE A COMITE DE LECTURE.....187

B - ARTICLES EN PREPARATION.....187

II - PARTICIPATION A DES CONGRES.....188

A - SEMINAIRE.....188

B - PRESENTATION DE POSTERS.....188

RESUME 189

LISTE DES ABREVIATIONS

ADN: acide désoxyribonucléique
ADNc: acide désoxyribonucléique complémentaire
ADP: adénosine 5'-diphosphate
ARN (polyA+ ou m): acide ribonucléique (polyadénylé ou messager)
ATP: adénosine 5'-triphosphate
BCIP: 5-bromo-4-chloro-3-indolyl-phosphate
BSA: bovin serum albumin
(d)dATP: (di)désoxyadénosine 5'-triphosphate
(d)dCTP: (di)désoxycytosine 5'-triphosphate
(d)dGTP: (di)désoxyguanosine 5'-triphosphate
(d)dTTP: (di)désoxythymidine 5'-triphosphate
DMSO: diméthyl sulfoxyde
DNase: désoxyribonucléase
dNTP: désoxynucléotide 5'-triphosphate
DTT: dithiothréitol
EDTA: acide éthylène diamine tétraacétique
HEPES: acide N-2-hydroxyéthypipérazine-N'-2-éthanesulfonique
HPLC: high performance liquid chromatography
IPTG: isopropyl-β-D-thiogalactopyranoside
kDa: kilodalton
kpb: kilopaire de base
m/v: masse/volume
MOPS: acide 3-(N-morpholino)propane sulfonique
NAD(H): nicotinamide adénine dinucléotide (forme oxydée/forme réduite)
NBT: nitrobleu de tétrazolium
pb: paire de bases
PIPES: piperazine-N, N'-bis(2-ethanesulfonic acid); 1,4-Piperazinediethanesulfonic acid
PMSF: phénylméthylsulfonyl fluoride
PVP: polyvinylpyrrolidone
RNase: ribonucléase
SDS: sodium dodecylsulfate
TEMED: NNN'N'-tétra-méthyl-1,2-diamino-éthane
TFA: acide trifluoroacétique
Tris: tris(hydroxyméthyl)-aminométhane hydrochloride
UV: ultraviolet
v/v: volume/volume
X-gal: 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

RESIDUS D'ACIDES AMINES

A (Ala): alanine
C (Cys): cystéine
D (Asp): acide aspartique
E (Glu): acide glutamique
F (Phe): phénylalanine
G (Glu): glycine
H (His): histidine
I (Ile): isoleucine
K (Lys): lysine
L (Leu): leucine
M (Met): méthionine
N (Asn): asparagine
P (Pro): proline
Q (Gln): glutamine
R (Arg): arginine
S (Ser): sérine
T (Thr): thréonine
V (Val): valine
W (Trp): tryptophane
Y (Tyr): tyrosine

INTRODUCTION

INTRODUCTION

Tous les organismes autotrophes, chimiotrophes ou phototrophes, sont capables de fabriquer des éléments carbonés à partir du gaz carbonique (CO_2). Chez l'algue la plus primitive comme chez les organismes les plus évolués tels que les angiospermes, se déroule grâce à l'énergie solaire, un processus permettant la conversion de l'énergie lumineuse en énergie chimique. Ainsi, la photosynthèse permet grâce à la réduction du CO_2, la formation de substances organiques et de molécules à potentiel énergétique élevé telles que l'ATP et le NADPH, indispensables à la croissance des organismes photosynthétiques.

Chez les plantes, les nombreuses variations des conditions extérieures (la température, l'humidité, la luminosité...) exercent une pression constante sur leur métabolisme, en particulier sur leur capacité à photosynthétiser. Face à cela, les végétaux réagissent de façon ponctuelle et rapide. En effet, dans des conditions de forte chaleur et de faible brassage de l'air, les stomates, situés à la face inférieure de l'épiderme foliaire et permettant les échanges gazeux, se ferment afin d'empêcher une évaporation excessive de l'eau. Ce phénomène entraîne au sein de l'appareil foliaire un affaiblissement de la teneur en CO_2. Dans ces conditions, les plantes à métabolisme de type C3 se mettent à incorporer de façon compétitive de l'oxygène à la place du CO_2, déclenchant ainsi le cycle de la photorespiration. La fonction essentielle de ce métabolisme est de dissiper en permanence l'excès d'énergie lumineuse reçue par les chloroplastes tout en permettant de recycler des molécules carbonées.

Ce cycle fait intervenir trois organites de la cellule végétale qui sont le chloroplaste, le péroxysome et la mitochondrie. C'est au sein de la matrice mitochondriale qu'a lieu une étape primordiale du cycle photorespiratoire: l'oxydation de la glycine par un complexe multienzymatique, la glycine décarboxylase (GDC). Ce complexe composé de 4 protéines (P, H, T et L), catalyse le clivage d'une molécule de glycine en libérant une molécule de CO_2 et de NH_3 et en fixant le groupement méthylène restant sur un cofacteur, le tétrahydrofolate. Une deuxième enzyme mitochondriale, la SHMT, intervient alors pour régénérer le tétrahydrofolate, en catalysant le transfert du groupement méthyl du méthylène tétrahydrofolate sur une deuxième molécule de glycine, formant ainsi une molécule de sérine.

Le complexe de la glycine décarboxylase a été largement étudié chez les procaryotes, les animaux et l'homme; de nombreuses équipes de recherche se sont penchées sur l'étude de son mécanisme réactionnel et de sa structure, notamment chez l'homme, car ce complexe peut être impliqué dans une maladie génétique grave.

Chez les végétaux, Neuburger et Douce, (1977) ont montré pour la première fois que les mitochondries provenant de tissus foliaires étaient capables d'oxyder très rapidement la glycine. En 1983, Sarojini et Oliver ont mis en évidence l'activité de la GDC par réaction d'échanges dans un extrait acétonique provenant de mitochondries de tissus foliaires de pois et montré que la GDC était fortement liée aux membranes. En 1986, Neuburger *et al.*, montraient qu'un extrait matriciel obtenu par ouverture ménagée des mitochondries de feuilles de pois était capable d'oxyder la glycine. Depuis, une part essentielle des efforts du laboratoire de Physiologie Cellulaire Végétale a été consacrée à la compréhension des mécanismes d'oxydation de la glycine par les mitochondries de tissus foliaires et à la structure des protéines impliquées. Cette étude a été rendue possible grâce à la forte représentation du complexe de la glycine décarboxylase au sein de la matrice mitochondriale associée à la maîtrise de la préparation de quantités importantes de mitochondries à partir de feuilles de pois. La caractérisation et la purification de toutes les protéines constituant ce complexe (Bourguignon *et al.*, 1988) ont permis de développer la plupart des outils (anticorps et ADNc) nécessaires à une étude moléculaire de l'expression des protéines du complexe de la glycine décarboxylase lors de la mise en route du processus photorespiratoire. En particulier, deux ADNc correspondant aux protéines H et L du complexe de la GDC ont été clonés au laboratoire (Macherel *et al.*, 1990 et Bourguignon *et al.*, 1992).

Les travaux présentés dans ce mémoire ont eu pour objectif d'approfondir nos connaissances concernant la protéine T impliquée dans le complexe de la glycine décarboxylase en faisant appel aux techniques de la biophysique et de la biologie moléculaire.

DONNEES BIBLIOGRAPHIQUES

DONNEES BIBLIOGRAPHIQUES

Après un rappel général sur la photorespiration dans le règne végétal et sur l'implication du complexe de la glycine décarboxylase dans le métabolisme des cellules eucaryotes et procaryotes, nous aborderons l'état des connaissances au niveau biochimique et moléculaire des différentes protéines du complexe de la glycine décarboxylase et en particulier de la protéine T, protéine à tétrahydrofolate jouant un rôle clef au sein du complexe de la glycine décarboxylase.

I - LA PHOTORESPIRATION CHEZ LES ORGANISMES CHLOROPHYLLIENS

A - RAPPEL SUR LA PHOTORESPIRATION CHEZ LES VEGETAUX A METABOLISME DE TYPE C3

Grâce à la photosynthèse, les plantes sont capables d'utiliser l'énergie lumineuse pour synthétiser des glucides à partir du gaz carbonique et de l'eau. La ribulose-1,5-bisphosphate carboxylase-oxygénase (Rubisco) joue un rôle fondamental dans ce processus en catalysant une étape clef du cycle de Benson et Calvin. Cette enzyme bifonctionnelle (susceptible de se comporter comme une oxygénase ou comme une carboxylase) est responsable de la fixation du gaz carbonique conduisant à la synthèse de 3-phosphoglycérate selon la réaction suivante:

$$CO_2 + \text{ribulose-1,5-bisphosphate} + H_2O \xrightarrow[\text{fonction } \textit{carboxylase}]{\textit{RUBISCO}} 2 \text{ acide 3-phosphoglycérique}$$

Dans des conditions atmosphériques normales, l'oxygène présent à forte concentration dans l'air (21%) est un substrat compétitif du gaz carbonique (0.03%). Ainsi, l'activité oxygénase de la Rubisco est déjà présente dans les conditions normales de température et de lumière. Cependant, dans des conditions de température élevée, les stomates de l'épiderme foliaire des plantes se ferment (afin d'éviter une évaporation excessive de l'eau) entraînant au sein de l'appareil foliaire une diminution de la concentration en CO_2. Dans ces conditions,

l'augmentation du rapport O_2/CO_2 dans le milieu cellulaire augmente l'activité oxygénase de la Rubisco qui incorpore de façon compétitive une molécule d'oxygène à la place d'une molécule de CO_2 (Boowes *et al.*, 1971) suivant la réaction:

$$O_2 + \text{ribulose-1,5-bisphosphate} \xrightarrow[\text{fonction } \textit{oxygénase}]{\textit{RUBISCO}} \text{2-phosphoglycolate} + \text{acide 3 - phosphoglycérique}$$

La synthèse de 2- phosphoglycolate, molécule en C2 et d'une seule molécule en C3 (l'acide 3-phosphoglycérique), entraînerait le désamorçage du cycle de Benson et Calvin (par défaut de molécule à 3 carbones), si les molécules de phosphoglycolate n'étaient pas recyclées vers le chloroplaste en molécules de glycérate-3-phosphate. Les plantes à métabolisme de types C3 comme vraisemblablement les cellules du fourreau des plantes à métabolisme de type C4 (Cf. paragraphe I-B) assurent ce recyclage par un type particulier de métabolisme oxydatif (non observé à l'obscurité): la photorespiration ou métabolisme de type C2 qui permet de dissiper en permanence l'énergie reçue lors de la photosynthèse, sans synthèse nette de molécules carbonées. En effet, ce métabolisme oxydatif qui s'est développé au cours de l'adaptation des plantes passant d'une atmosphère réductrice à un milieu de plus en plus riche en O_2, est indispensable à la survie des plantes. Ce mécanisme évite notamment, la formation de radicaux superoxydes et d'oxygène singulet, espèces très réactives qui provoquent une détérioration irréversible de l'appareil photosynthétique. Le cycle de la photorespiration est donc indispensable à la survie de la plante.

La première observation de la photorespiration a été effectuée en 1920 par Warburg qui a constaté que l'O_2 inhibait la fixation du CO_2 chez les algues vertes (*chlorella*). Mais c'est en 1950 que Benson et Calvin rapportent pour la première fois que le glycolate, la glycine et la sérine apparaissent successivement au cours de la photosynthèse des plantes supérieures à métabolisme de type C3. Par la suite, l'utilisation de molécules marquées au ^{14}C, notamment le glycolate, a permis de définir toutes les étapes impliquées dans la photorespiration (Tolbert, 1971a, b).

Trois compartiments cellulaires interviennent dans le cycle photorespiratoire: le chloroplaste, le péroxysome et la mitochondrie (figure 1).

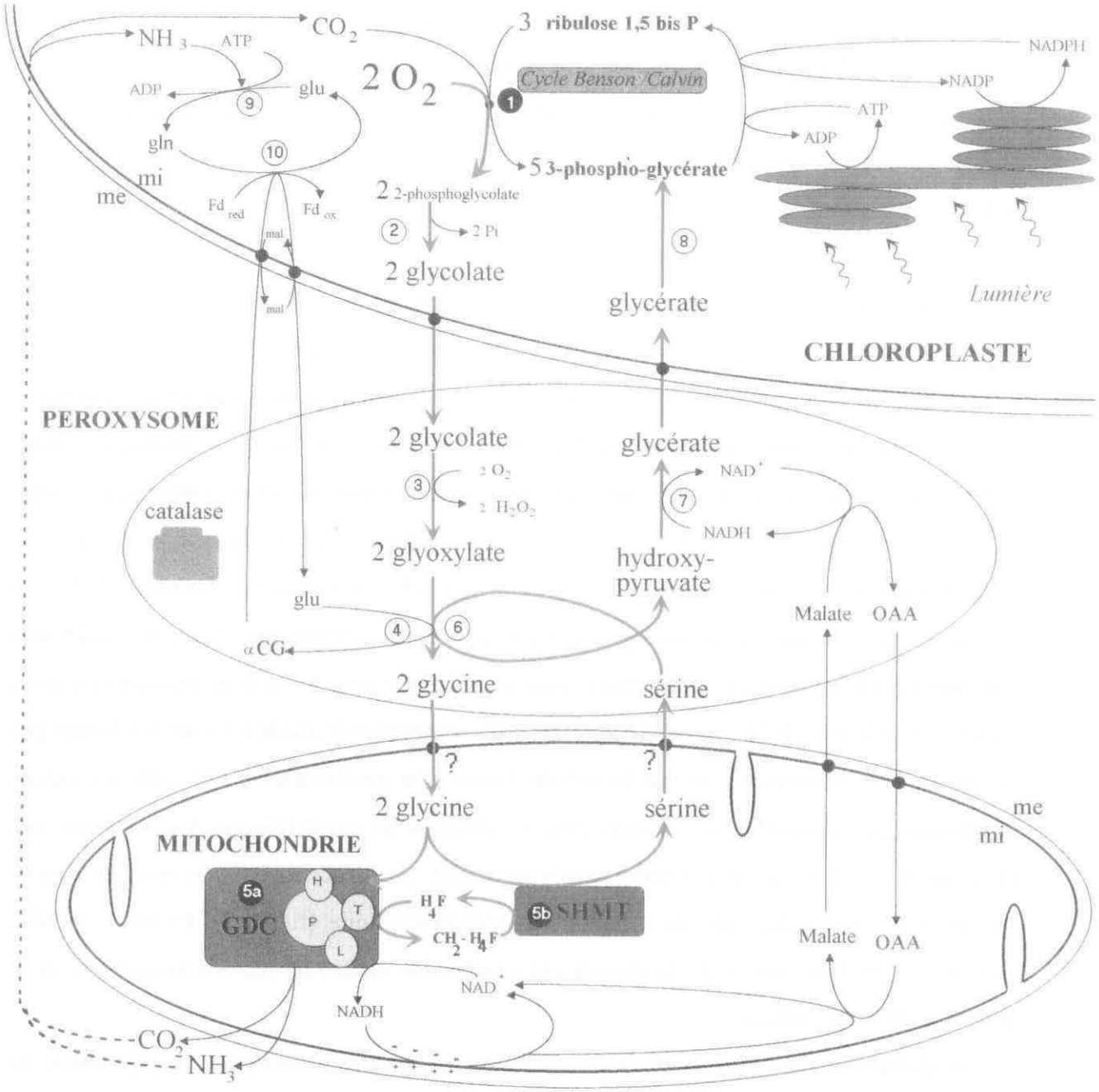


Figure 1 : Le cycle de la photorespiration chez les plantes à métabolisme de type C3
Ce schéma montre les trois compartiments: chloroplastique, péroxysomal et mitochondrial impliqués dans le cycle de la photorespiration dans les cellules de l'appareil foliaire végétal. L'oxydation de la glycine en sérine a lieu au sein de la matrice mitochondriale. (1) ribulose 1, 5-bisphosphate carboxylase/oxygénase (Rubisco); (2) P-glycolate phosphatase; (3) glycolate oxydase; (4) glutamate-glyoxylate aminotransférase; (5a) glycine-décarboxylase; (5b) sérine hydroxyméthyltransférase; (6) sérine-glyoxylate aminotransférase; (7) NADH hydroxypyruvate transférase; (8) glycérate kinase; (9) glutamine synthétase; (10) glutamate synthase. Les abrévations correspondent aux molécules suivantes: glutamate (glu); glutamine (gln); oxaloacétate (OAA); α-cétoglutarate (αCG); ferrédoxine (Fd); malate (mal.); membrane interne (mi); membrane externe (me).

Le cycle débute au sein du chloroplaste par la déphosphorylation du phosphoglycolate en glycolate, lequel quitte le chloroplaste pour pénétrer dans le péroxysome où il est converti en glycine par l'action de la glycolate oxydase et d'une glutamate-glyoxylate aminotransférase (Kisaki et Tolbert, 1969). La glycine migre alors vers la mitochondrie où elle est convertie en sérine, ammoniac (NH_3) et CO_2. La sérine regagne ensuite le péroxysome où une transamination et une réduction, catalysées respectivement par la sérine-glyoxylate aminotransférase et l'hydroxypyruvate réductase, la transforment en glycérate. Finalement, le glycérate retourne dans le chloroplaste où il est phosphorylé par la glycérate kinase pour donner l'acide 3-phosphoglycérique qui intègre le pool des trioses monophosphates impliqué dans le cycle de Benson et Calvin. Le flux de carbone au cours du cycle photorespiratoire représente 25 à 40 % du CO_2 fixé au cours de la photosynthèse (Tolbert, 1980; Keys, 1980; Lorimer et Andrews, 1981). Cette réaction est accompagnée par la libération de NH_3 en quantité équimolaire avec le CO_2. Selon Husic *et al.* (1987) et Lorimer et Andrews (1981), la libération de CO_2 par la photorespiration est considérée comme une perte de molécules carbonées pour la plante car seulement la moitié sont repris dans le cycle de Benson et Calvin avant de s'échapper de la feuille. Cependant, il est fort probable que le CO_2 ne s'échappe pas et soit immédiatement recyclé par la cellule. Quant à la production de NH_3, elle atteindrait rapidement des concentrations toxiques pour la cellule si les tissus foliaires n'avaient pas mis en place un système qui permet une assimilation rapide et efficace du NH_3: le cycle de l'azote qui est une étape importante du métabolisme photorespiratoire (figure 2). En effet, le NH_3 généré lors de l'oxydation de la glycine est rapidement réassimilé par l'intermédiaire de 2 enzymes chloroplastiques:

- la glutamine synthétase (GS; EC 6.3.1.2), ATP dépendante, catalyse l'incorporation de NH_3 sur une molécule de glutamate pour produire de la glutamine (Woo *et al.*, 1978; Tillberg *et al.*, 1977).

- la glutamine-oxoglutarate aminotransférase ou glutamate synthase (GOGAT; EC 1.4.7.1) qui utilise la ferrédoxine comme réducteur, catalyse le transfert du groupement aminé de la glutamine sur le 2-oxoglutarate formant ainsi le glutamate (Keys *et al.*, 1978).

Ainsi, la génération et l'assimilation de NH_3 associées à l'oxydation de la glycine lors de la photorespiration et l'utilisation de glutamate et de sérine comme donneurs de groupements aminés lors de la synthèse de glycine forment la base du cycle photorespiratoire de l'azote

(Woo *et al.*, 1978, 1982). Etant donné la production massive de NH_3 au cours de la photorespiration, ce cycle de l'azote doit fonctionner à une vitesse considérable.

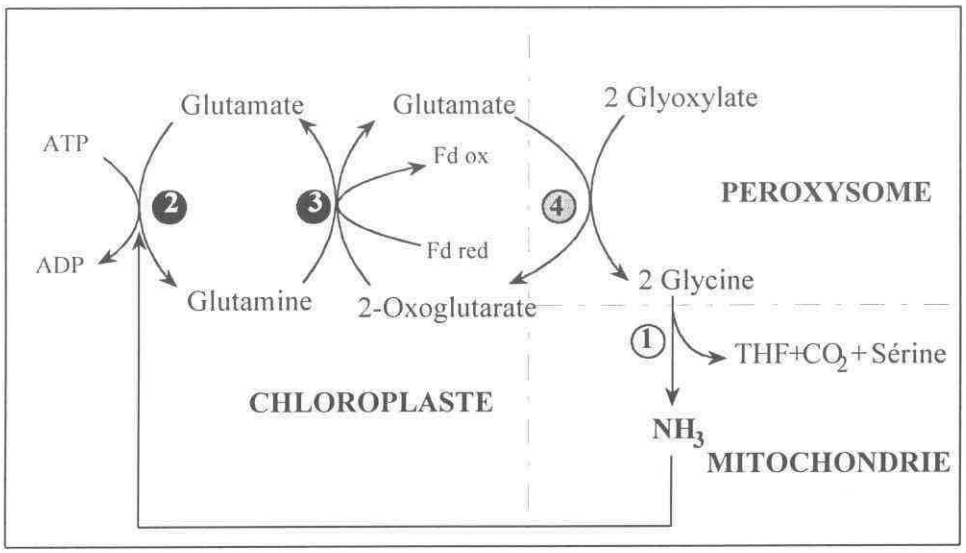


Figure 2: Schéma du cycle de l'azote: (1) glycine décarboxylase et sérine hydroxyméthyltransférase; (2) glutamine synthétase (GS); (3) glutamate synthase (GOGAT); (4) glutamate (alanine)-glyoxylate aminotransférase. THF: tétrahydrofolate.

B - ACTIVITE DE LA PHOTORESPIRATION CHEZ LES ORGANISMES POSSEDANT UN SYSTEME DE CONCENTRATION DU CO_2

Chez les organismes photosynthétiques primitifs, la concentration atmosphérique élevée en CO_2 (3000 ppm au milieu du crétacé) rendait négligeable l'activité oxygénase de la Rubisco (Lasaga *et al.*, 1985; Smith, 1986). Cependant, à la fin du crétacé, le taux de CO_2 dans l'atmosphère chute (alors que la concentration en O_2 augmente) pour se stabiliser durant le Paléocène et se rapprocher des valeurs actuelles (350 ppm). Ces changements de concentration en CO_2 ont excercé une forte pression sélective sur les plantes à métabolisme de type C3 afin d'améliorer leur rendement photosynthétique et de réduire la photorespiration. C'est durant le Paléocène qu'apparaissent les premiers végétaux à métabolisme de type C4. Ceux-ci disparaissent durant la période Eocène et Oligocène (à cause du doublement de la concentration atmosphérique en CO_2) pour réapparaître à partir du miocène et se développer jusqu'à nos jours (Ehleringer *et al.*, 1991).

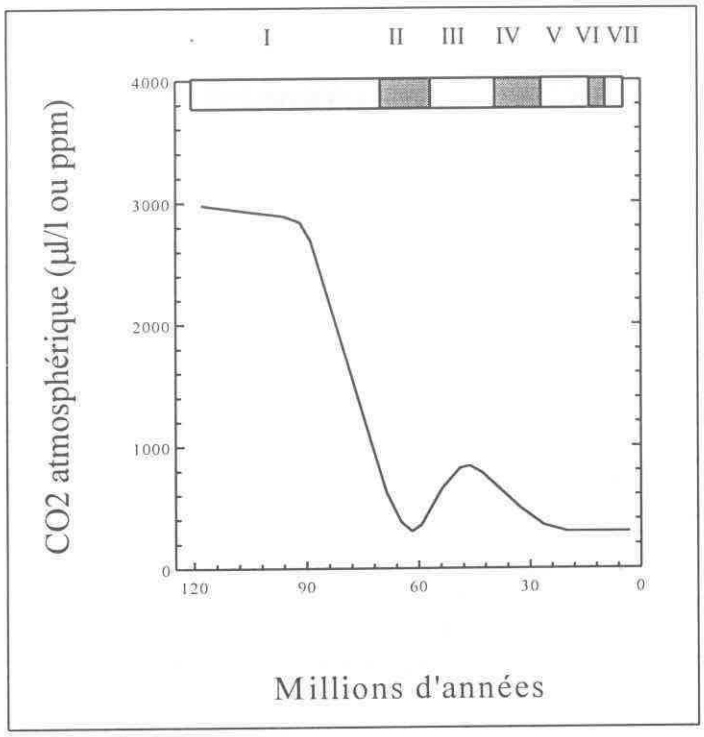


Figure 3: Evolution de la concentration en CO_2 depuis 100 millions d'années (d'après Ehleringer *et al.*, 1991). Les chiffres romains correspondent aux différentes périodes géologiques: I, Crétacé II, Paléocène; III, Eocène; IV, Oligocène; V, Miocène; VI, Pliocène; VII, Pléistocène.

Actuellement, une des causes principales du déclenchement de la photorespiration chez les plantes à métabolisme de type C3 est l'effondrement de la concentration de CO_2 au sein de l'appareil foliaire. Tel qu'on a pu le voir dans le paragraphe précédent, ce phénomène diminue le rendement de la photosynthèse au détriment du métabolisme de type C2. Cependant, une augmentation du taux de CO_2 au niveau du site actif de l'enzyme, inhibe la fonction oxygénase de la Rubisco et donc diminue le flux de carbone vers le cycle en C2. De nombreux organismes photosynthétiques ont développé différentes stratégies au cours de l'évolution, afin d'augmenter la concentration en CO_2 au niveau du site actif dans le but d'obliger la Rubisco à fonctionner comme une carboxylase et de réduire la photorespiration. Dans les paragraphes suivants, nous montrerons succintement que la photorespiration subsiste, malgré les nombreuses adaptations de certains végétaux (terrestres mais également marins) cherchant à augmenter le rendement de la photosynthése.

1 - Les plantes à métabolisme de types C4

Chez les plantes à métabolisme de type C4, le CO_2 atmosphérique est fixé, dans les cellules du mésophylle, sur un acide en C4 qui diffuse vers les cellules du fourreau pour subir une décarboxylation. Le CO_2 libéré est utilisé par la ribulose 1,5-bisphosphate carboxylase exclusivement présente dans les cellules du fourreau (Edwards and Walker, 1983; Hatch, 1987; Jenkins *et al.*, 1989; Henderson *et al.*, 1992). Ainsi, grâce à ce système, la fonction

oxygénase de la Rubisco est considérablement réduite. En effet, le flux de carbone vers le cycle en C2 est inférieur à 5% du CO_2 fixé au cours de la photosynthèse (Zelitch, 1975) alors qu'il est de 25% à 40% pour les plantes de type C3. Cependant, les phénomènes de diffusion du CO_2 à travers les cellules du fourreau et plus particulièrement la chute de la concentration extérieure en CO_2 (la capacité des cellules du mésophylle à concentrer le CO2 est limitée), sont autant de facteurs qui augmentent la photorespiration au sein des cellules du fourreau des plantes à métabolisme de type C4 (Ohnishi and Kanai, 1983; Ohnishi *et al.*, 1985). Il est à noter que les enzymes du cycle photorespiratoire sont présentes de façon prédominante dans les cellules du fourreau; les cellules du mésophylle étant déficientes en enzyme péroxysomale du cycle en C2.

Ainsi, les végétaux de type C4, malgré leurs systèmes de concentration en CO_2, sont des plantes qui photorespirent (pour une revue Dai *et al.*, 1993).

2 - Des espèces intermédiaires: les plantes à métabolisme de type C3-C4

Les plantes intermédiaires C3-C4 (telles que certaines crucifères du genre Moricandia) possèdent des caractéristiques se situant à mi-chemin entre les plantes en C3 et C4. Ces végétaux ont une anatomie similaire à celle des plantes en C4 et possèdent toutes les enzymes du cycle photorespiratoire. Elles sont distribuées de façon uniforme au sein des cellules photosynthétiques des feuilles et présentent une activité maximale comparable à celle présente chez les plantes de type C3 (Rawsthorne *et al.*, 1988b). Seules la glycine décarboxylase et la sérine hydroxyméthyltransférase sont confinées dans les cellules du fourreau (Hylton *et al.*, 1988). Ainsi, la glycine provenant des cellules du mésophylle lors de la photorespiration, diffuse vers les cellules du fourreau pour subir une décarboxylation (pour une revue voir Rawsthorne, 1992). L'étroite association des mitochondries avec les chloroplastes au sein des cellules du fourreau crée une augmentation locale de la concentration en CO_2 qui favorise sa recapture par la ribulose 1,5 bisphosphate carboxylase (Rawsthorne *et al.*, 1988a). Ainsi, ce système rudimentaire de concentration en CO_2 permet aux plantes intermédiaires C3-C4 de réagir plus efficacement à une diminution de la concentration en CO_2. En effet, l'anatomie des cellules et la distribution différentielle de la glycine décarboxylase dans les cellules du fourreau sont autant de modifications qui représentent une évolution vers un métabolisme de type C4.

3 - Quelques aspects de la photorespiration chez les algues

La littérature concernant la photorespiration chez les algues est peu abondante. Ce que l'on sait, c'est que la photorespiration chez les algues et les bactéries photosynthétiques, est similaire à celle observée chez les plantes supérieures (Tolbert, 1974). En effet, lorsque le taux d'O_2 augmente, une diminution de la photosynthèse est observée, alors que la production de glycolate, de glycine et de sérine croît (pour un revue voir Lorimer et Andrews, 1981). Il existe donc chez les algues comme chez les plantes terrestres une compétition entre le CO_2 et l'O_2. Toutefois, le cycle de la photorespiration chez les algues diffère des plantes supérieures de type C3 par quelques points particuliers.

- Chez certaines algues unicellulaires (comme les Zooxanthellae) une partie du glycolate est excrétée vers le milieu extérieur (Tolbert et Zill, 1956; Berry *et al.*, 1976). Ce phénomène a lieu lorsque la production de glycolate par la cellule excède la capacité d'oxydation du glycolate par la glycolate deshydrogénase (en faible concentration), notamment, dans un milieu trop riche en O_2 (Findinegg, 1977). Ainsi, le glycolate excrété (qui représente 10 à 15% du CO_2 fixé) est rapidement repris par d'autres algues capables de l'absorber du mileu extérieur et de le métaboliser par l'intermédiaire du cycle en C2 (Findinegg, 1977). En effet, le glycolate est capable de stimuler (Lord et Merrett, 1970) ou d'assumer entièrement la croissance de certaines algues vertes à la lumière. L'excédent de glycolate excrété par certaines algues est donc une source d'éléments carbonés pour d'autres algues qui utilisent pour cela le métabolisme en C2 (pour une revue voir Tolbert, 1980).

- Une autre particularité observée chez les algues est leur capacité à concentrer le carbone inorganique (notamment les ions HCO_3^-) lorsque la concentration en CO_2 est faible, afin d'induire la photosynthèse et de diminuer la photorespiration (Raven, 1985; Findinegg, 1976). Ce mécanisme de concentration est différent de celui observé chez les plantes à métabolisme de type C4, d'une part car l'activité de la PEP-carboxylase est très faible chez les algues et d'autre part parce que le mécanisme de fixation du CO_2 est similaire à celui des plantes à métabolisme de type C3. Ainsi, les algues ont mis en place un système qui utilise les ions bicarbonates comme source de carbone (Badger *et al.*, 1977, 1978). En effet, lorsque le CO_2 est en concentration limitante, une enzyme, l'anhydrase carbonique est rapidement induite catalysant la transformation des ions bicarbonates en CO_2 (Ingle et Colman, 1976; Reed et Graham, 1977, Badger et Dean Price, 1994). Grâce à cette enzyme et à une pompe à bicarbonate (cotransport avec un proton) la teneur en CO_2 au sein de la cellule augmente d'un

facteur 40 (voire plus) ce qui permet de diminuer la fonction oxygénase de la Rubisco (Husic *et al.*, 1987; Bowes, 1985; Colman et Cook, 1985; Badger et Dean Price, 1994).

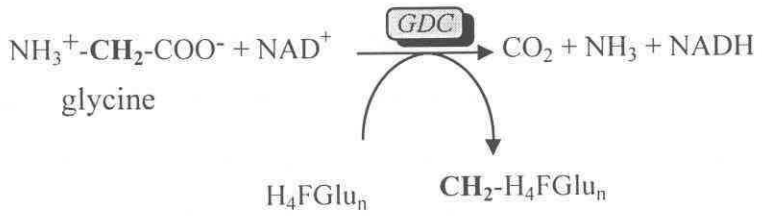
II - L'OXYDATION DE LA GLYCINE PAR UN COMPLEXE MULTIENZYMATIQUE: LA GLYCINE DECARBOXYLASE (GDC)

A - OXYDATION DE LA GLYCINE CHEZ LES VEGETAUX A METABOLISME DE TYPE C3

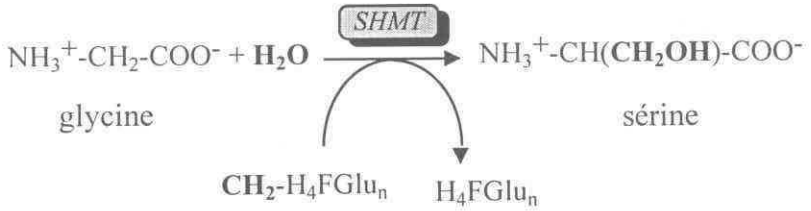
L'oxydation de la glycine chez les végétaux à métabolisme de type C3 est certainement une des étapes clef de la photorespiration. Cette réaction complexe catalysée par la glycine décarboxylase (GDC, EC 2.1.2.10) et la sérine hydroxyméthyltransférase (SHMT, EC 2.1.2.1) s'effectue au sein de la matrice mitochondriale des tissus foliaires et permet la conversion de 2 moles de glycine en une mole de sérine avec libération d'une mole de NH_3 et de CO_2 selon la réaction globale suivante (Tolbert, 1971a; Neuburger et Douce, 1977; Rathman, 1978; Arron *et al.*, 1979; Kisaki *et al.*, 1971):

$$\boxed{2\ \text{glycine} + NAD^+ \rightarrow \text{sérine} + CO_2 + NH_3 + NADH}$$

La glycine décarboxylase catalyse le clivage total d'une molécule de glycine pour libérer du CO_2 et du NH_3, alors que le carbone du groupement méthyle restant est transféré sur le tétrahydrofolate (H_4FGlu_n) pour former du N^5,N^{10} méthylènetétrahydrofolate (5,10-CH_2-H_4FGlu_n) (Neuburger *et al.*, 1986; Oliver *et al.*, 1990a, b). Le NAD^+ est l'accepteur final d'électrons de cette oxydation. Ce processus est considéré comme la source majeure de CO_2 produit au cours de la photorespiration (Kisaki et Tolbert, 1970; Kisaki *et al.*, 1971).



Afin de régénérer le tétrahydrofolate, la sérine hydroxyméthyltransférase transfère le groupement monocarboné provenant du 5,10-CH_2-H_4FGlu_n sur une deuxième molécule de glycine pour former la sérine.



Les vitesses d'oxydation de la glycine dans les mitochondries végétales par le complexe de la glycine décarboxylase sont considérables. Elles peuvent dépasser 1200 nmoles de glycine converties en sérine par mg de protéines mitochondriales et par minute. Contrairement au règne animal et aux procaryotes, cette vitesse d'oxydation rapide est due à la forte concentration matricielle du complexe enzymatique (la GDC représente plus d'un tiers des protéines solubles mitochondriales) (Bourguignon, 1988; Oliver *et al.*, 1990a, b).

Cette vitesse élevée d'oxydation de la glycine implique une vitesse tout aussi élevée du recyclage du tétrahydrofolate par la SHMT. Or, cela soulève un problème du couplage de la GDC avec la SHMT car cette dernière possède une constante d'équilibre qui est très en faveur de la formation de glycine et de méthylène tétrahydrofolate (Besson *et al.*, 1993). La SHMT, lors de l'oxydation de la glycine, catalyse en permanence la réaction dans le sens thermodynamiquement non favorable. Cette condition est cependant réalisée grâce à la forte activité du complexe de la glycine décarboxylase. En effet, la GDC a une activité maximale 2,5 à 3 fois plus grande que la SHMT (Bourguignon *et al.*, 1988; Rébeillé *et al.*, 1994) et la réaction catalysée par cette dernière est alors maintenue hors équilibre par la présence dans l'espace matriciel des mitochondries d'un rapport $CH_2H_4FGlu_n$/ H_4FGlu_n très élevé (Rébeillé *et al.*, 1994).

Un aspect particulièrement intéressant de ce système multienzymatique est que la GDC s'accumule principalement dans les mitochondries de feuilles vertes. En effet, dans les tissus étiolés ou non chlorophylliens, l'activité d'oxydation de la glycine est insignifiante (Moore *et al.*, 1977; Douce, 1985). Ainsi, chez l'épinard l'activité est présente au sein des mitochondries de feuilles mais absente des mitochondries de racines, de tiges ou de nervures (Gardeström *et al.*, 1980). De plus, Walker et Oliver (1986b) ont montré que l'activité glycine décarboxylase,

faible dans les feuilles étiolées augmentait considérablement à la lumière. L'augmentation de l'activité résulte de la synthèse « *de novo* » de nouvelles protéines dont la régulation semble s'effectuer au niveau transcriptionnel en ce qui concerne la protéine H (Macherel *et al.*, 1990; Kim et Oliver, 1990).

B - QUELQUES RAPPELS SUR L'OXYDATION DE LA GLYCINE CHEZ LES ORGANISMES NON CHLOROPHYLLIENS

Chez la plupart des organismes non chlorophylliens, animaux (Kikuchi, 1973; Hiraga et Kikuchi, 1980) et procaryotes (Klein et Sager, 1966; Vanden Boom *et al.*, 1991) ce complexe, présent en très faible concentration, est essentiellement impliqué dans le catabolisme de la glycine. Ainsi, au sein de la matrice mitochondriale des cellules hépatiques des animaux (pour une revue voir MacKenzie, 1984 et Cossins, 1980) et chez la plupart des bactéries telles que *Peptococcus glycinophilus* (Klein et Sagers, 1966), *Arthrobacter globiformis* (Kochi et Kikuchi, 1969) et *Escherichia coli* (Okamura-Ikeda *et al.*, 1992), la conversion d'une mole de glycine en CO_2, NH_3 et N5, N10 méthylène tétrahydrofolate est catalysée par le complexe de la glycine décarboxylase. Une exception concerne les bactéries du groupe des clostridium vivant en anaérobiose telles que *Clostridium acidiurici* (Karlson et Barker 1949) ou *Eubacterium acidaminophilum* (Zindel *et al.*, 1988) qui utilisent la glycine comme intermédiaire dans la synthèse d'acétate. Dans ce cas, deux enzymes sont nécessaires: la glycine décarboxylase et la glycine réductase qui transforment quatre molécules de glycine en une molécule d'acétate avec libération de 2 molécules de CO_2 et quatre molécules de NH_3 (Dürre et Andreesen, 1983; Zindel *et al.*, 1988).

Cependant, chez la plupart des organismes, le catabolisme de la glycine a pour fonction essentielle la production de métabolites primaires. Ainsi, la formation de méthylène tétrahydrofolate représente une importante source d'unités monocarbonées nécessaires à de nombreuses voies métaboliques impliquant notamment la synthèse de thymidylate, de purine et de méthionine (Cybulski et fisher, 1976).

Ce complexe est présent également dans d'autres tissus tels que les reins et le cerveau (Kikuchi, 1973; Yoshida et Kikuchi, 1973). Chez ce dernier, la glycine est un neurotransmetteur puissant qui intervient dans l'inhibition rapide des récepteurs à acétylcholine au sein des synapses des neurones.

Bien que la GDC soit présente en très faible concentration dans les mitochondries des cellules humaines, une réduction de l'activité du système de clivage de la glycine entraîne d'importants désordres physiologiques (Nyhan, 1978). En effet, les patients déficients en GDC présentent un niveau élevé de glycine non seulement dans les urines et dans le sang, mais également dans le liquide cérébro-spinal et dans le cerveau. Cette maladie congénitale, décrite sous le nom « d'hyperglycinémie non cétonique » (NKH: Nonketotic hyperglycinemia) provoque chez les patients qui ont survécu durant la période néonatale de graves troubles neurologiques (de Groot *et al.*, 1978), notamment un sévère retard mental, des convulsions, une hypotonie et une déficience respiratoire (Nyhan, 1978). Des traitements tels que le sodium benzoate (diminuant le taux de glycine dans le sang) (Krieger *et al.*, 1977) ou la strychnine (antagoniste puissant de la glycine qui possède une forte affinité pour les récepteurs à glycine dans le système nerveux central) (Gitzelmann *et al.*, 1978; Arneson *et al.*, 1979; Mac Dermot *et al.*, 1980; Warturton *et al*; 1980) permettent d'atténuer certains effets mais ne résorbent en aucun cas la maladie. On peut dire qu'il n'existe à ce jour aucune thérapie efficace pouvant traiter cette maladie autosomale récessive. Il semble que cette maladie soit due à des déficiences des deux protéines du complexe (P ou H).

A ce sujet, l'abondance de ce complexe au sein de la matrice mitochondriale des cellules végétales en fait un modèle d'étude « par excellence ».

III - STRUCTURE MOLECULAIRE ET MECANISME REACTIONNEL DU COMPLEXE DE LA GLYCINE DECARBOXYLASE

A - STRUCTURE MOLECULAIRE DE LA GDC

Chez les animaux (Kikuchi, 1973), les bactéries (Klein et Sagers, 1966; 1967a, 1967b; Sagers et Gonzalus, 1961) et les végétaux (Walker et Oliver, 1986a; Neuburger *et al.*, 1986) le mécanisme enzymatique impliqué dans l'oxydation de la glycine est tout à fait similaire. La glycine décarboxylase est un complexe multienzymatique qui nécessite l'action concertée de 4 protéines différentes: les protéines P, H, T et L (Walker et Oliver, 1986a; Bourguignon *et al.*, 1988). Ces quatre protéines ainsi que la SHMT ont été purifiées dans un très haut état d'homogénéité à partir de mitochondries de feuilles de pois (Bourguignon *et al.*, 1988). Leurs masses moléculaires ont ainsi pu être déterminées par analyse électrophorétique. On distingue:

présence d'une seconde molécule de glycine, en tétrahydrofolate lors de la réaction catalysée par la SHMT. Enfin, les groupements thiol du lipoate de la protéine H sont réoxydés par la protéine L (une dihydrolipoamide deshydrogénase) avec réduction séquentielle de FAD^+ et de NAD^+.

Ces réactions nécessitent donc la présence de cinq cofacteurs essentiels: le pyridoxal phosphate, le lipoate, le FAD, le NAD^+ et le tétrahydrofolate (Walker et Oliver, 1986a; Neuburger *et al.*, 1986; Bourguignon *et al.*, 1988; Oliver *et al.*, 1990a, 1990b).

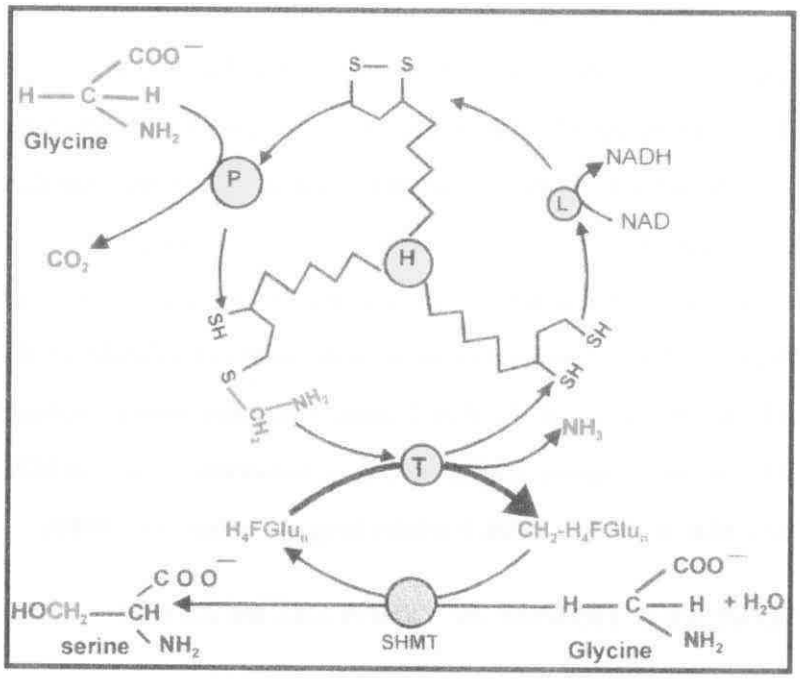


Figure 4: Mécanisme réactionnel du complexe de la glycine décarboxylase dans l'oxydation de la glycine. Dans le cas de l'oxydation de glycine par les mitochondries végétales, le méthylènetétrahydrofolate (H_4FGlu_n) réagit avec une deuxième molécule de glycine pour former la sérine.

La réaction globale de ce complexe peut être mesurée soit par la libération du $^{14}CO_2$ provenant d'une molécule de glycine marquée ^{14}C sur la fonction carboxylique (Sarojini et Oliver, 1983; Walker et Oliver, 1986a) soit par la réduction du NAD^+ en NADH (Neuburger *et al.*, 1986; Bourguignon *et al.* 1988).

2 - Régulations métaboliques

L'activité de la GDC n'est régulée ni par la lumière, ni par des modifications covalentes réversibles, ni par un contrôle protéique ou par une activation protéolytique. Cependant, il semble que le rapport NADH/NAD ait une influence prépondérante sur l'activité de la GDC. En effet, le NADH [inhibiteur puissant de la protéine L: K_i(NADH)=15μM] entre en

- la protéine P, homodimère de 2 x 97 kDa qui est une protéine à pyridoxal phosphate;
- la protéine H, monomère de 15,5 kDa qui possède l'acide lipoïque comme groupement prosthétique;
- la protéine T, monomère de 45 kDa qui utilise comme cofacteur le tétrahydrofolate;
- la protéine L, homodimère de 2 x 61 kDa qui est une dihydrolipoamide deshydrogénase,
- la SHMT, homotétramère de 4 x 53 kDa dont le cofacteur est aussi le tétrahydrofolate.

L'utilisation de techniques immunochimiques a montré que l'ensemble des protéines P, H, T et L représente plus de 30% des protéines solubles des mitochondries. De plus, chez les plantes le complexe est présent dans les proportions stoechiométriques suivantes:

2 protéines P / 27 protéines H / 9 protéines T / 1 protéine L (Oliver *et al.*, 1990).

La purification de la GDC de pois a montré l'éxistence de fortes similarités avec celles des procaryotes et des animaux. En effet, on retrouve les mêmes protéines avec cependant une différence chez certaines bactéries vivant en anaerobie telles que *Clostridium acidiurici* (Gariboldi et Drake, 1984) ou *Eubactérium acidaminophilum* (Zindel *et al.*, 1988). En effet, la protéine P3 ou P (les protéines H, T, P et L étant nommées respectivement chez les bactéries P1, P2, P3 et P4) est un tétramère $\alpha 2\beta 2$, de masse moléculaire de 230 kDa, où α = 65 kDa et β = 58 kDa (Gariboldi et Drake, 1984; Freudenberg et Andreesen, 1989).

B - MECANISME REACTIONNEL DU COMPLEXE DE LA GDC

1 - Description

Comme l'indique le schéma de la figure 4, la première étape de ce mécanisme enzymatique fait intervenir la protéine P dont le groupement carboxyle du pyridoxal phosphate forme une base de Schiff avec le groupement α-aminé d'une molécule de glycine, puis la protéine P ayant fixé une molécule de glycine, réagit avec la protéine H. Au cours de cette réaction, le groupement α-carboxyl de la glycine est libéré sous forme de CO_2. Simultanément, le groupe méthylamine (CH_2-NH_3) restant (après décarboxylation de la glycine) se fixe sur le groupement thiol du lipoate (cofacteur lié de façon covalente au groupement ε-aminé de la lysine 63 de la protéine H de pois). Le groupement méthylamine lié au lipoate de la protéine H réagit en tant que substrat avec la protéine T. Cette dernière catalyse la libération du groupement α-aminé de la glycine sous forme de NH_3 et surtout le transfert du groupe méthylène sur le tétrahydrofolate. Le méthylènetétrahydrofolate ainsi formé est recyclé, en

compétition avec le NAD^+ [$K_m(NAD^+)$=75µM] (Neuburger *et al.*, 1986; Bourguignon *et al.*, 1988). In vivo, il est très vraisemblable que l'oxydation du NADH matriciel formé au cours de l'oxydation de la glycine résulte d'une navette malate-oxaloacétate entre les mitochondries végétales et les péroxysomes qui se traduit par une réoxydation très rapide du NADH matriciel pour réduire l'hydroxypyruvate transitant dans les péroxysomes (Douce *et al.*, 1994).

De même, le rapport sérine/glycine joue un rôle important sur l'activité enzymatique du complexe. En effet, la sérine en compétition avec la glycine, inhibe de façon compétitive l'activité de la GDC [K_i(sérine) = 4mM; K_m(glycine) = 6mM] et semble se fixer sur la protéine P (Oliver et Sarojini, 1987). Ainsi, Oliver et Sarogini (1987) et Besson *et al,* (1993) suggèrent que la sérine doit être rapidement chassée de la mitochondrie afin de prévenir toute inhibition et de permettre à la SHMT de continuer à produire de la sérine sans risque de déplacer son équilibre dans un sens défavorable. Certains auteurs suggèrent la présence dans la membrane interne de protéines responsables du transfert de la glycine et de la sérine (Oliver, 1987; Walker et Oliver, 1982) alors que d'autres pensent que la glycine et la sérine diffusent à travers les membranes mitochondriales (Day et Wiskich 1980 et Shingles *et al.*, 1984). En effet, à ce jour aucun transporteur n'a été isolé.

Les deux autres produits de la réaction que sont le CO_2 et le NH_3 diffusent rapidement vers le chloroplaste pour être consommés respectivement par la Rubisco et la glutamine synthase. Par conséquent on voit que l'ensemble des produits lors du fonctionnement de la GDC et de la SHMT sont en fait rapidement métabolisés, empêchant de ce fait toute rétroinhibition des réactions. Ceci favorise le fonctionnement de l'ensemble GDC-SHMT dans le sens production de sérine, permettant ainsi à la photorespiration de se dérouler efficacement.

3 - Inhibiteurs spécifiques des protéines du complexe de la GDC

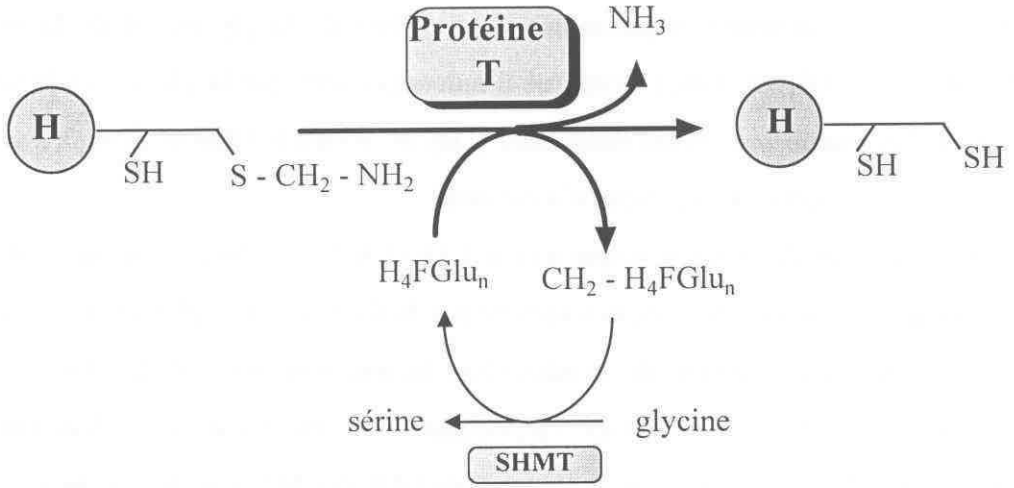
Plusieurs inhibiteurs chimiques spécifiques de certaines protéines de la GDC ont été utilisés afin d'étudier leurs effets sur l'activité de ce complexe. On distingue l'acide hydrazide isonicotinic (Pritchard *et al.*, 1962) et l'arsenite (Walker *et al.*, 1982) qui réagissent sur le lipoate de la protéine H, la carboxyméthoxylamine, la méthoxylamine et l'acéthydrazide qui agissent sur le pyridoxal-phosphate de la protéine P (Sarojini et Oliver, 1985). Plus récemment, il a été montré que la victorine, une toxine produite par un champignon provoquant la rouille de l'avoine (*Cochliobolus victoriae*), en se fixant sur la protéine P et/ou

sur la protéine H, est un inhibiteur puissant de la GDC (Wolpert et Macko, 1989; Wolpert *et al.*, 1994; Navarre et Wolpert, 1995).

IV - LA PROTEINE T: PROTEINE A TETRAHYDROFOLATE

A - REACTION CATALYSEE PAR LA PROTEINE T

Comme nous avons pu le voir précédemment, la protéine T est une des protéines appartenant au complexe de la GDC. Chez les plantes, la protéine T joue un rôle stratégique dans le cycle carboné photosynthétique car elle est responsable du relargage de NH_3 et catalyse de ce fait la première étape du cycle azoté de la photorespiration (Givan *et al.*, 1988; Keys *et al.*, 1978). De plus, elle est la source majeure (avec la SHMT) de la production de méthylènetétrahydrofolate. La protéine T catalyse la réaction suivante:



Ce sont Sager et Gonsalus (1961) qui ont montré pour la première fois que le tétrahydrofolate est nécessaire à l'oxydation de la glycine. En effet, la protéine T transfère le « CH_2 » du groupement méthylamine de la protéine H sur le tétrahydrofolate, formant le méthylènetétrahydrofolate et libérant le groupement amine sous forme d'ammoniac (NH_3). Le tétrahydrofolate est ensuite recyclé grâce à l'intervention de la SHMT (Bourguignon *et al.*, 1988). *In vitro*, l'oxydation de la glycine s'effectue dans des conditions de stricte anaérobiose (sous argon) afin d'éviter l'oxydation du tétrahydrofolate en dihydrofolate.

Il est à noter que le tétrahydrofolate nécessaire à la réaction est à l'origine du nom de la protéine T.

B - LE TETRAHYDROFOLATE

Le tétrahydrofolate fait partie de la famille chimique des folates, dont le nom dérive du mot latin *folia* « feuille». L'acide folique ou le folate est constitué de trois parties distinctes (comme l'indique la figure 5): une ptéridine hétérobicyclique substituée dont le précurseur est le GTP (Mitsuda *et al.*, 1966; Blakley, 1969), l'acide para-aminobenzoïque (PABA) et l'acide L-glutamique. Contrairement aux animaux, les végétaux ainsi que les micro-organismes sont capables d'assembler les différents éléments qui constituent le folate au cours d'une voie de synthèse complexe non entièrement élucidée dans la cellule végétale (pour une revue, voir Cossins, 1980).

Le folate peut exister sous différents états d'oxydation: entièrement oxydé, réduit sur les positions 7 et 8 (dihydrofolate) ou réduit sur les positions 5, 6, 7 et 8 (tétrahydrofolate).

Le tétrahydrofolate (également appelé tétrahydroptéroylglutamate), est un transporteur très polyvalent d'unités monocarbonées activées. Le groupe monocarboné porté par le tétrahydrofolate est lié à l'un des atomes d'azote N-5 ou N-10 (notés aussi N^5 et N^{10}) ou aux deux. Cette unité peut exister dans trois états d'oxydation (figure 5). La forme la plus réduite porte un groupe méthyle, tandis que la forme intermédiaire porte un groupe méthylène. Les formes les plus oxydées portent un groupe méthényle, formyle ou forminino. Chez les végétaux, le folate existe sous deux formes principales: les dérivés formyls ou méthyls du tétrahydrofolate qui sont compartimentés, avec un pool discret associé aux mitochondries (Clandinin et Cossins, 1972), et aux chloroplastes (Kim, 1970; Cossins et Shah, 1972). Généralement, le folate intracellulaire ne contient pas qu'une seule unité glutamate, mais plutôt une chaîne poly-γ-glutamylée. La longueur de la chaîne varie de 2 à 8 résidus glutamates et cette longueur varie en fonction des espèces. En effet, *E. coli* contient essentiellement des dérivés triglutamates (pour une revue, voir Cossins, 1984). Dans le foie de rat, on trouve des longueurs de chaîne allant jusqu'à l'heptaglutamate, mais les dérivés pentaglutamilés sont les plus représentés (Shane, 1982). Chez le porc et la souris, les dérivés portent quatre à neuf résidus glutamates, avec l'hexaglutamate comme forme majeure (Cichowics et Shane, 1987; Priest *et al.*, 1981). Enfin chez les plantes, Imeson *et al.*, (1990)

ont récemment démontré que les formes majeures des dérivés polyglutamylés chez le pois sont le diglutamate (46%), le pentaglutamate (38%) et le tétraglutamate (15%).

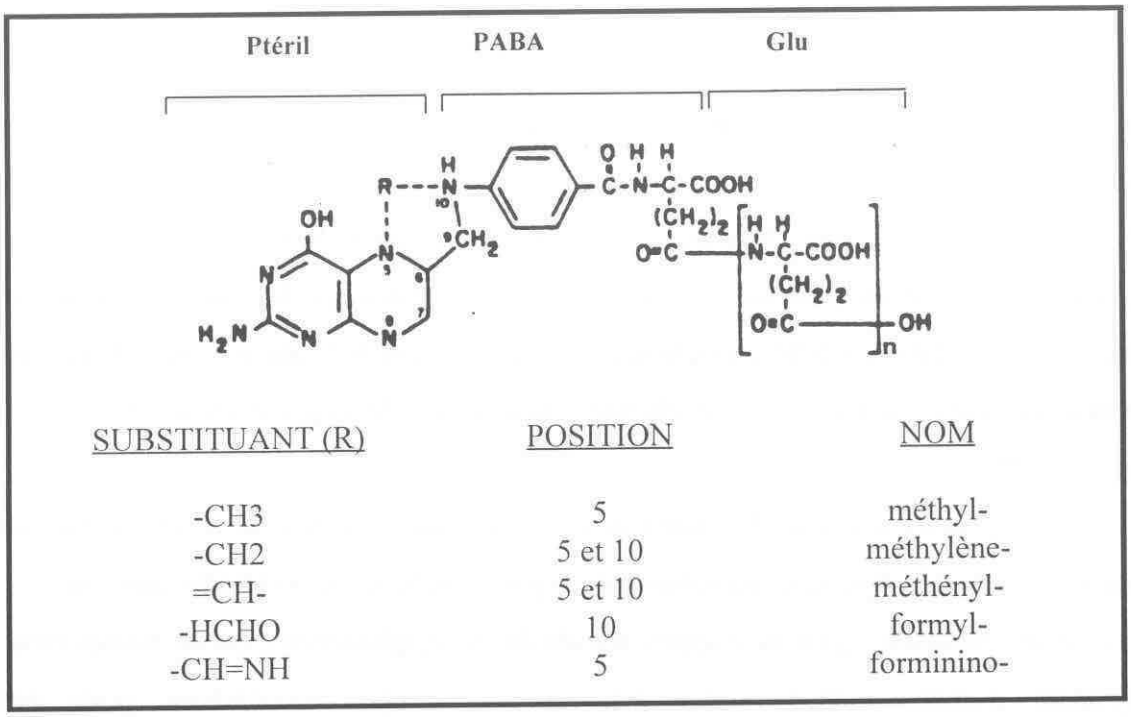


Figure 5: Structure du tétrahydrofolate et de ses dérivés. Le folate cellulaire existe essentiellement sous forme de dérivés γ-glutamylés du tétrahydroptéroyglutamate (Blakley, 1969; Cossins, 1980; Mac guire et Coward, 1984).

Il semblerait que la longueur de la chaîne joue un rôle prépondérant car elle influe sur l'affinité des enzymes associées au folate vis à vis de son cofacteur. Brody *et al.* (1976) estiment que la distribution des formes polyglutamylées varie au sein d'une même espèce en fonction de l'origine tissulaire des cellules.

C - INTERACTION ENTRE LA PROTEINE T ET LE TETRAHYDROFOLATE

Le tétrahydrofolate est l'un des cofacteurs nécessaire à la réaction d'oxydation de la glycine en assurant le lien cinétique entre la GDC (par l'intermédiaire de la protéine T) et la SHMT (Bourguignon *et al.*, 1988). Besson *et al.* (1993) ont démontré que la protéine T (au même titre que la SHMT) possède une grande affinité pour le tétrahydrofolate contenant notamment 3, 4 ou 5 résidus glutamates. Il a été établi que l'interaction du tétrahydrofolate avec la protéine T augmente avec la longueur de la chaîne poly-γ-glutamylée. Ce phénomène s'observe chez la plupart des protéines à folate (MacGuire et Bertino, 1981; Mac Kenzie, 1984; Schirch et Strong, 1989). Enfin, la constante de dissociation du tétrahydrofolate et du méthylène tétrahydrofolate pour la protéine T et la SHMT est identique (Rébeillé *et al.*, 1994).

Or, l'étape limitante de l'oxydation de la glycine par l'extrait matriciel des mitochondries est la régénération du tétrahydrofolate à partir du méthylène tétrahydrofolate par la SHMT (Oliver *et al.*, 1990b). En effet, l'activité de la GDC est beaucoup plus puissante que celle de la SHMT, entraînant un rapport CH_2-$H_4PteGlu_n$/$H_4PteGlu_n$ très élevé. Le système de clivage de la glycine maintient donc par l'intermédiaire de la protéine T une forte concentration de CH_2-$H_4PteGlu_n$, favorisant ainsi la réaction inverse catalysée par la SHMT.

V - OBJECTIFS DU TRAVAIL

Les mitochondries végétales se distinguent des mitochondries animales par un certain nombre de caractéristiques dont la plus spectaculaire est leur capacité à oxyder la glycine à des vitesses considérables. Durant ces dix dernières années, le laboratoire de Physiologie Cellulaire Végétale s'est investi dans l'étude des mécanismes d'oxydation de la glycine par les mitochondries de tissus foliaires ainsi que dans celle de la structure et de la régulation des protéines impliquées. La caractérisation et la purification des protéines impliquées dans l'oxydation de la glycine ont permis de développer les outils moléculaires nécessaires à l'étude de la régulation et de la structure du complexe de la glycine décarboxylase. Tous les ADNc correspondant aux protéines de la GDC ont été caractérisés chez les plantes, excepté l'ADNc correspondant à la protéine T.

L'objectif de ce travail a été dans un premier temps d'isoler l'ADNc correspondant à la protéine T de la GDC. Ceci a permis de déterminer la structure primaire de la protéine, qui a pu être confirmée de façon détaillée par la spectrométrie de masse et microséquençage. L'obtention de l'ADNc a permis d'étudier la régulation de l'expression du gène codant pour la protéine T au cours du développement du pois. Cette étude a été menée en parallèle avec celles des autres protéines du complexe de la GDC et de la SHMT.

Dans le cadre de cette étude sur la biogénèse du complexe de la GDC, nous avons également isolé le gène nucléaire codant pour la protéine T afin de comparer la zone promotrice de ce gène avec celles d'autres gènes dont l'expression dépend de la lumière.

Finalement, nous avons utilisé l'ADNc pour établir un système d'expression procaryotique permettant de produire de grandes quantités de protéine T mature. Ce système sera destiné à l'étude structurale ultérieure de la protéine T.

MATERIEL
ET METHODES

MATERIEL et METHODES

I - MATERIEL EXPERIMENTAL

A - MATERIEL VEGETAL

Nous avons utilisé une plante de la famille des légumineuses qui est le pois ou (*Pisum sativum.L* variété Douce provence). Cette variété de pois est vivace avec une tige grimpante. Les feuilles pennées sont disposées de façon alterne et composées d'une paire de stipules et de folioles terminées par une vrille.

Les feuilles de pois cultivées à la lumière sont prélevées, en fonction du stade de développement (figure 6), selon des critères prédéfinis; on distingue:

- le stade L1 (plants âgés de 4 jours): jeunes crosses, avec les 2 premières feuilles imbriquées.
- le stade L2 (plants âgés de 5 jours): les folioles de la première paire sont pliées.
- le stade L3 (plants âgés de 7 jours): les folioles de la première paire sont plus développées.
- le stade L4 (plants âgés de 9 jours): les folioles de la première paire sont en forme de coeur.
- le stade L5 (plants âgés de 12 jours): les folioles de la première paire présentent une morphologie de feuilles matures.

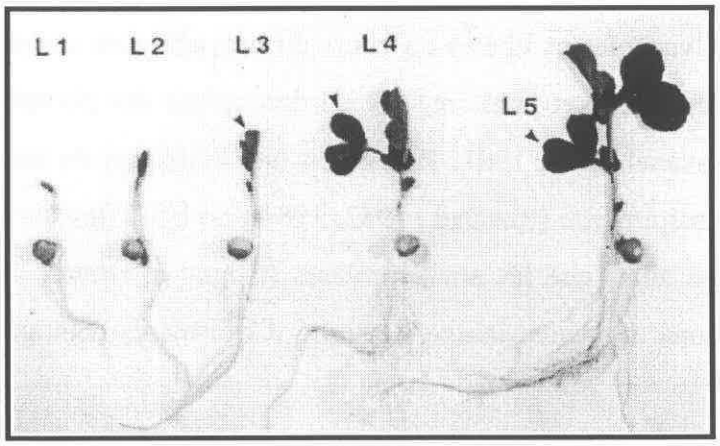


Figure 6: Aspect morphologique des plants de pois à différents stades du développement.

Les feuilles des plantes de pois étiolées sont prélevées généralement au 12^{iéme} jour. La morphologie des feuilles est comparable à celles des feuilles vertes aux stades L1 et L2.

B - CONDITIONS DE CULTURE

Les graines de pois sont imbibées pendant 12 h dans l'eau courante puis semées dans des bacs de vermiculite (20 à 30 graines/dm^2). Les bacs sont placés dans un phytotron équipé d'un éclairage de 170 µmol de photons m^{-2} s^{-1} (mesuré à l'aide de l'appareil Quantithum d'Hansatech) pendant 4 à 15 jours avec alternance d'une période de lumière de 14 h à 25°C suivie d'une période d'obscurité de 10 h à 18 °C. L'humidité relative est maintenue à 70%. Les plantes sont arrosées avec de l'eau chaque jour.

Pour l'obtention de pois étiolés, ceux-ci sont cultivés dans les conditions citées ci-dessus, mais en absence de lumière et avec seulement deux arrosages par semaine.

II - ISOLEMENT DES ORGANITES

A - EXTRACTION ET PURIFICATION DES MITOCHONDRIES DE FEUILLES DE POIS

1 - Composition des milieux

Le milieu de broyage est ainsi constitué: D-mannitol 0.3 M; EDTA 1 mM; pyrophosphate de sodium 20 mM (pH 7.5); β-mercaptoéthanol 3 mM; sérum albumine bovine (BSA) 2 g/l; polyvinyl-pyrrolidone 25 (PVP, Serva) 0,5% (m/v). Le milieu de lavage contient: D-mannitol 0.3 M, KH_2PO_4 10 mM (pH 7.2), EDTA 1 mM, BSA 1 g/l.

La présence d'agent réducteur tel que le β-mercaptoéthanol protège les groupements -SH des protéines contre toute oxydation. Le PVP a la propriété de fixer les composés phénoliques contenus dans les vacuoles et libérés au cours du broyage. Ces composés peuvent en effet réagir avec les membranes et produire, soit un découplage des phosphorylations oxydatives (Stenlid, 1970; Ravanel *et al.*, 1981, 1982), soit une inhibition du transfert des électrons le long de la chaîne respiratoire (Ravanel., 1982, 1984). La BSA fixe les acides gras libérés par l'action d'enzymes telles que les acylhydrolases (Bligny et Douce, 1978). L'EDTA, agent chélatant des cations divalents, inhibe l'activité d'hydrolases calcium-dépendantes libérées lors du broyage.

2 - Préparation des mitochondries brutes

Toutes les manipulations sont effectuées à + 4°C. Pour la purification de la protéine T, 1 à 4 kg de matière fraîche ont été nécessaires alors qu'il a fallu seulement 50 à 200g de feuilles pour étudier la biogénèse du complexe. Les feuilles de pois, une fois prélevées, sont broyées rapidement (3 pulsions de 2 secondes), à l'aide d'un broyeur à hélice "Waring blendor" tournant à faible vitesse (2 litres de milieu pour 1 kg de tissus). Le broyat est filtré sur 4 épaisseurs de gaze et une épaisseur de toile à bluter de 50 µm de vide de maille. Le filtrat est soumis à une centrifugation (rotor GS3, Sorvall) durant 15 min à 1500 g, ce qui permet de faire sédimenter les noyaux, les chloroplastes et autres fragments lourds cellulaires: ce culot est placé dans la glace et sera par la suite utilisé pour l'isolement des chloroplastes. Le surnageant est ensuite centrifugé pendant 20 min à 10 000 g (rotor GS3, Sorvall). Après élimination du surnageant, les culots contenant les mitochondries sont mis en suspension à l'aide d'un homogénéiseur de Potter dans du milieu de lavage. Les mitochondries subissent alors deux centrifugations: la suspension est centrifugée à 1500 g pendant 15 min (rotor SS34, Sorvall) afin d'éliminer les plastes intacts contaminants; le surnageant récupéré est centrifugé à 12 000g (rotor SS34, Sorvall) pendant 20 min afin de sédimenter les mitochondries brutes.

3 - Purification des mitochondries

Les culots de mitochondries sont encore largement contaminés par des organites (péroxisomes, plastes, les enzymes vacuolaires et particulièrement les thylakoïdes).

Les culots contenant les mitochondries brutes sont remis en suspension dans du milieu de lavage et sont purifiés par une centrifugation à travers un gradient de densité, réalisé à l'aide de Percoll (Pharmacia) selon la méthode de Neuburger *et al.* (1982) et Douce *et al.* (1987).

- pour des mitochondries de feuilles de pois étiolées ou de feuilles de pois aux stades L1, L2 et L3 exposées à la lumière, les tubes de centrifugation sont remplis avec 34 ml de milieu A contenant 28 % (v/v) de Percoll.

- pour des mitochondries des feuilles vertes correspondant aux stades L4 et L5, les tubes de centrifugation sont remplis à l'aide d'un générateur de gradient linéaire, par mélange progressif de 17 ml de milieu A "solution légère" à 17 ml de milieu B " solution lourde". Les milieux A et B sont de composition suivante:

<u>milieu A</u>: saccharose 0.3 M; phosphate 10 mM (pH 7.2), EDTA 1 mM, Percoll 28% (v/v), BSA 1 g/l.

milieu B: raffinose 0.3 M, phosphate 10 mM (pH 7.2), EDTA 1 mM, Percoll 28% (v/v), BSA 1 g/l, PVP 10% (m/v).

Trois gradients linéaires sont ainsi formés entre le haut et le bas du tube: un gradient 0.3-0 M en saccharose, un gradient 0-0.3 M en raffinose et un gradient 0-10% en PVP.

Les mitochondries brutes sont déposées au sommet du gradient puis centrifugées à 40 000 g pendant 45 min (rotor SS34 , Sorvall). Un gradient de densité continue en Percoll se forme dans les tubes (et se superpose au gradient de PVP), permettant la séparation des mitochondries (bande brune à la base du gradient) entre les membranes plastidiales situées au sommet du gradient et les péroxysomes au fond du tube. Cependant, les mitochondries de feuilles étiolées et de feuilles vertes aux stades L1, L2 et L3, se situent respectivement au tiers supérieur et à la moitié inférieure du gradient.

Les mitochondries recueillies sont diluées 10 fois dans du milieu de lavage sans BSA pour éliminer l'excès de Percoll, puis centrifugées 2 fois à 13 000g pendant 15 min (rotor SS34, Sorvall). Le culot final de mitochondries est repris dans un petit volume de milieu de lavage. Les mitochondries ainsi obtenues sont exemptes de contaminations détectables et présentent un état d'intégrité physiologique parfait durant plusieurs heures si elles sont conservées à la température de la glace fondante. Elles peuvent être stockées à -80°C en vue d'une analyse ultérieure de leurs composants.

Dans le but de comparer la densité des mitochondries isolées à différents stades de développement, nous réalisons dans un volume de 33 ml un gradient linéaire de Percoll (28% à 42% du haut vers le bas du tube) en présence de saccharose 0.3 M, K_2HPO_4 10 mM (pH 7.4), EDTA 1 mM et de la BSA 0.1 % (m/v). Trois ml de mitochondries isolées et lavées provenant de feuilles de pois étiolés et de feuilles vertes âgées de 7 et 12 jours sont déposés au sommet du gradient de Percoll. Des marqueurs de densité (Pharmacia) sont déposés sur un autre tube de gradient de Percoll afin de mesurer la densité le long du gradient. Les tubes sont centrifugés durant 50 min à 18 000 rpm (rotor SS 34, Sorvall) en utilisant un contrôleur automatique de freinage. Les gradients sont fractionnés en portions de 2 ml. Les 18 fractions obtenues sont diluées 15 fois dans du milieu de lavage puis centrifugées 20 min à 12 000 rpm (rotor SS 34, sorvall). Les culots sont resuspendus dans un petit volume de milieu de lavage. Le volume final de la suspension mitochondriale se situe entre 100 et 400 µl. Un cinquième à un dixième de ce volume est utilisé pour mesurer l'oxydation du NADH par les mitochondries à l'aide d'une électrode à oxygène de type Clark.

B - METHODES D'ETUDES DES MITOCHONDRIES ISOLEES

1 - Mesure des activités oxydatives

Les vitesses de consommation d'oxygène par les mitochondries sont mesurées à 25°C à l'aide d'une électrode à oxygène de type Clark (Hansatech Ldt, Angleterre) dans un volume final de 1 ml de milieu d'électrode [Mannitol 300 mM, KCl 10mM, $MgCl_2$ 5 mM, KH_2PO_4 5 mM pH 7.2, BSA 0,1% (m/v)] (Delieu et Walker, 1972). A cette température, la solubilité de l'oxygène est de 240 µM. Les valeurs des contrôles respiratoires et celles des rapports ADP/O sont mesurées selon la méthode d'Estabrook (1967). L'oxydation du NADH est mesurée en présence de 2 mM de NADH et d'une aliquote d'ADP (100 nmoles) (stade III) ajoutée dans le milieu de réaction. L'oxydation du pyruvate est mesurée dans le milieu de réaction en présence de: pyruvate 10 mM, thiamine pyrophosphate 0.4 mM, NAD^+ 0.2 mM, coenzyme A 0.4 mM, malate 0.5 mM et d'une aliquote d'ADP (100 µM) ajoutée dans le milieu. L'oxydation de la glycine est mesurée en présence de glycine 10 mM, NAD^+ 0.2 mM et d'une aliquote d'ADP (100 µM) ajoutée dans le milieu de réaction.

2 - Test d'intégrité des mitochondries

Ce test permet de mesurer le pourcentage d'organites endommagés lors de la préparation (Neuburger *et al.*, 1982). Son principe repose sur le fait que la membrane externe mitochondriale est imperméable au cytochrome *c*. De ce fait, on évalue l'intégrité de la membrane externe en mesurant l'oxygène consommé par la cytochrome oxydase en présence de cytochrome *c*, exogène réduit. Un enregistrement oxygraphique est donné à titre d'exemple dans la figure. Les vitesses de consommation d'oxygène par les mitochondries sont mesurées à 25°C comme décrit précédemment.

Une fraction aliquote de mitochondries purifiées (0,2-0,3 mg de protéine) est ajoutée à 1 ml de milieu d'électrode (trace A). Une fraction identique est ajoutée à 0,5 ml d'eau bidistillée, agitée durant environ 10 secondes, puis ramenée à des conditions isotoniques par addition de 0,5ml de milieu d'électrode en double concentration (trace B). En raison du choc osmotique, la membrane externe des mitochondries se rompt et le cytochrome *c* exogène, réduit par de l'ascorbate, peut venir au contact de la face externe de la membrane interne où il est oxydé par la cytochrome oxydase. L'addition de KCN 200 µM provoque l'inhibition de la cytochrome oxydase: la vitesse résiduelle mesure l'oxydation non enzymatique du cytochrome *c* par

l'oxygène du milieu. La vitesse de consommation d'O_2, V_A ou V_B, dépendante du cytochrome *c* et inhibée par le cyanure est entièrement imputable au fonctionnement de la cytochrome oxydase. Le pourcentage de mitochondries intactes est déduit des vitesses V_A et V_B selon la formule 100 x (1- V_A/V_B) (voir figure 7).

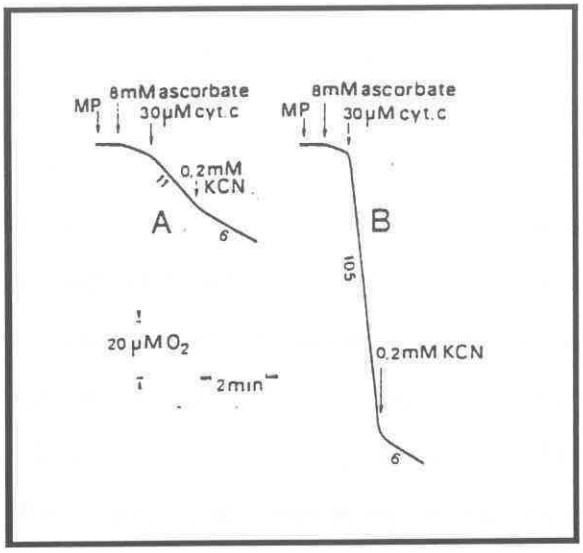


Figure 7: Mesure de l'intégrité des mitochondries purifiées par la consommation d'oxygène dépendante de l'ascorbate et du cytochrome *c*. Le test est effectué selon le protocole décrit dans le texte. Dans le cas présent le pourcentage de mitochondries purifiées (MP) intactes est égal à 100 x [1-(11-6)/(1054-6)], soit 95%. Les vitesses données le long des traces sont exprimées en nmoles d'O_2 consommées par minute et pour 200 µg de protéines mitochondriales. Le volume final est de 1 ml.

C - ISOLEMENT DE CHLOROPLASTES DE FEUILLES DE POIS (METHODE DE MOURIOUX ET DOUCE 1981)

1 - Composition du milieu

La constitution du milieu de lavage est la suivante: sorbitol 330 mM, tampon HEPES 50 mM (pH 7.8), BSA 1.5 g/l, EDTA 2 mM.

2 - Purification des chloroplastes

Les culots contenant les chloroplastes « bruts » obtenus au cours de la purification des mitochondries sont resuspendus délicatement dans le milieu de lavage. Les agrégats sont élués par passage sur une toile à blutter (50 µm de vide de maille) puis déposés sur un gradient de Percoll, préformé pendant 50 min à 20 000g sur rotor vertical SS90 [50 % de Percoll (v/v), dans du milieu de broyage]. Après une centrifugation de 10 min à 5000 g (rotor SS90), les chloroplastes intacts forment une large bande près du fond du tube, tandis que les thylakoïdes

ainsi que les autres systèmes membranaires restent au sommet. Afin d'éliminer la majeure partie du Percoll, les chloroplates intacts sont récupérés à l'aide d'une pipette dans un volume d'environ 5 ml puis dilués dans 40 ml de milieu de lavage. La suspension est centrifugée durant 5 min à 1500 g (rotor SS 34 sorvall). Le surnageant est éliminé par aspiration. Le culot est soigneusement repris dans un faible volume de milieu de lavage. La suspension est conservée à la température de la glace fondante ou éventuellement stockée à -80°C jusqu'à son utilisation.

D - ISOLEMENT DE NOYAUX

1 - Méthode de Watson et Thompson (1986)

Cette technique a été utilisée pour extraire de l'ADN génomique en vue d'un Southern.

2 - Méthode de Green *et al.* (1989)

a - Composition des milieux

La constitution du milieu de broyage est la suivante: héxylène glycol 1M, PIPES/KOH 10 mM (pH 7.0), $MgCl_2$ 10 mM, triton X-100 0.5% (v/v), β-mercaptoéthanol 5 mM, PMSF 0.8 mM. Le milieu de lavage est ainsi constitué: Hexylène glycol 0.5 M, PIPES/KOH 10 mM (pH 7.0) $MgCl_2$ 10 mM, β-mercaptoéthanol 5 mM, PMSF 0.8 mM, Triton X-100 0.1% (v/v). La présence d'agents réducteurs, tels que le β-mercaptoéthanol, protège les groupements -SH des protéines contre toute oxydation. Le PMSF est un inhibiteur de protéase.

b - Purification de noyaux

Toutes les manipulations sont menées à + 4°C. Les feuilles de pois (500 g de matière fraîche) sont broyées rapidement (3 pulsions de 2 secondes), à l'aide d'un broyeur à hélice "Waring blendor" dans environ 500 ml de milieu de broyage. Le broyat est ensuite filtré sur 4 épaisseurs de gaze et une épaisseur de toile à bluter de 50 µm de vide de maille. Le filtrat est soumis à une centrifugation (rotor GS3, Sorvall) durant 10 min à 3000 g. Après élimination du surnageant, les culots contenant les noyaux sont remis délicatement en suspension à l'aide d'un pinceau dans du milieu de lavage contenant du Triton X-100 (80 ml). La suspension est centrifugée à 3000 g pendant 5 min (rotor SS34, Sorvall), puis le culot est de nouveau resuspendu à l'aide d'un pinceau dans du milieu de lavage sans Triton X-100, en évitant de

resuspendre l'amidon qui est peu soluble. La suspension est centrifugée à 3000g durant 5 min (rotor SS34, Sorvall) afin de sédimenter les noyaux. Ceux-ci peuvent être conservés à -80°C dans du milieu de lavage (5 ml) contenant 20% de glycérol (v/v).

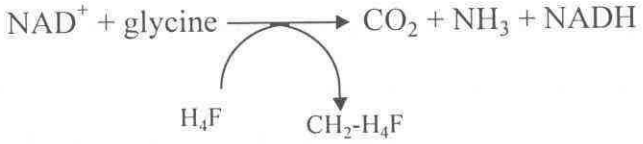
c - Préparation de l'extrait nucléaire

Les noyaux sont resuspendus (tube Falcon de 50 ml) dans 20 ml de tampon de lyse [KCl 110 mM, HEPES/KOH 15 mM (pH 7.5), $MgCl_2$ 5 mM, DTT 1 mM]. Un dixième du volume de sulfate d'ammonium 4M est alors progressivement ajouté puis délicatement mélangé (la solution devient immédiatement visqueuse). Après une incubation de 30 min sous agitation douce, la solution est centrifugée à 30 000 g durant 60 à 90 min (rotor SS 34, Sorvall) afin d'éliminer la chromatine ainsi que les histones. Le surnageant est repris puis son volume est mesuré. Du sulfate d'ammonium est alors progressivement ajouté pour atteindre une concentration finale de 0.3g/ml. La solution est ainsi maintenue sous agitation à +4°C durant 1 à 2 heures. Les protéines précipitées sont sédimentées dans un tube Corex à 10 000g durant 15 à 30 min puis séchées. Le culot est repris dans un tampon d'extraction [KCl 40 mM, HEPES/KOH 25 mM (pH 7.5), EDTA 0,1 mM, DTT 1mM, glycérol 10% (v/v)] puis la solution est dialysée dans un litre de tampon d'extraction durant 12 heures, avec 2 changements de tampon. Les éléments insolubles sont éliminés par centrifugation à 14 000 rpm pendant 10 min (centrifugeuse Eppendorff). Le surnageant (extrait nucléaire) est alors aliquoté et stocké à - 80°C jusqu'à utilisation.

III - MESURES D'ACTIVITES ENZYMATIQUES

A - MESURE D'ACTIVITE DE LA GLYCINE DECARBOXYLASE

L'extrait matriciel XM-300 est utilisé pour mesurer l'activité de la glycine décarboxylase par la méthode décrite par Neuburger *et al.* (1986) et Bourguignon *et al.* (1988). La glycine décarboxylase catalyse la réaction suivante:



L'oxydation de la glycine est mesurée, à l'aide d'un spectrophotomètre, en suivant la réduction du NAD^+ en NADH à 340 nm. La composition du milieu réactionnel est la

suivante: Tris 5 mM, Mops 5 mM (pH 7.4), β-mercaptoéthanol 1mM, EDTA 1 mM, pyridoxal-phosphate 10 µM, NAD^+ 2 mM, H_4F 350 µM. Ce milieu est placé sous argon pour éviter l'oxydation du tétrahydrofolate (H_4F) en dihydrofolate (H_2F). La réaction est initiée par addition de glycine (20 mM). La vitesse initiale, mesurée, est exprimée en nmoles de NADH formées par minute et par mg de protéines.

L'activité de la protéine T peut être mise en évidence par reconstitution du système enzymatique global.

B - MESURE DE L'ACTIVITE DE LA RIBULOSE 1,5 BISPHOSPHATE CARBOXYLASE/OXYGENASE (SELON LORIMER 1982)

Cinquante microlitres de chloroplastes sont mis en suspension dans 500 µl d'un milieu d'extraction [tris-HCl 150 mM (pH 7.8), $MgCl_2$ 5 mM, $MgSO_4$ 10 mM, KCl 50 mM, DTT 5 mM, PMSF 1 mM]. Trois séries rapides de congélation/décongélation dans un bain marie à 30°C permettent de libérer les protéines solubles de l'espace stromatique avec un bon rendement. Le dosage s'effectue de la façon suivante: tricine 50 mM (pH 8.2), $MgCl_2$ 20 mM, DTT 1 mM, Ribulose 1,5 bisphosphate 1 mM, $NaHCO_3$ 20 mM marqué au carbone 14 (2,5 µCi/ml). La réaction est stoppée à 1, 2 ou 3 mn par du HCL 2N. Les échantillons subissent une dessiccation toute la nuit à 60°C et la radioactivité du résidu est mesurée dans un compteur à scintillation Intertechnique SL 4000.

IV - PURIFICATION DE LA PROTEINE T IMPLIQUEE DANS LE COMPLEXE DE LA GLYCINE DECARBOXYLASE

A - PREPARATION DE L'EXTRAIT MATRICIEL

Les enzymes solubles sont libérées des mitochondries (420 mg de protéines totales mitochondriales) par choc osmotique et par 3 cycles de congélation/décongélation dans du tampon de composition suivante: Mops 5 mM, Tris/HCl 5 mM pH 7.4, EDTA 1mM, octylglycoside 1mM, β-mercaptoéthanol 1 mM et 20 µM de pyridoxal-5-phophate. Après élimination des membranes par ultracentrifugation, l'extrait est concentré par ultrafiltration à l'aide d'une membrane Amicon XM-300. Cet extrait matriciel (170 mg de protéines) contient toutes les protéines impliquées dans la conversion de la glycine en sérine (Neuburger *et al.*, 1986).

B - TECHNIQUE DE PURIFICATION DE LA PROTEINE T

La protéine T de pois est purifiée en deux étapes à partir d'un extrait matriciel:

- une chromatographie sur colonne Séphacryl S-300,
- une chromatographie d'échange d'anions sur colonne mono Q HR 5/5 (Pharmacia).

1 - Chromatographie sur colonne Séphacryl S-300

L'extrait matriciel (extrait XM-300) est déposé sur une colonne Séphacryl S-300 équilibrée avec du tampon A [Mops 5 mM, Tris 5mM (pH 7.5), KCl 50mM, β-mercaptoéthanol 1mM, EGTA 1mM, sérine 1mM, glycine 1 mM]. Les différentes protéines du système, impliquées dans l'oxydation de la glycine, sont éluées dans le même tampon, à 4°C, en fonction de leur masse moléculaire: les protéines P et SHMT (fraction lourde) sont éluées les premières, suivies de la protéine L (fraction intermédiaire) et des protéines T et H (fraction légère).

2 - Chromatographie sur colonne mono Q HR 5/5

Une aliquote de la fraction légère (6 mg de protéines) est déposée sur une colonne mono Q HR 5/5 (Pharmacia) équilibrée dans du tampon C [K-phosphate 10 mM (pH 7.5), EGTA 1 mM, β-mercaptoéthanol 1 mM]. Les protéines T et H sont éluées à 4°C et à pH 7.5 par un gradient linéaire croissant de K-phosphate (10-500 mM) dans du tampon C (Bourguignon *et al.*, 1988). La protéine T, chargée positivement à ce pH, n'est pas retenue par la colonne et est éluée immédiatement après le volume mort. Au contraire, la protéine H est éluée à une concentration environ de 210 mM de K-phosphate. Les fractions contenant les protéines H et T pures, sont concentrées à l'aide d'une membrane Amicon PM10 et dialysées contre du tampon dont la constitution est la suivante: KH_2PO_4 10 mM (pH 7.2), EDTA 1 mM, β-mercaptoéthanol 1 mM. Ces différentes fractions sont aliquotées puis conservées à -80°C.

Les chromatographies sont réalisées à 4°C avec le système « Fast Protein Liquid Chromatography » de Pharmacia (programmateur de gradient GP 250, pompe P 500, détecteur UV.2, enregistreur Rec 482, collecteur de fractions Frac 100). Tous les tampons sont filtrés sur membrane de 0,2 μm de diamètre de pore (Whatman).

V - TECHNIQUES D' ETUDE DES PROTEINES

A - EXTRACTION DES PROTEINES TOTALES DE FEUILLES

Les extraits totaux de protéines sont préparés selon la méthode de Hurkman et Tanaka, (1986). Après broyage dans l'azote liquide, les différents échantillons sont suspendus dans un tampon d'extraction [saccharose 0.7 M, Tris (pH 8.0), EDTA 5 mM, NaCl 0.1 M, β-mercaptoéthanol 2% (v/v) et PMSF 2 mM]. Un volume de phénol est ajouté à chaque mélange, l'ensemble est fortement agité pendant 30 secondes puis centrifugé (14 000 rpm, 15 min). La phase phénolique est récupérée à l'aide d'une pipette Pasteur; ensuite les protéines sont précipitées par l'addition de 5 volumes d'acétate d'ammonium (0,1 M) dissous dans du méthanol. Après centrifugation (14 000 rpm, 15 min), le culot de protéines est lavé deux fois avec 500 µl d'acétate d'ammonium/méthanol, puis une fois avec de l'acétone. On conserve à -80°C les protéines mises en suspension dans un faible volume de tampon dont la constitution est la suivante: Tris-HCl 50 mM (pH 6.8), DTT 100 mM, SDS 2% (m/v), bleu de bromophénol 0.1% (m/v), glycérol 10% (v/v).

B - DOSAGE DES PROTEINES

La concentration des protéines est déterminée suivant deux méthodes principales:

- la technique de Lowry *et al.* (1951).
- la méthode de Bradford (1976).

Dans les deux cas, le dosage est effectué par comparaison avec une gamme étalon réalisée pour chaque série de dosage avec une solution étalon de sérum albumine de boeuf (Sigma).

C - ELECTROPHORESE DES PROTEINES SUR GEL DE POLYACRYLAMIDE EN CONDITIONS DENATURANTES ET NON DENATURANTES

1 - Préparation des échantillons protéiques

Les échantillons (40 à 100 µg de protéines) sont repris dans un milieu dont la composition finale est la suivante: Tris 0,1 M (pH 6.8), glycérol 10% (v/v), SDS 2% (m/v), DTT 25 mM, Bleu de Bromophénol 0,1 % (m/v). Ce milieu dénature et charge négativement les protéines dont la dénaturation est parachevée par un chauffage de 2 min à 100°C.

2- Composition des gels

Les solutions d'acrylamide sont préparées dans un tampon Tris (pH 8.3 dans le gel de séparation, pH 6.5 dans le gel de concentration). Les protéines mitochondriales ont été séparées, par électrophorèse, à température ambiante, sur gel de polyacrylamide en gradient linéaire de 10% à 15% (m/v) en présence de SDS 0,1%. Un gel de concentration de polyacrylamide 5% (m/v) surmonte le tout.

Les gels pour électrophorèse en condition non dénaturante [acrylamide 5%, bisacrylamide 0.2% et Tris-HCl 40 mM (pH 8.0)] ont été préparés selon la technique de Laemmli (1970), le détergent étant omis dans les différents tampons.

L'électrophorèse est conduite sous un ampérage constant de 25mA en utilisant un tampon ainsi constitué: Tris 25 mM, glycine 0.192 M (pH 8.3), SDS 0.1 % (m/v) (Laemmli, 1970).

3 - Coloration des gels

Après électrophorèse, afin de révéler les polypeptides, on immerge le gel durant 2 heures dans une solution de coloration [acide acétique 10 % (v/v), éthanol 20% (v/v), bleu brillant de Coomassie (R 250 2,5 g/l)]. Le gel est ensuite décoloré par une succession de bains dans une solution de décoloration constituée d'acide acétique 10% (v/v) et d'éthanol 20 % (v/v)

D - IMMUNODETECTION DES PROTEINES PAR LA TECHNIQUE DU WESTERN-BLOT

Les anticorps ont été préparés par l'Institut Pasteur-NOVOTEC de Lyon à partir de la protéine T purifiée et séparée par électrophorèse en condition dénaturante sur gel de polyacrylamide (gradient de 7.5 à 15%). Les autres anticorps contre les protéines P, H et L étaient déjà disponibles au laboratoire.

Les protéines sont soumises à une éléctrophorèse en conditions dénaturantes sur gel de polyacrylamide (gradient de 10 à 15%) selon la technique de Laemmli (1970). Le gel ainsi que 4 feuilles de papier Whatman 3 MM et une feuille de nitrocellulose (Biorad) de la dimension du gel, sont incubés dans le tampon de transfert [Tris-HCl 48 mM (pH 8.3), glycine 39 mM, méthanol 20% (v/v)]. Le transfert est réalisé en utilisant un appareil de transfert semi-sec (Bio-Rad) composé de deux électrodes en graphite. On dépose successivement sur l'anode imprégnée de tampon de transfert: 2 feuilles de papier Whatman, la feuille de nitrocellulose, le gel de polyacrylamide et à nouveau 2 feuilles de papier

Whatman. Afin d'assurer un transfert efficace, on prend soin d'éliminer les bulles d'air entre les différentes couches. Le transfert est conduit sous un ampérage de 4 mA/cm^2 pendant 40 min. La membrane subit, par la suite, un traitement d'immunodétection décrit dans le paragraphe suivant.

Les polypeptides de la glycine décarboxylase réagissant avec les anticorps polyclonaux dirigés spécifiquement contre chacune des protéines sont révélés par un double marquage immunologique. Après transfert des protéines, la feuille de nitrocellulose est soumise à une série de traitements suivants (toutes les incubations sont effectuées à température ambiante et sous agitation):

- Durant 5 min, les sites de fixation libres de la membrane de nitrocellulose sont saturés dans du TBS-T 2% [Tris 50 mM pH 8.3, NaCl 250 mM, Tween 20 2% (v/v)].
- La membrane est ensuite incubée 2 heures avec des anticorps anti-protéine T, P, H ou L dilués au 1/20 000 dans du tampon TBS-T 0.05% [Tris 50 mM pH 8.3, NaCl 250 mM, tween 20 0.05% (m/v)].
- La membrane est lavée plusieurs fois pendant 25 min (5 x 5 min) dans du TBS-T 0.05% afin d'éliminer l'excès d'anticorps.
- La membrane est ensuite mise en présence d'un anticorps de chèvre, anti-IgG de lapin (dilution 1/3000), marqué à la phosphatase alcaline (Sigma). L' incubation de ce deuxième anticorps est conduite dans du tampon TBS-T 0,05% pendant une heure.
- La feuille de nitrocellulose est de nouveau rincée (5 lavages dans du tampon TBS-T 0.05%).

A ce stade, la membrane est incubée dans le tampon d'incubation de la phosphatase alcaline (Tris-HCl 100 mM pH 9.5, $MgCl_2$ 5 mM). Le complexe (protéine T/anticorps anti-protéine T/anti-IgG marqué à la phosphatase alcaline) est révélé par addition de 5-Bromo-4-chloro-3-indolyl-phosphate (BCIP) 0.06 mg/ml et de chlorure de bleu de tétrazolium nitré (NBT) 0,1 mg/ml. Le BCIP, déphosphorylé par la phosphatase alcaline réagit avec le NBT qui précipite sur le filtre et se colore en violet. La réaction enzymatique est stoppée par lavage de la membrane à l'eau distillée.

E - MICROSEQUENÇAGE DE LA PROTEINE T

Le microséquençage des peptides de la protéine T a été réalisé au sein du laboratoire de Jean Gagnon à l'IBS au CEA de Grenoble. La séquence des régions aminoterminales de la

protéine T purifiée ou des peptides de sa région interne, a été réalisée à l'aide d'un microséquenceur en phase gazeuse (Applied Biosystems, modèle 477-A) selon la méthode de dégradation d'Edman.

1 - Coupure au bromure de cyanogène

Le bromure de cyanogène coupe la chaîne peptidique du côté carboxyterminal des résidus méthionines en condition acide (Gross, 1967). Ces résidus sont alors transformés en homosérine ou en homosérine lactone (tableau 1). Les méthionines oxydées sous forme sulfoxyde méthionines ne peuvent pas réagir avec le bromure de cyanogène (tableau 1).

La réduction et la carboxyméthylation de la protéine T permettent d'éviter que les différents résidus cystéines interagissent entre eux. Ces deux réactions entraînent la rupture des ponts disulfures puis la transformation des groupes -SH libres en S-CH_2-COOH.

La protéine T est dénaturée dans une solution tampon [Chlorure de guanidine 6M, Tris-HCl 0.5M (pH 8.1), EDTA 2 mM] durant 30 min à 50 °C sous azote. Puis elle est réduite avec du dithiothréitol (300 moles de DTT par mole de protéine), pendant 3 heures, toujours à 50°C et sous azote. Pendant 20 min, à l'obscurité, la protéine est ensuite carboxyméthylée avec de l'acide iodoacétique (2 moles d'acide iodoacétique par mole de DTT précédemment ajoutées) à température ambiante. L'acide iodoacétique qui n'a pas réagi est neutralisé par un excès de DTT. Afin d'éliminer rapidement les réactifs, la solution est déposée sur une colonne G25 (Pharmacia) équilibrée avec du bicarbonate d'ammonium 0.01 M (pH 8.1). La protéine éluée, est ensuite lyophilisée puis stockée à - 80°C.

L'échantillon de protéine T réduite et carboxyméthylée est dissous à la concentration de 2.5 µg/µl dans une solution d'acide trifluoroacétique à 70% en volume. On rajoute ensuite du bromure de cyanogène en excès molaire de 1000 fois par rapport aux résidus méthionines de la protéine. Le mélange est incubé, sous azote, dans le noir, à 4°C pendant 24 heures. La réaction est arrêtée par l'addition de 10 à 20 volumes d'eau. L'excès de bromure de cyanogène est enlevé par évaporation sous vide (trompe à eau ou speed vac). L'échantillon est congelé puis lyophilisé.

	Formule chimique	Masse moyenne
Homosérine	CH_2-CH_2-OH | ~NH-CH-CO-	101,1051
Homosérine lactone (ester interne)	CO-O / | ~NH-CH \ | CH_2-CH_2	83,0902
Méthionine Sulfoxyde	CH_2-CH_2-SO-CH_3 | ~NH-CH-CO-	147,1920

Tableau 1: Les résidus méthionines produits après réaction avec le bromure de cyanogène sont transformés en Homosérines, Homosérine lactone ou Sulfoxyde.

2 - HPLC préparative

Les fragments de protéine T, provenant de la coupure au bromure de cyanogène, sont dissous dans une solution de chlorhydrate de guanidine 6M. La séparation par HPLC est réalisée à l'aide d'un appareil Applied Biosystem 140B. Soixante microgrammes de peptides (soit 1.5 nmol de protéine) sont injectés sur une colonne en phase inverse Browlee C8 (5µm) de 2,1x100 mm. Les deux éluants A et B sont composés d'un mélange eau/acétonitrile/TFA avec en volume A= 100/0/0.1 et B= 30/70/0.08. Le débit est de 200 µl/min. Le gradient linéaire est de 5 à 70% du solvant B en 50 min. Les peptides sont détectés en sortie de colonne par une cellule UV à 215 nm. Les différentes fractions sont recueillies manuellement à la sortie de la colonne. Elles sont par la suite partiellement évaporées sous vide de façon à éliminer l'acétonitrile.

Les fractions contenant un mélange de petits peptides ont été à nouveau purifiées par HPLC sur un colonne Browlee C18 (5 µm) de 2,1x100 mm avec un gradient de 15 à 70 % du solvant B en 50 min. Les échantillons issus de cette deuxième séparation sont concentrés. Un tiers des fractions est analysé par spectrométrie de masse en mode FAB. L'autre partie sera analysée par dégradation d'Edman.

Les fractions contenant des peptides de masse supérieure à 20 000 Da ont été directement soumises à l'analyse par dégradation d'Edman. On a pu isoler deux fragments [134-367] et [58-367] de masse élevée en effectuant une seconde purification par HPLC préparative avec 3 nmol (120 µg) du même échantillon de la protéine T coupée au bromure de cyanogène. Ces deux fragments ont ensuite subi des digestions par des endopeptidases.

VI - SPECTROMETRIE DE MASSE

Cette étude a été réalisée au laboratoire de Spectrométrie de Masse des Protéines de l'IBS en collaboration avec les Dr E. Forest et V. Mérand.

La spectrométrie de masse est une technique utilisée depuis longtemps par les chimistes pour l'analyse structurale, l'analyse élémentaire, et la détermination des rapports isotopiques de composés organiques. Depuis une vingtaine d'années, de nombreuses améliorations techniques ont rendu possible l'analyse de composés de masse moléculaire élevée, polaires et thermolabiles comme les protéines (pour plus d'information se référer à Beckey et Schulten, 1975; Barber *et al.*, 1981; Meng *et al.*, 1988; Karas et Hillenkamp, 1988). La spectrométrie de masse s'est donc ouverte à la biologie pour devenir une technique complémentaire à l'étude des protéines. Elle présente de nombreuses applications:

- la vérification d'une séquence,
- la détermination et la localisation de modifications post-traductionnelles,
- la localisation des ponts disulfures,
- l'étude des peptides bloqués du côté aminoterminal,
- la détermination de la pureté d'une protéine.

Ainsi, nous aborderons de façon succincte le principe de la spectrométrie de masse. Il s'agit d'une présentation, destinée à de non spécialistes, d'une technique couramment utilisée en biologie.

A - RAPPEL THEORIQUE

Considérons une particule de charge positive q et de masse m qui se déplace perpendiculairement à un champ magnétique uniforme B. La force magnétique ($F=qv \times B$) est perpendiculaire à la vitesse v, de sorte qu'elle en change la direction, mais pas la grandeur. Puisque le module de la force magnétique reste inchangé car v et B sont constants en grandeur et perpendiculaires, l'accélération ($a = F/m$) est constante en grandeur. L'accélération étant

toujours perpendiculaire à v, les ions se déplacent alors à une vitesse constante sur des trajectoires circulaires de rayon R ($R = mv/qB$), jusqu'à ce qu'ils atteignent un détecteur de masse, les distances augmentant proportionnellement à leur masse. Ainsi, le rayon de courbure croît avec la masse permettant la séparation des ions de même charge mais de masse différente (figure 8).

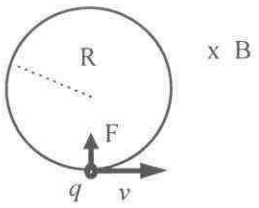


Figure 8: Une particule chargée se déplace avec une vitesse v, perpendiculairement à un champ magnétique uniforme B, orienté dans le sens entrant dans le plan du dessin. Comme la force et l'accélération sont constamment perpendiculaires à *v* et restent inchangées en module, la particule est en mouvement circulaire uniforme

B - DESCRIPTION TECHNIQUE

Un spectromètre de masse est un ensemble composé de cinq éléments maintenus sous vide:

- un système d'introduction des échantillons,
- une source qui permet l'ionisation des molécules,
- un analyseur (il trie les ions suivant leur rapport masse sur charge m/z),
- un détecteur,
- un système informatique qui gère l'enregistrement et les données.

L'étude de la protéine T a été réalisée grâce aux techniques de spectrométrie de masse FAB et ESI et aux techniques de couplage HPLC/ESI/MS.

1 - La spectrométrie de masse FAB (Fast Atom Bombardment)

a - Principe

La FAB a été introduite par Barber *et al.* (1981) au début des années 1980. Cette technique permet l'analyse en routine des peptides et des petites protéines de masse inférieure à 15 000 Daltons. Pour des molécules de masse plus élevée d'autres techniques sont alors utilisées.

Le composé à analyser est dissous dans un solvant (solution d'acide acétique 5% ou un mélange eau/acétonitrile/TFA 50/50/0.1). Ensuite, sur la cible FAB contenant 1 à 2 µl d'une matrice liquide, on dépose 0.1 à 5 nmol de protéines ou peptides dissous [Quelques centaines de picomoles seulement en LSIMS (Liquid Secondary Ion Mass Spectrometry)]. Sous un vide de 10^{-6} mbar, le mélange est soumis à un bombardement d'un faisceau d'atomes d'argon ou

de xénon [actuellement les atomes de gaz rares sont de plus en plus souvent remplacés par des ions métalliques de césium, on parle alors de LSIMS (Alberth *et al.*, 1982)], accélérés avec une énergie de 2 à 8 keV, expulsant ainsi de la matrice des ions positifs, négatifs et des particules neutres. En effet, les ions positifs ont capté soit un proton $(M + H)^+$, soit un cation métallique $(M + Na)^+$, alors que les ions négatifs sont des molécules qui ont soit perdu un proton $(M - H)^-$, soit fixé un anion $(M - Cl)^-$ (Figure 9). Les molécules chargées sont ensuite accélérées sous une tension de 2 à 10 kV pour être détectées et analysées par un analyseur de masse.

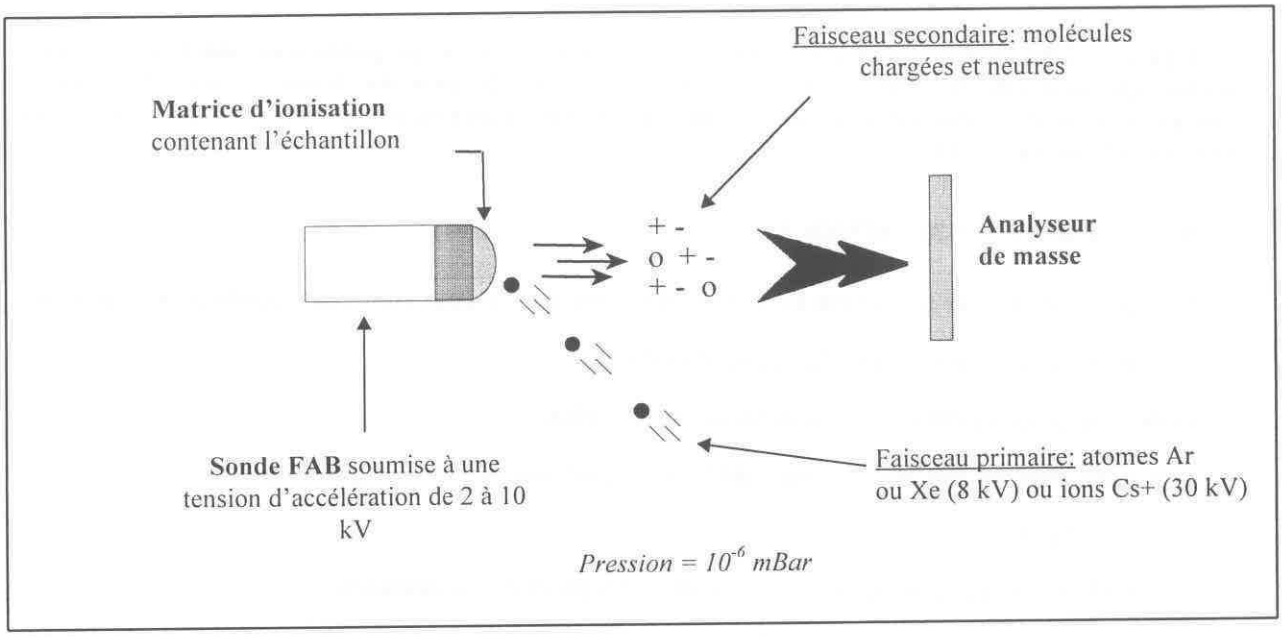


Figure 9: Désorption et ionisation d'un échantillon dans la source FAB ou LSIMS

La sensibilité de cette technique peut être améliorée de 20 à 150 fois (pour une molécule de 200 à 2000 Da) par la technique de FAB à flux continu (CF/FAB) (Caprioli, 1988a; 1988b; Wolf *et al.*, 1992; Tomer et Parker, 1992; Caprioli et Suter, 1992). Pour plus d'information voir la thèse de Mérand (1994).

b - Domaine d'application

L'application la plus courante de cette technique reste l'étude de la cartographie de protéine (Suter *et al.*, 1991). Ainsi, les peptides issus de la digestion enzymatique d'une protéine (par la trypsine ou la chymotrypsine) sont injectés directement sur une colonne HPLC et analysés en masse par la technique du couplage.

2 - La spectrométrie de masse par électrospray (ESI)

a - Principe

C'est en 1968, que l'on a mis en évidence, pour la première fois, la formation d'ions à partir de la surface d'un liquide sous l'effet d'un fort champ électrique (Dole *et al.*, 1968). Cette technique a eu des retombées importantes dans le domaine d'étude des protéines.

La première étape de l'ionisation par électrospray consiste à vaporiser une solution d'échantillons sous forme de gouttelettes chargées. L'échantillon en solution est amené dans la source par l'intermédiaire d'un tube capillaire de silice fondue de 75 µm de diamètre prolongé par un tube en acier inoxydable (Thomson et Irbane, 1979). La nébulisation et l'ionisation sont réalisées par l'action d'un champ électrostatique (environ 3000 V/cm) sur la solution, à pression atmosphérique. Le débit est alors limité à quelques microlitres par minute.

Fenn *et al.* (1990) et Mann *et al.* (1989) expliquent le mécanisme d'émission d'ions en phase gazeuse à partir de gouttelettes chargées: l'excès de charges électriques dans chaque gouttelette, combiné à l'évaporation progressive des solvants, conduit à un point où les répulsions coulombiennes excèdent les forces de cohésion des gouttelettes. Ce phénomène induit une cascade d'explosions coulombiennes au cours desquelles des micro-gouttelettes formées permettent la désorption de molécules multichargées (figure 10).

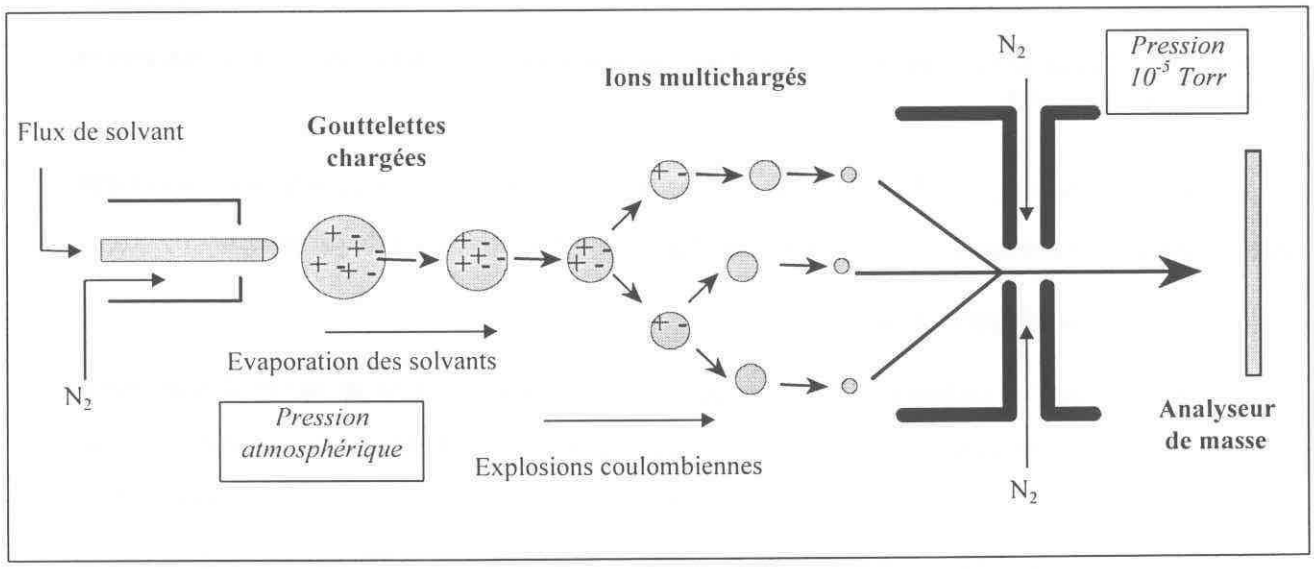


Figure 10: Source ionspray. Schéma de formation des gouttelettes par electrospray. Sous l'effet de l'évaporation du solvant, l'excès de charges en solution provoque la désorption des molécules mono ou multichargées.

Les ions multichargés formés dans la source à pression atmosphérique sont introduits dans le vide du spectrophotomètre de masse par l'intermédiaire d'un orifice d'échantillonnage (diamètre 100 µm) puis les gouttelettes sont desséchées à travers un rideau d'azote. Le mode d'ionisation peut être, soit positif, dans ce cas les charges sont données par les protons et les ions sont alors du type $(M+nH)^{n+}$, soit négatif, alors les charges proviennent de cations (Na^+, K^+, NH_4^+) ou de la perte de contre ions (Cl^-).

Ces espèces multichargées permettent de ce fait l'analyse de masse moléculaire élevée, les rapports masse sur charge (m/z) se situant entre 10 et 3000 uma (unité de masse atomique). Si l'on connaît la masse des molécules responsables de l'apparition des charges (H^+, Na^+, NH_4^+ ...), deux états de charges sont suffisants pour calculer la masse d'une molécule (figure 11) (Mann *et al.*, 1989).

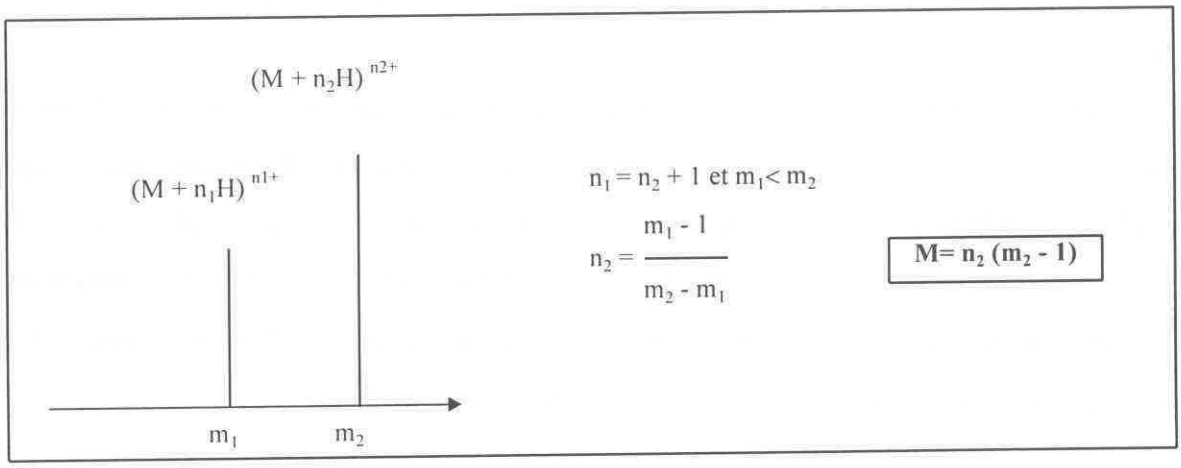


Figure 11: Calcul de la masse moléculaire d'un composé à partir de deux espèces multichargées en mode d'ionisation positif.

De plus, le couplage HPLC/MS avec l'interface électrospray a été rapidement développé pour l'analyse de mélanges de peptides et de digestions enzymatiques de protéines.

b - Domaine d'application

L'ionisation par les techniques de spray permet l'analyse en routine de molécules de masse de 100 à 200 000 Da avec une précision de l'ordre de 0.01% et évite les fragmentations (Van Dorsselaer *et al.*, 1990; Smith *et al.*, 1990). La technique est surtout utilisée pour l'étude structurale des protéines. En particulier, elle sert à la vérification de séquences, à l'identification et la localisation de modifications post-traductionnelles, à la localisation de ponts disulfures. C'est une technique fort utilisée pour l'étude des produits de digestion

enzymatiques de protéines (Covey *et al.*, 1991), pour l'analyse de nucléotides (Covey *et al.*, 1988), de stéroïdes (Hemling *et al.*, 1990) et d'autres molécules chimiques non volatiles (Banno et Horimoto, 1991).

3 - Principe d'un analyseur de masse

Une fois les ions formés dans la source, il reste à les séparer et à les analyser en fonction de leur rapport m/z. Pour ce faire, il faut extraire les ions de la source et les envoyer dans le tube analyseur: les ions sont repoussés vers la fente de sortie dès leur formation par le faible potentiel positif (quelques volts) du repousseur. Ils sont alors accélérés vers la partie analytique de l'appareil par une série de champs électrostatiques créés par des plaques successives. Seuls les ions positifs sont analysés; les ions négatifs sont piégés. En principe, toutes les espèces ioniques reçoivent la même énergie potentielle (tension d'accélération: 3 à 8 kv). En conséquence, elles sont censées avoir toutes la même énergie cinétique au moment d'entrer dans l'analyseur. Cette énergie cinétique est égale à zV telle que:

$$1/2mv^2=zV$$

Placés dans un champ magnétique B perpendiculaire à la direction du mouvement, les ions décrivent un arc de cercle de rayon R tel que:

$$zvB=mv^2/R$$

De ces deux relations, on en déduit:

$$m/z=B^2R^2/2V$$

Ainsi, pour une tension d 'accélération et une valeur du champ magnétique données, seules les espèces ayant un rapport m/z déterminé décriront la trajectoire de rayon R et seront détectées à la sortie de l'aimant. Un balayage de V ou de B (le plus fréquent) permet alors de détecter les différents rapports m/z des ions.

C - PROTOCOLE OPERATOIRE POUR L'ETUDE DE LA PROTEINE T

1 - FAB/MS

Les fractions contenant les petits peptides élués à partir d'une colonne HPLC C18 sont analysées en utilisant un spectromètre de masse de type VG ZAB SEQ (Fisons Instrument) équipé d'une source d'ions FAB et d'un canon à cesium. Le volume de l'échantillon est réduit sous vide jusqu'à 2 µl, puis acidifié par addition de 1 µl de 1% d'acide trifluoracétique.

Chaque échantillon (2 µl ≈ 100 pmol) est déposé sur la cible FAB de matrice (2 µl) contenant 50 % (v/v) de glycérol et 50 % (v/v) de thioglycérol. L'énergie des ions césium est de 30 KeV, la tension d'accélération des ions dans la source, de 8 kV et la résolution, de m/z = 800. L'appareil est balayé exponentiellement de la masse 6000 à la masse 500 avec une vitesse de 10 sec/ balayage.

2 - HPLC/ESI/MS

La protéine T a été analysée sous forme: native, et dénaturée, réduite, carboxyméthylée puis coupée au bromure de cyanogène.

Le spectromètre de masse utilisé est un API III (Sciex) à triple quadripôle, équipé d'une source à pression atmosphérique et d'une interface SCIEX Ionspray. L'appareil HPLC (Applied Biosystems 130A) est modifié pour travailler à un débit de 5 µl/min. Il est équipé d'une colonne capillaire en silice fondue fabriquée au laboratoire du Dr E. Forest, remplie de particules Nucléosil C8, 5 µm, de dimensions 0.25 x 150 mm. Les solvants contiennent 0,1% de TFA. Les paramètres de la séparation HPLC et de l'analyse par spectrométrie de masse sont donnés pour chaque échantillon dans le tableau 2.

Protéine T	gradient d'acétonitrile	durée du gradient	quantité injectée	m/z enregistré	vitesse de balayage
native	9-90%	30 min	2,5 µg (60 pmol)	700-1700	2 msec pour 0,5 uma
dénaturée, réduite, carboxyméthylée et coupée au CNBr	7-70%	45 min	1,5 µg (35 pmol)	600-1600	2 msec pour 0,5 uma

Tableau 2: Analyses HPLC/ESI/MS de la protéine T

VII - TECHNIQUES GENERALES D' ETUDE DES ACIDES NUCLEIQUES

Nous ne présenterons dans ce chapitre que les techniques particulières de ce travail. Les techniques les plus couramment utilisées telles que, le dosage de l'ADN et des ARN, la préparation et la digestion des plasmides, la préparation des gels d'agarose, la purification des ARN messagers sur colonne d'oligo (dT)-celluloses et le sous-clonage de fragments d'ADN dans des plasmides sont décrites dans le manuel de laboratoire Sambrook *et al.* (1989).

A - TECHNIQUES D'ANALYSE DES ACIDES DEOXYRIBONUCLEIQUES

1 - Extraction de l'ADN nucléaire de pois

L'ADN nucléaire a été préparé à partir de noyaux de feuilles de pois purifiés selon la technique décrite par Watson et Thompson (1986). L'ADN est purifié par deux centrifugations sur gradients de CsCl, suivies d'une dialyse avec 10mM-Tris/HCl (pH8.0), 1mM EDTA. L'ADN est précipité en présence de 2.5 volumes d'éthanol et de 0,1 volume d'acétate de sodium pH 5.3. Le culot (rincé à l'éthanol 70%, pour les grandes quantités d'acide nucléique ou à l'éthanol 80%, pour les petites quantités) est repris dans du tampon TE [10mM Tris-HCl (pH8.0) et 1mM EDTA (pH 8.0)].

2 - Techniques d'hybridation

a - Marquage de l'ADN

Les marquages sont effectués avec le kit Boehringer "Ramdom Primed Labelling". Un mélange de nucléotides 9-mers (1,25 unité de D.O. à 260 nm) contenant des séquences aléatoires est utilisé comme amorce pour la polymérase de Klenow (provenant de E. *coli*). L'ADN (25 ng) est dénaturé 5 min à 95-100°C, puis refroidi jusqu'à la température de 20°C. On ajoute alors successivement:

- 5 µl de tampon 10 fois concentré [Tris-HCl 0.4M pH 7.2, $MgCl_2$ 0.1M, DTT 1mM, SAB 0,2% (m/v)],
- 1nmole de dATP, 1nmole de dGTP, 1nmole de dTTP,
- 50 µCi de [α-^{32}P]- dCTP (10mCi / ml, 3000 Ci/mmole, Amersham),
- 2 unités de polymérase de Klenow.

La réaction est conduite dans un volume final de 50 µl à 37°C pendant 30 min. Elle est stoppée par l'ajout d'EDTA 10mM, pH 8.0.

b - Transfert de l'ADN sur membrane de nylon (Thomas ,1980)

L'ADN est transféré sur un filtre de nylon Hybond-N^+ (Amersham) selon la méthode de Southern (1975). Le gel est placé sur deux couches de papier Whatman 3 MM dont les extrémités baignent dans une solution de SSC 10 fois concentrée: SSC 10X [citrate de sodium (pH 7.0) 15 mM, NaCl 150 mM]. On dépose successivement sur le gel: le filtre de nylon, trois feuilles de papier Whatman, imbibées de SSC 10X, 5 cm (environ) de papier absorbant

[QuickDraw blotting paper (SIGMA CHEMICAL)], une plaque de verre surmontée d'une masse de 500 g environ. Après 12 heures de transfert, l'ADN subit une fixation alcaline. Durant 20 min, le filtre est déposé sur 3 couches de papier Whatman imbibées de NaOH 0.4 M (la face du filtre où se trouve fixé l'ADN étant dirigée vers le haut) puis rincé dans du SSC 5X. Il est possible de fixer de façon plus efficace les acides nucléiques, en exposant la membrane aux ultra-violets. Dans ce cas, la membrane séchée durant 10 min à 80°C, est exposée (la face contenant l'ADN ou l'ARN dirigée face aux lampes UV) à un rayonnement UV de 70 000 µjoules/cm_2 pendant 10 à 15 sec (Stratalinker, Stragène).

c - Préhybridation, hybridation et rinçage

Les sites de fixation libres de la membrane de nylon sont saturés au cours d'une incubation de 30 min à 43°C dans 10 ml de tampon de préhybridation [50% (v/v) de formamide déionisée, BSA 0,2% (m/v), PVP 40 000 0,2% (m/v), Ficoll 40 000 0,2% (m/v), Tris 50 mM (pH 7.5), Na_2pyrophosphate 0,1% (m/v), SDS 1% (m/v), sulfate de dextran 10% (m/v), NaCl 1.2 % (m/v), ADN de sperme de saumon (0.1 mg/ml)]. La sonde radioactive marquée au P^{32} est dénaturée durant 5 min à 95-100°C, puis ajoutée au milieu de préhybridation. L'hybridation est effectuée pendant 12 heures à 43°C (appareil à hybridation "mini hybridisation oven", Appligene). Le filtre nylon subit ensuite, sous agitation, plusieurs cycles de lavage:

- Trois cycles de lavage de 20 min dans du 2 x SSC concentré, et du SDS 0.1% (m/v) effectués à température ambiante,
- Un cycle de 20 minutes à 65°C dans du 0.5 x SSC concentré, et du SDS 0.1% (m/v) suivi de deux rinçages dans du 0.5 x SSC.

Les filtres humides sont ensuite disposés entre 2 feuilles de film plastique (Saranwrap) et exposés en présence d'un film autoradiographique βmax (Amersham) à -80°C.

3 - Amplification *in vitro* de fragments d'ADN

Afin d'amplifier les fragments d'ADN clonés dans le phage λgt11 ou dans le plasmide pUC18, nous avons utilisé l'enzyme *taq* polymérase (provenant d'un organisme procaryote nommé *Thermophilus aquaticus* vivant dans les sources d'eau chaude) associée à la technique d'amplification par réaction en chaîne ("Polymérase Chain Réaction" ou PCR).

Nous avons utilisé cette technique pour préparer spécifiquement et en quantité importante, certains fragments d'ADN, afin d'atteindre les objectifs suivants:

- déterminer la taille d'un insert cloné dans un phage,
- sous cloner un fragment d'ADN dans le plasmide pUC18,
- préparer une sonde nucléotidique pour réaliser une analyse d'hybridation...

Le principe de la méthode consiste à amplifier, *in vitro*, un fragment d'ADN par une suite de cycles d'hybridation/polymérisation/dénaturation des brins d'ADN:

1) Dénaturation à haute température de l'ADN double brin à amplifier,

2) Diminution rapide de la température permettant l'hybridation d'amorces nucléotidiques mono-brin (entre 17 et 35 mers selon le choix) complémentaires aux deux extrémités (5' et 3') du fragment d'ADN à amplifier.

3) Copie de l'ADN matrice par la *taq* polymérase (polymérisation dans le sens 5'-3' à partir de deux amorces), puis retour à l'étape de dénaturation de l'ADN.

L'ADN matrice (1 à 10 ng) est dilué dans 100 µl de milieu de PCR [Tris-HCl 20 mM (pH 8.3), Tween-20 0.5% (v/v), nucléotides triphosphates 50 µM, KCl 50 mM, gélatine 1 mg/ml et 40 pmoles de chaque amorce spécifique]. Après centrifugation durant quelques secondes, le milieu est placé à 96°C durant 5 min dans un incubateur programmable (Crocodile/Appligene). Une quantité de 0.4 U de *Taq* polymérase (Appligene) est ensuite ajoutée au milieu. Puis le milieu est recouvert de 100 µl d'huile minérale (afin d'éviter toute évaporation). L'amplification est réalisée après trente cycles, chaque cycle comportant une période:

- de 30 secondes à 96°C permettant la séparation des 2 brins de l'ADN,
- de 30 secondes à 50-55°C conduisant à l'hybridation des amorces. La température d'hybridation dépend de la nature de amorces (longueur, nombre de purine).
- de 1 min 30 à 70°C permettant l'élongation de l'ADN par la *Taq* polymérase.

La réaction se termine par une incubation à 95°C durant 30s puis à 50-55°C durant 30s et enfin à 72°C durant 5 min.

L'analyse des fragments amplifiés est réalisée en soumettant 10 µl du mélange d'amplification à une électrophorèse sur gel d'agarose.

Les amorces spécifiques des ADNc correspondant aux protéines P et rbcS sont repectivement (5'-CTGTAACAGCATTCATGGAACGTG-3'; 5'-GATAGCAGCATCTGCA ATTGCTTT-3') et (5'-CAATGGTGGAAGAGTAAAGTGCA-3'; 5'-CCATTTCCATTT CACAGTTCG-3').

4 - Technique de séquençage de l'ADN

a - Réaction de séquençage

Toutes les séquences ont été réalisées selon la méthode de Sanger *et al.*(1977) avec le kit "DNA sequencing kit" (United States Biochemical). Les séquences ont été déterminées sur les deux brins de l'ADN des plasmides recombinant pUC18 en utilisant des amorces appropriées. Le protocole de séquençage est composé de 4 étapes successives:

- Dénaturation: L'ADN purifié (1 à 4 µg) est dénaturé 15 min à 37°C dans du tampon (NaOH 0.2M, EDTA 0.2mM), puis précipité à l'éthanol en présence d'acétate de sodium 0.3M (pH 5.2). L'ADN est repris par 7 µl d'eau, 2 µl de tampon d'hybridation [Tris-HCl 200mM (pH 7.5), $MgCl_2$ 100mM, NaCl 250mM] et 1 µl d'amorce (3 à 5 pmol/µl d'oligonucléotides de 17 bases).

- Hybridation: La réaction d'hybridation est réalisée par une incubation de 2 min à 65°C dans un bloc chauffant. Le bloc est alors extrait de son support puis placé à température ambiante, afin de laisser progressivement (environ 30 min) la température décroître vers 35°C.

- Elongation: La réaction d'élongation est conduite pendant 5 min à 37°C dans un volume de 15 µl avec 3 µl de tampon de réaction (dithiothréitol 20mM, 1.5 µM dGTP, 1.5 µM dCTP, 1.5µM dTTP), 0.5 µl d'[α-^{35}S]-dATP, 20 µCi/ µl (5 µCi Amersham) et 1.6 unité d'ADN polymérase T_7.

- Terminaison: Les réactions de terminaisons sont réalisées par l'ajout simultané de 3.5µl de milieu réactionnel dans 4 tubes Eppendorf dont le contenu, dans l'ordre, est le suivant: 2.5 µl de nucléotides de terminaison ddGTP (8 µM), ddATP (8 µM), ddTTP (8 µM) et ddCTP (8 µM). En outre, la composition de ces tampons de terminaison est la suivante: dGTP 80 µM, dATP 80 µM, dTTP 80 µM, dCTP 80µM et du NaCl 50mM.

La réaction de terminaison est conduite à 37°C pendant 5 min puis stoppée par 4 µl de solution d'arrêt [formamide 95% (v/v), EDTA 20 mM (pH 7,5), bleu de bromophénol 0.05% (m/v) et du xylène cyanol FF 0.05% (m/v)].

Les milieux réactionnels sont alors conservés à - 20°C avant analyse par électrophorèse sur gel d'acrylamide.

b - Electrophorèse sur gel de polyacrylamide

Le mélange réactionnel généré au cours de chacune des 4 réactions de terminaison est analysé sur gel d'acrylamide en conditions dénaturantes. On prépare, dans l'eau, une solution d'acrylamide 40 % (m/v)-N, N'-méthylènebisacrylamide 2 % (m/v), à laquelle on ajoute 5 % (m/v) d'amberlite. La solution est ensuite filtrée puis conservée à 4°C à l'abri de la lumière.

Le gel d'acrylamide (6%) -bisacrylamide (0,3 %) est préparé dans 80 ml de tampon TBE-urée (Tris-base 100 mM, acide borique 86 mM, EDTA 2 mM, urée 8.7 M). Après addition de 480 µl de persulfate d'ammonium 10% (m/v) et de 62 µl de TEMED (Bio-Rad), le gel est coulé entre deux plaques de verre siliconisées (33,5 cm x 39 cm), séparées par deux espaceurs de 0,4 mm d'épaisseur.

L'électrophorèse est réalisée à l'aide d'un appareil BRL (modèle PS3003) qui, par l'intermédiaire d'une plaque d'aluminium en étroit contact avec une des deux plaques de verre renfermant le gel, permet de redistribuer uniformément la chaleur et donc d'assurer une migration uniforme en tout point du gel.

L'électrophorèse est conduite dans le tampon TBE à une puissance de 65 W et une tension de 1600 volts. Avant dépôt des échantillons, le gel est préchauffé par une électrophorèse de 45 min. Les échantillons sont alors dénaturés par une incubation de 2 min à 75°C, puis déposés dans les puits préalablement rincés par du tampon TBE.

Pour chaque réaction, quatre temps de migration différents 1 heure 45 min, 3 heures 30 min, 5 heures 30 min et 7 heures 30 min permettent de déterminer jusqu'à 400 bases dans la séquence nucléotidique.

Le gel est ensuite séché sur papier Whatmann 3 MM (sécheur de gel modèle SE 1160, Hoeffer Scientific Instrument) pendant une heure.

Les bandes, marquées au ^{35}S, sont révélées par une autoradiographie d'une durée de 12 à 24 heures avec un film Amersham (Hyper film-MP) à température ambiante. Les séquences sont lues et analysées à l'aide du logiciel PC/GENE Release 6.8 (IntelliGenetics, Inc).

5 - Synthèse des oligonucléotides

Les synthèses d'oligonucléotides ont été réalisées dans le laboratoire de Biochimie (CENG) par le Dr Issartel, à l'aide d'un synthétiseur automatique d'oligonucléotides (DNA synthetizer modèle 381A, Applied Biosystem Inc) utilisant des colonnes de 40 nmoles de capacité. L'oligonucléotide synthétisé est décroché de la colonne par un rinçage avec 1,5 ml

d'hydroxyde d'ammonium 30 %, puis incubé toute une nuit à 65°C. La précipitation s'effectue dans l'éthanol en présence d'acétate de sodium 0,3 M. La quantification est réalisée au spectrophotomètre à 260 nm (1 unité de D.O.= 20 µg/ml). Les oligonucléotides sont alors stockés à -20°C jusqu'à leur utilisation.

B - TECHNIQUES D'ANALYSE DES ACIDES RIBONUCLEIQUES

1 - Extraction des ARN totaux

Cinq grammes de tissus (feuilles, tiges et racines), provenant de plants de pois âgés de 8 jours, sont broyés dans un mortier en présence d'azote liquide. La poudre obtenue est homogénéisée à température ambiante dans 20 ml de tampon de lyse [Tris-HCl 50mM pH 9.0, SDS 1% (m/v)] et 20 ml de phénol équilibré à pH 8 par du Tris-HCl [au phénol est ajoutée de l'hydroquinoline 0.1% (m/v)]. Après une centrifugation de 10 min à 3000 rpm, les ARN contenus dans le surnageant subissent de nouveau une extraction au phénol (dans un rapport volumique de 1/1), puis une extraction au phénol/chloroforme/alcool isoamylique (25/24/1) et enfin une extraction au chloroforme/alcool isoamylique (équilibré à pH 8 par du Tris-HCl) qui permet d'éliminer le phénol restant. Les ARN sont ensuite précipités sélectivement par addition de chlorure de lithium (LiCl) à la concentration finale de 2M, à 0°C pendant 3 heures. Après centrifugation à 18 000 rpm durant 30 min, le culot contenant les ARN est solubilisé dans 10 ml d'eau bidistillée et autoclavée. Les ARN sont, par la suite, précipités par ajout d'acétate de sodium 0.3 M dans 2 à 2.5 volumes d'éthanol. L'échantillon est maintenu à 4°C pendant 1 heure, ou à -80°C durant 10 à 15 min. Après centrifugation à 14 000 rpm pendant 30 min (centrifugeuse Eppendorff), le culot est ensuite lavé à l'éthanol 70% (v/v). Le culot est repris dans 0.1 à 0.2 volume d'eau, puis les ARN sont quantifiés au spectrophotomètre à 260nm.

Les ARN messagers sont purifiés sur colonne oligo(dT)-cellulose selon la technique décrite dans Sambrook *et al.*, 1989.

2 - Electrophorèse des ARN sur gel d'agarose et transfert sur membrane de nylon

Les électrophorèses sont réalisées sur gels d'agarose horizontaux à la concentration de 1.3% (m/v) dans du tampon phosphate 10 mM pH 7.0 (4 mM de NaH_2PO_4 et 6 mM de Na_2HPO_4). L'agarose est dissous, dans le tampon, par ébullition au four micro-ondes puis

refroidi jusqu'à une température de 50°C-60°C avant d'être coulé. Les ARN (2 µg pour les ARN messagers et 20 µg pour les ARN totaux) sont dénaturés, dans un volume final de 30 µl, par un traitement au glyoxal [glyoxal 1M, DMSO 53% (v/v) et 0.01 M NaPi (pH 7.0)]. L'ensemble est incubé pendant 60 min à 50°C, puis centrifugé et mis dans la glace. Les échantillons sont déposés ensuite dans le gel à l'aide de 6 µl d'un tampon de charge glyoxal/DMSO [50% glycérol (m/v ou v/v), 10mM NaPi (pH 7,0), bleu de bromophénol 0.25% (m/v), xylène cyanol FF 0.25% (m/v)]. L'électrophorèse est conduite à un voltage de 100 volts, durant 3 heures, avec recyclage permanent du milieu de migration [0.01M de NaPi (pH 7.0)].

La taille des ARN sera déterminée par rapport à une échelle d'ARN dont la longueur est connue [RNA Molecular Weight Marker III: 0.3 à 7.4 kb (Boehringer)] et qui migrent parallèlement aux échantillons.

Après électrophorèse, les ARN sont transférés, par capillarité, sur un filtre de nylon (Hybond N^+, Amersham), selon la méthode décrite pour les ADN par Southern (1975) (Cf. paragraphe VII. A. 2). Le tampon de transfert est composé de 10 x SSC. Les ARN sont fixés par traitement au four à UV (Stratalinker, Stratagène).

3 - Hybridation, quantification et coloration des ARN

Les sondes d'ADN sont marquées au P^{32} comme indiqué précédemment. Les techniques de préhybridation et d'hybridation sont identiques à celles utilisées pour l'ADN (Cf. paragraphe VII. A. 2).

La membrane de nylon, humide, est disposée entre deux feuilles de film plastique (type Saran Wrap) et exposée en présence d'un film autoradiographique β-max (Amersham) à -80°C pendant 12 à 72 heures.

La quantification relative des signaux est effectuée grâce au Phosphorimager. La membrane de nylon est placée au contact d'une plaque contenant des cristaux de $BaFBrEu^{2+}$. Le rayonnement β émis par les noyaux ^{32}P de la sonde provoque la formation d'un complexe $BaFBr^-/Eu^{3+}$ qui, après excitation par un rayon laser, émet une lumière bleue quantifiée par l'appareil.

Après autoradiographie, la membrane de nylon est colorée au bleu de méthylène par une incubation de 15 min dans l'acide acétique 5% (v/v) suivie d'une incubation de 10 min

dans une solution de coloration contenant de l'acétate de sodium 0.5M pH 5.2 et du bleu de méthylène 0.04% (m/v). La membrane est rincée à l'eau et les ARN sont visualisés.

4 - Détermination du site d'initiation de la transcription par la méthode d'extension d'amorce

Cette méthode est utilisée afin de déterminer le site d'initiation d'un transcrit. En effet, lors de la fabrication de la banque les ADNc, synthétisés à partir d'une amorce oligo-dT, ont souvent leur région 5' manquante. De ce fait, pour localiser le site de démarrage de la transcription, nous avons utilisé des oligonucléotides (proches de la région 5'de l'ADNc isolé) comme amorce pour la synthèse de fragments d'ADNc correspondant à la région 5' du transcrit. Ainsi, 3 oligonucléotides de 40-mers PE1 (5'-TCGCCGCGCAACAGCCTT CTTATCTCCTCCGTTCGCAAGA), PE2 (5'-AGTGATCGATTGGCCAAGTTGCCACAA ACCCCCTCTCAT) et PE3 (5'- ACCCCCTCTCATTTTCTAAGCAAAATATCTCAAA TTCA) complémentaires de l'ARNm correspondant à la protéine T sont utilisés comme amorce pour la réverse transcriptase. Les ARNm proviennent de feuilles de pois âgés de 7 jours. Chaque oligonucléotide (3 pmol) est hybridé avec 1 µg d'ARNm de feuille de pois dans un volume de 10 µl de tampon [$MgCl_2$ 16 mM, KCl 60 mM, DTT 20 mM et du Tris/HCl 100 mM (pH 8.3)]. Après une incubation de 3 min à 75°C, les échantillons sont doucement ramenés à température ambiante puis mis dans la glace. Cinq microlitres d'une solution contenant 2 µM dATP + dGTP + dTTP, 20 µCi de [α-^{32}P]dCTP (3000 Ci/mmol), 100 µg de BSA (sérum albumine de boeuf) et 25 unités de réverse transcriptase (M-MuLV) sont ajoutés à chaque échantillon et incubés 5 min à température ambiante. Huit microlitres d'une solution contenant 60 mM KCl, 200 µM de dATP + dGTP + dTTP + dCTP sont alors additionnés dans chaque échantillon et l'incubation est prolongée durant 15 min à 42 °C. La réaction est stoppée par ajout de 0.1 volume de 3 M d'acétate de sodium (pH 5.3) et de 2.5 volumes d'éthanol. Après une centrifugation (15 000 g, 20 min), le culot d'ADN est lavé avec de l'éthanol à 70 % (v/v) et resuspendu dans du TE. Le produit d'élongation est séparé, sur un gel de séquençage, en condition dénaturante, avec, en parallèle, la séquence du gène codant pour la protéine T (pGT) (en utilisant le kit Sequenase ver. 2.0; USB) réalisée à l'aide du même oligonucléotide.

C - ISOLEMENT ET CARACTERISATION D'UN ADNc ET D'UN GENE

1 - Criblage, à l'aide d'anticorps polyclonaux et d'une sonde, d'une banque d'expression dans lambda gt11

La banque que nous avons utilisée a permis d'isoler, auparavant, l'ADNc codant pour la protéine H du complexe de la glycine décarboxylase (Macherel et *al.*, 1990). La banque originale a été préparée à partir d'ARN messagers de jeunes feuilles de pois ayant poussé à la lumière. Elle contenait 1.7 10^6 bactériophages/ml avec 96% de recombinants.

L'amplification de la banque a été réalisée à partir de 600 000 recombinants environ. Sept boîtes de pétri (12 x 12 cm) contenant du milieu LB (Bacto Tryptone 10 g/l, Bacto Yeast 5 g/l, NaCl 170 mM) avec de l'ampicilline 50 µg/ml sont réchauffées à 42°C. Pour chaque boîte 50 µl de phages λgt 11 (85 000 phages recombinants) sont incubés 15 min à 37°C avec 100 µl d'une suspension bactérienne Y1090 R dans 1/5 volume de $MgSO_4$ 10mM. Cette suspension est ensuite diluée dans 3 ml de top-agarose (maintenu à 45°C), puis coulée sur une boîte de pétri. Les boîtes sont incubées 6 heures à 42°C (Les plages de lyse ne doivent pas être confluentes) puis toute la nuit à 4°C avec 10ml de tampon SM (NaCl 0.1 M, $MgSO_4$ 8 mM, Tris base 50 mM, 2% gélatine pH 7.5). Cette suspension est centrifugée à 7000 rpm durant 10 min, le surnageant est ensuite stocké en présence de chloroforme 4% (v/v). Le titre de la banque amplifiée (L3) a été estimé à 2 10^{10} phages/ml.

Le criblage de la banque est réalisé sur une souche bactérienne de type *E. Coli* Y1090 R qui possède le répresseur de lac z (inhibé par addition d'IPTG), ainsi qu'une déficience de la protéase lon (qui augmente la stabilité de la protéine de fusion) et de sup F (qui supprime les mutations qui bloquent la lyse).

La croissance des bactéries s'effectue à 37°C sous agitation pendant environ trois heures dans 5 ml de milieu LB additionné en maltose 0.4% (m/v) et de 100 µg/ml d'ampicilline. Les bactéries sont récupérées en phase de croissance (A_{600} nm = 0,5), centrifugées 10 min à 3000g, puis suspendues dans 1/5 volume de $MgSO_4$ 10 mM. Cent microlitres de cette suspension bactérienne sont incubés avec 100 µl de suspension de phages (10^4 phages) durant 15 min à 37°C. Cette suspension est diluée avec 3 ml de top agarose (maintenus à 45°C) puis coulée dans des boîtes de pétri contenant du milieu LB-Agar avec de l'ampicilline 50 µg/ml.

La formation des plages de lyses est initiée à 42°C sans IPTG (inducteur artificiel de la transcription). Dans ces conditions, le développement des plages de lyses n'est pas affecté par la présence des protéines de fusion.

a - Utilisation d'anticorps

Après 2 ou 3 heures d'incubation, les plages de lyses sont visibles et l'expression des protéines de fusions est initiée à 37°C par ajout sur chaque boîte d'un filtre de nitrocellulose (Biorad) imbibé d'isopropyl-β-D-thiogalactopyranoside (IPTG) 10 mM et partiellement séché. Après une incubation de 3 heures à 37°C pour permettre l'expression des protéines de fusions, les membranes sont retirées des boîtes qui sont conservées à 4°C. Les sites de fixation libres de la membrane de nitrocellulose sont saturés durant 5 min dans du TBS-Tween 2% (m/v). La révélation des protéines par les anticorps dirigés spécifiquement contre la protéine T, s'effectue selon la méthode décrite dans le paragraphe V. D (Techniques d'immunodétection). Les phages présentant une réponse positive à l'anticorps sont ensuite purifiés par criblages successifs.

b - Utilisation d'une sonde

Après la période d'incubation à 42 °C, qui révèle les plages de lyses, on dépose sur la boîte, pour une durée de 90 secondes, un filtre nylon Hybond-N^+ (Amersham). Les filtres sont ensuite retirés et déposés (la face où sont fixés les phages dirigée vers le haut) sur 3 couches de papier Whatmann 3 MM imbibé de tampon de dénaturation (NaOH 0.5 M, NaCl 1.5 M). Après 7 min, le filtre est transféré 2 fois de suite pendant 3 min sur 3 couches de papier Watmann imbibé d 'une solution neutralisante [NaCl 1.5 M, Tris-HCl 0.5 M (pH 7.2), EDTA 1 mM] puis rincé dans du 2 x SSC. L'ADN des phages subit alors une fixation alkaline (NaOH 0.4M) pendant 20 min puis est rincé dans du 5 x SSC. Les clones positifs sont révélés par autoradiographie après hybridation avec une sonde marquée au P^{32} selon le protocole décrit dans le paragraphe VII. A. 2.

Les plages de lyses ainsi révélées (contenant les clones positifs) sont localisées sur les boîtes de pétri, extraites avec une pipette Pasteur, puis incubées 12 heures à 4°C dans 1 ml de tampon SM. Plusieurs cycles de criblages successifs permettent d'isoler les phages positifs.

La suspension de phages provenant de régions ayant donné une réponse positive (anticorps ou sonde P^{32}) est étalée sur des bactéries Y 1090 R^- à plusieurs dilutions différentes (20, 100,

1000 et 10 000 phages par boîte de pétri de 87 mm de diamètre). Des cycles de criblages sont réalisés à nouveau jusqu'à l'obtention d'un pourcentage de 100% de positifs. Quatre à cinq purifications successives sont nécessaires. Les clones purifiés sont ensuite amplifiés selon la méthode décrite dans le paragraphe VII. C. 1.

2 - Criblage d'une banque génomique dans lambda EMBL 3 à l'aide d'une sonde

Une banque génomique (Clontech) contenant $3x10^9$ pfu/ml a été criblée à l'aide de l' ADNc correspondant à la protéine T.

Le vecteur EMBL 3 utilisé ne possède ni le gène du répresseur (cI) ni le système d'intégration dans le génome bactérien, de ce fait le phage est obligé de suivre la voie lytique formant ainsi des plages de lyses après infection de la bactérie.

La croissance des phages recombinants s'effectue selon le même protocole décrit dans le paragraphe VII. C. 1. avec deux différences:

- la souche bactérienne est de type *E. Coli* NM 539.
- la formation des plages de lyses est initiée à une température de 37 °C.

Après 10 à 12 heures d'incubation, on dépose, sur les plages de lyses, un filtre de nitrocellulose en prenant soin d'orienter le filtre. L'ADN des phages subit le traitement décrit dans le paragraphe VII. C. 1. b (dans ce cas la sonde utilisée est l'ADNc TB2 correspondant au transcrit de longueur complète). Les clones positifs révélés par autoradiographie sont localisés sur les boîtes de pétri, extraits, puis incubés 12 heures à 4 °C dans 1 ml de tampon SM (NaCl 0.1 M, $MgSO_4$ 8 mM, Tris base 50 mM, 2% gélatine pH 7.5).

3 - Isolement de l'ADN des phages

La suspension du clone purifié (environ 15 10^7 phages dans 150 µl) est incubée avec 150 µl de bactéries Y 1090 R^- dans 425 µl de SM pendant 15 min à 37°C. Cette suspension est ensuite diluée dans 250 ml de milieu CYMG [casamino acid 1% (m/v), extrait de levure 0.5 % (p/v) , NaCl 0.3 % (m/v), KCl 0.2 % (m/v) additionné en $MgSO_4$ 10 mM et du Tris-base 25 mM (pH 7.4)]. La croissance s'effectue sous agitation à 37 °C durant 4 heures. A cette température, les phages de type λgt11 sont en cycle lysogénique et se multiplient avec les bactéries. La lyse des bactéries est déclenchée en augmentant la température à 42 °C, provoquant ainsi le cycle lytique des phages. Dans le cas des phages de type EMBL3, la lyse est initiée à 37°C. Des agglomérats de bactéries lysées sont observables 2 heures plus tard. Du

chloroforme à 0.2 % (v/v) est ajouté afin de compléter la destruction des bactéries. L'ADN et l'ARN bactériens sont détruits par une incubation à température ambiante pendant 45 min en présence de DNase (10 µg/ml) et de RNase A (10 µg/ml). Le volume est ajusté à 300 ml et les phages sont précipités par ajout de polyethylèneglycol 6000 2 % (m/v) et de NaCl 8 % (m/v) puis incubés pendant 12 heures à la température de la glace fondante. Les phages sont centrifugés 30 min à 10 000 g (rotor GS 3) puis repris dans 4 ml de SM. Les résidus d'ADN et d'ARN bactériens sont digérés à 37 °C, durant 15 min, par ajout de DNase (100 µg/ml) et de RNase A (10 µg/ml). La capside des phages est ensuite détruite par une incubation à 70 °C, pendant 15 min en présence de protéinase K (100 µg/ml), de SDS 0.5 % (m/v) et d'EDTA 20 mM. L'ADN viral est purifié par deux extractions successives au phénol, au phénol-chloroforme (1/1) et au chloroforme, suivies d'une précipitation à l'isopropanol (1 volume) en présence d'acétate de sodium (pH 5.3) 0.3 M. Le culot est lavé 2 ou 3 fois à l'éthanol 70% et 100%, suspendu dans 100 µl de TE et laissé une nuit à +4 °C. L'ADN des phages est conservé à - 20 °C jusqu'à utilisation.

D - SOUS-CLONAGE DE L'ADNc

Le vecteur lambda gt11 ne permet pas l'obtention de quantités importantes d'ADNc. Par ailleurs, il est mal adapté pour le séquençage. Il est donc nécessaire de sous-cloner l'ADNc dans un plasmide. Notre choix s'est donc porté sur le sous clonage de l'ADNc dans le plasmide pUC 18 [ce plasmide fait partie de la série pUC, développée par Messing et ses collègues (Veira et Messing, 1982)].

Toutes les conditions de digestion et de ligation ont été effectuées selon le manuel de Sambrook *et al.*, (1989).

1 - Préparation des bactéries compétentes

a - Première méthode

Les cellules *Escherichia coli* DH5α compétentes sont préparées par traitement au $CaCl_2$ (Sambrook *et al.*, 1989).

b - Deuxième méthode

Cette méthode a été utilisée pour rendre compétente les cellules Escherichia coli BMH71-18 lors de la mutagénèse dirigée (Cf. paragraphe VIII. B.). Les bactéries sont cultivées à

37°C, dans 4 ml de milieu LB (contenant comme antibiotique de la tétracycline à 15 µg/ml), durant 12 heures. Une aliquote de 1 ml de la suspension bactérienne est prélevée puis ajoutée à 100 ml de milieu LB en présence de tétracycline à 15 µg/ml. Les bactéries sont de nouveau cultivées à 37°C jusqu'à une absorbance se situant entre 0.25 et 0.5 mesurée à 600 nm. Les cellules sont collectées par centrifugation à 2500 rpm, pendant 15 min, à 4°C. Après avoir éliminé complètement le surnageant, les cellules sont resuspendues délicatement dans 1/3 de volume initial avec la solution I [KCl 100 mM, acétate de potassium 30 mM, CaCl 60 mM, glycérol 15 % (v/v) (amené à pH 5.8 avec de l'acide acétique)] puis incubées une heure à 4°C. Les bactéries sont sédimentées par centrifugation, à 3000 rpm, durant 15 min, à 4°C. Le surnageant est éliminé puis les cellules sont resuspendues délicatement dans 1/25 du volume initial avec la solution II [MOPS 10 mM, KCl 10 mM, CaCl 75 mM, glycérol 15 % (v/v) (amené à pH 6.8 avec une solution de NaOH 10 N)] puis incubées dans la glace durant 15 min. Des fractions aliquotes de 200 µl sont alors rapidement congelées dans l'azote liquide en présence de 20 % (v/v) de glycérol puis stockées à - 80°C jusqu'à utilisation.

2 - Transformation des bactéries compétentes

La réaction de ligation (10 µl) est ajoutée à une aliquote de 200 µl de bactéries compétentes DH5α puis le tube est placé dans la glace pendant 40 min. Un choc thermique de 2 min à 42°C permet l'introduction des plasmides dans les bactéries. La suspension bactérienne est alors incubée dans 0.8 ml de milieu LB durant une heure à 37 °C.

On étale, ensuite, 100 µl de cette suspension (ou un plus grand volume ramené à 100 µl en concentrant les bactéries par centrifugation) sur une boîte de pétri contenant du milieu LB, de la carbénicilline (50 $\mu g.ml^{-1}$) et sur laquelle on a étalé 60 µl de X-gal (5-Brom-4-Chlor-3-Indolyl-β-D-galactopyrannosid) 2% (m/v). Les boîtes sont ensuite placées à 37 °C, pendant 12-15 heures. Dans ces conditions, seules les bactéries ayant incorporé un plasmide peuvent se développer. La β-galactosidase produite par le plasmide se complémente ou non (en absence ou en présence de l'ADNc) avec la forme défective de l'enzyme codée par la cellule, produisant respectivement soit une colonie bleue non recombinante, soit une colonie blanche recombinante. Les colonies blanches sont sélectionnées à l'aide de cure-dents stériles et répiquées sur une boîte de LB plus carbénicilline (50 $\mu g.ml^{-1}$).

3 - Préparation de l'ADN des plasmides

Les plasmides ont été préparés selon les techniques décrites dans le manuel de Sambrook *et al.*, (1989). Pour la préparation d'ADN de plasmide en grande quantité, nous avons utilisé le kit Qiagen-midi (Qiagen).

VIII - TECHNIQUE D'EXPRESSION PROCARYOTIQUE

A - CONSTRUCTION DU VECTEUR D'EXPRESSION PROCARYOTIQUE

Nous avons choisi le vecteur pET-3d développé par Studier *et al.* (1990) car il se distingue par la force du promoteur situé en amont du site d'insertion de l'ADN codant pour la protéine à exprimer (Cf. chapitre IV).

L'ADNc, correspondant au précurseur de la protéine T de la GDC de mitochondries de feuilles de pois, a été utilisé comme matrice pour amplifier par PCR un fragment correspondant à la protéine T mature portant les sites de restriction nécessaires à son sous-clonage dans le vecteur d'expression pET-3d. Un oligonucléotide sens 5'-GGGGGAATTCCATG**GCTACAGAATCTGAGCTGA** composé d'un site *Eco*RI suivi du site *Nco*I (souligné) est synthétisé. Les nucléotides en caractères gras correspondent aux premiers acides aminés de la région amino-terminale de la protéine T mature (ATESEL). L'oligonucléotide complémentaire de synthèse 5'-GGGGGGATCCCTA**TGAAGGCTTATA GTA** comporte un site de restriction BamHI (souligné) en aval du codon stop de la protéine T. Les nucléotides en caractères gras sont complémentaires à la séquence de l'ADNc correspondant à la protéine T. Le fragment d'ADN, amplifié à l'aide de ces oligonucléotides spécifiques, a été purifié sur gel d'agarose selon la technique du GENECLEAN (BIO 101 Inc., California). Le fragment est digéré à l'aide des enzymes de restriction *Eco*RI et *Bam*HI, de nouveau purifié sur agarose puis sous-cloné dans le vecteur pUC 19 digéré avec les même enzymes. Les vecteurs recombinants sont transformés dans les souches E. coli DH5α puis sélectionnés sur boîtes de pétri contenant de la carbénicilline et du X-gal (2%). Une colonie blanche (clone positif) est amplifiée dans du LB en présence de carbénicilline puis, le plasmide est purifié et entièrement séquencé selon la technique de Sanger *et al.*, (1977), en utilisant le kit (Sequenase ver. 2.0, USB, Cleveland, OH, USA), afin de détecter d'éventuelles mutations effectuées par la taq polymérase.

B - MUTAGENESE DIRIGEE

Le séquençage complet du fragment de 1137 pb correspondant à la protéine T mature dans le plasmide pUC 19 (pUC-TM) a permis de découvrir une mutation se situant à 1054 pb: une adénine a été remplacée par une thymidine. Le rétablissement de la base modifiée a été réalisé par mutagénèse dirigée à l'aide d'un kit de mutagénèse (U.S.E. mutagenesis Kit, Pharmacia Biotech). Deux oligonucléotides de synthèse sont simultanément hybridés sur un des deux brins du plasmide pUC-TM préalablement dénaturé. Un des deux oligonucléotides est utilisé pour introduire la mutation désirée, alors que le deuxième est utilisé pour muter un site unique de restriction situé sur le plasmide (dans notre cas le site de restriction *Sca* I a été remplacé par *Mlu* I) ceci en vue d'une sélection. Les détails de la construction de l'ADNc correspondant à la protéine T mature sont illustrés dans le chapitre IV.

Une digestion, à l'aide de l'enzyme de restriction *Sca* I, permet de linéariser uniquement les plasmides sauvages, alors que ceux possédant le site de restriction muté resteront circulaires. La solution contenant les plasmides (mutés ou non) est utilisée pour transformer une souche de *E. coli* (BMH71-18 *mut*S) déficiente en réparation de l'ADN. L'efficacité de la transformation des plasmides circulaires est, de ce fait, augmentée d'un facteur 100 à 1000 au détriment des plasmides linéarisés. Les souches transformées sont amplifiées durant 12 heures à 37°C dans 1,5 ml de LB en présence de carbénicilline (100 µg/ml). Les plasmides sont purifiés puis soumis à une deuxième digestion à l'aide de la même enzyme de restriction afin d'accroître le pourcentage des plasmides mutés. Ces derniers sont, par la suite, transformés dans une souche sauvage de *E. coli* (DH5α). La région mutée du plasmide pUC-TM est vérifiée par séquençage. Les constructions des vecteurs pET-TM et pET-HTA et pET-HTU sont décrits dans le chapitre IV.

C - CLONAGE DANS LES VECTEURS D'EXPRESSION DU GENE argU

La présence, dans la protéine T, de nombreux codons rares AGA, AGG codant pour l'arginine, nous a amenés à insérer dans nos vecteurs d'expression le gène *arg*U codant pour l'ARNt déficient. Nous avons obtenu, par l'intermédiaire du National Institute of Genetics, Mishima, Japan, une souche d'*E.coli* JA 200 contenant le plasmide PLC 22-1 (13,4 kpb) porteur d'un fragment génomique d'*E. coli* de 7 kpb environ dans lequel se trouve le gène *arg*U. La souche d'*E.coli* JA 200 est cultivée dans 10 ml de LB en présence de chloramphénicol durant 15 heures à 37°C (Maniatis *et al.*, 1982) puis les plasmides sont

purifiés. Une séquence de 276 pb contenant le gène *argU* est amplifiée par PCR à l'aide de 2 amorces spécifiques possédant à leur extrémité 5' le site EcoRI (YF: 5'-CGGAATTCGATCA AAAGCCATTGACTCAGCAAGGG-3'; YR:5'-CGGAATTCCCTTGCGCCTAATCATT TGACAGAGC-3'). Les conditions de PCR sont les suivantes: 500 ng de plasmide PLC 22-1 sont dilués dans 100 µl de milieu de PCR [Tris-HCl 20 mM (pH 8.3), Tween-20 0.5% (v/v), nucléotides triphosphates 100 µM, KCl 50 mM, gélatine 1 mg/ml et 100 pmoles de chaque amorce spécifique]. Après centrifugation durant quelques secondes, le milieu est placé à 95°C durant 2 min dans un incubateur programmable (Robotcycler, Stratagène). Une quantité de 2.5 U de *Taq* polymérase (Boehringer) est ensuite ajoutée au milieu. Puis le milieu est recouvert de 100 µl d'huile minérale (afin d'éviter toute évaporation). L'amplification est réalisée après 10 cycles, chaque cycle comportant une période:

- de 45 secondes à 94°C permettant la séparation des 2 brins de l'ADN,
- de 45 secondes à 45°C conduisant à l'hybridation des amorces
- de 1 min à 72°C permettant l'élongation de l'ADN par la *Taq* polymérase.

Le fragment amplifié, correspondant au gène *arg*U, est purifié sur gel d'agarose 1%. Ce fragment est inséré dans les vecteurs (préalablement linéarisés avec l'enzyme Eco RI et déphosphorylés) pET-TM (vecteur contenant l'ADNc codant pour la protéine T mature), pET-HTA et pET-HTU (vecteurs contenant l'ADNc codant pour la protéine H mature) dans 2 orientations possibles, sens et anti-sens par rapport à la transcription du gène T (Cf. Chapitre IV). L'orientation du gène *arg*U au sein des vecteurs a été déterminée par PCR à l'aide des amorces reverse/YF ou YR et YF/YR. Nous avons ainsi obtenu les plasmides pETY-TM1 (sens), pETY-HTA1 (sens), pETY-HTU1 (sens) et pETY-TM1 (antisens), pETY-HTA1, (antisens), pETY-HTU1 (antisens).

D - CONDITIONS D'EXPRESSION

Deux souches ont été utilisées: la souche BL21 (DE3) dont les principales caractéristiques sont présentées dans le chapitre IV et la souche B834(DE3)pLysS (souche parentale de BL21) qui présente de nombreux avantages par rapport à la souche BL21 (Doherty *et al.*, 1995). En effet, cette souche est capable d'exprimer une plus grande variété de gènes sous contrôle du promoteur de la T7 polymérase. Sa vitesse de croissance est deux fois plus rapide que celle de la souche BL21. Enfin, la souche B834(DE3)pLysS contient le plasmide pACY184(pLysS) qui confère la résistance au chloramphénicol et code pour le lysozyme T7. Cette enzyme

inhibe le niveau basal d'expression de la T7 polymérase avant induction par l'IPTG (Studier & Moffat; 1986). Cette précaution évite les problèmes dus à une éventuelle toxicité de la protéine cible. En effet, le risque est que des bactéries mutées non productrices puissent se multiplier au détriment des cellules dont la croissance dans le milieu de culture sera limitée par la présence de la protéine toxique.

La souche de *E. coli* [BL21 (DE3) ou B834(DE3)pLysS], transformée par le plasmide pET-TM, pET-HTA ou HTU (stockée à -80°C selon la méthode décrite dans le paragraphe VII. D. 1. b), est étalée sur une boîte de pétri contenant du LB agar et de l'ampicilline à 200 µg/ml puis incubée à 37°C durant 12 heures. Une colonie est mise en préculture dans 5 ml de milieu LB en présence d'ampicilline (50 µg/ml) et incubée durant 12 heures à 37°C. Une aliquote sera prélevée pour réaliser l'expression de la protéine T, dont les conditions de culture sont définies de la façon suivante:

- 20 ml de milieu LB,
- culture à 37 °C (250 rpm) jusqu'à ce que l'absorbance à 600 nm atteigne 0,5,
- carbénicilline (150 µg/ml),
- induction de la synthèse de la protéine recombinante par addition d'IPTG (0.4 mM final),
- culture à 37 °C durant 2 à 3 heures.

Des échantillons de 1 ml de culture sont prélevés à différentes périodes après induction. Les bactéries sont centrifugées 2 min à 3000 g (Eppendorff) puis retirées dans 100 µl de tampon de dépôt (Laemmli 1970). Après chauffage à 100 °C (2 min), 10 µl d'échantillon sont analysés par électrophorèse sur gel de polyacrylamide en présence de SDS. Les protéines sont révélées au bleu de Coomassie ou analysées par Western-blot.

IX - EXPERIENCE DE GEL RETARD

La technique de gel retard permet la mise en évidence de la formation d'un complexe ADN-protéine par la diminution de sa mobilité électrophorétique par rapport à l'ADN seul. La réaction s'effectue par l'incubation d'un fragment d'ADN (au maximum de 300 pb) marqué au phosphore 32 avec un extrait nucléaire. Le ou les complexes formés sont séparés de la sonde libre sur un gel de polyacrylamide en conditions non dénaturantes. Il est à noter que les sondes utilisées doivent être de taille inférieure à 300 pb afin de détecter toute variation de leur mobilité électrophorétique et d'éviter la présence de plusieurs sites potentiels d'interaction avec des protéines de feuille de pois selon la technique de Green *et al.* (1987).

A - FABRICATION ET MARQUAGE DES SONDES

1 - Hybridation des oligonucléotides monomériques complémentaires

Les oligonucléotides double brin utilisés comme sonde, correspondent à des fragments de la région promotrice du gène codant pour la protéine T. Chaque oligonucléotide possède, à son extrémité 5', trois thymidines qui seront nécessaires au marquage de la sonde. Ainsi, dans un volume de 10 µl, les oligonucléotides sens sont hybridés, dans des proportions équimolaires, avec leurs complémentaires, dans un tampon d'hybridation [Tris-HCl 40 mM (pH 7.5), $MgCl_2$ 20 mM, NaCl 50 mM]. La réaction d'hybridation est réalisée par une incubation de 2 min à 85°C dans un bloc chauffant (Bioblock). Puis le bloc est placé à température ambiante afin d'approcher la température finale de 35°C.

2 - Marquage et purification de l'oligonucléotide double brin

Vingt ng d'oligonucléotides sont marqués dans des conditions citées ci-dessous:

- 4 µl de tampon 5 fois concentré [Tris-HCl 200 mM (pH 7.5), $MgCl_2$ 100 mM, NaCl 250 mM, DTT 5 mM].
- 50 µCi de [α-^{32}P]- dCTP (10mCi / ml, 3000 Ci : mmole, Amersham),
- 8 unités d'ADN polymérase T_7.

La réaction est conduite dans un volume final de 20 µl à 37°C pendant 30 min. Après avoir augmenté le volume avec 20 µl d'eau, la réaction est stoppée par ajout d'un volume de phénol/chloroforme/alcool isoamylique (25/24/1) et un volume de chloroforme/alcool isoamylique (équilibré à pH 8 par Tris-HCl). La sonde est alors purifiée sur une colonne Séphadex G-25 (Boehringer). Après avoir correctement resuspendu la colonne, le tampon STE [Tris 10 mM (pH 8.0), EDTA 1 mM, NaCl 0,1 M] est élué, dans un premier temps, par simple gravité puis, dans un second temps, par une centrifugation de 2 min à 1100g (rotor GS-4). L'éluat est alors éliminé tandis que la colonne est maintenue verticalement. La sonde est alors déposée délicatement sur le sommet de la colonne, puis centrifugée durant 4 min à 1100g (rotor GS-4). Les oligonucléotides non incorporés sont retenus par la colonne alors que la sonde est récupérée dans l'éluat. L'ADN marqué ainsi purifié est stocké à -20°C. La quantité de radioactivité incorporée est mesurée par effet Cerenkov.

B - PROTOCOLE EXPERIMENTAL D'UN GEL RETARD

Les expériences de gel retard sont réalisées suivant le protocole établi par Fried et Crothers, (1981). Le mélange réactionnel de 10 µl est effectué de la façon suivante:

A la sonde marquée (20 000 à 40 000 cpm), sont ajoutés successivement, du tampon [HEPES/KOH 25 mM (pH 7.9), EDTA 1 mM, DTT 1 mM, KCl 50 mM, glycérol 10 % (v/v)], 2 µg de poly (dI-dC).poly (dI-dC) (compétiteur aspécifique) puis 3 µg d'extrait protéique pour un volume de 10 µl.

L'ensemble est incubé à température ambiante durant 20 min. Il est à noter qu' un témoin ne contenant par d'extrait nucléaire est toujours réalisé. De plus, pour des expériences de compétition avec une sonde froide spécifique ou aspécifique en excès (10 à 100 fois), une préincubation de 5 min à température ambiante est auparavant effectuée sans la sonde chaude. Les complexes nucléoprotéiques sont immédiatement analysés par séparation électrophorétique sur un gel de polyacrylamide en condition non dénaturante.

Le gel de polyacrylamide 6% (rapport acrylamide/bisacrylamide 29/1) est préparé dans du tampon TBE 0.5 fois concentré [Tris-base 50 mM, acide borique 43 mM, EDTA 1 mM (pH 8.0)]. Le gel est préchauffé durant 20 min, à 80 volts. Les échantillons ainsi que du bleu témoin de migration [bleu de bromophénol 0.025% (m/v), xylène cyanol FF 0.025% (m/v), glycérol 3 % (v/v)] sont déposés dans les puits du gel pour subir une électrophorèse durant 2 heures, sous une tension de 140 volts. Le gel est alors séché sur papier Watmann puis exposé en présence d'un film autoradiographique βmax (Amersham) à -80°C.

CHAPITRE I

CHAPITRE I
DETERMINATION ET ANALYSE DE LA STRUCTURE PRIMAIRE DE LA PROTEINE T

Au cours de ce chapitre, nous décrirons comment il a été possible de déterminer la structure primaire de la protéine T du complexe de la glycine décarboxylase. Dans un premier temps, nous décrirons brièvement les étapes chromatographiques qui ont conduit à l'obtention de la protéine T purifiée. Celle-ci a permis la fabrication d'anticorps polyclonaux dirigés spécifiquement contre la protéine T et également d'entreprendre une étude approfondie de la protéine à l'aide en particulier de la spectrométrie de masse. L'obtention d'anticorps spécifiques a permis le criblage d'une banque d'ADNc dans λgt11 qui a abouti à l'isolement et à la caractérisation d'ADN complémentaires correspondant à la protéine T.

I - PURIFICATION DE LA PROTEINE T

A - ETAPE DE PURIFICATION

La purification de la protéine T a été réalisée selon le protocole mis au point au laboratoire (Bourguignon *et al.*, 1988) à partir de mitochondries purifiées de feuilles de pois qui sont capables d'oxyder très rapidement la glycine à des vitesses de l'ordre de 250 nmoles d'O_2 consommées par min et par mg de protéines mitochondriales. L'ouverture ménagée de ces mitochondries suivie d'une ultracentrifugation, permet de libérer les protéines matricielles (protéines solubles) qui comprennent toutes les enzymes impliquées dans l'oxydation de la glycine et de les séparer des protéines membranaires (Neuburger *et al.*, 1986). Les protéines de masse moléculaire élevée qui sont retenues et concentrées sur une membrane d'ultrafiltration Amicon XM 300, oxydent rapidement la glycine en présence de NAD^+ et de H_4PteGlu (Neuburger *et al.*, 1986). L'oxydation de la glycine par l'extrait matriciel est mise en évidence à 25°C en mesurant la formation de NADH sous argon en présence de glycine 20 mM, de NAD^+ 2 mM, de pyridoxal-phosphate 20 μM et d'un large excès de H_4PteGlu 400 μM.

A partir de cet extrait matriciel, la purification de la protéine T du complexe de la glycine décarboxylase se déroule en deux étapes, une chromatographie sur colonne Séphacryl S-300 suivie d'une chromatographie d'échanges d'anions sur colonne mono Q HR 5/5 (Pharmacia).

L'extrait matriciel filtré est déposé sur une colonne Séphacryl S-300 (Pharmacia) équilibrée avec un tampon contenant 50 mM KCl (Cf. Matériel et méthodes). Au cours de cette chromatographie sur gel de filtration, les protéines sont éluées en fonction de leur masse moléculaire. Ainsi, contrairement à la pyruvate deshydrogénase qui est un complexe relativement stable, élué dans les toutes premières fractions, le complexe de la GDC se dissocie car les protéines P, L, T et H sont éluées, à 4°C, en 3 fractions (figure 12A). Ainsi, en se basant sur l'analyse d'un gel de polyacrylamide en présence de SDS (figure 12B) on distingue successivement la fraction lourde contenant la SHMT et la protéine P, puis la fraction intermédiaire avec la protéine L et enfin la fraction légère qui contient les protéines T et H.

Pour retrouver l'activité d'oxydation de la glycine, il est nécessaire de recombiner les trois fractions éluées. Par la suite, l'activité de la protéine T est mesurée au spectrophotomètre par réassociation de celle-ci avec les autres protéines du complexe.

La protéine T est purifiée à homogénéité en déposant une aliquote de la fraction légère (6 mg de protéines) sur une colonne chromatographique mono Q HR 5/5 (Pharmacia) échangeuse d'anions (Cf. Matériel et méthodes). Les protéines T et H sont éluées à 4°C et à pH 7,5 par un gradient linéaire croissant de K-phosphate. A ce pH, la protéine T chargée positivement, n'est pas retenue par la colonne et est éluée immédiatement après le volume mort. Par contre, la protéine H est éluée à une concentration de K-phosphate plus élevée (environ 210 mM).

L'analyse de l'éluat contenant la protéine T, sur un gel de polyacrylamide dans des conditions dénaturantes (en présence de SDS), montre une masse moléculaire de 43 ± 1 kDa. De plus, l'étude par chromatographie sur gel de filtration et sur gel de polyacrylamide dans des conditions non dénaturantes (la masse moléculaire de la protéine T native obtenue est de 45 kDa) indique que la protéine T est composée d'un seul polypeptide (monomère). La protéine T ainsi purifiée à homogénéité (en quantités importantes), a permis d'une part de fabriquer des anticorps spécifiques dirigés contre cette enzyme et d'autre part, d'aborder l'étude de sa structure primaire.

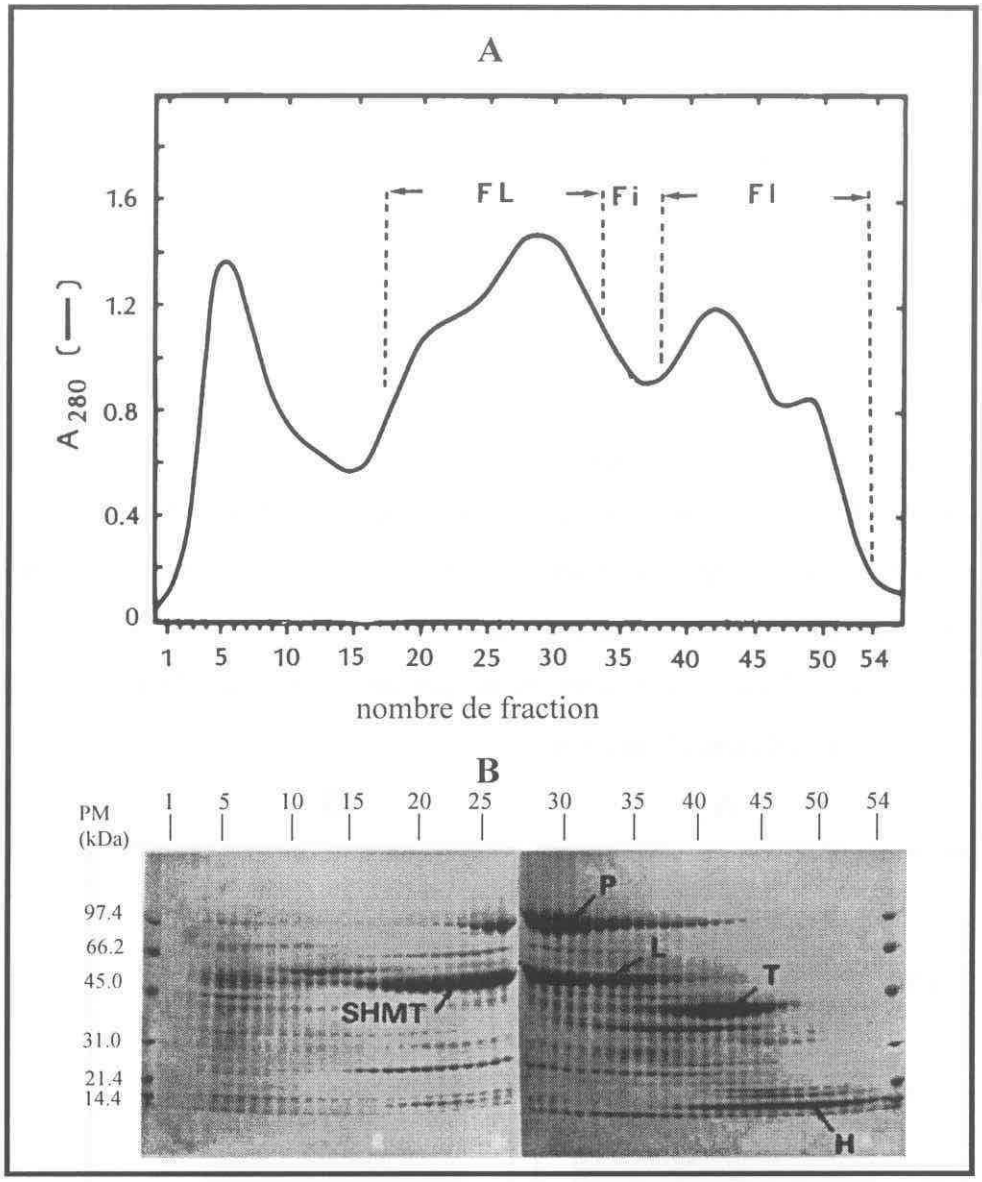


Figure 12:

A) Fractionnement des protéines matricielles de feuilles de pois par chromatographie sur une colonne de Séphacryl S-300. Approximativement 170 mg de protéines ont été déposées sur la colonne (2,5 cm x 90 cm; débit: 0,3 ml/min). Le volume des fractions collectées est de 1,5 ml. Fl (fractions légères), Fi (fractions intermédiaires), FL (fractions lourdes).

B) Analyse par gel dénaturant de polyacrylamide (7,5-15 %) coloré au bleu de Coomassie des fractions éluées de la colonne Séphacryl. 15 µl de chaque fraction ont été déposés sur le gel. P, L, T et H représentent les différentes protéines de la glycine décarboxylase; SHMT, la sérine hydroxyméthyltransférase.

Il est à noter que la protéine T à forte concentration (> 5 mg/ml) est instable en raison de phénomènes d'agrégations irréversibles. Cette protéine semble être stabilisée lorsqu'elle se trouve en présence de la protéine H (Cf. paragraphe V). Plus récemment, nous avons montré que le milieu de dilution de la protéine T joue un rôle extrêmement important. En effet, en

présence de tampon d'acétate d'ammonium 0,1 M (pH 7.5), la protéine T peut être concentrée jusqu'à des valeurs de 10 mg/ml sans qu'elle ne s'agrège.

B - PREPARATION D'ANTICORPS POLYCLONAUX DIRIGES CONTRE LA PROTEINE T ET ANALYSE IMMUNOLOGIQUE

Les anticorps de lapin dirigés contre la protéine T ont été obtenus à partir de la protéine T purifiée extraite des mitochondries de feuilles de pois. La protéine T pure et dénaturée (300 µg par injection) est injectée à des lapins 4 fois de suite à 3 semaines d'intervalle sous la forme d'une émulsion contenant un adjuvant de Freud.

L'analyse selon la technique de « Western blot » a permis de tester la spécificité du sérum obtenu. Pour ce faire, trois préparations contenant respectivement un extrait total de protéines de feuilles de pois obtenu par extraction phénolique (Cf. Matériel et méthodes), un extrait matriciel et la protéine T pure, sont séparées sur gel de polyacrylamide puis transférées sur une membrane de nitrocellulose (figure 13).

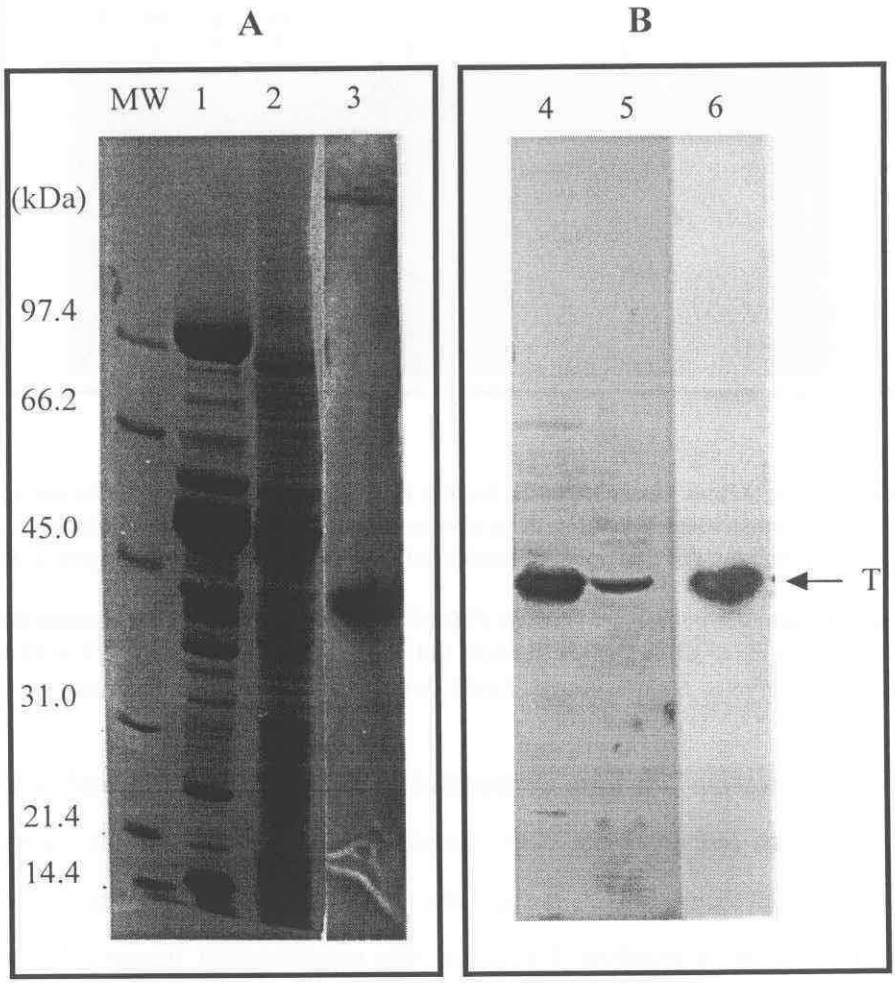


Figure 13: Analyse par Western-blot de la spécificité des anticorps anti-protéine T. Migration sur gel dénaturant de polyacrylamide (7,5-15%), d'un extrait matriciel de mitochondries de feuilles de pois: 50 µg (1 et

4), d'un extrait total de protéines de feuilles de pois: 100 µg (2 et 5), de la protéine T purifiée: 2 µg (3 et 6) **A)** Gel ayant subi une coloration au bleu de Coomassie. **B)** Membrane de nitrocellulose après transfert et incubation avec un sérum dirigé contre la protéine T (dilution 1/20 000 dans un milieu TBS, contenant du Tween 20 à 0.05g/l)

Les anticorps réagissent avec une protéine dont la masse correspond à celle de la protéine T (43 kDa). Cette étude démontre également que la protéine T n'a pas subi d'attaque protéolytique au cours des différentes étapes de purification puisque aucun fragment provenant d'une éventuelle protéolyse n'est observé. Il faut souligner que le bruit de fond est faible, notamment au niveau de l'extrait total de protéines de feuilles, preuve d'une bonne spécificité des anticorps.

C - MICROSEQUENÇAGE DE LA PROTEINE T

Afin d'obtenir des informations concernant la structure primaire de la protéine T, nous avons entrepris de séquencer la région aminoterminale de la protéine T purifiée. Dix microgrammes environ de protéine T ont été analysés par dégradation d'Edman. Nous avons ainsi obtenu la séquence des 40 premiers résidus acides aminés: Ala-Thr-Glu-Ser-Glu-Leu-Lys-Lys-Thr-Val-Leu-Tyr-Asp-Phe-His-Val-Ala-His-Gly-Gly-Lys-Met-Val-Pro-Phe-Ala-Gly- Trp-Ser-Met-Pro-Ile-Gln-Tyr-Lys-Asp-Ser-Ile-Met-Asp. De plus, la protéine T a subi un clivage chimique par le bromure de cyanogène qui coupe la chaîne peptidique du côté carboxyterminal des résidus méthionines (Cf. paragraphe III. B). Les 10 peptides obtenus (60 µg) ont été purifiés par HPLC préparative. Certains de ces fragments, notamment un peptide de 11 acides aminés, ont été entièrement séquencés par dégradation d'Edman. La séquence obtenue est la suivante: Pro-Phe-val-pro-Thr-Lys-Tyr-Tyr-Lys-Pro-Ser. Les résultats présentés dans ce paragraphe ne sont que partiels. Une étude beaucoup plus approfondie de la structure primaire de la protéine T a été réalisée. Celle-ci sera présentée ultérieurement.

La purification de la protéine T a permis d'obtenir d'une part, des anticorps spécifiques et d'autre part, des informations concernant la structure primaire de la protéine (51 acides aminés de la protéine). Nous avons donc en notre possession tous les outils et les renseignements nécessaires pour nous permettre d'entreprendre à l'aide de ces anticorps, le criblage d'une banque d'expression dans λgt 11 afin d'isoler l'ADNc correspondant à la protéine T.

II - ISOLEMENT DE L'ADN COMPLEMENTAIRE (ADNc) CORRESPONDANT A LA PROTEINE T

A - CRIBLAGE D'UNE BANQUE D'EXPRESSION DANS λGT 11 A L'AIDE DES ANTICORPS DIRIGES SPECIFIQUEMENT CONTRE LA PROTEINE T.

1 - Caractérisation d'un ADNc de taille incomplète correspondant à la protéine T (TD3)

L'utilisation d'anticorps polyclonaux monospécifiques, dirigés contre la protéine T, a servi à cribler une banque amplifiée d'ADNc dans λgt 11. Cette dernière avait été fabriquée à partir d'ARN messagers provenant de jeunes feuilles de pois âgées de 8 jours (Macherel *et al.*, 1990). Le criblage a été réalisé sur $8x10^5$ phages recombinants. L'immunodétection nous a permis de déceler 13 plages de lyses positives. Les clones correspondant à ces plages ont ensuite, été purifiés par trois cycles successifs de criblage. A ce stade, la totalité des plages de lyses présente alors une réponse positive à l'immunodétection. L'ADN phagique des différents clones positifs a été purifié. Malheureusement, la digestion de ceux-ci par les enzymes de restriction E*co*RI ou *Not*I n'a pas toujours révélé l'insert qu'ils contenaient. Un tel résultat était surprenant car ces deux sites sont présents théoriquement dans des adaptateurs ajoutés aux deux extrémités de chaque insert lors de la construction de la banque d'ADNc dans λgt 11 en vue de faciliter les sous-clonages ultérieurs. Par la suite, en séquençant certains clones nous avons constaté que l'un des adaptateurs était absent rendant de ce fait impossible le sous-clonage des inserts. La taille approximative des ADNc a pu, cependant, être déterminée par amplification à l'aide de la taq polymérase et de deux oligonucléotides complémentaires spécifiques de λgt 11 situés de part et d'autre de l'insert. Le plus grand insert obtenu (nommé TD3) a une taille approximative de 500 pb (figure 14). Cette taille légèrement surestimée est due à la portion du phage λgt 11 amplifié avec l'insert. Ce dernier a été sous-cloné dans le plasmide pUC 18.

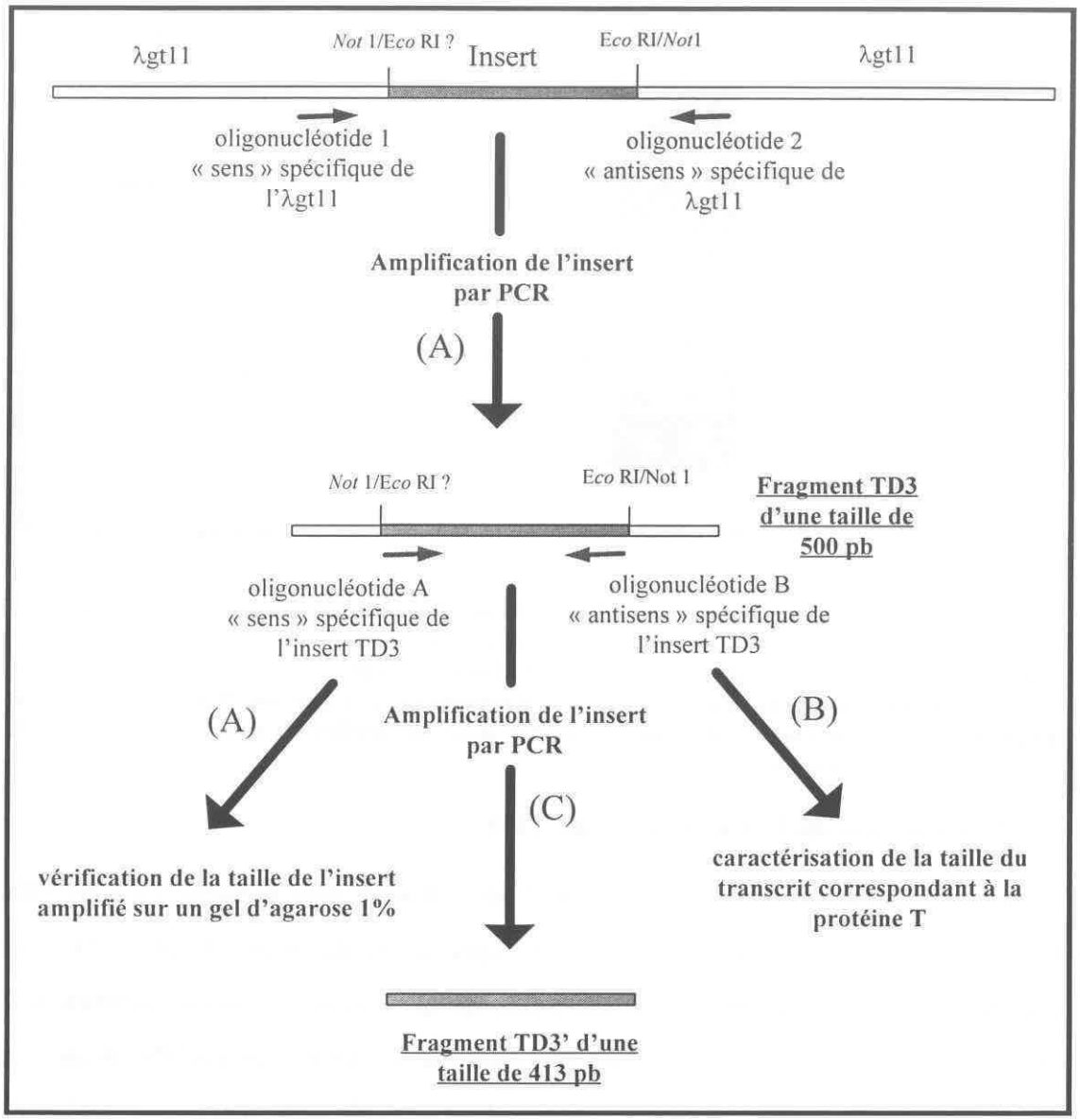


figure 14: Stratégie d'étude du clone isolé. (A) Détermination de la taille des inserts clonés dans λgt 11 par PCR. (B) L'insert TD3 a permis de déterminer par « Northern-blot » la taille du transcrit correspondant à la protéine T. (C) L'insert TD3' fabriqué par PCR sera utilisé comme sonde spécifique afin d'isoler l'ADNc de taille complète correspondant à la protéine T.

Des expériences d'hybridation (Northern-blot), sur des ARN messagers extraits de feuilles de pois, ont été effectuées avec l'insert TD3 marqué au [α-P^{32}]-dCTP. Une population d'ARN d'une taille de 1,6 kb a été détectée (figure15). Cette taille est en accord avec la masse moléculaire de la protéine T de pois déterminée par électrophorèse.

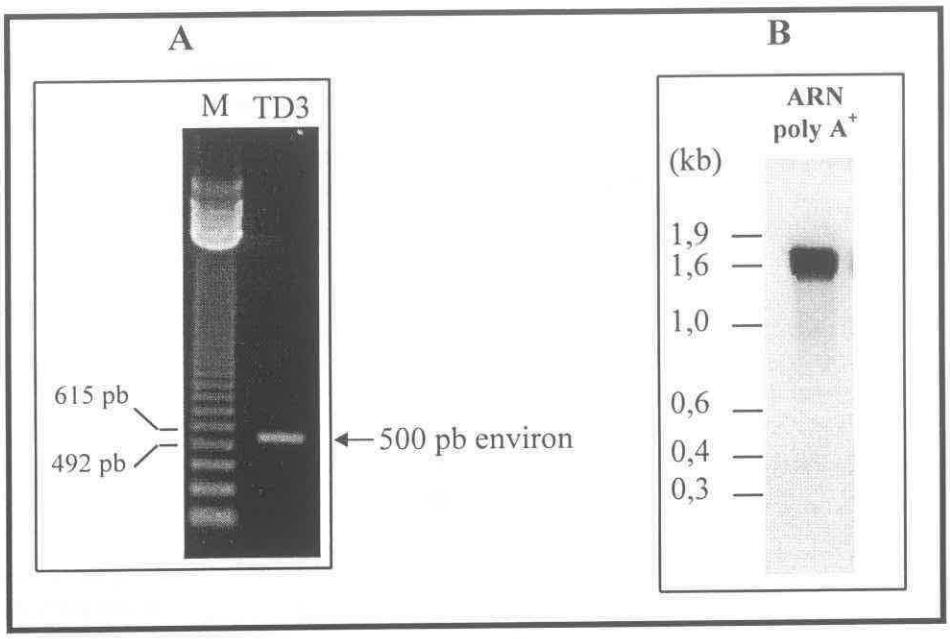


Figure 15:

A) Détermination par PCR de la taille de l'insert TD3 cloné dans le phage. Le fragment TD3 a été amplifié à l'aide de deux amorces complémentaires spécifiques du phage λgt11 selon une technique décrite dans « Matériel et méthodes ». M: marqueurs de taille des ADN (n x123 pb).

B) Détermination de la taille du transcrit correspondant à la protéine T de pois par hybridation du fragment amplifié in vitro (TD3) sur les ARNm isolés de feuilles de pois. Les ARN poly A^+ (3 µg) proviennent de feuilles chlorophylliennes.

2 - Validation de l'identité du clone TD3

Le séquençage complet de l'ADNc (413 pb), présenté dans la figure 16 a été effectué sur les deux brins en utilisant deux amorces spécifiques du plasmide pUC 18 (« universal primer » et « reverse primer »). La déduction de sa séquence en acides aminés, à l'aide du programme « translate » de PCGENE (Intelligenetics), montre que l'ADNc possède un cadre de lecture ininterrompu de 144 pb suivi d'un site de terminaison de traduction TAG caractéristique. L'identité de ce clone est confirmée par la comparaison de la séquence peptidique correspondant à l'ADNc, avec celle d'un peptide interne de 11 résidus acides aminés obtenus après digestion chimique au bromure de cyanogène et microséquençage de la protéine T purifiée (Cf. paragraphe I-C). Le fragment peptidique microséquencé correspond à la partie carboxyterminale de la protéine T déduite de l'ADNc TD3. Ces résultats confirment que l'ADNc (TD3) isolé correspond bien à la protéine T. De ce fait, cet ADNc sera utilisé par la suite comme sonde, afin de réaliser une autre série de criblages qui nous permettra d'isoler un ADNc de taille complète.

```
          10        20        30        40        50        60
           |         |         |         |         |         |
AAGAAGAACATAGCTATTGGATATGTCAAATCTGGATTGCACAAAGCAGGTACCAAAGTAAAGATCATT
LysLysAsnIleAlaIleGlyTyrValLysSerGlyLeuHisLysAlaGlyThrLysValLysIleIle

70        80        90       100       110       120       130
 |         |         |         |         |         |         |
ATTCGAGGTAAACAGAATGAAGGAGTTGTCACGAAAATGCCGTTCGTACCCACAAAATACTATAAGCCT
IleArgGlyLysGlnAsnGluGlyValValThrLysMETProPheValProThrLysTyrTyrLysPro

140       150       160       170       180       190       200
 |         |         |         |         |         |         |
TCATAGTTTGCTGCAGCATTTCCCATTTCTTGACTAATTTTCCCATAAACTGAGTGTTACATAAATGTT
Ser *

210       220       230       240       250       260       270
 |         |         |         |         |         |         |
AATGTCGTTATTATGTTTCTCATGTCAAGTACAAGGCTTCATCTTGATTCATAACTCAATTTCTTTTCA

280       290       300       310       320       330       340
 |         |         |         |         |         |         |
AACATGTATTGCTTGAACTTGATGTTATCATGAATTTTTTCCAATATTTCCCCCTGTGTGTACTGTGCT

350       360       370       380       390       400       410
 |         |         |         |         |         |         |
ATTTTGAGGCCTGTTTTTCATTCAATAGTAGCTATTCATATAAGACTAAATCTGAATAATATAATAAT
```

Figure 16: Séquence nucléotidique de l'ADNc du clone TD3 correspondant à la région carboxy-terminale de la protéine T impliquée dans le complexe de la glycine décarboxylase de feuilles de pois ainsi que sa séquence protéique déduite. La séquence peptidique soulignée correspond à la région C-terminale de la protéine T, obtenue après digestion au bromure de cyanogène de la protéine T pure et après analyse d'un de ses peptides par dégradation d'Edman. La séquence nucléotidique non traduite correspond à la région 3' non codante de l'ADNc correspondant à la protéine T.

B - CRIBLAGE DE LA BANQUE D'EXPRESSION DANS λGT 11 A L'AIDE DE L'ADNc TD3'

1 - Caractérisation d'un ADNc de taille complète correspondant à la protéine T

Le criblage de la banque a été réalisé sur 8 000 phages par des expériences d'hybridation avec l'insert TD3' marqué au [α-P^{32}]-dCTP (figure 14). Deux clones positifs nommés TA1 et TB2 ont ainsi pu être décelés. Ces clones correspondant aux plages de lyses ont ensuite été purifiés par trois cycles successifs de criblage. A ce stade, toutes les plages de lyses s'hybrident avec la sonde TD3'. Les inserts contenus dans ces clones ont été amplifiés par PCR à l'aide de primer spécifiques de λgt11 et leurs tailles, estimées sur gel d'agarose. TA1 et TB2 ont une taille respectivement de 1,5kpb et de 1,8kpb (figure 17). Ces tailles, bien que

légèrement surestimées (du fait de l'amplification d'une partie du phage), s'approchent de la taille du transcrit déterminée auparavant par Northern blot.

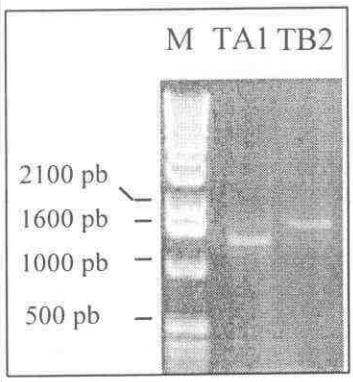


Figure 17: Détermination de la taille des inserts TA1 et TB2 clonés dans le phage par amplification in vitro (PCR). Les fragments TA1 (1,5 kpb) et TB2 (1,8 kpb) ont été amplifiés à l'aide de deux amorces complémentaires spécifiques du phage λgt11 selon une technique décrite dans « Matériel et méthodes ». M: marqueur de taille des ADN (1 kpb).

Par la suite, les ADN des phages recombinants ont été isolés puis digérés par l'enzyme de restriction E*co*RI. Le résultat montre que les clones TA1 et TB2 ont une taille réelle de 1,2 kpb et 1,5 kpb. La taille de l'ADNc TB2 étant proche de celle de l'ARN messager (1,6 kb) détectée par hybridation avec TD3, il est vraisemblable que l'ADNc correspondant à la protéine T soit complet.

2 - Validation et détermination de la séquence primaire de la protéine T déduite de l'ADNc

a - Stratégie de séquençage de l'ADNc codant pour la protéine T

Les fragments d'ADN des clones TA1 et TB2 obtenus par digestion enzymatique (E*co*RI) ont été sous-clonés dans le vecteur pUC 18, puis séquencés. Le séquençage des extrémités 5' et 3' de l'ADNc a été effectué à l'aide des amorces du plasmide pUC 18. Le reste de la séquence a été déterminé progressivement à l'aide d'oligonucléotides de synthèse complémentaires de l'ADNc. Les deux brins complémentaires de chaque clone ont ainsi été entièrement séquencés. La stratégie de séquençage ainsi que les sites de restriction présents dans la séquence nucléotidique sont présentés dans la figure 18.

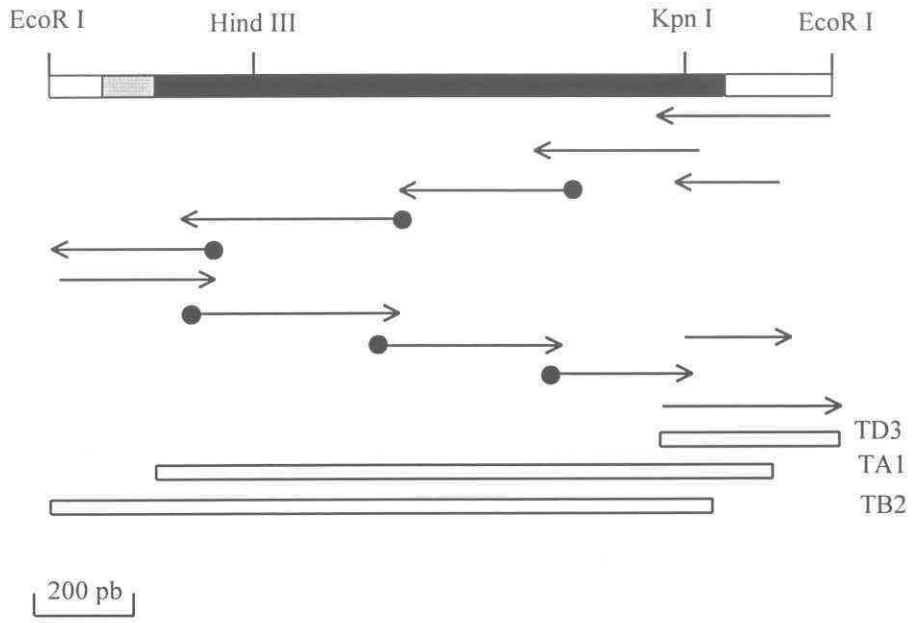


figure 18: Carte de restriction et stratégie de séquençage de l'ADNc correspondant à la protéine T de pois (var. Douce provence). La région codant pour la protéine T mature est représentée en noir. La région codant pour le peptide leader est en gris alors que les régions non codantes sont non colorées. Les flèches indiquent la direction et la taille du fragment séquencé. Les flèches symbolisées par ⟶ représentent respectivement les régions séquencées à l'aide des amorces M13 universelles et complémentaires et, par ●⟶ celles séquencées à l'aide d'amorces de synthèses, complémentaires de l'ADNc correspondant à la protéine T. La position des différents clones (TD3, TA1 et TB2) est également représentée.

b - Analyse de la séquence de l'ADNc

La séquence complète des ADNc TA1 (1224 pb) et TB2 (1430 pb) a été analysée. La séquence présentée dans la figure 19 est la combinaison des trois séquences correspondant aux clones TD3, TA1 et TB2. L'ADNc correspondant à la protéine T a une longueur de 1562 pb et possède un cadre de lecture de 1223 pb. La séquence peptidique aminoterminale de 40 acides aminés (commençant avec un résidu Alanine) obtenue par le microséquençage de la protéine T mature purifiée, est parfaitement homologue à une portion de la séquence primaire de la protéine déduite de l'ADNc. Le premier résidu méthionine dans le cadre de lecture se positionne 30 acides aminés (nucléotides 67-69) en amont du début de la protéine mature (résidu alanine). Cela indique que l'ADNc présente une région 5' non traduite de 66 nucléotides et que la protéine est synthétisée probablement avec une préséquence aminoterminale de 30 acides aminés (PM: 3347 Da). De plus, la présence à 2 codons en amont de ce résidu méthionine, d'un codon stop confirme la taille de cette préséquence.

Contrairement à de nombreuses protéines de plantes (Lûtcke *et al.*, 1987), le résidu méthionine n'est pas suivi d'un résidu alanine. En effet, la séquence nucléotidique située autour du codon d'initiation (GAAA**ATG**AG) est légèrement différente de la séquence consensus proposée (AACA**ATG**GC) pour l'initiation de la traduction chez les plantes (Lûtcke *et al.*, 1987). Le cadre de lecture ininterrompue code pour un polypeptide de 40961 Da. Ce résultat est en accord avec la valeur de 43000 Da obtenue avec la protéine T mitochondriale sur un gel de polyacrylamide (Bourguignon *et al.*, 1988). Toutefois, il est à noter que la région 3' non codante de 269 nucléotides ne comporte ni la séquence de polyadénylation typiquement eucaryote (AATAAA) (Proodfoot et Browlee, 1976), ni la queue poly (A). Dans ces conditions, il est probable que la région 3' non codante de l'ADNc soit incomplète.

```
            10        20        30        40        50        60
             |         |         |         |         |         |
     ATTCACTCCGAACTCAATCAACAAACCAATCTCTTGTTTGAATATTTGAGATATTTTGCTTAGAAAATG
                                                                      MET
    70        80        90       100       110       120       130
     |         |         |         |         |         |         |
     AGAGGGGGTTTGTGGCAACTTGGCCAATCGATCACTCGCCGTCTTGCGAACGGAGGAGATAAGAAGGCT
- 29 ArgGlyGlyLeuTrpGlnLeuGlyGlnSerIleThrArgArgLeuAlaAsnGlyGlyAspLysLysAla

    140       150       160       170       180       190       200
     |         |         |         |         |         |         |
     GTTGCGCGGCGATGTTTCGCTACAGAATCTGAGCTGAAGAAGACAGTTCTGTATGACTTCCATGTTGCT
  -6 ValAlaArgArgCysPheAlaThrGluSerGluLeuLysLysThrValLeuTyrAspPheHisValAla

    210       220       230       240       250       260       270
     |         |         |         |         |         |         |
     CATGGTGGAAAGATGGTTCCTTTTGCTGGTTGGAGTATGCCTATTCAATACAAAGACTCAATCATGGAC
 18  HisGlyGlyLysMETValProPheAlaGlyTrpSerMETProIleGlnTyrLysAspSerIleMETAsp

    280       290       300       310       320       330       340
     |         |         |         |         |         |         |
     TCAACGTTAAACTGTAGACAGAACGGTAGTCTTTTCGATGTTTCTCATATGTGTGGTCTTAGTCTCAAG
 41  SerThrLeuAsnCysArgGlnAsnGlySerLeuPheAspValSerHisMETCysGlyLeuSerLeuLys

    350       360       370       380       390       400       410
     |         |         |         |         |         |         |
     GGAAAGGATGTTATTTCGTTCCTTGAGAAGCTTGTCATTGCGGATGTTGCTGCTCTTGCCAACGGAACC
 64  GlyLysAspValIleSerPheLeuGluLysLeuValIleAlaAspValAlaAlaLeuAlaAsnGlyThr

    420       430       440       450       460       470       480
     |         |         |         |         |         |         |
     GGGACTTTGACTGTTTTCACTAATGAAAAGGGAGGGGCGATCGATGATTCTGTGATTACTAAGGTGACA
 87  GlyThrLeuThrValPheThrAsnGluLysGlyGlyAlaIleAspAspSerValIleThrLysValThr
```

```
        490       500       510       520       530       540       550
         |         |         |         |         |         |         |
    GAGGATCATCTTTACTTGGTTGTCAATGCTGGGTGCAGAGATAAAGATTTGGCTCATATTGAGGAGCAT
110 GluAspHisLeuTyrLeuValValAsnAlaGlyCysArgAspLysAspLeuAlaHisIleGluGluHis

        560       570       580       590       600       610       620
         |         |         |         |         |         |         |
    ATGAAGGCATTCAAGGCCAAAGGCGGTGATGTGTCGTGGCACATTCACGATGAGAGATCTCTACTCGCT
133 METLysAlaPheLysAlaLysGlyGlyAspValSerTrpHisIleHisAspGluArgSerLeuLeuAla

        630       640       650       660       670       680       690
         |         |         |         |         |         |         |
    CTTCAGGGTCCTCTTGCTGCTCCTGTTCTTCAACATCTGACAAAAGAGGATTTGAGCAAGCTGTACTTT
156 LeuGlnGlyProLeuAlaAlaProValLeuGlnHisLeuThrLysGluAspLeuSerLysLeuTyrPhe

        700       710       720       730       740       750
         |         |         |         |         |         |
    GGGGAGTTCCGCGTGTTGGACATCAATGGCTCGCAGTGCTTTCTCACACGGACAGGGTACACTGGTGAA
179 GlyGluPheArgValLeuAspIleAsnGlySerGlnCysPheLeuThrArgThrGlyTyrThrGlyGlu

        760       770       780       790       800       810       820
         |         |         |         |         |         |         |
    GATGGATTTGAGATCTCAGTTCCTTCAGAGCATGGAGTTGAACTTGCCAAGGCACTGCTGGAAAAATCT
202 AspGlyPheGluIleSerValProSerGluHisGlyValGluLeuAlaLysAlaLeuLeuGluLysSer

        830       840       850       860       870       880       890
         |         |         |         |         |         |         |
    GAAGGGAAAATAAGATTGACAGGACTCGGTGCTAGAGATAGTCTACGTCTCGAAGCTGGACTGTGCTTA
225 GluGlyLysIleArgLeuThrGlyLeuGlyAlaArgAspSerLeuArgLeuGluAlaGlyLeuCysLeu

        900       910       920       930       940       950       960
         |         |         |         |         |         |         |
    TACGGAAATGACCTGGAACAGCACATTACACCTATTGAAGCAGGACTGACATGGGCTATAGGTAAGAGA
248 TyrGlyAsnAspLeuGluGlnHisIleThrProIleGluAlaGlyLeuThrTrpAlaIleGlyLysArg

        970       980       990      1000      1010      1020      1030
         |         |         |         |         |         |         |
    AGGAGAGCAGAAGGTGGTTTTCTAGGAGCTGATGTTATCCTGAAACAGCTCGCAGACGGTCCTTCCATT
271 ArgArgAlaGluGlyGlyPheLeuGlyAlaAspValIleLeuLysGlnLeuAlaAspGlyProSerIle

       1040      1050      1060      1070      1080      1090      1100
         |         |         |         |         |         |         |
    AGGCGTGTCGGTTTCATTTCTTCTGGTCCACCTCCAAGAAGCCACAGCGAGATTCAAGATGAAGGAGGA
294 ArgArgValGlyPheIleSerSerGlyProProProArgSerHisSerGluIleGlnAspGluGlyGly

       1110      1120      1130      1140      1150      1160      1170
         |         |         |         |         |         |         |
    AACAACATTGGTGAAGTGACCAGTGGTGGATTCAGTCCTTGTCTCAAGAAGAACATAGCTATTGGATAT
317 AsnAsnIleGlyGluValThrSerGlyGlyPheSerProCysLeuLysLysAsnIleAlaIleGlyTyr

       1180      1190      1200      1210      1220      1230      1240
         |         |         |         |         |         |         |
    GTCAAATCTGGATTGCACAAAGCAGGTACCAAAGTAAAGATCATTATTCGAGGTAAACAGAATGAAGGA
340 ValLysSerGlyLeuHisLysAlaGlyThrLysValLysIleIleIleArgGlyLysGlnAsnGluGly
```

```
        1250      1260      1270      1280      1290      1300      1310
          |         |         |         |         |         |         |
    GTTGTCACGAAAATGCCGTTCGTACCCACAAAATACTATAAGCCTTCATAGTTTGCTGCAGCATTTCCC
363 ValValThrLysMETProPheValProThrLysTyrTyrLysProSer *

        1320      1330      1340      1350      1360      1370      1380
          |         |         |         |         |         |         |
    ATTTCTTGACTAATTTTCCCATAAACTGAGTGTTACATAAATGTTAATGTCGTTATTATGTTTCTCATG

        1390      1400      1410      1420      1430      1440
          |         |         |         |         |         |
    TCAAGTACAAGGCTTCATCTTGATTCATAACTCAATTTCTTTTCAAACATGTATTGCTTGAACTTGATG

1450      1460      1470      1480      1490      1500      1510
  |         |         |         |         |         |         |
  TTATCATGAATTTTTTCCAATATTTCCCCCTGTGTGTACTGTGCTATTTTGAGGCCTGTTTTTCATTCA

1520      1530      1540      1550      1560
  |         |         |         |         |
  ATAGTAGCTATTCATATAAGACTAAATCTGAATAATATAATAAT
```

Figure 19: Séquence nucléotidique de l'ADNc correspondant à la protéine T impliquée dans le complexe de la glycine décarboxylase et sa séquence en aminoacides déduite de la séquence nucléotidique. La séquence nucléotidique est la combinaison de la séquence des 3 clones: TB2 (nucléotides 1-1430), TA1 (nucléotides 232-1456), et TD3 (nucléotides 1150-1562) (voir figure 18). La séquence peptidique déduite, est écrite sous forme de trois lettres positionnées sous la séquence nucléotidique et, numérotée à partir du premier acide aminé de la protéine mature. Les acides aminés soulignés représentent les séquences obtenues par dégradation d'Edman de la région aminoterminale de la protéine T mature et de peptides internes purifiés après digestion au bromure de cyanogène (Cf. paragraphe III. B). Les résidus méthionines sont en majuscule et le codon stop est symbolisé par un astérisque (*).

c - Composition globale en acides aminés de la protéine T mature

La composition en acides aminés de la protéine T mature (résidus 1 à 378) déduite de l'ADNc correspond parfaitement à celle déterminée par hydrolyse acide de la protéine T mitochondriale purifiée (Cf. tableau 3). Nous pouvons observer la présence d'un grand nombre de résidus glycine (11,38 %), particulièrement représentés dans la moitié carboxyterminale de la protéine et une représentation assez importante des acides aminés hydrophobes (39,42 %). Les résidus chargés représentent 27,23 %, avec notamment une forte présence de résidus basiques (15,3%). Les autres résidus totalisent seulement 81 représentants, soit 21,44 %.

Acides aminés				
Code		Nombre	Pourcentage	
			Analysé	Calculé
Asx	Asn (N)	11		
	Asp (D)	21	8.96	8.46
Glx	Gln (Q)	9	9.25	8.74
	Glu (E)	24		
Thr (T)		20	5.53	5.29
Ser (S)		26	6.81	6.88
Pro (P)		14	4.23	3.70
Gly (G)		43	11.60	11.38
Ala (A)		26	7.36	6.38
Cys (C)		6	0.63	1.59
Val (V)		25	6.49	6.61
Met (M)		6	1.44	1.59
Ile (I)		24	5.93	6.35
Leu (L)		39	10.62	10.30
Tyr (Y)		9	2.40	2.38
Phe (F)		14	3.78	3.70
His (H)		13	3.27	3.43
Lys (K)		30	7.78	7.94
Arg (R)		15	3.93	3.97
Trp (W)		3	non détecté	0.79

Tableau 3: Composition en acides aminés de la protéine T du complexe de la glycine décarboxylase. La composition en acides aminés a été analysée après hydrolyse acide de la protéine T mature. La composition calculée est basée sur la séquence en acides aminés déduite d'ADNc correspondant à la protéine T.

Les techniques biochimiques et de biologie moléculaire nous ont donc permis d'isoler et de caractériser un ADNc dont le cadre de lecture code pour un précurseur de la protéine T de 408 acides aminés. Ce dernier comporte une préséquence de 30 acides aminés suivie d'une séquence de 378 acides aminés qui correspond à la protéine T mature dont la masse moléculaire calculée est de 40961 Da.

III - ETUDE DE LA PROTEINE T A L'AIDE DE LA SPECTROMETRIE DE MASSE

Afin de confirmer la structure primaire de la protéine T et de détecter d'éventuelles modifications post-traductionnelles, la protéine T a fait l'objet d'une étude en spectrométrie de masse. En effet, la spectrométrie de masse s'est avérée, à plusieurs titres, un outil précieux pour la caractérisation des différentes protéines du complexe de la GDC. Ainsi, elle a permis de lever une ambiguïté survenue entre la séquence nucléotidique de la protéine L établie par

Turner *et al.*, (1992b) et celle établie par Bourguignon *et al.*, (1992). De plus, cette technique est intéressante pour la détermination de modifications post-traductionnelles, notamment elle a permis de localiser le site de fixation du lipoate sur la protéine H Mérand *et al.* (1993). Ces travaux ont été réalisés, en collaboration avec Véronique Mérand Louvet du Laboratoire de Spectrométrie de Masse des Protéines de l'Institut de Biologie Structurale.

A - DETERMINATION DE LA MASSE MOLECULAIRE DE LA PROTEINE T NATIVE

La protéine T native a été analysée par couplage de la chromatographie haute performance avec la spectrométrie de masse par électrospray (HPLC/ESI-MS) (Cf. Matériel et méthodes). L'ionisation par les techniques de spray est une technique qui permet une analyse en routine de molécules de masse de 100 à 200 000Da. De plus, son mode d'ionisation douce ne provoque pas de fragmentations et permet l'identification de modifications post-traductionnelles. Le spectre de masse obtenu (figure 20) montre clairement une série d'ions multichargés correspondant à des molécules portant de 25 à 58 charges. Le spectre de masse, reconstruit à partir de la série de pics multichargés, donne une masse moyenne de 40 966 ± 5 Da. Cette masse mesurée est similaire à la masse calculée (40 961 Da) à partir de la séquence en acides aminés déduite de l 'ADNc (le calcul étant réalisé avec les cystéines sous forme réduite). La masse moléculaire de la protéine T avait été estimée par chromatographie de filtration sur gel à 45 kDa. L'analyse par électrophorèse sur gel de polyacrylamide donnait une masse moléculaire de 43±1 kDa. La mesure par spectrométrie de masse est beaucoup plus précise que les estimations faites par les méthodes biochimiques classiques. Elle confirme la séquence de la protéine T. De plus, ce résultat nous montre que la protéine T ne subit apparemment aucune modification post-traductionnelle. Ce résultat indique que le cofacteur de la protéine T (le tétrahydrofolate), n'est pas fixé de façon covalente à l'enzyme.

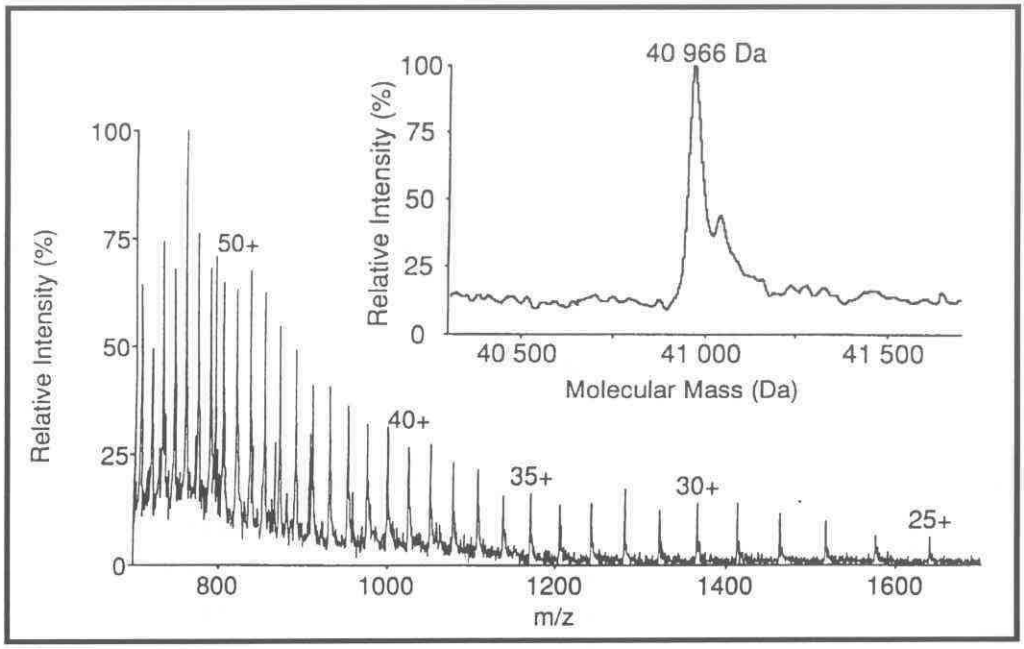
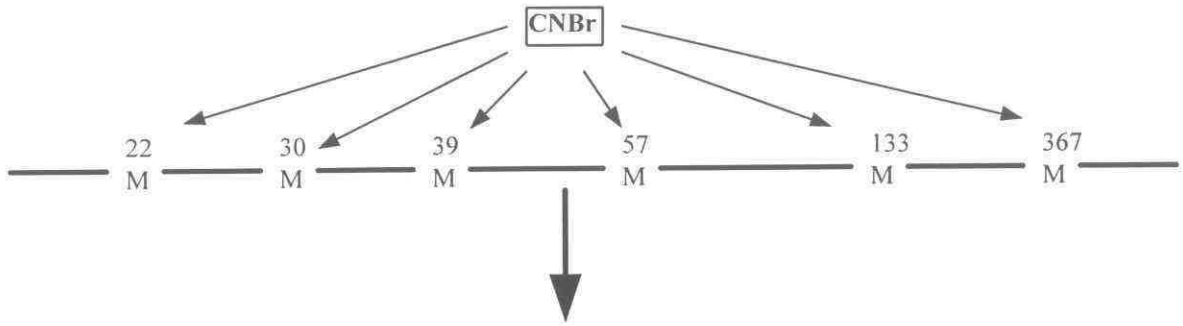


Figure 20: Spectre de masse obtenu par analyse HPLC/ESI-MS de la protéine T native purifiée à partir de mitochondries de feuilles de pois. La charge nette de quelques ions est indiquée au-dessus des pics qui correspondent à la forme multiprotonnée de la protéine T. Le profil de la masse de la protéine T a été calculé par déconvolution algorithmique.

B - ANALYSE DES PRODUITS DE LA DIGESTION AU BROMURE DE CYANOGENE DE LA PROTEINE T

Le tableau 4 montre les coupures théoriques de la protéine T au bromure de cyanogène (CNBr) (Cf. Matériel et méthodes). Dans un premier temps, la protéine T a subi une réduction et une carboxylation, afin d'éviter que les différents résidus cystéines n'interagissent entre eux ou avec d'autres résidus au cours des manipulations suivantes. En effet, ces deux réactions entraînent la rupture des ponts disulfures et la transformation des groupements thiols (S-H) libres en S-CH_2-COOH (Cf. Matériel et méthodes). La masse moléculaire des fragments a été calculée par informatique (program MacBioSpec: Perkin Elmer Sciex) d'après la séquence déduite de l'ADNc (tableau 4). Pour le calcul des masses, nous avons tenu compte des groupes carboxyméthyles greffés sur les cystéines.



Fragment	*résidu carboxy-terminal*	*masse (Da)*
1-22	m	2432
	m*	2413,8
23-30	m	864,3
	m*	846,1
31-39	m	1064,5
	m*	1046,3
40-57	m	2052,5
	m*	2034,3
58-133	m	8172,2
	m*	8153,3
134-367	m	25 328
	m*	25 309
368-378	-	1326,6

Tableau 4: Fragmentation théorique de la protéine T par le bromure de cyanogène déduite de la séquence de l'ADNc. Les valeurs données correspondent aux masses moléculaires moyennes des peptides en Daltons. m: homosérines; m*: homosérines lactones.

1 - Analyse des fragments de la protéine T par couplage de la chromatographie liquide avec la spectrométrie de masse par électrospray ou ionspray (HPLC/ESI-MS)

Les polypeptides obtenus après digestion au CNBr de la protéine T réduite et carboxyméthylée ont, tout d'abord, été analysés par HPLC en phase inverse (colonne C8) couplée à un spectromètre de masse électrospray [Electrospray ionization (ESI) mass spectrometer]. Cette technique permet d'une part, de séparer les différents peptides d'autre part, de mesurer leurs masses directement en sortie de colonne HPLC. Les masses des différents peptides ainsi que leurs temps de rétention sur la colonne sont récapitulés sur le

tableau 5. Ainsi, l'analyse par HPLC/ESI-MS des fragments de protéine T issus de la coupure au bromure de cyanogène a permis d'obtenir de nombreuses informations:

- D'après la séquence déduite de l'ADNc, la protéine T comporte 6 résidus méthionines. Nous avons retrouvé les 7 peptides issus de la coupure au bromure de cyanogène de la protéine T (1-22, 23-30, 31-39, 40-57, 58-133, 134-367, 368-378) avec cependant plusieurs temps de rétentions différents pour chaque peptide. En effet, le clivage chimique modifie la méthionine carboxyterminale sous forme d'homosérine (notée m) et sous forme d'homosérine lactone (notée m*). Ces deux formes sont éluées avec des temps de rétention légèrement différents durant la chromatographie. Quant au fragment carboxyterminal de la protéine (368-378), il ne contient pas de méthionine. Il se distingue donc des autres par sa masse moléculaire unique.

- Le clivage chimique n'est pas complet. En effet, les résidus méthionines en position 22, 39 et 133 ne réagissent que partiellement avec le bromure de cyanogène. Ces résidus apparaissent dans les fragments 1-30, 31-57, 58-367. Ces résultats expliquent pourquoi le nombre de peptides obtenu n'est pas de 7 mais de 10. Dans ces conditions, les résidus méthionines non coupés peuvent être sous trois formes différentes: Met (M), sulfoxyde méthionine (M*) ou homosérine (m). Toutes les combinaisons possibles sont résumées dans le tableau 5. Les formules chimiques des homosérines, des homosérines lactones et des méthionines sulfoxydes sont présentées dans le chapitre Matériel et méthodes.

- Enfin, l'analyse par HPLC/ESI-MS a permis d'établir une cartographie de la protéine T. En effet, les résultats obtenus montrent que les masses moléculaires de tous les fragments peptidiques mesurées par cette technique confirment les masses théoriques obtenues à partir de la séquence primaire de la protéine T déduite de l'ADNc (voir tableau 5).

fragments peptidiques	temps de rétention	Masses moléculaires				Séquences obtenues par microséquençage
		Masses mesurées	Masses moyennes théoriques	Masses mesurées en FAB		
	min	Da				
1-22	23,2	2 431,1	2 432	2 432,0	m	ATESELKKTVLYDFHVAHGG
	23,5	2 412,5	2 413,8	2 413,8	m*	K
23-30	18,3	863,5	864,3	863,4	m	VPFAGWS
	20,3	845,5	846,1	846,4	m*	
1-30	26,8	3 277,0	3 278,3		mm	
	26,8	3 322,2	3 324,1		M*m	
	27,5	3 258,7	3 260,1	3 260,1	mm*	peptide non séquencé
	27,5	3 305,2	3 308,1		M*m*	
31-39	7,8	1 063,5	1 064,5	1 063,5	m	PIQYKDSI
	11,7	1 045,5	1 046,3	1 046,0	m*	
40-57	18,5	2 050,7	2 052,5	2 051,9	m	peptide non séquencé
	19,6	2 034,0	2 034,3	2 033,5	m*	
31-57	22,0	3 097,7	3 098,9		mm	
	22,0	3 144,0	3 144,7		M*m	
	22,7	3 079,2	3 080,7		mm*	peptide non séquencé
	22,7	3 125,6	3 128,7		Mm	
58-133	42,6	8 151,4	8 153,3		m*	peptide non séquencé
58-367	41,0	33 495 ± 10	33 493		Mm*	C*GLSLKGKDVISFLEKLVIAD
			33 511		Mm	VAALANGTGXLTVFXNEKXX
			33 451		mm	AID...
134-367	36,9	25 311 ± 5	25 309		m*	KAFKAKGGDVSWHIHDERSL
			25 327		m	LALQGPLAAPVLQHLTKEDL..
368-378	16,4	1 325,7	1 326,6	1 325,5		PFVPTKYYKPS

Tableau 5: Analyse physique des fragments obtenus après clivage de la protéine T au CNBr. Sur le tableau sont inscrites les masses des différents fragments mesurées par HPLC/ESI-MS et leur temps de rétention. Les masses de tous les fragments, à l'exception des fragments 58-133, 58-367 et 134-367, sont confirmées par FAB-MS après séparation par HPLC. Sont inscrites également les masses calculées des peptides déduites de l'ADNc et de la dégradation d'Edman. Les symboles m (homosérine) et m* (homosérine lactone) représentent la nature des méthionines transformées du côté carboxy-terminal des peptides (voir tableau 1). Les peptides ayant subi un clivage incomplet sont indiqués par un double symbole (mm, M*m, mm*, Mm*); le premier symbole correspond à la méthionine interne (modifiée ou non) alors que le second représente la nature de la méthionine modifiée se situant dans la région carboxy-terminale du fragment. M et M* symbolisent respectivement la méthionine et la sulfoxide méthionine (voir tableau 1). C* (dans la région amino-terminale du fragment 58-367) représente une cystéine carboxyméthylée suite au traitement à l'iodoacétate. X représente des acides aminés non déterminés lors du séquençage.

2 - Analyse par FAB (Fast Atom Bombardment) et dégradation d'Edman des fragments purifiés par HPLC

Malheureusement, l'analyse par HPLC/ESI-MS ne permet pas de récupérer les peptides, car l'HPLC est directement couplée au spectromètre de masse. L'analyse en mode FAB complète à ce titre les analyses précédemment effectuées car elle permet d'une part, de microséquencer une partie des fractions obtenues en sortie de colonne HPLC et d'autre part, de déterminer avec une assez grande précision les peptides de faibles masses moléculaires. Afin de séparer les différents fragments peptidiques provenant de la coupure chimique au bromure de cyanogène, nous avons effectué une chromatographie en HPLC de type préparative en utilisant une colonne de capacité plus grande. Soixante microgammes environ de protéine T traitée au CNBr ont été injectés sur cette colonne, puis les différents peptides ont été élués selon les conditions discutées dans le chapitre Matériel et méthodes. Le profil d'élution des fragments peptidiques est tout à fait comparable à celui obtenu lors de l'expérience précédente (HPLC/ESI-MS). Les 2 grands peptides facilement identifiables ont été directement microséquencés. Les petits fragments ont subi une autre chromatographie en utilisant une colonne de type C18. Ainsi, un tiers des fractions obtenues ont été identifiées par spectrométrie de masse de type FAB alors que les fractions restantes ont été microséquencées par dégradation d'Edman. Les séquences obtenues correspondant aux régions aminoterminales de ces fragments ainsi que les masses des différents fragments sont présentées dans le tableau 5.

Toutes les séquences, déterminées par microséquençage des fragments générés par coupure chimique au CNBr de la protéine T, sont totalement identiques à la séquence en acides aminés déduite à partir de l'ADNc (figure 19). Notamment, le peptide PFVPTKYYKPS (368-378) qui correspond à la partie carboxyterminale de la protéine T, nous a permis d'identifier le clone TD3 lors de la première série de criblages (Cf. paragraphe II-A-1). De plus, les masses obtenues sont parfaitement identiques à celles déduites de la séquence primaire de la protéine T.

Cette technique a donc permis:

- de confirmer la masse de la protéine T native déduite de son ADNc,
- de confirmer 35% de la séquence peptidique de la protéine T (soit 131 acides aminés sur 378),
- de démontrer l'absence de modification post-traductionnelle de la protéine T mature.

IV - ANALYSE DE LA SEQUENCE PRIMAIRE DE LA PROTEINE T

A - QUELQUES PARAMETRES BIOCHIMIQUES DE LA PROTEINE T, DEDUITS DE LA SEQUENCE EN ACIDES AMINES

1 - Estimation du pH isoélectrique (pI) et de la charge globale de la protéine T mature au pH matriciel.

Le pI de la protéine T, estimé à l'aide du logitiel PC Gene, est de 7,35. On pourrait donc en déduire que la protéine T est neutre. Cependant, ces résultats calculés sont en complète contradiction avec ceux observés par les techniques biochimiques. En effet, comme on a pu le voir dans le paragraphe I-A, la protéine T est éluée dans le volume mort lors de la purification sur une colonne Mono-Q HR 5/5 à pH 7,5. A ce pH (proche du pH matriciel), la protéine est chargée positivement. Ce résultat est appuyé par la mesure du point isoélectrique de la protéine T qui, chez le boeuf est de 9,8 et chez le pois semble être supérieur à 9 (Okamura-Ikeda *et al.*, 1982). L'interprétation que l'on peut apporter est que le logiciel ne tient pas compte des structures secondaires et tertiaires de la protéine. Certains domaines de la protéine riches en protéines basiques peuvent être en surface alors que d'autres renfermant des résidus acides pourraient être enfouis au sein de la protéine. Par exemple, si on considère chaque moitié de la séquence de la protéine T comme un peptide à part entière, [la moitié aminoterminale allant de l'acide aminé 1 à 189, l'autre moitié constituant la partie carboxyterminale (190 à 378)] nous remarquons que le pI de chaque moitié est très différent de celui de la protéine globale. La moitié aminoterminale présente un pI de 5,78 avec une charge globale (calculée au pH matriciel) négative (le nombre de charges négatives est d'environ de 6), alors que celui de la moitié carboxyterminale de 9,66 est fortement basique (le nombre de charges positives est estimé à 6). Cette dernière valeur est proche de celle mesurée biochimiquement. Ainsi, globalement, la protéine T allie deux tendances électriques opposées, basique pour la moitié aminoterminale et acide pour la moitié carboxyterminale. En regardant la répartition électrique de la protéine, on pourrait de façon spéculative supposer que la partie carboxyterminale constitue la zone périphérique de la protéine T.

2 - Analyse du degré d'hydrophobicité de la protéine T

Le profil d'hydrophobicité de la protéine T de pois (selon la méthode de Kyte et Doolittle, 1982) s'effectue à l'aide d'un logiciel qui calcule, pour un intervalle de 19 acides aminés, le degré d'hydrophobicité le long de la séquence primaire. La ligne en pointillés détermine la frontière entre les domaines hydrophobes et hydrophiles respectivement situés au-dessus et au-dessous de la ligne (figure 21).

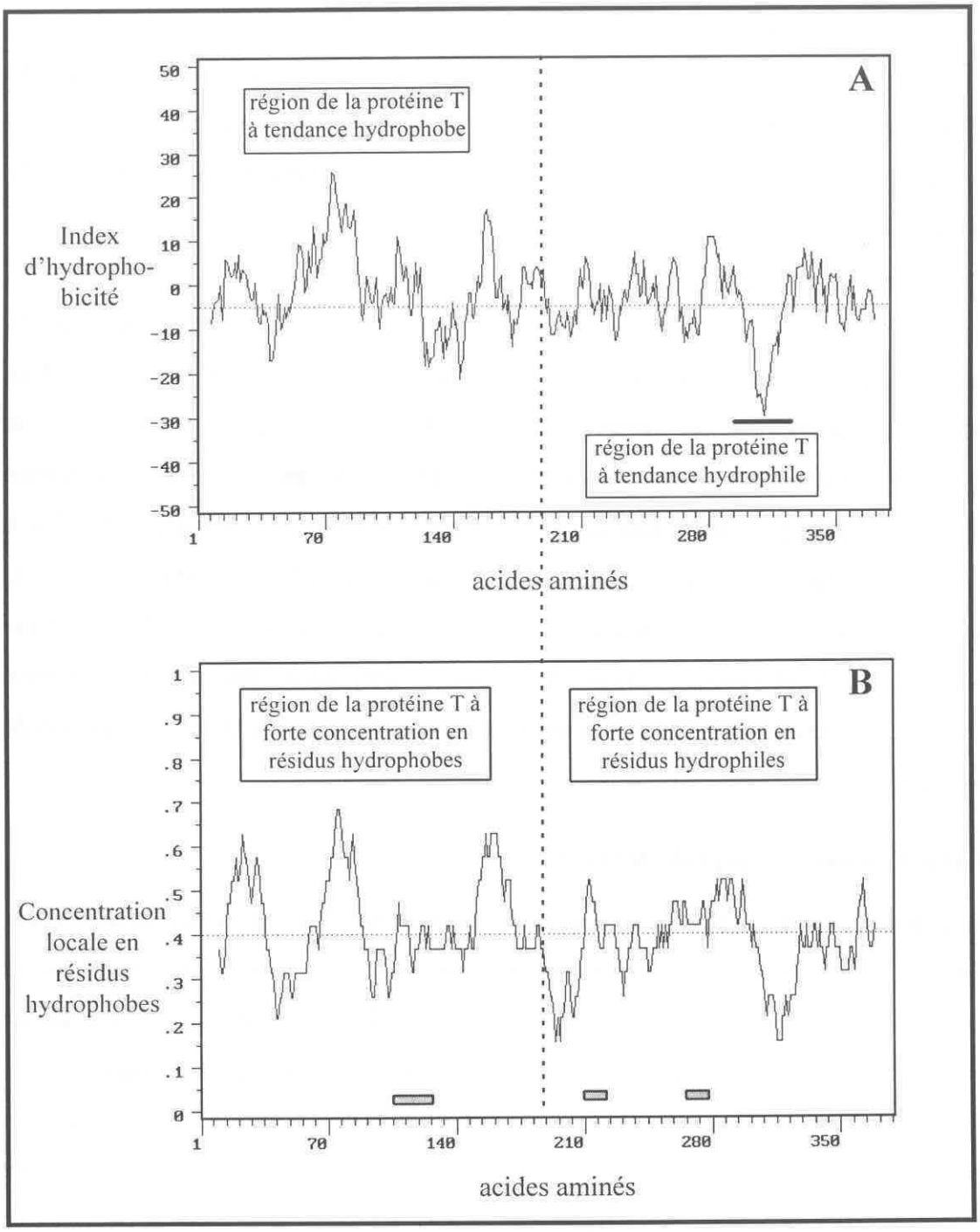


Figure 21:
A) Profil d'hydrophobicité établi selon la méthode de Kyte et Doolittle (1982) de la séquence primaire de la protéine T. Le principal pic hydrophile est souligné par un trait (▬▬).
B) Distribution des acides aminés hydrophobes le long de la séquence primaire de la protéine T pour un intervalle de 19 acides aminés. Les déterminants antigéniques (▭) sont établis selon la méthode de Hopp et Woods (1981) et se situent entre les acides aminés 269-274, 222-227 et 110-125 (avec pour les deux derniers cités un pourcentage d'erreur de 30 %).

Une première vue d'ensemble sur le profil d'hydrophobicité de la protéine T nous montre que l'on retrouve les deux domaines distincts de la protéine:

- la moitié aminoterminale comporte principalement 3 pics hydrophobes (résidus 63 à 91 et 140 à 168) séparés par un domaine plus hydrophile (courbe A). De plus, il semble que la concentration en résidus hydrophobes (A, I, F, L, M, P,V,W) le long de la structure primaire de la protéine soit beaucoup plus importante dans cette partie aminoterminale que dans la deuxième moitié de la protéine (courbe B).

- la moitié carboxyterminale semble avoir une tendance hydrophile plus marquée, notamment par la présence d'un pic important (résidus 301 à 322) qui confirme le profil hydrophile de cette portion de la protéine. Enfin, les prédictions antigéniques possibles déterminées selon la méthode de Hopp et Woods (1981) sur la séquence primaire de la protéine T, montrent que les trois régions les plus hydrophiles susceptibles d'être antigéniques se situent, pour les deux premières dans la région carboxyterminale hydrophile [résidus 269 à 274 (Lys-Arg-Arg-Arg-Ala-Glu) et 222 à 227 (Glu-Lys-Ser-Glu-Gly-Lys)] et pour la troisième, dans la partie aminoterminale hydrophobe [de l'acide aminé 120 à 125 (Gly-Cys-Arg-Asp-Lys-Asp)]. Il est à noter que la première région citée possède une très forte tendance antigénique potentielle du fait de sa forte proportion en résidus polaires et de la répétition de l'acide aminé arginine.

3 - Caractérisation du peptide de transit

Les mitochondries et les chloroplastes importent la plupart de leurs protéines (codées par le génome nucléaire) sous la forme d'une préprotéine possédant un peptide de transit au niveau de sa région aminoterminale. Ce peptide de transit permet à la protéine de pénétrer au sein de l'organite correspondant. Dans la plupart des cas, immédiatement après l'import de la protéine, ce peptide de transit est éliminé grâce à des protéases se trouvant dans les organites (pour une revue voir Moore *et al.*, 1994). Ces peptides se caractérisent par des propriétés communes, à savoir un enrichissement typique en acides aminés hydrophobes et l'absence de résidus acides. Toutefois, les peptides de transit mitochondriaux et chloroplastiques possèdent

des domaines structuraux spécifiques, selon la destination du peptide, qui permettent de les distinguer. Ainsi, par leur simple contenu en acides aminés mais également par leur contenu relatif en Ser et Arg (déterminé selon un calcul statistique: $f_{Ser} > 0{,}07 + 1{,}4 \times f_{Arg}$), von Heijne *et al.* (1989) ont pu distinguer (avec 90 % de réussite) les peptides de transit mitochondriaux et chloroplastiques.

Les peptides de transit mitochondriaux se caractérisent par une très faible proportion de glutamate, d'aspartate, d'isoleucine et de lysine et par une forte proportion d'alanine, de leucine, d'arginine et de sérine (von Heijne *et al.*, 1989). La séquence de transit de la protéine T obtenue à partir de l'ADNc semble posséder ces caractéristiques:

Le peptide de transit de la protéine T possède, effectivement, un pourcentage élevé en arginine (16,6%), en alanine (10%) et en leucine (10%) et ne contient pas ou peu de glutamate (0%), d'aspartate (3,33%), d'isoleucine (3,33%) et de lysine (6,66%). Il est à noter, cependant le très faible pourcentage de sérine et la forte présence de glycine (dont la plupart des molécules sont disposées par deux). De plus, le résidu arginine fréquemment rencontré aux positions -2 ou -10 du site de clivage matriciel (von Heijne *et al.*, 1989) est présent en position -3 (figure 22). En effet, chez les plantes la position -3 du résidu arginine est généralement préférée à la position -2 tandis que la position -10 n'est jamais présente (Moore *et al.*, 1994). On remarque notamment que la plupart des protéines impliquées dans le complexe de la GDC possèdent à la position -3 une arginine à l'exception de la protéine P qui la présente en position - 2.

protéines matricielles	Préséquences -3
Psgdc-T	MRGGLWQLGQSITRRLANGGDKKAVAR**R**CF
Psgdc-H	MALRMWASSTANALKLSSSSRLHLSPTFSIS**R**CF
Psgdc-L	MAMANLARRKGYSLLSSETLRYSFSLRS**R**AF
PsHSP22	MASSLALKRFLSSGLLSSSFLRPVASSAS**R**SF
PsHSP70*	MAATLLRSLQRRNLSSSSVSAFRSLTGSTKTSYATHKLASLTRPFSS**R**PA
CvmMDH	MKASILRSVRSAVRSSSSNRLLS**R**SF
IbdF1ATPase	MAMTGRARSMGFSILQKALSSAQRSNAHRSILCPTLSNSELL**R**NY
Atcpn60*	MYRFASNLASKARIAQNARQVSSRMSWS**R**NY

figure 22: Préséquences de certaines protéines mitochondriales de plantes. Psgdc-T: protéine T de pois; Psgdc-H: protéine H de pois (Macherel *et al.*, 1990); Psgdc-L: protéine L de pois (Bourguignon *et al.*, 1992); PsHSP22: protéine HSP22 de pois (Lenne *et al.*, 1995); PsHSP70*: protéine HSP70 de pois (Watts *et al.*, 1992); CvmMDH: malate deshydrogénase de pastèque (Gietl *et al.*, 1990); IbdF1ATPase: sous-unité F1 de l'ATPase de

patate douce (Kimura *et al.*, 1990a); Atcpn60*:Protéine cpn 60 d'arabidopsis (Prasad *et al.*, 1992). L'acide aminé en caractère gras correspond au résidu arginine rencontré en position - 3 au niveau du peptide de transit des plantes.

Les peptides de transit mitochondriaux sont composés généralement par 2 domaines structuraux distincts: une région aminoterminale à caractère amphiphile contenant une hélice α et un domaine hydrophobe du côté carboxyterminal. On retrouve ces 2 domaines caractéristiques au sein du peptide de transit de la protéine T mais disposés de façon inverse. En effet, il semble que, contrairement à la plupart des peptides de transit mitochondriaux (et à ceux des protéines H et P du complexe de la GDC), la région aminoterminale du peptide de transit de la protéine T est une région hydrophobe suivie de son côté carboxyterminal par un domaine capable de former une hélice α. En effet, le moment hydrophobique de la région aminoterminale (pour un angle de 95° entre les différents acides aminés) est plus élevé que celui de la partie carboxyterminale. Ce qui suggère que la région aminoterminale pourrait correspondre à une structure en feuillet-β à caractère amphiphile (Eisenberg *et al.*, 1984). En général, cette structure est trouvée plutôt dans la région carboxyterminale des peptides de transit mitochondriaux.

L'ensemble de ces observations montre que le peptide de transit de la protéine T de pois possède un certain nombre de caractéristiques similaires à celles définies par von Heijne *et al.* (1989) à partir de l'analyse de 37 peptides de transit mitochondriaux. Le peptide de transit de la protéine T se différencie, néanmoins, par une région fortement hydrophobe du côté de sa région aminoterminale.

B - RECHERCHE D'HOMOLOGIES DE SEQUENCES

La séquence en acides aminés de la protéine T de pois (variété Douce provence), déduite de la séquence nucléotidique de l'ADNc, a été alignée selon la méthode de Higgins et Sharp (1988), avec les séquences primaires de protéines T appartenant à *Escherichia coli* (Okamura-Ikeda *et al.*, 1993) et à quatre espèces eucaryotes différentes: la pomme de terre, le boeuf, le poulet et l'homme (Kopriva et Bauwe, 1994; Okamura-Ikeda *et al.*, 1991, 1992; Hayasaka *et al.*, 1993) (figure 23).

```
                                 +  +                  +           +
POIS(DP)      ATESELKKTVLYDFHVAHGGKMVPFAGWSMPIQYKDSIMDSTLNCRQNGS   50
P. de TERRE   ASDADLKKTVLYDFHVVNGGKMVPFAGWSMPIQYKDSIMDSTVNCRENGS   50
POULET        SAPEGLKQTPLDALHRARGGRMVPFAGWSLPVQYGRGHLESHLHTRRHCS   50
BOEUF         -AQDVLHRTPLYDFHLAHGGKMVAFAGWSLPVQYRDSHVNSHLHTRQHCS   49
HUMAIN        -AQEVLRRTPLYDFHLAHGGKMVAFAGWSLPVQYRDSHTDSHLHTRQHCS   49
               . . *..* * ..⊕ ..**.**.*****.*.**  .  .* .. * . *

                                 +     +    +      +        ++  +
E.coli        MAQ----QTPLYEQHTLCGARMVDFHGWMMPLHY-GSQIDEHHAVRTDAG   45
               .     .* * . ⊕   *..** * ** .*..*  .  ..    * . .

                    +      + +       +                       +
POIS(DP)      LFDVSHMCGLSLKGKDVISFLEKLVIADVAALANGTGTL-TVFTNEKGGA   99
P. de TERRE   LFDVSHMCGLSLKGKDTIPFLEKLVIADVAGLAPGTGSL-TVFTNEKGGA   99
POULET        LFDVSHMLQTRVYGRDRVRFLESLVVGDIAELRPGQGTL-TLLTNERGDI   99
BOEUF         LFDVSHMLQTKIFGCDRVKLMESLVVGDIAELKPNQGTL-SLFTNEAGGI   98
HUMAIN        LFDVSHMLQTKILGSDRVKLMESLVVGDIAELRPNQGTL-SLFTNEAGGI   98
              *****⊕*   .. * * . ..*.**..*.*.*  . *.* ...*** *.

                    +      +  + +   +        +  +  +
E.coli        MFDVSHMTIVDLRGSRTREFLRYLLANDVAKLTKSGKALYSGMLNASGGV   95
              .****⊕*    . *     ..  *. .*.* *  .  .* . . *. *.


                      +    +           +   +   + +  + +      + +
POIS(DP)      IDDSVITKVTEDHLYLVVNAGCRDKDLAHIEEHMKAFKAKGGDVSWHIHD  149
P. de TERRE   IDDSVVTKVTNDHIYLVVNAGCRDKDLAHIEEHMKSFKSKGGDVSWHIHD  149
POULET        VDDLIVTNTAEDHLYVVSNAGCADKDRAVMEGRAAELRAAGGDVHLEVSG  149
BOEUF         LDDLIVTSASEGHLYVVSNAGCREKDLTLMQDKVRELQNKGSDVALEVMD  148
HUMAIN        LDDLIVTNTSEGHLYVVSNAGCWEKDLALMQDKVRELQNQGRDVGLEVLD  148
              .** ..*.....*.*.* **** .⊕* . ....  .... * **  .. .

                                 +       + +       +               +
E.coli        IDDLIVYYFTEDFFRLVVNSATREKDLSWITQHAEPF-----GIEITVRD  140
              .** ..   ... . .* *..  .⊕* . .  .   .     ..   . .


               +                + +    +      +             +
POIS(DP)      ERSLLALQGPLAAPVLQHLTKEDLSKLYFGEFRVLDINGSQ-CFLTRTGY  198
P. de TERRE   ERSLLALQGPLAAPVLQYLTKDDLSKMYFGEFRVLDINGAP-CFLTRTGY  198
POULET        QR-AAGVQGPSMAQVLQAGLPDDLTKLTFMTSTATTVFGVPGCRVTRCGY  198
BOEUF         NA-LLALQGPTAAQVLQAGVADDLRKLPFMTSAVMEVFGVSGCRVTRCGY  197
HUMAIN        NA-LLALQGPTAAQVLQAGVADDLRKLPFMTSAVMEVFGVSGCRVTRCGY  197
              .    ..***  *.***    .** ⊕. * .  . .. *   * .*⊕ **

                                 +         +      +
E.coli        DLSMIAVQGPNAQAKAATLFNDAQRQAVEGMKPFFGVQAGDL-FIATTGY  189
              .    ..***  ..  .    ..  .         .. .     ..  **
```

```
                                +     +    +   + +      +   +
POIS(DP)      TGEDGFEISVPSEHGVELAKALLEKSEGKIRLTGLGARDSLRLEAGLCLY  248
P. de TERRE   TGEDGFEISVPSENALDLAKALLEKSEGKIRLTGLGARDSLRLEAGLCLY  248
POULET        TGEDGVEISVPAGRAVELAERLLGCPE--VWPAGLAARDSLRLEAGLCLY  246
BOEUF         TGEDGVEISVPAAEAVHLAAALLKNPE--VKLAGLAARDSLRLEAGLCLY  245
HUMAIN        TGEDGVEISVPVAGAVHLATAILKNPE--VKLAGLAARDSLRLEAGLCLY  245
              ***** ***** . ...**  .*  .*  .  .**.*⊕***⊕********

                                +     +          +      +   +
E.coli        TGEAGYEIALPNEKAADFWRALV---EAGVKPCGLGARDTLRLEAGMNLY  236
              ***.* **..* . . ..   ..   *  .   **.*⊕*.*⊕****. **

                     +              ++++             +        ++
POIS(DP)      GNDLEQHITPIEAGLTWAIGKRRRAEGGFLGADVILKQLADGPSIRRVGF  298
P. de TERRE   GNDMEQHTTPVEAGLTWAIGKRRRAEGGFLGAEVILKQIEEGPKIRRVGF  298
POULET        GNDIDESTTPVEAGLLWTLGKRRRTAMDFPGAAIIMEQVKEKPKRKRVGL  296
BOEUF         GNDIDEHTTPVEGSLSWTLGKRRRAAMDFPGASVIVPQLKSKAQRRRVGL  295
HUMAIN        GNDIDEHTTPVEGSLSWTLGKRRRAAMDFPGAKVIVPQLKGRVQRRRVGL  295
              ***... .**.*..* *..*⊕⊕⊕⊕.. .* ** .*. *. .    .⊕**.

                                          +    +      + +
E.coli        GQEMDETISPLAANMGWTIA-WEPADRDFIGREALEVQREHGTEKLVGLV  285
              *..... ..*..... *... . ... .* *   .  *

                     + +                            ++       +   +
POIS(DP)      ISSGPPPRSHSEIQ---DEGGNNIGEVTSGGFSPCLKKNIAIGYVKSGLH  345
P. de TERRE   FSSGPPPRSHSEIQ---DSNGQNIGEITSGGFSPCLKKNIAMGYVKTGNH  345
POULET        TSVGPPLRPPAAIL---GPEGTPVGTVTSGCPSPSLGKNIAMGYVQAAHS  343
BOEUF         MCDGAPVRAQSPIL---SPEGTVIGAVTSGCPSPCLKKNVAMGYVPYEYS  342
HUMAIN        MCEGAPMRAHSPIL---NMEGTKIGTVTSGCPSPSLKKNVAMGYVPCEYS  342
               . *.* *... *    . .*  .*..***  **.* ⊕*.*.***  .

                 +   +     +        +                    +
E.coli        MTEKGVLRNELPVRFTDAQGNQHEGIITSGTFSPTLGYSIALARVPEGI-  334
                     *    .    . ..   * .***  ** *  ..*.. *  .

              +   + +   + +       +      +  +
POIS(DP)      KAGTKVKIIIRGKQNEGVVTKMPFVPTKYYKPS   378
P. de TERRE   KAGTNVKIVIRGKSYDGVVTKMPFVPTKYYKP-   377
POULET        RPGTTLTVEVRKKQHPALVTKMPFVPTHYYMAK   376
BOEUF         RPGTPLLVEVRRKQQPAVVSKMPFVSTNYYILK   375
HUMAIN        RPGTMLLVEVRRKQQMAVVSKMPFVPTNYYTLK   375
              ..** . . .⊕ ⊕   ..*.⊕****.*.**

                        + +    +  +    +  +
E.coli        --GETAIVQIRNREMPVKVTKPVFVRNGK-AVA   364
                *.   . .⊕ .     *.⊕  **..
```

figure 23: Comparaison de la séquence en acides aminés de la protéine T mature provenant de feuilles de pois (DP: Douce provence), de pomme de terre, de foie de boeuf, de poulet, d'homme et d'E. coli. Ces séquences proviennent respectivement de Kopriva et Bauwe, 1994; Okamura-Ikeda *et al.*, 1991, 1992; Hayasaka *et al.*, 1993; Okamura-Ikeda *et al.*, 1993. Le symbole * indique que les résidus sont identiques alors que le symbole . indique qu'ils sont similaires. Le symbole + montre la présence de résidus basiques chez la protéine T de pois. Le symbole ⊕ indique la position en commun de résidus chargés positivement. Les cadres en gris indiquent les régions fortement homologues.

La comparaison de la séquence primaire de la protéine T de pois présente une très forte homologie d'environ 87 % avec celle de la protéine T de pomme de terre. De même, les degrés d'identité obtenus avec les séquences primaires de la protéine T provenant du règne animal sont assez importants: 52% en moyenne. Par contre, nous observons une faible homologie avec la séquence primaire de la protéine T d'*E coli*: 30% environ (Okamura-Ikeda *et al.*, 1993). Ces résultats témoignent donc de la forte conservation de cette protéine chez les eucaryotes. En effet, il est fort probable que certaines des séquences conservées soient impliquées dans le site catalytique de la protéine T, mais également dans l'interaction de la protéine T avec son cofacteur le tétrahydrofolate et la protéine H.

V - DISCUSSION-CONCLUSION

La purification en deux étapes de la protéine T du complexe de la glycine décarboxylase, à partir d'un extrait matriciel de mitochondries de feuilles de pois, a permis d'obtenir des informations (par digestion chimique et microséquençage) sur la séquence peptidique des régions NH_2-, COOH- et interne de la protéine T native et, d'obtenir des anticorps polyclonaux de lapin dirigés spécifiquement contre cette protéine. A l'aide de ces anticorps nous avons criblé une banque d'ADNc (dans λgt 11) fabriquée à partir d'ARNm de feuilles de pois et, isolé deux ADNc proches de la taille complète (1,4 kpb). Ainsi, nous avons pu déduire, de la combinaison des séquences des différents clones, la structure primaire de la protéine T. L'identité de l'ADNc codant pour la protéine T ainsi que l'exactitude de la structure primaire de cette protéine ont pu être confirmées:

- par la présence, dans la séquence déduite de l'ADNc, de la séquence amino et carboxyterminale de la protéine T purifiée,
- par la similitude de la masse moléculaire du polypeptide mature déduite de la séquence nucléotidique, avec celle de la protéine purifiée analysée par gel d'électrophorèse,

- par une très bonne corrélation entre la composition en acides aminés déduite de la séquence nucléotidique et la composition obtenue par hydrolyse acide de la protéine T mitochondriale purifiée à partir du pois,

- enfin par l'analyse de la protéine T, à l'aide de la spectrométrie de masse et par dégradation d'Edman. En effet, la masse de la protéine T native mesurée par spectrométrie de masse (40966±5 Da) correspond à celle déduite de l'ADNc (40961 Da). Ce résultat a été conforté par la mesure des masses des différents peptides obtenus par digestion chimique de la protéine suivie d'une analyse par HPLC/ESI-MS. De plus, le microséquençage par dégradation d'Edman de certains peptides a permis de confirmer 35% de la structure primaire de la protéine T (soit 131 acides aminés sur 378).

La recherche d'homologies de séquences au sein de la protéine T a permis de localiser des régions hautement conservées, notamment une région comportant un alignement de 4 résidus basiques (K^{269}RRR). Certaines de ces régions jouent vraisemblablement un rôle important au niveau du site catalytique de la protéine T. En particulier, elles pourraient être impliquées dans l'interaction de la protéine T avec son cofacteur le tétrahydrofolate et la protéine H. Cependant, nous n'avons jamais détecté par spectrométrie de masse la présence de tétrahydrofolate lié de façon covalente à la protéine T. Ainsi, contrairement à la protéine H qui subit une lipoylation au niveau de la lysine 63, la protéine T ne subirait pas de modification post-traductionnelle. Il est donc fort probable que le tétrahydrofolate réagisse avec la protéine T par des interactions de type électrostatique et/ou hydrogène. Ainsi, les résidus chargés positivement précédemment décrits pourraient être impliqués dans la liaison avec le tétrahydrofolate. En effet, Rebeillé *et al.* (1994) et Besson *et al.* (1993) ont montré que l'affinité du tétrahydrofolate pour la GDC augmente avec le nombre de résidus poly-γ-glutamylés (Kd≈0.5 μM pour $H_4PteGlu_5$) pour atteindre un maximum lorsque le nombre de glutamates est supérieur à trois (80% du folate situé dans l'espace matriciel des mitochondries est sous forme $H_4PteGlu_4$ et $H_4PteGlu_5$) suggérant ainsi d'éventuelles interactions ioniques entre les charges négatives de la queue polyglutamate et les charges positives de la protéine T. Par ailleurs, Usha *et al.*, 1992 ont pu mettre en évidence de telles interactions entre le tétrahydrofolate et certains résidus arginine de la SHMT de foie de brebis. Des expériences de mutagénèse dirigée sur les différents résidus basiques pourraient être envisagées afin de détecter une éventuelle diminution de l'affinité du tétrahydrofolate. Il faut également noter que le renforcement des interactions du tétrahydrofolate avec la protéine T, protège la

molécule de tétrahydrofolate de toutes oxydations (oxydation du tétrahydrofolate en dihydrofolate).

L'alignement des séquences primaires de la protéine T provenant de différentes espèces (pois, poulet, boeuf et homme) avec celles d'enzymes utilisant le $H_4(H_2)PteGlu_n$ comme substrat: la thymidylate synthase (Maley *et al.*, 1979; Hardy *et al.*, 1987), la sérine hydroxyméthyltransférase (Turner *et al.*, 1992c; Martini *et al.*, 1989), la dihydrofolate réductase (Blakley, 1984), la C1-H_4PteGlu synthétase (Staben et Rabinowitz, 1986) n'a pas permis de mettre clairement en évidence des séquences consensus conservées pouvant interagir avec le tétrahydrofolate. En particulier, nous ne retrouvons pas dans la séquence de la protéine T, le domaine de liaison du folate identifié dans la thymidylate synthase par Maley *et al.* (1982).

Ainsi, comme le suggèrent de nombreux auteurs (Blakley, 1984; Thompson and Freisheim, 1991; Usha *et al.*, 1992), nous pensons que les enzymes « folate dépendante » possèdent leurs propres sites de liaison et que de nombreux acides aminés présents au sein des différentes structures secondaires de la protéine pourraient être impliqués dans la liaison du cofacteur. Comme le proposent Jennings et Cotton (1990), nous pouvons imaginer que, s'il existe des homologies entre les protéines à folates, elles seraient essentiellement d'ordre conformationnel.

Actuellement, de nombreuses études sont en cours afin de mieux comprendre les interactions de la protéine T avec la protéine H et en particulier avec la protéine H chargée en méthylamine. En plus de son rôle catalytique, la protéine H semble jouer un rôle de «chaperon» pour la protéine T. En effet, nous avons constaté que la protéine T subit des phénomènes d'agrégations irréversibles lorsqu'elle se trouve à des concentrations supérieures à 1 mg/ml. Par contre, lorsque la protéine T se trouve en présence de la protéine H, ces phénomènes d'instabilité disparaissent.

Récemment, la cristallisation de la protéine H et la détermination de sa structure tridimentionnelle obtenues en collaboration avec le laboratoire de Cristallographie Macromoléculaire de l'Institut de Biologie Structurale de Grenoble, nous ont apporté de précieux renseignements sur la conformation de la protéine H (figure 24). La protéine H est une protéine qui existe sous trois formes différentes: oxydée (H*ox*), chargée en méthylamine (H*met*) et réduite (H*red*), et c'est sous ses formes oxydée et méthylamine qu'elle a été cristallisée. La forme méthylée s'avère particulièrement intéressante car c'est dans cette

conformation que la protéine H interagit avec la protéine T et le tétrahydrofolate. En effet, la réaction catalysée par la protéine T permet le transfert du groupement $-CH_2-$ présent sur le lipoate vers le tétrahydrofolate.

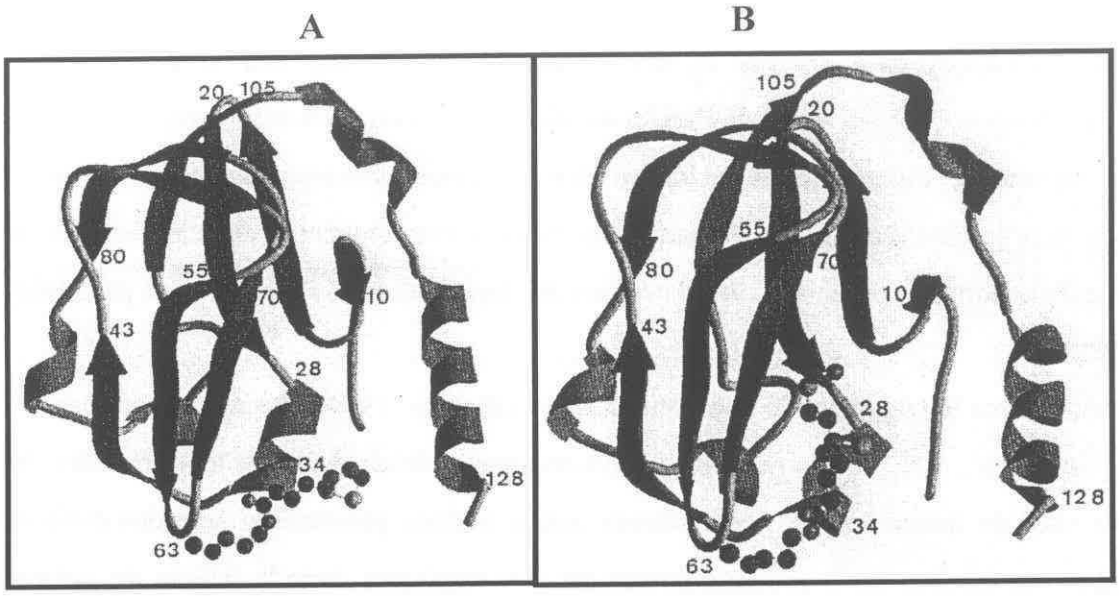


Figure 24: Structure tridimensionnelle de la protéine H*ox* et H *met*. Structure en ruban de la protéine H*ox*. (A) H *met*. (B) La structure de la protéine H *met* révèle la cavité dans laquelle vient s'enchasser le résidu lipoate ayant fixé le groupement méthylamine. Noter la rotation du bras lipoate lors de la fixation du groupement méthylamine (Cohen-Addad *et al.*, 1995).

L'étude par diffraction des rayons X a permis, après traitement informatique, de montrer que le coeur de la protéine H est formé de deux « feuillets » β formant une structure en sandwich. L'un des feuillets est constitué de quatre brins β, l'autre feuillet est formé de trois brins β dont deux sont disposés de façon antiparallèle. Ces deux brins adjacents sont raccordés par une boucle en épingle à cheveux très fermée portant la lysine 63 sur laquelle est lié le résidu lipoate. Le résidu lipoate se situe dans une situation très externe à la surface de la protéine et dans une conformation relativement flexible, ce qui rend le lipoate très accessible à d'éventuelles réactions. La fixation du groupement méthylamine provoque une rotation de 90° du bras lipoate par rapport à sa position initiale. Cette rotation permet l'ancrage du groupe méthylamine dans une poche hydrophobe située en surface de la protéine. De ce fait, le groupement CH_2 ainsi enfoui est protégé de l'hydrolyse (attaque nucléophile par une molécule d'eau) qui provoquerait la décomposition spontanée de la méthylamine en formaldéhyde et NH_3. Actuellement, l'hypothèse émise est la suivante: la protéine T qui interagit étroitement avec la protéine H viendrait positionner le cofacteur tétrahydrofolate dans la poche

hydrophobe. Dans ces conditions, l'attaque nucléophile par l'azote basique (N5) du cycle ptérine du tétrahydrofolate conduirait à un transfert direct du groupement méthylène (-CH2) sur le tétrahydrofolate sans passer par l'intermédiaire formaldéhyde libre.

Ainsi, les résultats obtenus sur la structure de la protéine H et la détermination de la structure primaire de la protéine T nous ont permis d'envisager, en collaboration avec Claudine Cohen-Addad, la cristallisation de la protéine T ou sa co-cristallisation avec la protéine H*met*. Des expériences de mutagénèse dirigée sont également envisagées afin de mieux comprendre les interactions des différentes protéines du complexe de la glycine décarboxylase. Dans cette optique et afin d'obtenir facilement de grandes quantités de protéine T (la préparation à partir d'extrait de pois de protéine T pure en grosse quantité étant lourde et laborieuse), nous avons entrepris un projet visant à produire la protéine T dans un système procaryote (Chapitre IV). Ce programme de recherche s'intègre dans une étude plus globale du complexe de la glycine décarboxylase des mitochondries végétales. En effet, depuis 1990 de nombreux efforts ont été entrepris au laboratoire dans le but de réaliser une étude moléculaire de l'expression des protéines du complexe de la GDC. Ainsi, notre premier objectif va être d'étudier la régulation ainsi que la biogénèse des protéines de ce complexe.

CHAPITRE II

CHAPITRE II

ETUDE DE L'EXPRESSION DE LA PROTEINE T ET DE LA MISE EN PLACE DU COMPLEXE DE LA GLYCINE DECARBOXYLASE AU COURS DU DEVELOPPEMENT

Grâce à l'isolement et la caractérisation de l'ADNc correspondant à la protéine T, nous avons abordé l'étude de l'expression de cette protéine dans différents tissus en parallèle avec les autres protéines de la glycine décarboxylase. De plus, comme la Ribulose 1,5-bisphosphate carboxylase oxygénase et le complexe de la GDC catalysent les étapes clefs du cycle de la photorespiration, il nous a paru intéressant de suivre l'apparition de ces deux complexes au cours du développement foliaire des plantules de pois.

I - DISTRIBUTION TISSULAIRE ET INFLUENCE DE LA LUMIERE SUR L'EXPRESSION DES DIFFERENTES PROTEINES DU COMPLEXE DE LA GLYCINE DECARBOXYLASE

A - ANALYSE PAR WESTERN-BLOT DE LA PRESENCE DE LA PROTEINE T ET DES AUTRES PROTEINES DE LA GDC DANS DES MITOCHONDRIES ISOLEES DE DIFFERENTS TISSUS

Des protéines mitochondriales isolées de tissus foliaires de pois (exposés durant 12 jours ou non à la lumière) et de racines ont été séparées par électrophorèse sur gel de polyacrylamide, transférées sur une membrane de nitrocellulose (Cf. Matériel et méthodes) et analysées à l'aide de sérums dirigés spécifiquement contre les protéines T, P, H et L de la GDC. La figure 25 montre clairement que la protéine T ainsi que les protéines P et H sont présentes en très faible quantité au sein des mitochondries de tissus étiolés. Par contre, une induction très importante (d'un facteur 8 à 10) est observée lorsque les feuilles sont exposées

à la lumière. En ce qui concerne la protéine L, nous pouvons constater que cette protéine est beaucoup moins abondante dans les tissus chlorophylliens et semble ne subir qu'une faible induction à la lumière (facteur 2 à 3). De plus, nous pouvons remarquer que les protéines T, P et L sont présentes en faible quantité au sein des mitochondries de tissus racinaires, alors que la protéine H n'est pas détectée.

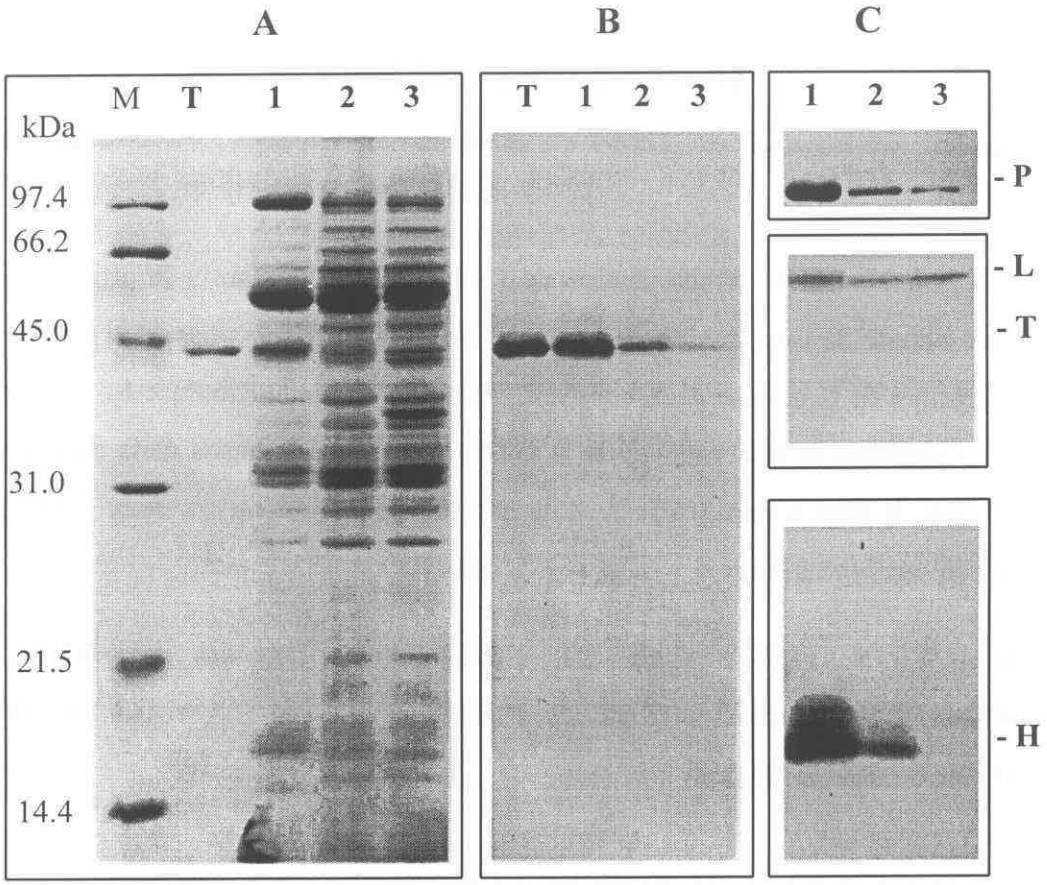


Figure 25: Analyse électrophorétique (A) et immunodétection (B, C) des protéines T, P, L et H du complexe de la glycine décarboxylase. A: Les protéines ont été séparées sur gel de polyacrylamide en présence de SDS puis visualisées par coloration au bleu de Coomassie R250. M, marqueurs de masse moléculaire. B: Immunodétection de la protéine T sur une membrane de nitrocellulose après transfert des protéines séparées sur gel de polyacrylamide en présence de SDS. C: Immunodétection des protéines P, L et H sur une membrane de nitrocellulose, après transfert des protéines séparées sur gel de polyacrylamide en présence de SDS. Ligne T: 2 µg de protéines T purifiées; lignes 1, 2 et 3: 30 µg de protéines mitochondriales extraites respectivement de feuilles de pois chlorophylliennes, de feuilles de pois étiolées et de racines de pois. Les anticorps utilisés ont été préparés à partir des protéines purifiées T, P, L et H de mitochondries de feuilles de pois cultivées à la lumière. Les lettres à droite indiquent la position des protéines T, P, H et L.

B - ANALYSE PAR NORTHERN-BLOT DU NIVEAU DES TRANSCRITS CORRESPONDANT A LA PROTEINE T

Les ARN totaux isolés de feuilles de tiges et de racines (provenant de plantules de pois exposées ou non à la lumière) ont été séparés sur un gel d'électrophorèse en conditions

dénaturantes et transférés sur une membrane Hybond N^+ (Cf. Matériel et méthodes). L'ADNc correspondant à la protéine T a été utilisé comme sonde pour étudier le niveau d'expression du gène codant pour cette protéine dans les différents tissus. Cette étude a été menée en parallèle avec celle de la protéine H en utilisant son ADNc correspondant comme sonde. L'autoradiographie présentée dans la figure 26 montre que les transcrits correspondant à la protéine T au même titre que ceux correspondant à la protéine H s'expriment essentiellement dans les tissus foliaires chlorophylliens. La quantification des signaux, par analyse au densitomètre, montre une induction d'un facteur 10 environ entre les feuilles chlorophylliennes et les feuilles étiolées. Notons que les messagers correspondant à la protéine T sont observés dans les tiges et les racines, alors que ceux correspondant à la protéine H ne sont pas détectés.

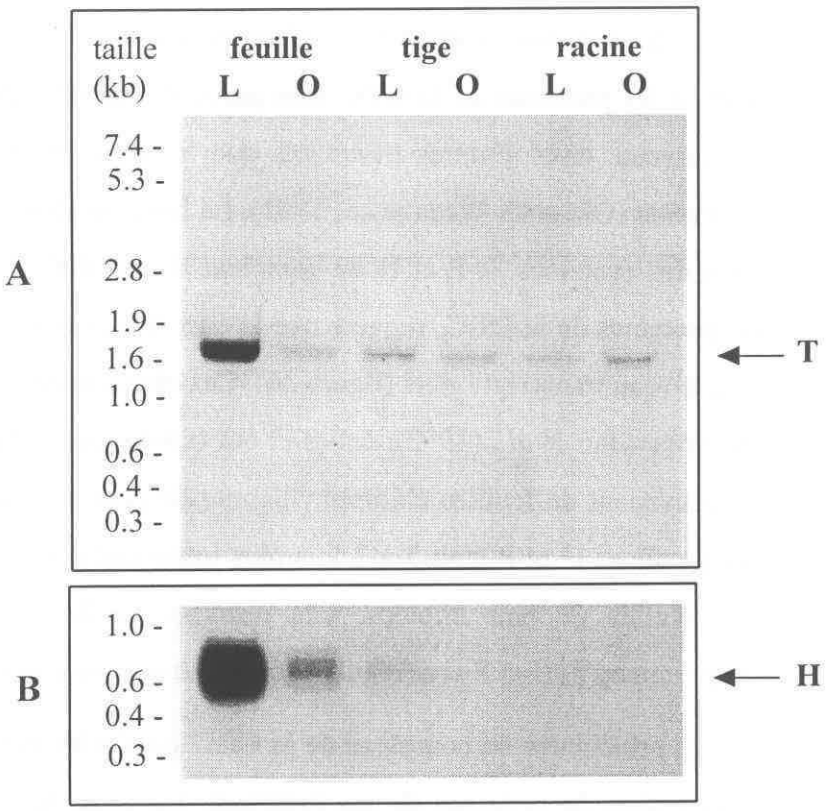


Figure 26: Distribution tissulaire des ARNm correspondant à la protéine T (A) et à la protéine H (B). 7,5 µg d'ARN totaux isolés à partir de tissus foliaires, de tiges et de racines de pois cultivés à la lumière (L) ou à l'obscurité (O) ont été séparés sur gel d'agarose, transférés sur une membrane de nylon, puis hybridés avec une sonde radioactive réalisée à partir de l'ADNc TB2 (voir figure 18) correspondant à la protéine T (A) et Pgh12 correspondant à la protéine H (Cf. Macherel *et al.*, 1990) (B). Les marqueurs d'ARN sont à gauche.

C - DISCUSSION

Les expériences de « Northern blot » et de « Western-blot » montrent que le gène codant pour la protéine T est exprimé essentiellement dans les feuilles chlorophylliennes. L'induction que l'on observe à la lumière au niveau protéique est également observée au niveau des transcrits. Ce rapport de l'ordre de 8 à 10 est similaire à celui observé au niveau des transcrits correspondant aux protéines H, P et SHMT (Macherel *et al.*, 1990; Turner *et al.*, 1992a; Turner *et al.*, 1992c). Toutefois, de façon surprenante, nous avons remarqué qu'un taux non négligeable de messagers correspondant à la protéine T sont présents dans les racines et dans les tiges, alors qu'aucun messager correspondant à la protéine H n'a été détecté. Le même phénomène a été observé au niveau de l'expression des protéines. En effet, il est possible que la cellule maintienne un certain niveau de protéines P et T et que la réaction globale catalysée par la GDC soit contrôlée par la présence de la protéine H. On retrouve, dans le règne animal les mêmes variations de distribution entre les différents tissus avec, cependant, une petite différence. En effet, si les protéines de la GDC sont particulièrement représentées et actives dans le foie et le cerveau, dans d'autres tissus tels que le coeur et les poumons, c'est la protéine P qui est absente (Okamura-Ikeda *et al.*, 1991). La forte corrélation qui existe entre le niveau d'expression des protéines T, P et H du complexe de la GDC et celui des transcrits correspondant aux protéines de la GDC, suggère que le taux d'expression des protéines de la GDC est modulé au niveau transcriptionnel (figure 26). Ces observations ont été appuyées par les expériences de Srinivasan *et al.*, (1992); celles-ci ont montré que le taux de transcription des noyaux isolés provenant de feuilles chlorophylliennes de pois pour des ARN messagers correspondant aux protéines H et P était 5 à 7 fois plus important que celui obtenu avec des noyaux isolés de feuilles de pois étiolées. Ces résultats confirment que les transcrits correspondant aux protéines T, H et P sont régulés à la lumière au niveau transcriptionnel.

Une des propriétés principales du complexe de la GDC est l'effet inducteur de la lumière sur le niveau d'expression de certaines de ses protéines. En effet, les protéines T, P et H subissent la même influence de la lumière que d'autres enzymes chloroplastiques; l'exemple le plus étudié étant la petite sous-unité de la Rubisco (rbcS) (Gallagher et Ellis, 1982; Smith et Ellis, 1979). En effet, Kim et Oliver, (1990); Kim *et al.* (1991) et Srinivasan *et al.* (1992; 1993) ont montré que durant le verdissement des feuilles de pois étiolées, le taux des transcrits correspondant aux protéines de la GDC et celui des transcrits de la rbcS sont exprimés de

façon similaire. Ces deux protéines ont une masse moléculaire proche (15 kDa) et s'accumulent à des concentrations millimolaires dans la partie soluble de leurs organites respectifs. Elles sont codées par le génome nucléaire et subissent une régulation à la lumière au niveau transcriptionnel. Ces observations suggèrent, donc, que des mécanismes de contrôles similaires s'exercent sur ces deux gènes et probablement sur ceux codant pour les protéines T et P du complexe de la GDC. Cependant, Turner *et al.* (1993) ont montré que le taux de transcription des messagers correspondant aux protéines de la GDC ainsi que le niveau d'expression de leurs protéines ne sont pas induits de façon synchronisée lors d'une exposition de plantules étiolées à la lumière. En effet, les transcrits correspondant à la protéine P sont plus rapidement exprimés que les transcrits correspondant aux protéines T et H. Ainsi, on peut se demander s'il n'existe pas un ordre de mise en place des différentes protéines de la GDC pour former un complexe actif.

En ce qui concerne la protéine L, la lumière semble avoir un effet beaucoup plus modéré sur son niveau d'expression (Bourguignon *et al.*, 1992; Turner *et al.*, 1992b, 1993). Il semble que cette dihydrolipoamide déshydrogénase ait une régulation différente des autres protéines du complexe. En fait, récemment, Bourguignon *et al.* (1995) ont démontré que la protéine L impliquée dans le complexe de la GDC et la protéine E3 associée au complexe mitochondrial de la pyruvate deshydrogénase étaient la même protéine. Le fait qu'une unique dihydrolipoamide déshydrogénase soit impliquée dans différents complexes multienzymatiques chez les mitochondries des plantes supérieures, soulève d'intéressantes questions sur la régulation et l'aiguillage de cette protéine au sein de ces complexes. Il est donc fort probable que la dihydrolipoamide déshydrogénase soit régulée par un mécanisme différent de celui des protéines P, H et T du complexe de la GDC (Bourguignon *et al.*, 1992).

II - ETUDE DE LA MISE EN PLACE DU COMPLEXE DE LA GLYCINE DECARBOXYLASE AU COURS DU DEVELOPPEMENT DES FEUILLES DE POIS.

Tout organisme pluricellulaire possède au sein de son patrimoine génétique toutes les informations nécessaires à l'édification d'un organisme complet. Il existe un ordre chronologique dans la transcription des gènes et dans l'expression de ceux-ci, au cours du développement; certains gènes sont transcrits alors que d'autres sont réprimés (pour une revue voir Albert *et al.*, 1983). Dans notre cas, il nous a semblé intéressant d'étudier comment un complexe enzymatique mitochondrial, catalysant une étape importante de la photorespiration,

se mettait en place au cours du développement du pois. C'est pourquoi, afin de suivre l'évolution du complexe de la glycine décarboxylase, nous avons déterminé arbitrairement 5 stades successifs nommés L1, L2, L3, L4 et L5 correspondant respectivement à 4, 5, 7, 9 et 12 jours de croissance (Cf. Matériel et méthodes). Ainsi, pour chacun de ces stades, nous avons isolé, à partir de la première feuille, les mitochondries, les chloroplastes et les ARNs totaux, dans le but de mesurer le taux d'oxydation de la glycine par les mitochondries et d'étudier l'apparition des protéines et des messagers de la GDC en parallèle avec l'apparition de la ribulose 1,5-bisphosphate/carboxylase oxygénase (Rubisco) dans les chloroplastes. Cette étude a été réalisée en collaboration avec Nicolas Diallo de l'université de Dakar.

A - ETUDE DE LA MISE EN PLACE DES PROTEINES IMPLIQUEES DANS L'OXYDATION DE LA GLYCINE AU COURS DU DEVELOPPEMENT

1 - Evolution de l'oxydation de la glycine par des mitochondries isolées aux différents stades de développement.

A chacun des stades de développement, les mitochondries ont été isolées et purifiées sur gradient de Percoll, afin d'étudier leurs capacités oxydatives. Les mitochondries isolées à partir de la première feuille de plantules âgées de 4, 5 et 7 jours ont été purifiées sur un gradient de Percoll 28 % (ce gradient est généralement utilisé pour purifier les mitochondries de pommes de terre). Par contre, les mitochondries de plantules âgées de 9 et 12 jours ont été purifiées sur un gradient de Percoll autogénéré de 28% combiné avec un gradient linéaire de PVP 25, saccharose et raffinose selon la technique décrite par Douce *et al.*, (1987). La figure 27, montre que les mitochondries provenant de jeunes feuilles correspondant aux stades L1, L2 et L3 oxydent faiblement la glycine (3,2 à 12 nmoles d'O_2 consommées/min/mg de protéines mitochondriales en présence d'ADP). Cependant, à partir du stade L4, la vitesse d'oxydation de la glycine devient beaucoup plus élevée pour atteindre un maximum au stade L5 (138-232 nmol d'O_2 consommées/min/mg de protéines mitochondriales en présence d'ADP). Il semble donc qu'il existe au cours du développement une période charnière (entre le 7ième et le 9ième jour: stades L3 et L4) durant laquelle le complexe se met en place.

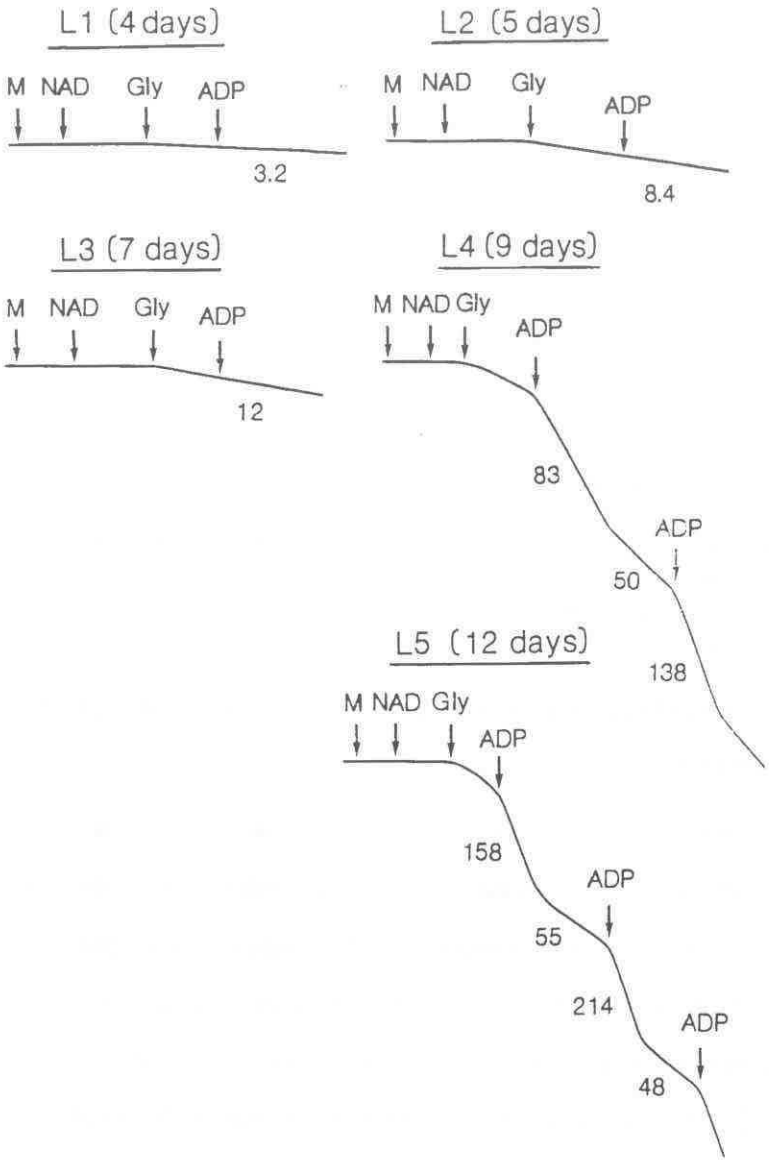


Figure 27: Oxydation de la glycine par des mitochondries purifiées de feuilles de pois, selon le stade de développement. La concentration des substrats et des cofacteurs dans le milieu d'électrode (Cf. Matériel et méthodes) est: glycine 10 mM, NAD 0,2 mM, ADP 0,1mM. Les valeurs indiquées le long des traces, correspondent au nombre de nmoles d'oxygène consommées/mn/mg de protéines mitochondriales. M: Mitochondrie.

Contrairement à ce que l'on observe au niveau de l'oxydation de la glycine, les vitesses d'oxydation des substrats tels que le NADH, le pyruvate ou le malate sont extrêmement élevées quel que soit le stade de développement (tableau 6).

	L1	L2	L3	L4	L5
Glycine	3.2	8.4	9.5	149	201
NADH	245	263	244	258	363
Pyruvate	125	113	135	112	115
Malate	143	186	181	134	154

Tableau 6: Consommation d'oxygène au stade III par des mitochondries purifiées de feuilles de pois à différents stades de développement. La concentration des substrats et des cofacteurs dans le milieu d 'électrode (Cf. Matériel et méthodes) est:

- glycine 10 mM, NAD^+ 200 µM
- NADH 2 mM
- pyruvate 10 mM, TPP 400 µM, NAD^+ 200 µM, malate 500 µM, Coenzyme A 400 µM
- malate 10 mM, NAD^+ 200 µM.

Le stade III est déclenché en ajoutant 100 µM d'ADP au milieu réactionnel. Les valeurs indiquées correspondent à des consommations d'oxygène au stade III en nmoles O_2/min/mg de protéines.

2 - Analyse électrophorétique et détection immunologique des protéines impliquées dans l'oxydation de la glycine

Les protéines mitochondriales isolées à chacun des stades de développement ont été analysées d'une part sur gel de polyacrylamide et colorées au bleu de Coomassie puis d'autre part par Western-blot à l'aide d'anticorps polyclonaux spécifiquement dirigés contre les protéines P, T, H et la SHMT. La figure 28 montre clairement que les polypeptides spécifiques du complexe de la GDC et de la SHMT sont faiblement décelables aux très jeunes stades (L1 et L2) et commencent à apparaître à partir du stade L3, pour augmenter considérablement aux stades L4 et L5 (figure 28). En revanche, la protéine L est présente dès les premiers stades à un taux relativement élevé. Au-delà du 7ième jour de croissance, le taux de protéine L augmente également, mais pas dans les mêmes proportions que celles observées pour les autres composants de la GDC (résultat non montré).

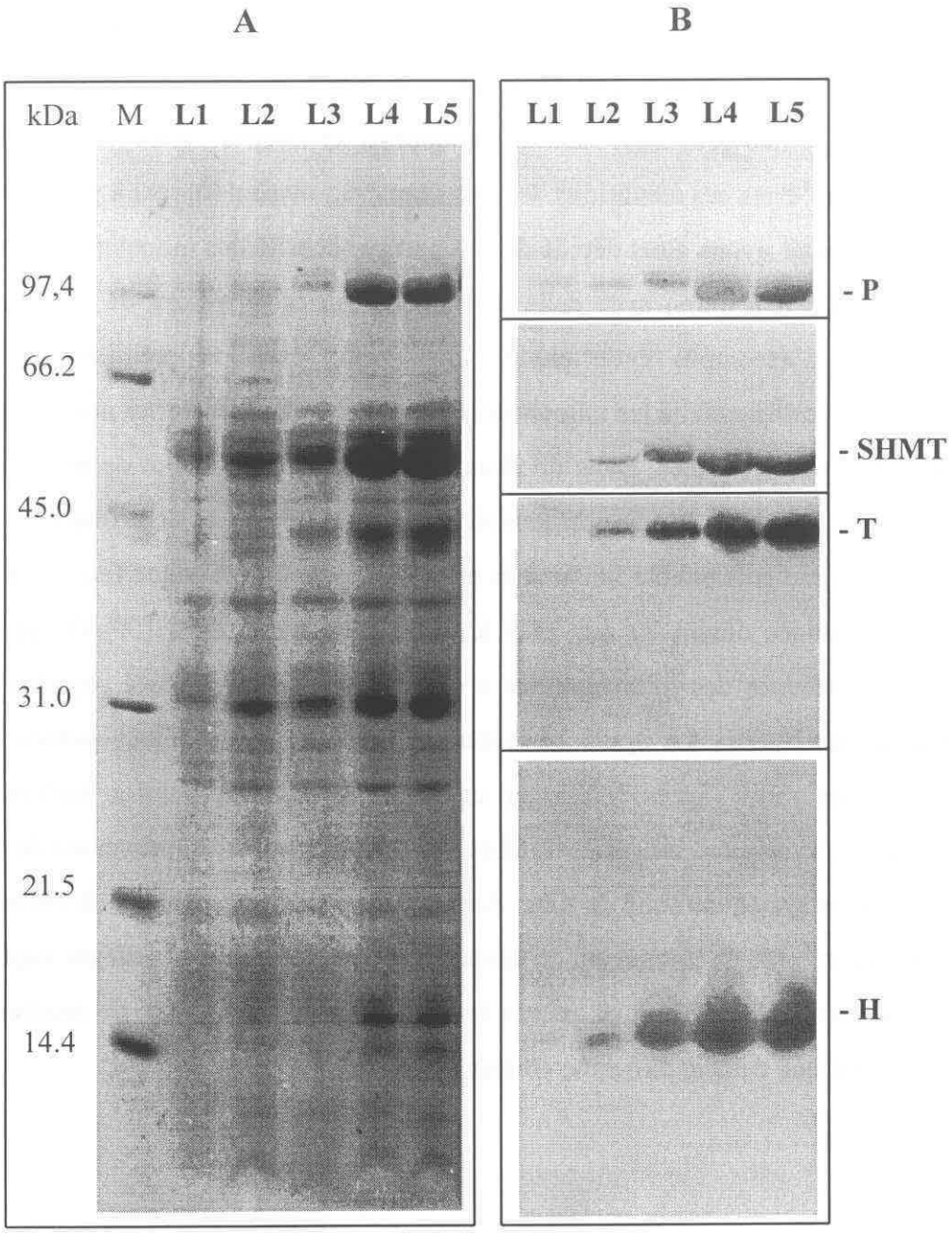


Figure 28: Séparation électrophorétique des protéines mitochondriales (A) et immunodétection des protéines T, P, L, H et SHMT à différents stades de développement (L1 à L5) (B). Les protéines mitochondriales (50 µg par puits) séparées par électrophorèse sur gel de polyacrylamide 10-15% en conditions dénaturantes sont A: soit colorées au bleu de Coomassie R-250, B: soit transférées sur une membrane de nitrocellulose. La membrane a été découpée, puis les morceaux de membrane ont été incubés en présence d'anticorps dirigés spécifiquement contre les protéines T, P, H et SHMT. M: désigne les marqueurs de masse moléculaire (15 µg) exprimés en kDa.

3 - Modification de la densité des mitochondries au cours du développement du pois

Lorsque nous avons commencé ce travail nous avons eu quelques difficultés à purifier les mitochondries de très jeunes feuilles, âgées de 4, 5 et 7 jours à l'aide du gradient de Percoll généralement utilisé pour isoler des mitochondries de feuilles vertes matures (Douce *et al.*, 1987). En effet, dans ces conditions, les mitochondries avaient du mal à pénétrer dans ce type de gradient. Nous avons alors décidé de comparer la densité des mitochondries aux stades L3 (7 jours) et L5 (12 jours) avec celle des mitochondries provenant de feuilles étiolées (12 jours). Afin d'accomplir cette étude, nous avons dans un premier temps, isolé par centrifugation différentielle les mitochondries à partir de feuilles vertes aux stades L3 et L5 et à partir de feuilles étiolées. Dans un deuxième temps, chaque culot de mitochondries a été déposé sur un gradient linéaire de Percoll (28-42%). Après une centrifugation de 50 min à 35 000g, les gradients ont été fractionnés en aliquotes de 2ml. Chaque fraction a été diluée 15 fois et centrifugée durant 20 min puis le culot a été repris dans 100-400 µl de milieu de lavage. La présence des mitochondries a été détectée à l'électrode à oxygène en mesurant l'oxydation du NADH. La figure 29 montre clairement que les mitochondries, provenant de jeunes feuilles vertes (âgées de 7 jours) et celles provenant de feuilles étiolées, sédimentent dans la partie supérieure du gradient, alors que les mitochondries, provenant de feuilles âgées de 12 jours, se situent au fond du tube. Ainsi, la densité apparente des mitochondries au stade L5, estimée à l'aide de marqueurs de densité, est de 1,096. Cette valeur est supérieure à celle des mitochondries provenant de jeunes feuilles vertes (stade L3) et de feuilles étiolées (12 jours) qui ont une densité moyenne voisine de 1,049.

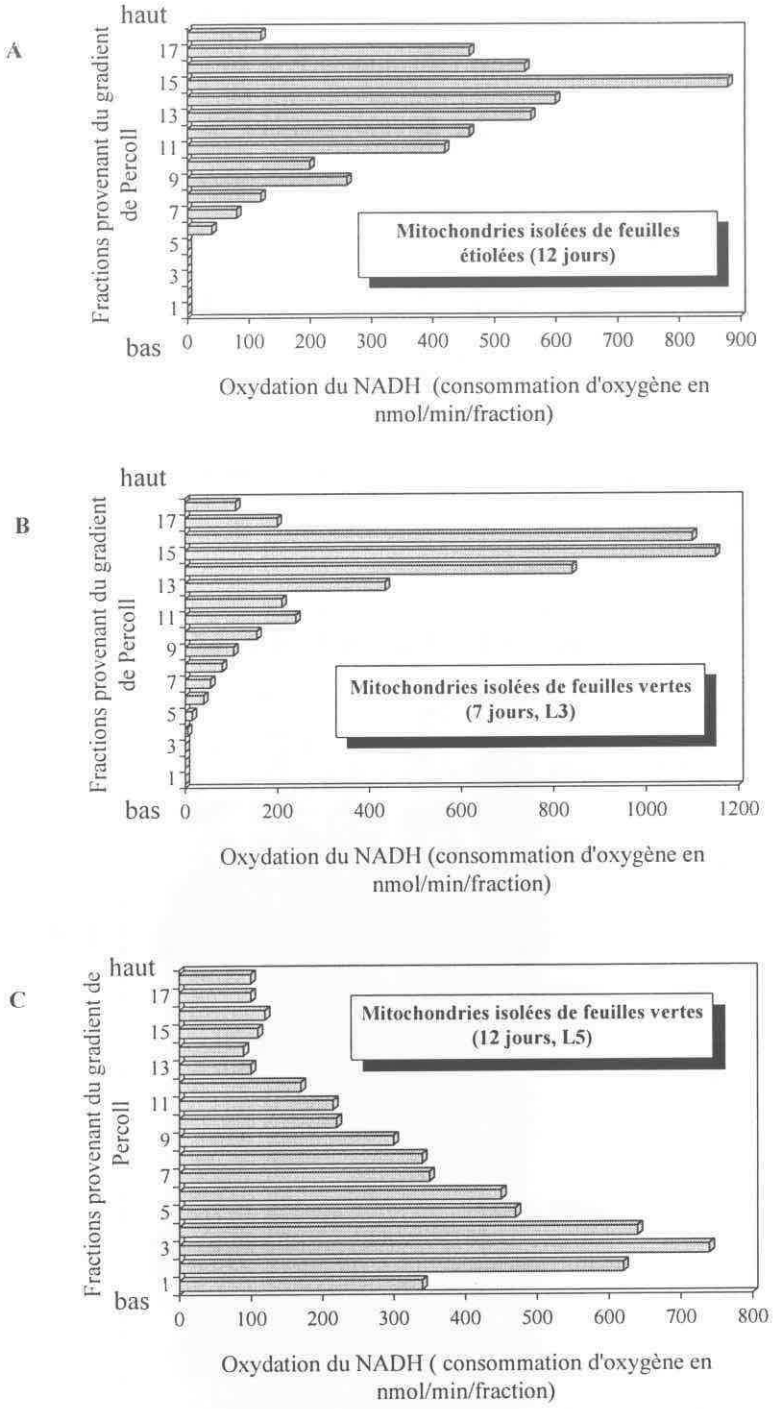


Figure 29: Localisation des mitochondries provenant de 3 types différents de feuilles [étiolées (12 j), jeunes feuilles (7 j) et feuilles matures(12 j)] le long d'un gradient de Percoll (28-42%). Chaque fraction de 2 ml a subi une centrifugation et les culots ont été remis en suspension dans 100-400 µl de milieu de lavage dont le 1/10e est testé à l'électrode à oxygène en présence de NADH 1 mM et d'ADP de 100 µM.

B - ETUDE DE LA MISE EN PLACE DE LA RIBULOSE 1,5 BISPHOSPHATE CARBOXYLASE/OXYGENASE DANS LES CHLOROPLASTES

Dans le but de comparer l'apparition des peptides de la GDC et de la SHMT avec celle de la Rubisco, nous avons isolé des chloroplastes intacts aux différents stades de développement

précédemment définis. L'activité de la Rubisco a été mesurée selon la technique décrite par Lorimer (1982) et les valeurs obtenues sont présentées dans le tableau 7. Il apparaît que les chloroplastes, isolés à tous les stades du développement de la feuille, sont capables d'incorporer le CO_2 via la Rubisco à des vitesses considérables (de 68 à 160 nmoles de CO_2/min/mg de protéines chloroplastiques).

La présence de la Rubisco, en grosse quantité, à tous les stades de développement, est confirmée par analyse des protéines chloroplastiques sur gel d'électrophorèse en conditions dénaturantes. En particulier, la figure 30 montre que les 2 sous-unités de la Rubisco (petite et grande sous-unités) sont déjà présentes dans les chloroplastes isolés dès le 4ième jour de développement de la feuille (L1).

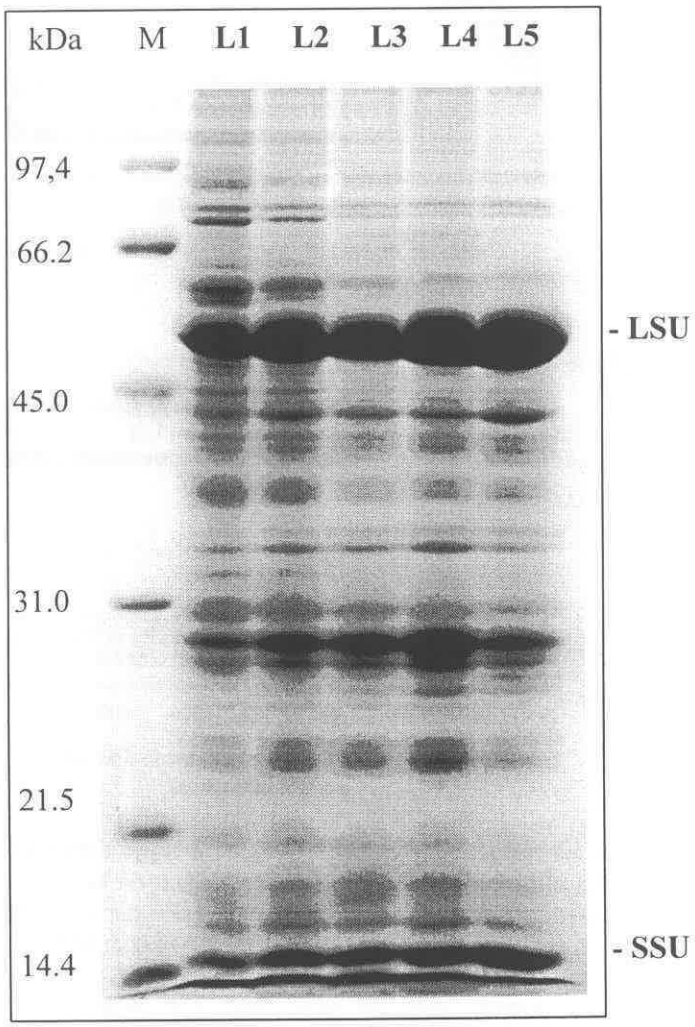


Figure 30: Séparation électrophorétique des polypeptides chloroplastiques selon le stade de développement (L1 à L5). Les protéines chloroplastiques (50 µg par puits) sont séparées par électrophorèse sur gel de polyacrylamide 10-15% en conditions dénaturantes puis colorées au bleu de Coomassie R-250. LSU et SSU; grande et petite sous-unité de la Rubisco. M désigne les marqueurs de masse moléculaire (15 µg) exprimés en kDa.

	L1	L2	L3	L4	L5
nmoles de CO_2/min/mg de protéine	68	161	162	157	156

tableau 7: Activité spécifique de la Rubisco dans les chloroplastes purifiés selon le stade de développement. 35 µg d'extrait chloroplastique prélevés à chaque stade du développement sont incubés durant 1, 2 et 3 min en présence d'un tampon contenant $NaH^{14}CO_3$ (2,5 µCi/ml) (Cf. Matériel et méthodes). L'incorporation du $^{14}CO_2$ est mesurée à l'aide d'un compteur à scintillation. Les valeurs correspondent au nombre de nmoles de CO_2 incorporées/min/mg de protéine chloroplastique.

C - EVOLUTION AU COURS DU DEVELOPPEMENT DE L'EXPRESSION DES TRANSCRITS CORRESPONDANT AUX DIFFERENTES PROTEINES DU COMPLEXE MULTIENZYMATIQUE DE LA GDC, DE LA RUBISCO ET DE LA PROTEINE cab

Le niveau des ARN messagers correspondant aux protéines du complexe de la GDC a été analysé au cours du développement de la première feuille. Les ARN totaux ont été isolés à chaque étape et analysés selon la technique de Northern blot en utilisant comme sondes les ADNc correspondant aux protéines H, T, P et L (figure 31). Comme on peut l'observer sur la figure 31, les profils d'expression des gènes codant pour les protéines H, T et P sont relativement similaires. Le niveau des transcrits correspondant à ces protéines augmente progressivement durant les premiers jours avec un pic d'accumulation au stade L3 (7^{ieme} jour). En effet, au stade L1, les signaux correspondant à la présence des ARN messagers des protéines H, T et P représentent respectivement 25, 35 et 40 % du maximum d'intensité détectée au stade L3 pour ces trois protéines. Au-delà du stade L3, le taux des transcrits diminue pour atteindre au stade L5, un niveau nettement inférieur à celui du stade L1. L'évolution des transcrits correspondant à la protéine L est tout à fait différente de ce que l'on peut observer pour les trois autres protéines. En effet, la quantité des transcrits correspondant à la protéine L est, dès les premiers jours, très élevée. Le taux d'ARN messagers au stade L1 représente 70 % du taux maximum obtenu dans ce cas au stade L2.

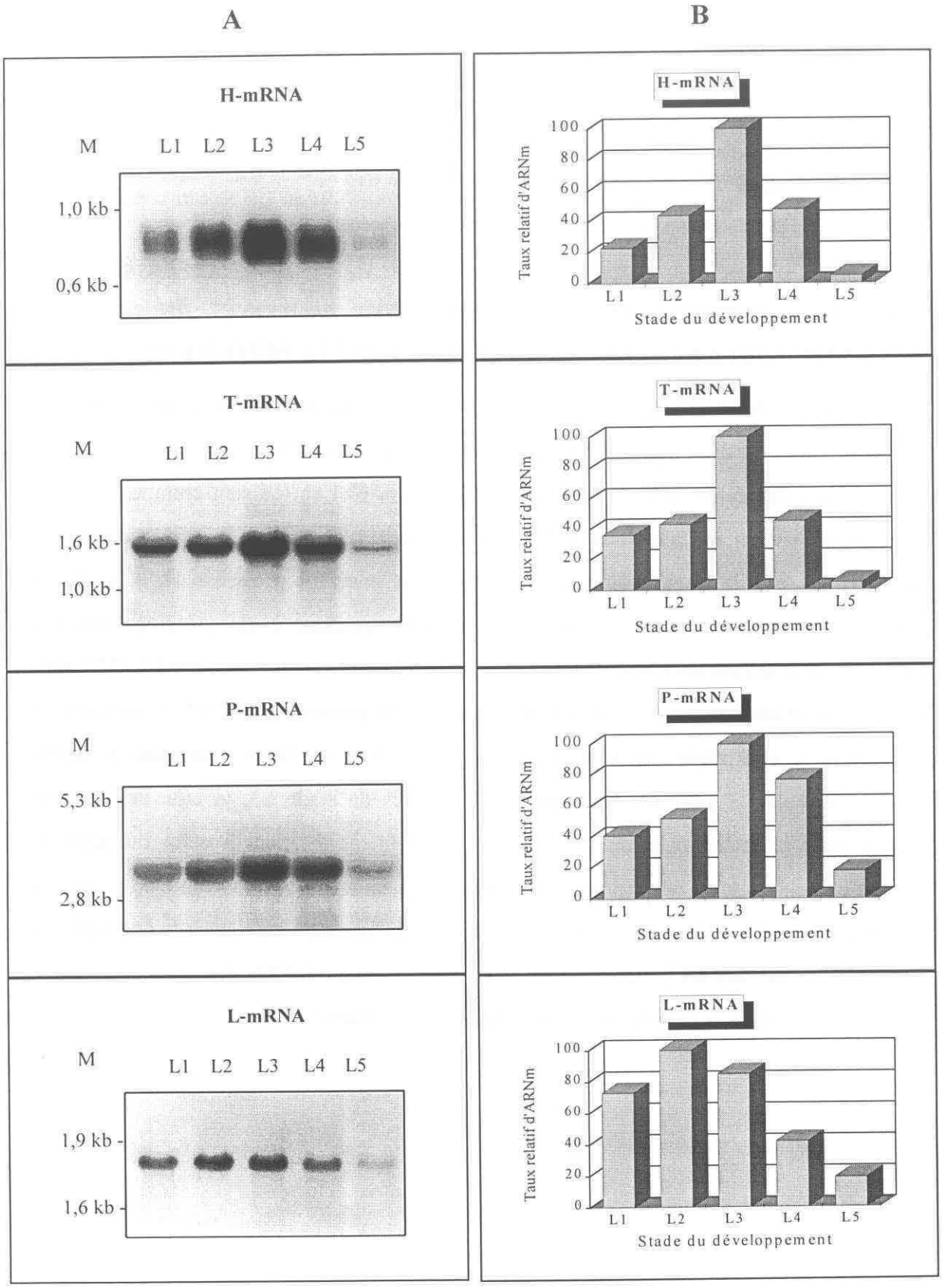

A
B
H-mRNA
M
L1 L2 L3 L4 L5
1,0 kb
0,6 kb
T-mRNA
1,6 kb
1,0 kb
P-mRNA
5,3 kb
2,8 kb
L-mRNA
1,9 kb
1,6 kb
Taux relatif d'ARNm
0
20
40
60
80
100
Stade du développement

Figure 31: Analyse du taux des transcrits correspondant aux protéines de la GDC au cours du développement par hybridation moléculaire (A). 20 µg d'ARN totaux isolés à partir de feuilles récoltées aux divers stades de développement (L1, L2, L 3, L4, L5) sont déposés par ligne puis séparés sur gel d'agarose en conditions dénaturantes et transférés sur une membrane Hybond N^+. Les ADNc correspondant aux différentes protéines de la GDC sont marqués au phosphore 32 et hybridés spécifiquement avec leurs ARN messagers correspondants. Les résultats sont révélés après exposition d'un film autoradiographique avec la membrane de nylon. M désigne en kb les marqueurs de taille. La quantité d'ARN dans chaque ligne est contrôlée, après autoradiographie, par coloration de la membrane au bleu de méthylène dans l'acide acétique. **(B)** La quantification des ARNm des différentes protéines est obtenue en scannant les membranes au Phosphorimager (Molecular Dynamics). La quantification des rayonnements β émis par l'aire de chaque bande permet d'avoir des indications précises sur les taux relatifs de transcription. Un taux relatif de 100 correspond à la valeur maximum détectée au Phosphorimager.

Enfin, nous avons également analysé le niveau des ARN messagers correspondant aux protéines de la petite sous-unité de la Rubisco et à la protéine liant les chlorophylles a et b (cab) au cours du développement. La figure 32 montre que le profil d'évolution des transcrits correspondant à ces protéines est similaire à celui de la GDC. Cependant, la rbcS semble avoir au stade L3 une induction plus élevée de ces transcrits (de l'ordre de 10 par rapport au stade L1).

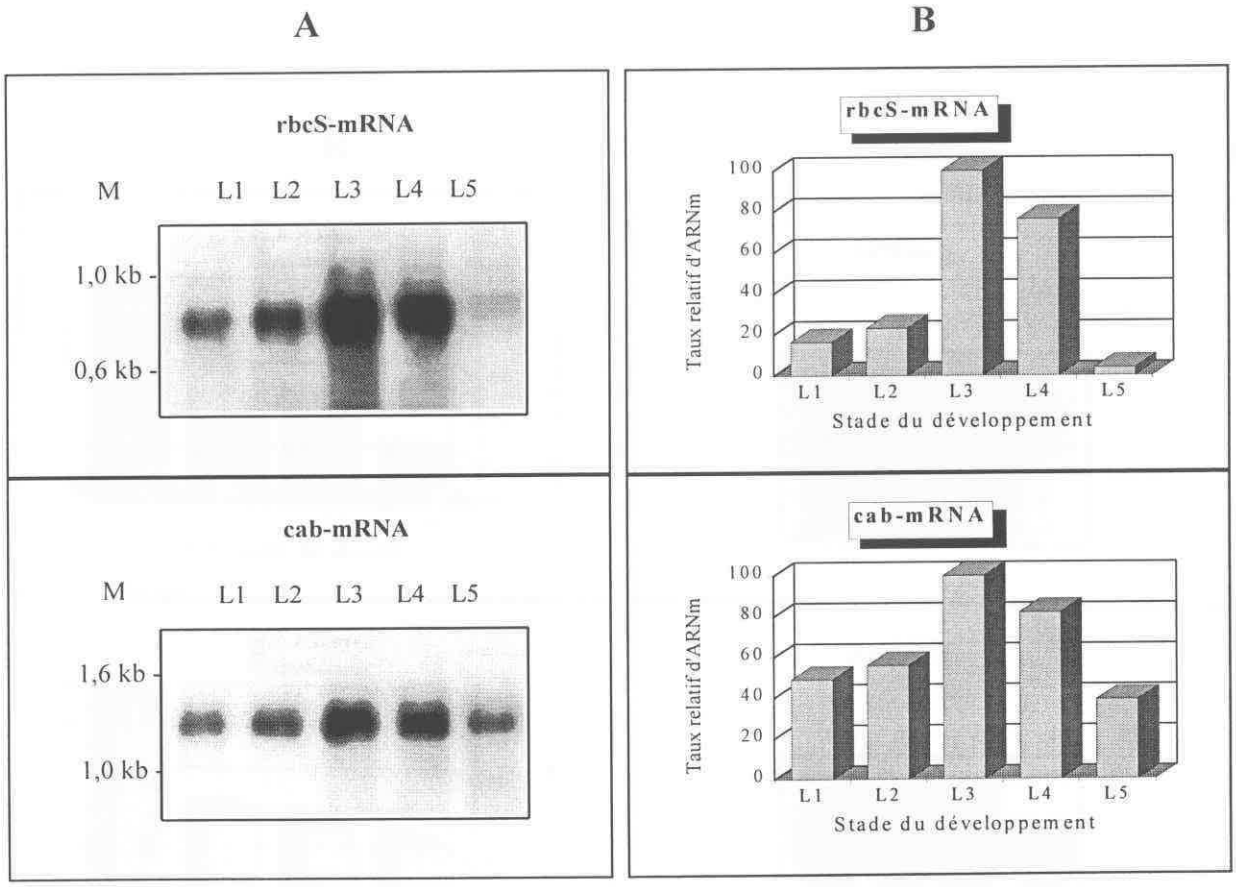
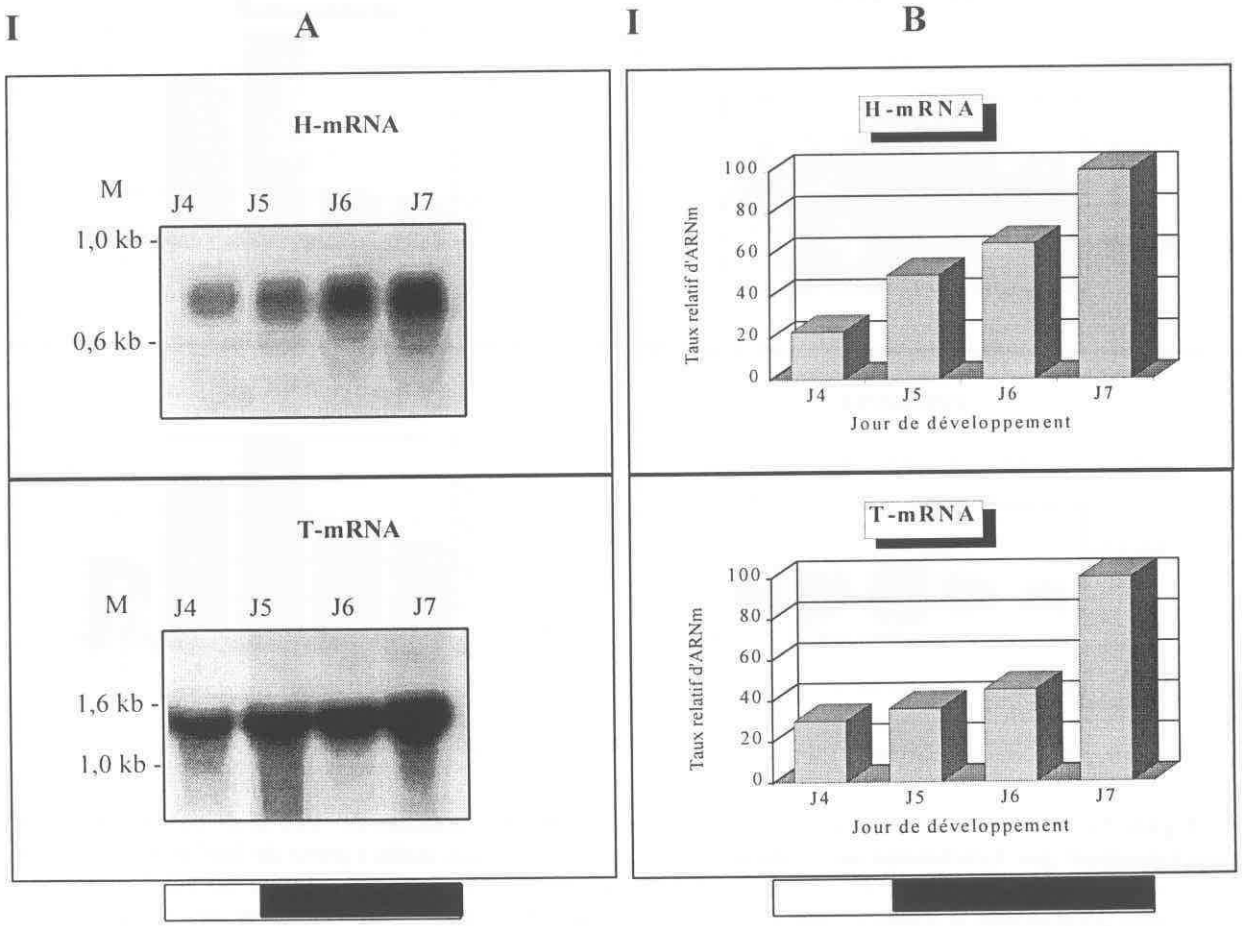


Figure 32: Analyse du taux des transcrits correspondant aux protéines de rbcS et de cab au cours du développement par hybridation moléculaire (A). 20 µg d'ARN totaux isolés à partir de feuilles récoltées aux divers stades de développement (L1, L2, L 3, L4, L5) sont déposés par ligne puis séparés sur gel d'agarose en conditions dénaturantes et transférés sur une membrane Hybond N^+. Les ADNc correspondant aux différentes

protéines de la rbcS et de cab sont marqués au phosphore 32 et hybridés spécifiquement avec leurs ARN messagers correspondants. Les résultats sont révélés après exposition d'un film autoradiographique avec la membrane de nylon. M désigne en kb les marqueurs de taille. La quantité d'ARN dans chaque ligne est contrôlée, après autoradiographie, par coloration de la membrane au bleu de méthylène dans l'acide acétique. **(B)** La quantification des ARNm des différentes protéines est obtenue en scannant les membranes au Phosphorimager (Molecular Dynamics). La quantification des rayonnements β émis par l'aire de chaque bande permet d'avoir des indications précises sur les taux relatifs de transcription. Un taux relatif de 100 correspond à la valeur maximum détectée au Phosphorimager.

D - EVOLUTION DES TRANSCRITS AU COURS DE LA CROISSANCE DU POIS APRES EXPOSITION DES FEUILLES A L'OBSCURITE A DIFFERENTS STADES DU DEVELOPPEMENT

L'expérience consiste, dans un premier temps, à faire pousser des plantules de pois dans des conditions normales (14h de lumière et 10h d'obscurité) jusqu'au stade L2 ou L3, puis dans un deuxième temps, à les placer à l'obscurité. Nous avons alors analysé et quantifié l'évolution des transcrits selon la méthode décrite précédemment. Les résultats présentés sur la figure 33 montrent que les transcrits correspondant aux protéines H et T suivent une évolution parallèle quel que soit le type d'expérience.



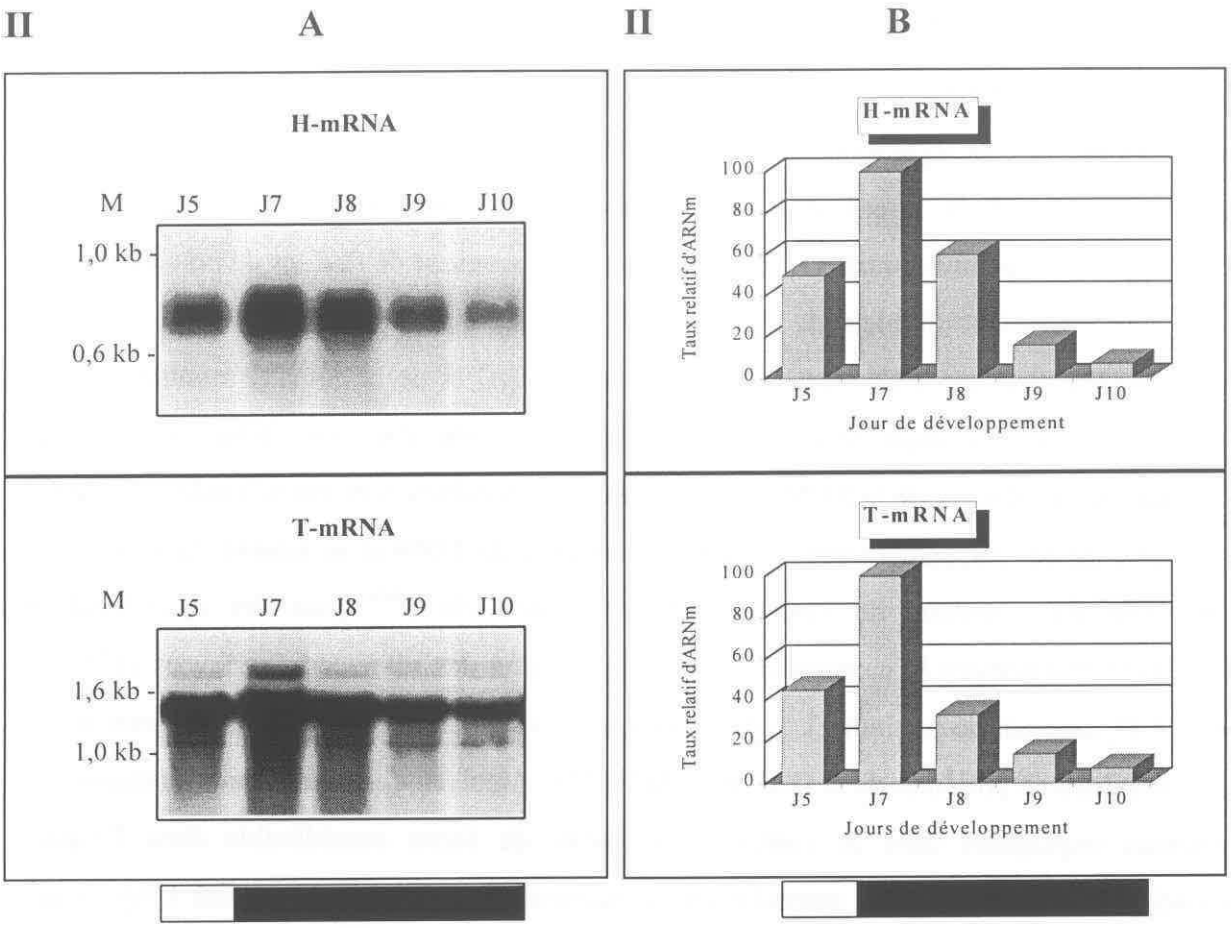


Figure 33: Analyse du taux des transcrits correspondant aux protéines T et H de la GDC au cours du développement par hybridation moléculaire après exposition à l'obscurité des feuilles de pois âgées (I) de 5 jours ou (II) de 7 jours. 20 µg d'ARN totaux isolés à partir de feuilles de pois [(I) du 4ième au 7ième jours et (II) du 7ième au 10ième jours] sont déposés par ligne et séparés sur gel d'agarose en conditions dénaturantes, puis transférés sur une membrane Hybond N^{+}. Les ADNc correspondant aux protéines T et H de la GDC sont marqués au phosphore 32 et hybridés spécifiquement avec leurs ARN messagers correspondants. **A)** Les résultats sont révélés après exposition d'un film autoradiographique avec la membrane de nylon. M désigne en kb les marqueurs de taille. La quantité d'ARN dans chaque ligne est contrôlée, après autoradiographie, par coloration de la membrane au bleu de méthylène dans l'acide acétique. **B)** La quantification des ARNm des différentes protéines est obtenue en scannant les membranes au Phosphorimager (Molecular Dynamics). La quantification des rayonnements β émis par l'aire de chaque bande permet d'avoir des indications précises sur les taux relatifs de transcription. Un taux relatif de 100 correspond à la valeur maximum détectée au Phosphorimager.

Dans le cas où les pois sont mis à l'obscurité au stade L3, nous remarquons que le taux des transcrits diminue rapidement pour atteindre, au bout de 72 heures d'obscurité, 5% du maximum détecté au début de l'expérience. Par contre, lorsque les plantules de pois sont placées à l'obscurité à partir du stade L2, nous remarquons qu'après 48 heures d'obscurité, les niveaux des transcrits continuent d'augmenter. Ce profil est similaire à celui observé chez les plantules cultivées en présence de lumière, notamment nous retrouvons une accumulation massive des ARN messagers au 7ième jour de croissance.

E - DISCUSSION-CONCLUSION

Dans le but d'obtenir des informations sur la biogenèse du complexe de la GDC et de mieux comprendre sa régulation, nous avons entrepris d'étudier l'expression des gènes codant pour les protéines P, H , T et L de la GDC durant le développement de la première feuille de pois. Ainsi, nous avons montré que les mitochondries provenant de feuilles de pois âgées de 4 à 7 jours oxydent faiblement la glycine. L'analyse sur un gel d'acrylamide et l'immunodétection des différents constituants de la GDC et de la SHMT, montrent que la faible activité d'oxydation de la glycine est en corrélation avec une faible synthèse des protéines de la GDC et de la SHMT au sein des mitochondries. Les mitochondries isolées à ces stades de développement ont une densité moyenne de 1,049 et se situent dans la moitié supérieure d'un gradient de Percoll 28-42%. Au-delà du 7ième jour, les mitochondries acquièrent brutalement la capacité d'oxyder la glycine à des vitesses très élevées (138-232 nmol d'O_2 consommées/min/mg). A ces stades, on assiste véritablement à un chargement des mitochondries en protéines de la GDC et de la SHMT. Ainsi, en quelques heures, la masse des protéines impliquées dans ce complexe augmente de façon considérable dans l'espace matriciel mitochondrial pour atteindre une concentration de 0,15 à 0,20 g/ml (Day *et al.*, 1985; Bourguignon *et al.*, 1988). Cette surabondance des protéines de la GDC et de la SHMT chez les mitochondries de feuilles matures semble provoquer une augmentation du rapport protéines/lipides, ce qui leur confère une densité plus élevée. Ceci se traduit par une augmentation de la densité des mitochondries qui atteint 1,096. On peut constater que pour des feuilles matures (stades supérieurs à L3) l'ensemble de la population de mitochondries a vu sa densité augmenter, ce qui indique que le chargement en protéines de la GDC s'effectue dans les mitochondries existantes, et qu'il n'y a pas apparition d'une nouvelle population de mitochondries spécifiques de la photorespiration.

Notons que les mitochondries isolées entre le 7ième et le 9ième jour présentent une grande capacité à importer les protéines mitochondriales. A ce titre, elles pourraient être un outil tout à fait remarquable pour étudier les phénomènes de translocation des protéines à travers la membrane mitochondriale.

L'analyse du niveau des ARN messagers correspondant aux protéines de la GDC au cours du développement, montre que les gènes s'expriment dès le 4ième jour avec un pic au niveau du 7ième jour de développement. Durant cette période, les ARN messagers sont présents dans la feuille en quantité importante, alors que les protéines de la GDC sont faiblement

représentées au sein des mitochondries. Il semblerait donc que les transcrits ne soient pas traduits. De telles observations suggèrent fortement qu'un contrôle de type post-transcriptionnel puisse avoir lieu au cours de la biogénèse de ce complexe. Des études récentes ont montré que le contrôle de l'initiation de la traduction peut être modulé soit par la présence de ribonucléoprotéines qui, en séquestrant les transcrits (mRNP) empêcheraient toute initiation de la traduction soit par modification de la longueur de la queue polyadénylée (pour une revue voir Curtis *et al.*, 1995). Bien entendu, des expériences supplémentaires sont nécessaires pour confirmer l'existence de telles régulations. Il apparaît clairement que la période séparant le stade L3 de L4 soit le temps nécessaire pour permettre la traduction et la mise en place de la majorité des protéines de la GDC au sein de la matrice mitochondriale. Au-delà du 7ième jour, le taux des transcrits correspondant aux protéines T, P et H diminue pour quasiment disparaître au 12ième jour.

Nous constatons de nouveau que la protéine L semble avoir un comportement différent des autres protéines; la protéine L ainsi que ses transcrits étant nettement représentés dès les premiers stades du développement. Ceci s'explique par le fait que la protéine L est associée au complexe de la pyruvate deshydrogénase présent dans les mitochondries dès les premiers stades de développement. En effet, les mitochondries isolées à ces stades ont la capacité d'oxyder rapidement le pyruvate.

Contrairement à ce que l'on observe au niveau de la mise en place de la GDC dans les mitochondries, les chloroplastes isolés dès les tous premiers stades de développement, renferment en quantité importante la Rubisco. Ainsi, bien que la Rubisco et les protéines impliquées dans l'oxydation de la glycine interviennent ensemble dans l'acte photorespiratoire, leur biogénèse ne semble pas synchronisée. En effet, alors que la Rubisco s'accumule et fonctionne dès les tous premiers stades du développement de la feuille, il faut attendre le stade L4, stade où les feuilles se comportent comme de véritables capteurs solaires, pour que les mitochondries soient capables d'oxyder la glycine.

Par ailleurs, nous remarquons que les profils d'évolutions des transcrits correspondant aux protéines rbcS et cab sont quasiment identiques à ceux des transcrits correspondant aux protéines T, P et H de la GDC. Cette observation nous amène à penser que, comme le suggéraient Turner *et al.* (1993), des mécanismes de régulation similaires pourraient exister au niveau de la transcription des gènes codant pour la GDC et la Rubisco. Ces observations seront discutées de manière plus approfondie dans le chapitre III. Cependant, contrairement

aux protéines P, H et T de la GDC aucun décalage ne semble exister entre la présence des transcrits et celle des protéines de la rbcS. En effet, les transcrits semblent être traduits sans délai. Toutefois, les valeurs présentées dans le tableau 7 montrent que l'activité spécifique de la Rubisco semble être à son maximum dès le stade L2 alors que le taux des transcrits correspondant à rbcS continue à croître jusqu'au stade L3. L'interprétation que l'on peut donner, afin d'expliquer les profils d'évolutions des transcrits correspondant à rbcS et cab au cours du développement, sont les suivantes. Des expériences réalisées au laboratoire par Nicolas Diallo ont montré que l'intensité photosynthétique mesurée sur des protoplastes provenant des stades L1 à L5 augmente au cours du développement de la première feuille. Cependant, si l'on ramène l'évolution de la photosynthèse à celle de la quantité de chlorophylle par protoplaste, on constate que l'intensité photosynthétique est pratiquement constante du stade L2 à L5. En fait, l'augmentation de la photosynthèse semble être étroitement liée à l'augmentation de la masse de chlorophylle, c'est-à-dire au processus de maturation des chloroplastes au sein de la feuille au cours du développement. Ainsi, il semble que l'évolution du taux des ARN messagers correspondant à la Rubisco et à la protéine cab suit l'évolution et la maturation de l'appareil photosynthétique au sein de la feuille. De tels résultats sont parfaitement en accord avec ceux obtenus par Dubell et Mullet, (1995). En effet, ces auteurs ont montré que le niveau de transcription des chloroplastes augmentait fortement du 4ième au 6ième jour de croissance de la plantule de pois. Cette activité transcriptionnelle est suivie en parallèle par une accumulation jusqu'au 7ième jour des transcrits correspondant à différentes protéines impliquées dans l'appareil photosynthétique pour diminuer au-delà du 7ième jour et être quasiment nulle au 14ième jour de croissance. Ces résultats suggèrent donc qu'il existe durant les premiers jours du développement une augmentation commune du niveau de transcription des gènes codant pour les protéines chloroplastiques.

De même, ce manque de corrélation directe entre le taux de transcription et l'activité d'oxydation de la glycine ne semble pas être observé avec des expériences classiques de transfert à la lumière de plantes étiolées (Turner *et al.*, 1992; 1993). En effet, ces auteurs ont montré que l'exposition à la lumière de pois étiolés âgés de 6 jours induisait de façon synchrone les protéines de la GDC et leurs transcrits correspondant. Il semblerait donc que dans ces conditions, la lumière soit le seul facteur mis en jeu. Cependant, il est fort probable qu'à l'effet de la lumière, il faille ajouter la différence d'organisation cellulaire entre les tissus étiolés et les tissus chlorophylliens (Macherel *et al.*, 1990). De même, le stade auquel les pois

étiolés ont été exposés à la lumière pourrait avoir une importance. En d'autres termes, la lumière, bien que jouant un rôle primordial, ne serait pas la seule responsable de la mise en place du complexe de la GDC. Ces observations ont pu être confirmées par des expériences de transfert à l'obscurité de plantules de pois âgées de 5 ou 7 jours. Ainsi, 2 comportements différents ont pu être observés dans ces conditions. Dans le cas où les pois sont mis à l'obscurité à partir du 7ième jour de croissance, nous observons une décroissance rapide du taux des messagers correspondant aux protéines T et H. Ces résultats sont parfaitement en accord avec ceux observés par Turner *et al.*, (1992) sur des plantes matures. Cependant, dans le cas où de jeunes plantules de pois sont cultivées à l'obscurité à partir du 5ième jour de développement, nous constatons, contrairement aux observations effectuées par Turner *et al.* (1992), que les transcrits continuent à s'accumuler jusqu'au 7ième jour de croissance de façon analogue à ceux des plantes cultivées à la lumière. Ces résultats suggèrent donc que la lumière n'est pas la seule responsable de l'évolution du taux de transcription des protéines T et H, mais qu'il existe également des phénomènes liés au développement impliqués dans la régulation des messagers de la GDC.

Notre étude sur la mise en place du complexe de la GDC au cours du développement de la première feuille de plantule de pois, a permis de confirmer que la biogénèse de la GDC est en étroite relation avec celle de l'appareil photosynthétique. Cependant, même si au cours du développement, le niveau de transcription des gènes codant pour les protéines P, H et T de la GDC, suit celui des gènes codant pour rbcS et cab, nous avons pu montrer qu'il existe un temps de latence entre la présence, dès le 4ième jour, d'une Rubisco active dans les chloroplastes et l'apparition des protéines de la GDC au sein de la matrice mitochondriale (7ième jour). Ce délai indique donc l'existence d'un contrôle posttranscriptionnel qui se lève lorsque la feuille est pleinement ouverte. L'interprétation physiologique que l'on peut apporter à ce phénomène est que la photorespiration semble apparaître uniquement lorsque la photosynthèse a atteint un niveau d'activité qui nécessite la présence de fortes concentrations d'enzymes de la photorespiration, afin de recycler les molécules en C2 qui affluent. Néanmoins, les fortes similitudes de comportement entre les gènes codant pour les protéines P, H et T de la GDC et ceux codant pour rbcS et cab, suggèrent l'existence d'un mode de régulation transcriptionnel commun. Ceci nous a amenés à envisager le clonage du gène codant pour la protéine T, afin de déceler d'éventuels éléments de régulation.

CHAPITRE III

CHAPITRE III
CLONAGE ET CARACTERISATION DU GENE NUCLEAIRE CODANT POUR LA PROTEINE T

Afin de poursuivre l'étude de la régulation des gènes codant pour les protéines du complexe mitochondrial de la GDC, nous avons décidé d'isoler et de caractériser le gène nucléaire codant pour la protéine T du complexe de la GDC. Comme nous l'avons montré au cours des travaux précédents, il semble que l'expression des protéines T, P et H du complexe s'effectue de manière coordonnée. Notamment, au même titre que certaines protéines impliquées dans la photosynthése telles que la petite sous-unité de la Rubisco (rbcS) et la protéine liant les chlorophylles a/b dans le photosystème II (cab), ces protéines sont principalement exprimées dans les tissus foliaires exposés à la lumière. Nous avons donc envisagé d'isoler le gène nucléaire codant pour la protéine T et de caractériser sa région promotrice dans le but de la comparer avec celle du gène codant pour la protéine H (Macherel *et al.*, 1992), mais également avec celles des gènes rbcS et cab, ceci afin de mettre en évidence des éléments de régulation communs.

I - ISOLEMENT ET IDENTIFICATION DU GENE CODANT POUR LA PROTEINE T

A - DETECTION PAR SOUTHERN-BLOT DU GENE CODANT POUR LA PROTEINE T DANS LE GENOME DE POIS

L'utilisation d'ADN nucléaire de pois, préalablement digéré avec différentes enzymes de restriction, séparé sur gel d'agarose 0,8%, transféré sur membrane puis hybridé avec l'ADNc correspondant à la protéine T, nous a permis d'estimer le nombre de gènes codant pour cette protéine.

Nous avons utilisé deux sondes spécifiques de tailles différentes: la première de 1,5 kpb correspond à la longueur totale du transcrit codant pour la protéine T (TB2), la seconde de 413 pb (TD3') correspond à la partie 3' de l'ADNc incluant principalement la région 3' non codante (Cf. chapitre I).

Les sondes, préalablement marquées au phosphore 32, sont hybridées avec de l'ADN nucléaire de pois digéré entièrement par différentes endonucléases de restriction. L'autoradiographie obtenue à l'aide de l'ADNc TD3' comme sonde (figure 34) montre que pour toutes les digestion testées (*Xba*I, *Dra*I, *Ssp*I et *Cla*I), un seul fragment de restriction est marqué. Ces résultats suggèrent que le gène de la protéine T est unique. Cependant, comme l'ADNc TD3' ne représente qu'une portion du transcrit correspondant à la protéine T, nous avons donc confirmé ce résultat par hybridation d'une sonde (TB2) correspondant à l'ADNc entier avec de l'ADN nucléaire digéré par les enzymes *Eco*RI, *Hind*III, *Kpn*I et *Sac*I.

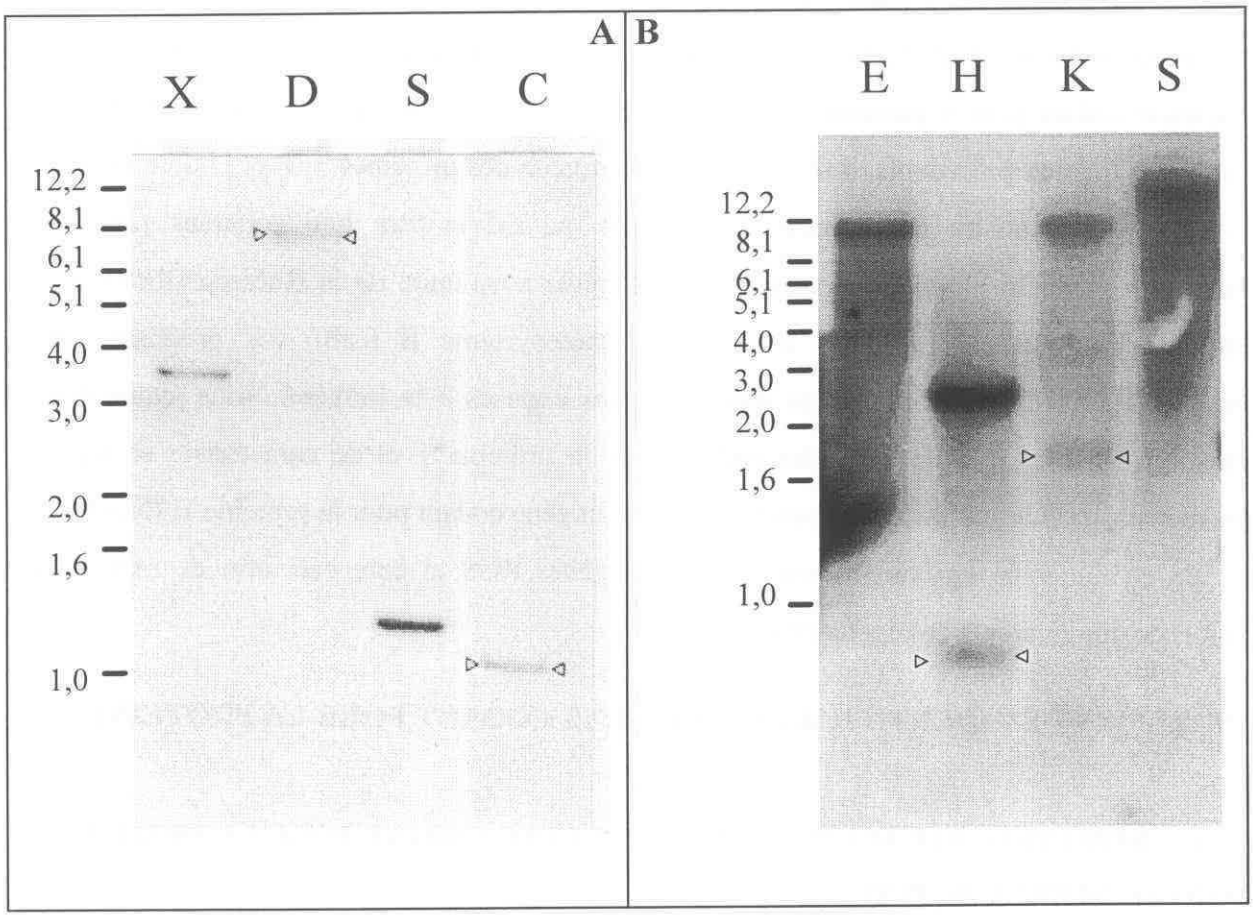


figure 34: Analyse de l'ADN génomique selon la technique de Southern. 10 µg d'ADN nucléaire de pois ont été digérés avec diverses endonucléases de restriction: **(A)** XbaI (X), DraI (D), SspI (S) et ClaI (C); **(B)** EcoRI (E), HindIII (H), KpnI (K), et SacI (S). Les produits de digestion sont séparés sur un gel d'agarose 0,8%, transférés sur une membrane de nylon et hybridés respectivement avec la sonde radioactive TD3 **(A)**, et TB2 **(B)**.

Une seule bande marquée est observée lorsque la digestion est effectuée à l'aide des enzymes de restriction *Eco*RI et *Sac*I. Les deux bandes marquées provenant de la digestion avec les endonucléases *Hind*III et *Kpn*I, s'expliquent par la présence de ces sites de restrictions au sein de l'ADNc. On peut donc conclure que la protéine T est

vraisemblablement codée par un gène unique dans le génome du pois au même titre que les gènes codant pour les protéines H et L (Macherel *et al.* 1992 et Bourguignon *et al.*, 1992). Par contre la protéine P semble être codée par une famille de gènes (Turner *et al.*, 1992). Les chromosomes portant les gènes de la GDC ont été localisés par Turner *et al.* (1993). Le chromosome 7 possèderait les gènes codant pour les protéines T et L; le chromosome 1 porterait la famille de gènes codant pour la protéine P alors que le gène codant pour la protéine H se situerait sur le chromosome T1/4. Enfin, des analyses similaires effectuées chez l'homme (Hayasaka *et al.*, 1991) et chez le boeuf (Okamura-Ikeda *et al.*, 1992) ont montré que la protéine T est également codée par un gène nucléaire unique.

B - CRIBLAGE DE LA BANQUE GENOMIQUE DANS λEMBL3 A L'AIDE DE L'ADNc TB2: PURIFICATION ET AMPLIFICATION

Le gène codant pour la protéine T impliquée dans le complexe de la GDC, a été isolé à partir d'une banque génomique de pois (*Pisum sativum*, var. Alaska) fabriquée par les laboratoires Clontech (Palo Alto, CA). Cette librairie est constituée de fragments d'ADN nucléaires de 8 à 21 kpb (provenant de la digestion avec l'enzyme de restriction *Mbo*I), clonés au niveau du site *Bam*HI dans un vecteur du bactériophage λ EMBL3. Approximativement, 80 000 phages recombinants ont été criblés à l'aide de l'ADNc correspondant à la protéine T de pois (*Pisum sativum* var. Douce provence). Ainsi, un clone positif nommé GT a pu être isolé et purifié après trois cycles successifs de criblage.

1 - Caractérisation du gène codant pour la protéine T (*gdcT*)

L'ADN du phage recombinant a été isolé, digéré par l'enzyme de restriction *Sal*I, puis analysé sur un gel d'agarose [la perte du site BamHI a eu lieu lors de la ligation du vecteur EMBL3 (digéré *Bam*HI) avec les fragments d'ADN nucléaires (digérés *Mbo*I)]. Trois produits de digestion de tailles approximatives 9,2 kpb, 10,3 kpb et 19,3 kpb ont été révélés. L'hybridation de ces trois fragments d'ADN avec l'ADNc TB2 marqué au phosphore 32 (α-P^{32}), indique que l'insert d'une taille de 10,3 kpb renferme le gène *gdcT*. Les deux autres fragments d'ADN correspondent aux « bras » droits et gauches du vecteur du bactériophage EMBL3. Ce fragment a été sous-cloné dans le vecteur pUC 18 (pGT) puis de nouveau digéré par l'enzyme de restriction *Sal*I et hybridé avec la sonde TB2. Le résultat obtenu confirme le

sous-clonage du fragment ainsi caractérisé. Par la suite, l'utilisation de différentes enzymes de restriction a permis de dresser la carte de restriction du clone.

2 - Cartographie du clone

La simple digestion de l'insert (avec ou sans son plasmide pUC 18) par diverses enzymes de restriction (*Bam*HI, *Eco*RI, *Kpn*I, *Pst*I et *Hin*dIII) et l'analyse des fragments de restriction sur gel d'agarose permettent de dresser la carte de restriction de l'insert (figure 35). De plus, l'hybridation de ces fragments d'ADN avec l'ADNc TB2 permet la localisation précise du gène au sein de ces fragments (figure 35 et tableau 8). En effet, le gène codant pour la protéine T est présent dans la région 3' de l'insert. La présence des deux sites de restriction *Kpn*I et *Pst*I, (distants de 100 pb) situés dans la région 3' de l'ADNc, permet de déterminer l'orientation du gène. Celui-ci, se situe approximativement à 1 kb du site de sous-clonage *Sal*I en 3', la région 5'promotrice du gène étant dans la région la plus longue de l'insert (6,5 kpb). Enfin, la comparaison des distances, entre les sites de restriction de l'ADNc et ceux des inserts hybridés, permet d'estimer la taille de la région codante du gène à 2,7 kpb. Le gène possèderait 1000 pb d'introns. On peut conclure que l'insert pGT renferme la totalité du gène *gdcT* avec sa région promotrice. Cependant, on remarque qu'un site de restriction *Hind*III présent dans l'ADNc a disparu au sein du gène (figure 36). Ce phénomène, confirmé lors du séquençage du gène *gdcT*, est probablement dû à la variété du pois et sera expliqué plus en détail dans le paragraphe II. A. 2.

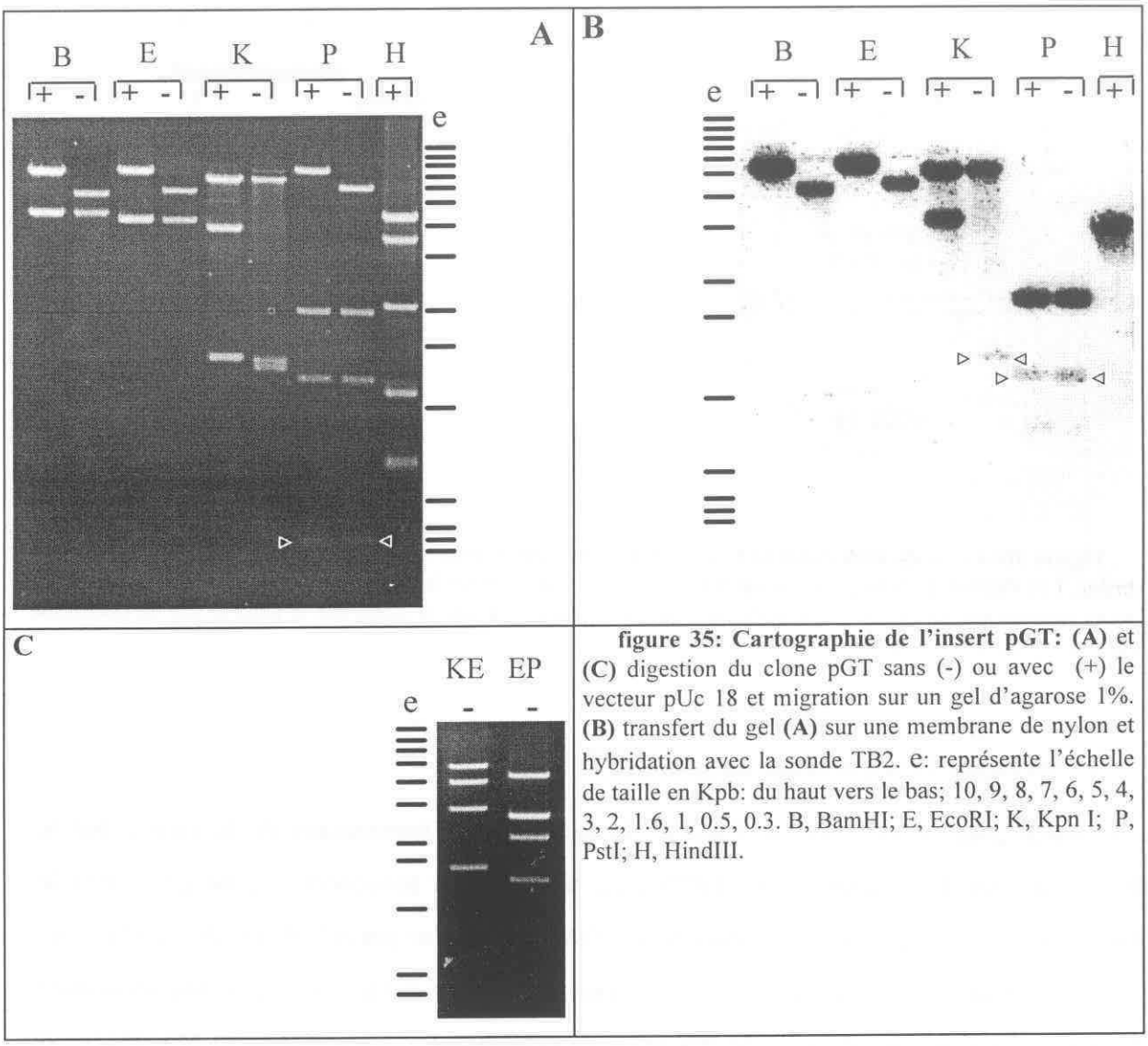


figure 35: Cartographie de l'insert pGT: (A) et **(C)** digestion du clone pGT sans (-) ou avec (+) le vecteur pUc 18 et migration sur un gel d'agarose 1%. **(B)** transfert du gel **(A)** sur une membrane de nylon et hybridation avec la sonde TB2. e: représente l'échelle de taille en Kpb: du haut vers le bas; 10, 9, 8, 7, 6, 5, 4, 3, 2, 1.6, 1, 0.5, 0.3. B, BamHI; E, EcoRI; K, Kpn I; P, PstI; H, HindIII.

ENZYME DE RESTRICTION		*BamHI*	*Eco*RI	*Kpn*I	*Pst*I	*Hind*III	*Kpn*I/*Eco*RI	*Eco*RI/*Pst*I
Longueur en Kpb des fragments d'ADN obtenus	avec plasmide +	8,7* 4,3	8,8* 4,2	7,4* 4,1* 1,5	9,2 2,1* 1,3* 0,4	4,9 3,9* 2,2 1,2 0,8		
	sans plasmide -	6,0* 4,3	6,1* 4,2	7,4* 1,4* 1,5	6,5 2,1* 1,3* 0,4		4,7 4,1 2,7 1,5	4,2 2,5 2,1 1,3 0,3

Tableau 8: Longueur (Kpb) des fragments d'ADN digérés et hybridés (*) avec la sonde TB2.

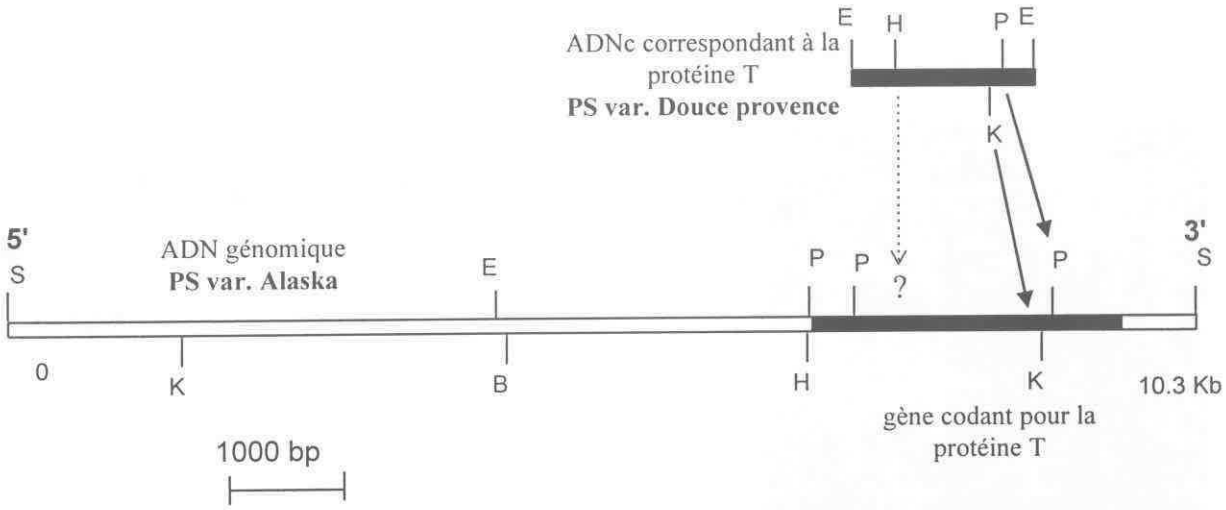


Figure 36: Carte de restriction de l'insert GT contenant le gène codant pour la protéine T de pois var. Alaska. Les flèches en trait plein indiquent la correspondance entre les sites de restriction de l'ADNc et de l'ADN génomique. La flèche en pointillés indique la région où devrait se situer sur le gène le site de restriction Hind III. PS: *Pisum sativum*.

3 - Stratégie de séquençage du gène *gdcT*

Comme pour la séquence de l'ADNc, la technique de terminaison de la chaîne par les didéoxynucléotides (Sanger *et al.*, 1977) a été utilisée pour séquencer le gène *gdcT* dans les deux sens. La stratégie choisie consiste en l'utilisation d'oligonucléotides de synthèse qui servent d'amorces pour la réaction d'élongation. Les exons contenus dans la région codante du gène ont été séquencés à l'aide des amorces synthétiques utilisées lors du séquençage de l'ADNc. Les introns et la région promotrice ont été séquencés de proche en proche. La séquence du produit d'élongation d'une amorce, sert de support pour la synthèse d'un oligonucléotide complémentaire qui permettra de poursuivre le séquençage initié par l'amorce précédente (figure 37).

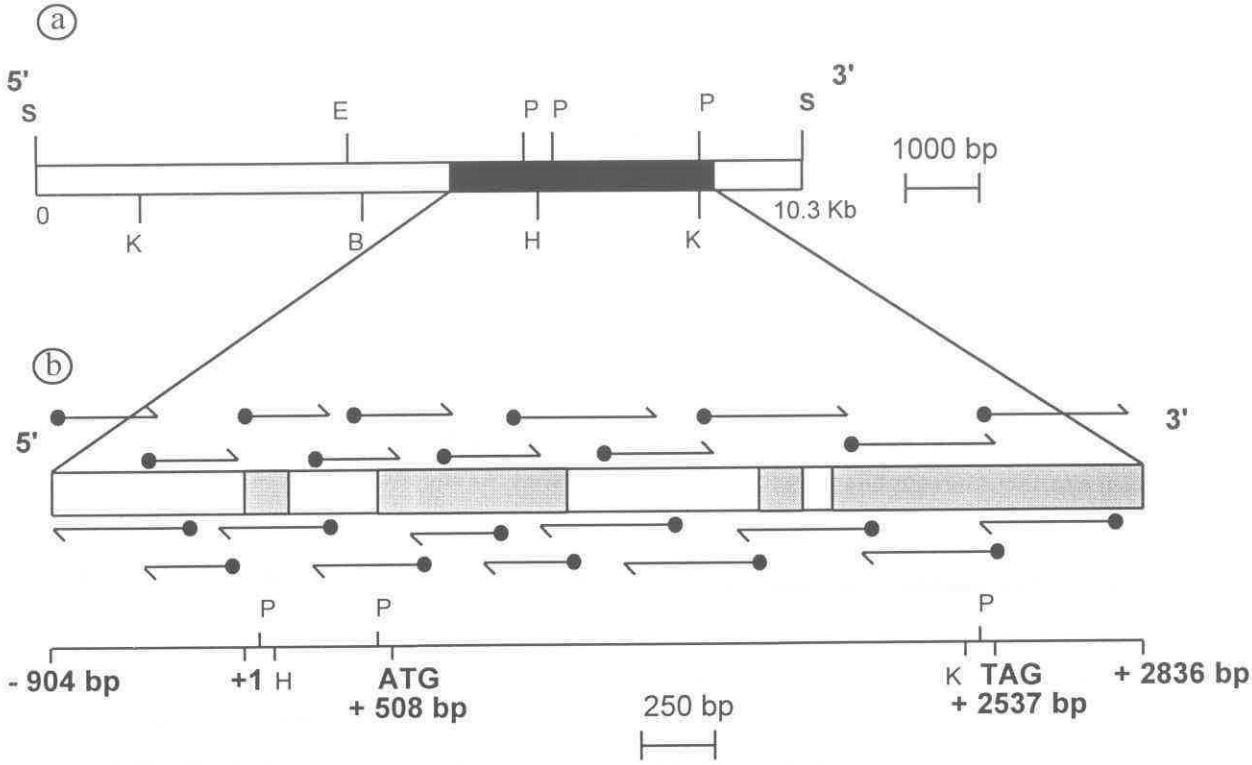


Figure 37: Stratégie de séquençage de l'insert GT. a: Site de restrictions de l'insert GT contenant le gène codant pour la protéine T (en noir). Les sites de restrictions indiqués sont B, BamHI; E, EcoRI; K, KpnI; P, PstI et S, SalI. b: Le gène codant pour la protéine T de pois var. Alaska a été séquencé sur les deux brins à l'aide d'oligonucléotides de synthèse. La numérotation positive commence à +1 qui correspond au premier démarrage de transcription déterminé par extension d'amorce. Les exons sont représentés en gris.

II - ANALYSE DE LA SRUCTURE DU GENE CODANT POUR LA PROTEINE T

A - ORGANISATION DU GENE gdcT

1 - Détermination de la séquence nucléotidique

La séquence complète du gène codant pour la protéine T impliquée dans le complexe de la GDC mitochondriale, est présentée sur la figure 38. La comparaison entre la longueur complète de l'ADNc correspondant à la protéine T et la séquence du gène révèle la présence de trois introns arbitrairement nommés introns 1, 2 et 3 de taille très variable. Le premier intron d'une longueur de 321 pb est localisé dans la région 5' non codante de l'ARN messager correspondant à la protéine T, alors que les deux autres de taille respective 691 pb et 114 pb se situent dans la partie codante.

- 905attttgcataattgaaacgaatgctgagagaagtggacgaagggaaggagagagaagagagagactaaaat
- 835 gaaagaaagaacttatcttttagttatcttatagagtgaagaaataaaatgtagttatgaaggagccaag
- 765 gaaagaatgagacataatgatgatttaaggagattgagataagatgatggacgagattctggtgtggaat
- 695 agatgatcttgcacgtggaggaaagctcctgtagtaaaaaaatttgatttgaccaagaatgatgaaaaat
- 625 tagcgtgaagaagagatatatgattagaatttaattcttctaataccatgttaaatttagtctattggac
- 555 accttaaccaatatataatttttataacaaaaaaagaagcaaaaaacaattaaaattgaaataatttttt
- 484 cttcttattattatatttactacttagtgggcttatttatatatccatgttgttaacaaaaagcccacga
- 414 tatggatagattgttacaacaaacaccatagagaagtagcaattaattaaaattaaagaatacactaata
- 344 atgatgtgtaacaaatgcattacatagcacttcatatctacttcataagaaaaaaaaattattttttaaa
- 274 tttattaaagtttattaaataaacttggttggtggagaaaagaaaaagaaaaaaaagaatgggaaagagg
- 204 taatggtaatgttgctaccactcactaaggatggggtatcataggctctcttcatttggggcactcactc
- 134 aatcaagatgaggttaagttgttctcaatctcacatccaccaactccctctttttatattttttatttaa

+1 ⌐▷

- 64 attaaataaatatattaatatattatttttatttattttattcatttatgtccaacaaattcatTTGATA

+31 ⌐▷

7 AACGAGGTTCTCTCTTCTCGTTTCTTGTCTCTCGCAGTGTTTCATAAAAACAAAAACAAGTTACCTAAGC
77 TTCTCCATTGTTTTGTTCCTGCAGATGCGTACTTGATTTTCCTCATTCACTCCGAACTCAATCAACAAAC
147 CAATCTCTgtcagttctctatccttttcattactctcaatgttgttgcacttttctataacattatcttg
217 atctatatacatcttgttattaacttcttcatcatcaataatttagatctatgaaatgctatacttttct
287 ttttctcattctgcctgcgtcacttcttgtttctggttgtgatgtgatgaacaatcttgattcagttttg
357 tcattgaaatttttgtctattttgtggtgtttaatgtatttaacagtgtcataaccatagtagctaatta
427 gttaagtaacaaaatgattcatttatcttggaactgtttattactgcagTGTTTGAATATTTGAGATATT
497 TTGCTTAGAAAATGAGAGGGGGTTTGTGGCAACTTGGCCAATCGATCACTCGCCGTCTTGCGAACGGAGG
- 30 *M R G G L W Q L G Q S I T R R L A N G G*
567 AGATAAGAAGGCTGTTGCGCGGCGATGTTTCGCTACAGAATCTGAGCTGAAGAAGACAGTTCTGTATGAC
- 10 *D K K A V A R R C F* **A T E S E L K K T V L Y D**
637 TTCCATGTTGCTCATGGTGGAAAGATGGTTCCTTTTGCTGGTTGGAGTATGCCTATTCAATACAAAGACT
14 F H V A H G G K M V P F A G W S M P I Q Y K D
707 CAATCATGGACTCAACGTTAAACTGTAGACAGAACGGTAGTCTTTTCGATGTTTCTCATATGTGTGGTCT
37 S I M D S T L N C R Q N G S L F D V S H M C G L
777 TAGTCTCAAGGGAAAGGATGTTGTTTCGTTCCTTGAGAAGCTCGTCATTGCGGATGTTGCTGCTCTTGCC
61 S L K G K D V V S F L E K L V I A T V A A L A
847 CACGGAACCGGGACTTTGACTGTTTTCACTAATGAAAAGGGAGGGGCGATCGATGATTCCGTGATTACTA
84 H G T G T L T V F T N E K G G A I D D S V I T
917 AGGTGACAGATGATCATCTTTACTTGGTTGTCAATGCTGGGTGCAGAGATAAAGATTTGGCTCATATTGA
107 K V T D D H L Y L V V N A G C R D K D L A H I E
987 GGAGCATATGAAGGCATTCAAGGCCAAAGGCGGTGATGTGTCGTGGCACATTCATGATGAGAGATCTCTA
131 E H M K A F K A K G G T V S W H I H D E R S L
1057 CTCGCTCTTCAGgtcatcatttatcactctattatacattagaatcatccaagtataacactcactctta
154 L A L Q
1127 acttgttttgtttctttgatatctaacacctatctcttgcttctgtattttagttgttaagatctatagc
1197 attgtttgttttgtttcttccccatgtgcctgcaccctacgctgtgtatactttggtgtacttagtgttg

```
1267 tcaaatagtggatatagcatagtgaaatttgagcaaaatgctactgttgtgcaatacgctatttggtaag
1337 acatgttatgaaatagcggtgctatagcactgctgcgtagtggagttttaacaaacttttgtaattccgc
1407 gattgacagcatgaggtgtagtgtgagttgtcacacattgtctattgtggtcataaaatgattctaattc
1477 acttcttcaggtggtttgtaccaggtttgtcttaatttttcggagactccgtgtaagttagccacgattt
1547 aatcgtgactaatattttatgatgccatatttaatcacaatttaacaatgtcaaattttgggccctgttc
1617 ggtagcctgccttccacgccctcaagacaccttggtttgtatccatgttaaagttgacaaataattgagt
1687 gagttttaaattaatgatatgcacctaatcgtaactaacactaatgaagttttcttcttttctctggaag
1757 cagGGTCCTCTTGCTGCTCCTGTTCTTCAACATCTGACAAAAGAGGATTTGAGCAAGCTGTACTTTGGGG
158       G  P  L  A  A  P  V  L  Q  H  L  T  K  E  D  L  S  K  L  Y  F  G
1827 AGTTCCGCGTGTTGGACATCAATGGCTCGCAGTGCTTTCTCACACGGACAGGgtatgcttcatttttgaa
180  E  F  R  V  L  D  I  N  G  S  Q  C  F  L  T  R  T  G
1897 tgtgttatcatgtattgtgcatttggatatttgttagatatgatttgagtttgttcgatgcatacatgat
1967 aacacgctctttaatgagattttcagGTACACTGGTGAAGATGGATTTGAGATCTCAGTTCCTTCAGAGC
198                                Y  T  G  E  D  G  F  E  I  S  V  P  S  E
2037 ATGGAGTTGAACTTGCCAAGGCACTGCTGGAAAAATCTGAAGGGAAAATAAGATTGACAGGACTCGGTGC
212  H  G  V  E  L  A  K  A  L  L  E  K  S  E  G  K  I  R  L  T  G  L  G  A
2107 TAGAGATAGTCTACGTCTCGAAGCTGGACTGTGCTTATACGGAAATGACCTGGAACAGCACATTACACCT
236    R  D  S  L  R  L  E  A  G  L  C  L  Y  G  N  D  L  E  Q  H  I  T  P
2177 ATTGAAGCAGGACTGACATGGGCTATAGGTAAGAGAAGGAGAGCAGAAGGTGGTTTTCTAGGAGCTGATG
257   I  E  A  G  L  T  W  A  I  G  K  R  R  R  A  E  G  G  F  L  G  A  D
2247 TTATCCTGAAACAGCTCGCAGACGGTCCTTCCATTAGGCGTGTCGGTTTCATTTCTTCTGGTCCACCTCC
282  V  I  L  K  Q  L  A  D  G  P  S  I  R  R  V  G  F  I  S  S  G  P  P  P
2317 AAGAAGCCACAGCGAGATTCAAGATGAAGGAGGAAACAACATTGGTGAAGTGACCAGTGGTGGATTCAGT
306    R  S  H  S  E  I  Q  D  E  G  G  N  N  I  G  E  V  T  S  G  G  F  S
2387 CCTTGTCTCAAGAAGAACATAGCTATTGGATATGTCAAATCTGGATTGCACAAAGCAGGTACCAAAGTAA
329   P  C  L  K  K  N  I  A  I  G  Y  V  K  S  G  L  H  K  A  G  T  K  V
2457 AGATCATTATTCGAGGTAAACAGAATGAAGGAGTTGTCACGAAAATGCCGTTCGTACCCACAAAATACTA
352  K  I  I  I  R  G  K  Q  N  E  G  V  V  T  K  M  P  F  V  P  T  K  Y  Y
2527 TAAGCCTTCATAGTTTGCTGCAGCATTTCCCATTTCTTGACTAATTTTCCCATAAACTGAGTGTTACATA
376    K  P  S  -
2597 AATGTTAATGTCGTTATTATGTTTCTCATGTCAAGTACAAGGCTTCATCTTGATTCATAACTCAATTTCT
2667 TTTCAAACATGTATTGCTTGAACTTGATGTTATCATGAATTTTTTCCAATATTTCCCCCTGTGTGTACTG
2737 TGCTATTTTGAGGCCTGTTTTTCATTCAATAGTAGCTATTCATATAAGACTAAATCTGAATAATATAATA
2807 ATGTgaaaatatAATAAAAAAACAGTATAC
```

figure 38: Séquence nucléotidique du gène codant pour la protéine T du complexe de la glycine décarboxylase de pois (var. Alaska) avec sa séquence en aminoacide déduite. ⟶ correspond à un démarrage de transcription détecté par extension d'amorce. La séquence soulignée en trait plein est une région promotrice probable de type boîte TATA. Le site de polyadénylation aataaa est représenté en caractères gras.

Les trois introns possèdent les courtes séquences consensus GT et AG (respectivement en 5' et 3'de l'intron) impliquées dans le processus d'épissage des transcrits au cours de leur maturation, au niveau de leur jonction avec les exons (Breathnach et Chambon, 1981). De

plus, les séquences polypeptidiques des deux séquences nucléotidiques complémentaires de chaque intron, ne révèlent pas de cadre ouvert de lecture significatif.

On remarque que la séquence correspondant au peptide de transit mitochondrial est contenue dans le deuxième exon. De plus, comme chez la plupart des gènes eucaryotes, les exons ont un pourcentage en adénine et thymidine (de 49,6% à 61,5% avec une moyenne de 56,7%) moins élevé que les introns (de 63,5% à 70,4% avec une moyenne de 67,7%) (tableau 9).

GENE T	Exon I	Exon II	Exon III	Exon IV	Intron 1	Intron 2	Intron 3
% AT	61.5	54.8	49.6	60.9	70.4	63.5	69.3

Tableau 9: Composition en adénine et en thymidine des exons et des introns de *gdcT*.

Enfin, le dernier exon contient le signal de polyadénylation typiquement eucaryote (AATAAA) (Proodfoot et Browlee, 1976). Ce dernier absent de la région non codante de l'ADNc correspondant à la protéine T, se situe à 340 pb en aval du signal de terminaison de traduction (figure 38). Cette séquence permet à une endonucléase de cliver l'ARN de 15 à 30 nucléotides en aval. Ainsi, c'est au niveau de ce site de coupure que la poly(A) polymérase ajoutera une queue polyadénylée (Whitelaw et Proudfoot, 1986; Logan *et al.*, 1987; Connely et manley, 1988).

2 - Analyse de la structure primaire de la protéine T

La comparaison entre la séquence nucléotidique des exons du gène *gdcT,* obtenue à partir de la variété « Alaska » de pois, et celle de l'ADNc provenant de la variété « Douce provence » (Cf. chapitre I) révèle six différences localisées au niveau de l'exon 1. Comme nous l'avons présenté dans le chapitre I, la séquence peptidique déduite de l'ADNc (variété « Douce provence ») fut confirmée biochimiquement par la spectrométrie de masse et par microséquençage des peptides obtenus grâce au clivage chimique de la protéine T purifiée. Ces variations observées entre les deux variétés sont présentées dans le tableau 10:

nucléotides modifiés		**acides aminés modifiés**	
Gène; variété «Alaska»	ADNc; variété «Douce provence»	Gène; variété «Alaska»	ADNc; variété «Douce provence»
Guanine 799	Adénine	valine 68	Isoleucine
Cytosine 819	Thymidine	pas de modification	
Cytosine 847	Adénine	Histidine 84	Asparagine
Cytosine 906	Thymidine	pas de modification	
Thymidine 927	Guanine	Aspartate 110	Glutamate
Thymidine 1041	Cytosine	pas de modification	

Tableau 10: Localisation des quelques différences nucléotidiques et d'acides aminés trouvées entre la partie codante du gène codant pour la protéine T provenant de la variété de pois « Alaska » et le transcrit correspondant à la protéine T provenant de la variété « Douce provence ».

Trois de ces mutations ont modifié la séquence en acides aminés de la protéine. En effet, sur six mutations, trois sont silencieuses car elles se situent sur la troisième base du codon. Une mutation peut parfois se traduire par le remplacement d'une pyrimidine par une purine à cette position. Par exemple, la troisième base du codon GA**C** (aspartate 110) chez le gène correspond au niveau de l'ADNc à une purine GA**G** (glutamate). Ces acides aminés aux propriétés similaires sont représentés par des codons voisins, ce qui permet de minimiser les effets des mutations. La même observation peut s'effectuer sur les deux autres mutations qui n'entraînent pas de modifications importantes. De plus, l'alignement de la structure primaire de la protéine T de pois avec celle des autres séquences publiées, situe ces mutations dans des régions très peu conservées. On peut donc penser que ces régions (contrairement à d'autres décrites dans le chapitre I) ne sont pas impliquées dans le domaine catalytique de l'enzyme. En effet, des mutations spontanées se produisent continuellement au cours de l'évolution sans aucun rapport avec une éventuelle nécessité immédiate, ou une quelconque utilité. D'ailleurs, la plupart sont anodines lorsqu'elles ne sont pas nuisibles dans les conditions environnementales où elles apparaissent. Deux hypothèses peuvent être avancées quant à l'origine de ces mutations:

- Ces mutations pourraient être d'ordre allélique. Si tel était le cas, elles seraient récessives et nous aurions détecté en spectrométrie de masse 2 populations de protéines.

- Ces mutations seraient propres à la variété du pois. Cette hypothèse semble être la plus plausible. En effet, une espèce, qu'elle soit animale ou végétale, est constituée en son sein par

une famille plus ou moins grande d'individus (variété) avec des modifications qui leur sont propres (ceci étant dû à la pression sélective) mais qui n'entraînent en aucun cas de profonds changements au niveau de leur métabolisme et de leur structure. Les mutations n'ont aucun effet phénotypique apparent et ne semblent pas toucher de régions importantes impliquées dans l'activité de la protéine T.

B - DETERMINATION DU SITE D'INITIATION DE LA TRANSCRIPTION DU GENE CODANT POUR LA PROTEINE T

L'identification du site d'initiation de transcription du gène codant pour la protéine T est réalisée par extension de la séquence nucléotidique vers l'extrémité 5' de l'ARNm correspondant à la protéine T. Cette technique d'extension d'amorce a été choisie préférentiellement à la technique de protection à la nucléase S1. La présence, dans la région non codante, d'un intron de 321 pb risquerait de gêner l'hybridation du gène avec l'ARN messager, augmentant de ce fait le risque d'éventuels artéfacts lors de la digestion par la nucléase S1.

Le principe de l'extension d'amorce est basé sur la synthèse d'ADNc de faible taille, copies de la région en 5' de certains ARNm. Cette réaction catalysée par une enzyme de rétrovirus, la transcriptase réverse, est initiée à l'aide d'un oligonucléotide de synthèse, complémentaire à la séquence nucléotidique de la protéine étudiée. Le marquage de l'ADNc est réalisé par incorporation de nucléotides marqués (α-^{32}P-dCTP). Le choix de la position de l'amorce tient compte du risque de la pause prématurée de la transcriptase réverse au niveau d'éventuelles structures secondaires de l'ARNm. En règle générale, l'oligonucléotide se situe à proximité de l'extrémité 5' du transcrit.

1 - Résultat

Nous avons fait synthétiser trois oligonucléotides de 40 pb (PE1, PE2 et PE3) situés à 116, 72 et 44 pb de la région 3' du premier intron. Ceux-ci sont hybridés avec des ARNm issus de tissus foliaires de pois âgés de 7 jours. Les résultats obtenus, par l'action de la transcriptase réverse, montrent deux bandes majeures, avec les trois amorces utilisées (deux extensions d'amorces sont présentées dans la figure 39). Ceci suggère fortement que le gène codant pour la protéine T possède deux sites de démarrage de la transcription. Ces deux transcrits,

correspondant à la protéine T, n'ont pu être détectés par les techniques d'hybridation classiques, car leurs tailles diffèrent par un trop petit nombre de nucléotides.

Le produit d'élongation le plus important s'arrête à une thymidine (on attribue la position +1 à ce nucléotide) et correspond au site d'initiation (TCAT**T**TGA) situé à environ 43 pb en aval d'une boîte TATA détectée par le programme PC GENE (Eukprom). En effet, la distance séparant la boîte TATA du site d'initiation est supérieure aux distances les plus couramment rencontrées (-39, -9) (Bucher, 1990). Cependant, la région de la boîte TATA, déterminée par le programme PC GENE, se situe dans une région riche en thymidine et en adénine, rendant délicate la localisation précise du coeur de la boîte TATA.

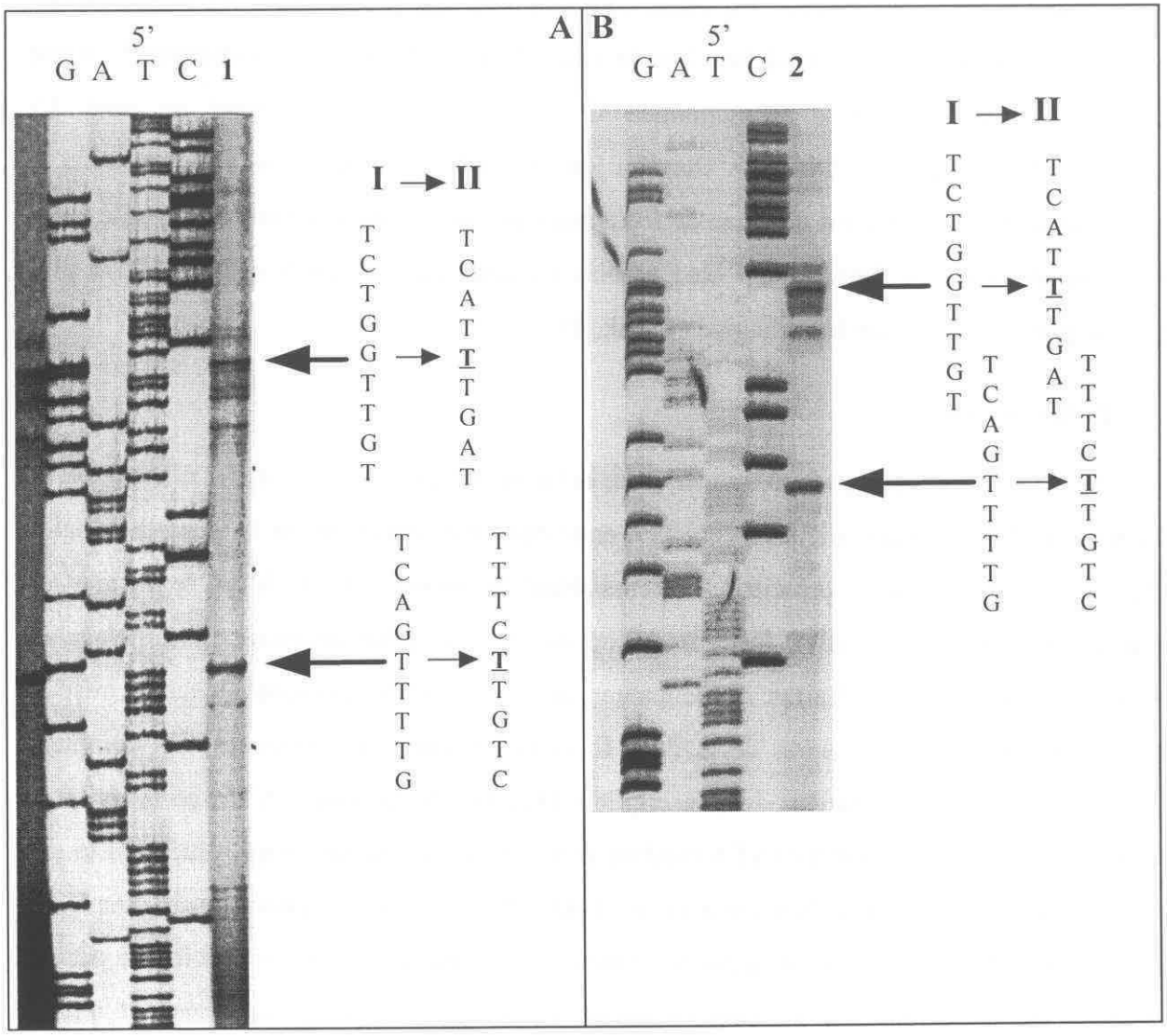


Figure 39: Identification des sites d'initiation de transcription de *gdcT* par extension d'amorce. Des ARN polyadénylés sont hybridés avec les amorces PE2 (nucléotides 480 à 519) (A-1) et PE3 (nucléotides 508 à 547) (B-2) (Cf. Matériel et méthodes), complémentaires chacune d'une séquence nucléotidique située en 5' de *gdcT* . La réverse transcriptase est alors utilisée pour synthétiser les ADNc correspondant à la région 5' des trancrits de la protéine T. Les lignes GATC représentent la séquence d'un fragment de *gdcT* obtenu à l'aide de

ces mêmes amorces. Comme un intron se situe entre le site d'initiation et les différentes amorces, les séquences lues sur le gel ne correspondent pas à celles localisées près du site d'initiation. La séquence lue a donc été utilisée comme simple échelle pour localiser les sites d'initiation situés en amont de l'intron. **I:** séquences lues. **II:** séquences déduites.

Il est à noter que ce programme a déterminé un autre site d'initiation se situant à 29 pb de la boîte TATA précédemment déterminée. Mais, aucune bande correspondant à ce site n'a pu être détectée. Le second site d'initiation (+31) déterminé par extension d'amorce du gène codant pour la protéine T, se situe en aval du premier site et commence à une thymidine (TTC**T**TGT). Ce site est apparemment trop éloigné de la boîte TATA pour être sous son influence. Ce résultat sera développé par la suite. Ces sites d'initiation jouent un rôle important lors de la traduction. En effet, contrairement aux procaryotes, les transcrits des eucaryotes sont rapidement modifiés dans le cytosol, par addition à leur extrémité 5' d'une coiffe (cap) sous forme d'un nucléotide G (guanine) méthylé, formant un pont 5'-5'triphosphate (Shatkin, 1985). Ce chapeau joue un rôle important dans l'initiation de la synthèse protéique, car il est un signe de reconnaissance pour l'appariement de la petite sous-unité du ribosome sur l'extrémité 5' de l'ARNm. Celui-ci migrant vers le codon AUG le plus proche permettra d'initier la traduction (Kosak, 1987).

2 - Discussion

Chez la plupart des organismes eucaryotes, la faible distance séparant le site d'initiation de la boîte TATA est nécessaire à l'interaction entre deux éléments formant un complexe rigide ADN-protéine, les deux principaux acteurs étant la protéine POL II et le facteur de transcription TFIID. En effet, la transcription basale d'un gène nécessite l'assemblage de nombreuses protéines nucléaires. Actuellement, quelques modèles qui schématisent la mise en place du complexe d'initiation de la POL II, commencent à être proposés. Un complexe « initial », composé du facteur de transcription TFIID et de la boîte TATA, se forme pour constituer l'ensemble [TBP] (TATA-binding protein) qui couvre une région incluant le site d'initiation (de -45 à -10) (Nakajima *et al.*, 1988). Pour une transcription basale simple, ce complexe suffit pour être reconnu par les autres facteurs, mais ne peut pas être régulé par des facteurs de transcription se trouvant en amont. Par la suite, ce complexe devient le site de fixation d'un deuxième facteur de transcription TFIIB qui a pour rôle de recruter la POL II et un troisième facteur de transcription TFIIF, formant ainsi un complexe stable. Il est à noter que dans certaines conditions ce simple complexe est capable d'initier la transcription. Par la

suite, deux autres facteurs TFIIE et TFIIH viendront s'associer à ce complexe et permettront d'initier la transcription par phosphorylation du domaine carboxyterminal de la POL II à l'aide du facteur TFIIH (qui possèderait une activité kinase). Ainsi constituée, la transcription basale ATP dépendante peut débuter (Buratowski, 1994). La boîte TATA et le site d'initiation du gène seraient donc deux éléments indispensables à la reconnaissance et à la mise en place du complexe de transcription. Cependant, il existe d'autres éléments initiateurs pouvant induire la transcription lorsque la boîte TATA est absente ou trop éloignée pour avoir une influence.

En effet, le second site d'initiation (+31) déterminé par extension d'amorce du gène codant pour la protéine T, ne peut être sous l'influence de la même boîte TATA, car la distance entre les deux éléments (75 pb) est largement supérieure à celle souvent observée (30 pb) chez les gènes transcrits par l'ARN polymérase II (POL II) (Breathnach et Chambon, 1981; Bucher, 1990). Une situation analogue a été observée chez le gène codant pour la protéine H (Macherel *et al.* 1992). Ainsi, les gènes codant pour les protéines H et T possèdent deux sites d'initiation de transcription, dont l'un serait sous le contrôle de la boîte TATA alors que le second site (distant de 29 pb du premier site d'initiation dans le cas du gène codant pour la protéine T et de 33 pb pour le gène codant pour la protéine H), se situerait dans une région riche en pyrimidine. Les comparaisons de séquences nucléotidiques représentées sur la figure 40, révèlent que cette région riche en pyrimidine encadrant le second site d'initiation de transcription des gènes *gdc T* et *gdc H*, est similaire aux sites d'initiations de plusieurs gènes tels que: le gène codant pour la terminale déoxyribonucléotide transférase (Smale et Baltimore, 1989), la région promotrice du virus SV 40 (Ayer et Dynan, 1988; Smale *et al.*, 1990) et le gène codant pour la sous-unité α2 de l'Intégrine humaine (Zutter *et al*, 1994). Smale et Baltimore (1989) ont appelé cette séquence d'initiation « Initiator Element » (Inr) car cette séquence joue un rôle de promoteur en relation ou non avec la boîte TATA.

```
                                                  ┌──▷
gdc T S.2     -6  C G T T T C T T G T C T C T   +8
                  |   |   | | |   |           | | |
gdc H S.2     -6  C A T T T C A T T C T T C T   +8
                  •           | | | | | |       • •
TdT  Inr      -6  G C C C T C A T T C T G G A   +8
                  |   |   | | |   |   | | •
AdML Inr      -6  G T C C T C A C T C T C T C   +8
                  |       | | |   |     •   •   •
A2 Int. Inr   -6  G C T C T C A C C G G G C G   +8
                            └──┘
```

figure 40: Comparaisons de la région riche en pyrimidine encadrant le second site d'initiation des gènes codant pour les protéines T (*gdcT* S.2) et les protéines H (*gdcH* S.2) avec d'autres séquences initiatrices connues (Inr). Les séquences initiatrices du gène codant pour la terminale déoxynucléotidyltransférase (*TdT*

Inr), du promoteur tardif principal de l'adénovirus (AdML) et la séquence initiatrice du gène codant pour l'α2 intégrine (A2 Int. Inr) proviennent de Smale et Baltimore (1989), Ayer et Dynan (1988) et Zutter *et al.*, (1994). Le symbole | indique que les nucléotides sont identiques alors que le symbole . indique qu'ils sont similaires.

Plusieurs hypothéses ont été proposées pour justifier le fonctionnement de ce type de promoteur.

Dans le cas où la boîte TATA est absente, O'Shea-Greenfield et Smale (1992) ont supposé que cette région riche en pyrimidine (nommée élément Inr) aurait pour rôle soit de guider la POL II vers un site précis d'initiation de transcription, soit d'avoir une influence sur la direction de la transcription (O'Shea-Greenfield et Smale, 1992).

Si on considère le cas où la région promotrice possède la boîte TATA, l'élément Inr agirait comme un activateur. Cependant, pour que cette coopération entre la boîte TATA et l'élément Inr puisse s'effectuer, les deux séquences doivent être distantes au maximum de 30 pb. Au-delà de cette distance, les deux éléments agissent de façon indépendante; deux démarrages de transcription peuvent alors être observés (O'Shea-Greenfield et Smale, 1992).

Zenzie-gregory *et al.* (1992), en se basant sur des expériences de cinétiques (réalisées pour comparer les mécanismes d'initiation de transcription en présence de la boîte TATA ou de l'élément Inr), ont démontré que sur des extraits de cellule, le mécanisme d'initiation de transcription en présence de l'élément Inr ne laisse pas apparaître la période de latence présente lorsque la transcription est sous l'influence de la boîte TATA. En effet, la mise en place du complexe de préinitiation (PIC) semble s'effectuer plus rapidement, lorsque la transcription est dirigée par l'élément Inr, car la machinerie impliquée dans la transcription est moins complexe. Certains facteurs de transcription se substituent à un ensemble d'autres facteurs. Ainsi, récemment, Usheva et Shenk (1994) ont démontré que la seule présence des facteurs de transcription YY1, TFIIB et de la POL II suffisait à initier une transcription basale sur un ADN contenant l'élément Inr. Dans des conditions appropriées, YY1 jouerait la fonction de la TBP (TATA binding protein) dont le rôle est de recruter la POL II (Seto *et al.*, 1993). D'autres auteurs ont également démontré l'implication du complexe TFIID ou du facteur de transcription TFII-I dans l'initiation de la transcription à partir de l'élément Inr (Smale *et al.*, 1990; Pugh et Tjian, 1991; Roy *et al.*, 1991; Carcamo *et al.*, 1991; Zenzie-Gregory *et al.*, 1993). Chez les plantes, Maas *et al.* (1990) ont également localisé dans la région promotrice du gène *Shrunken* de maïs (*Sh*: gène codant pour la saccharose synthase), un site d'initiation situé dans une région riche en pyrimidine et capable d'initier la

transcription en absence de la boîte TATA. De plus, Lugert et Weer (1994) ont isolé, en criblant une banque d'expression λgt 11 (technique du « South-Western »), une protéine (IBP 1) qui se lie de façon spécifique à cet élément Inr. Cependant, il existe peu d'homologies entre la séquence du démarrage de transcription du gène *Sh* et celle des animaux. Toutefois, en ce qui concerne la région riche en pyrimidine, déterminée chez le gène codant pour la protéine T, une étude plus approfondie s'avère indispensable pour confirmer l'influence de cet élément initiateur dans l'initiation de la transcription du gène *gdcT*. En effet, cet élément Inr est apparemment beaucoup trop éloigné de la boîte TATA pour avoir un effet activateur sur le niveau de transcription. Il serait donc intéressant de comprendre le rôle physiologique de ces deux sites d'initiation présents également dans le gène codant pour la protéine H (Macherel *et al.*, 1992).

III - ETUDE DE LA REGION PROMOTRICE

Les résultats obtenus sur l'expression des protéines T, H et P impliquées dans le complexe de la GDC montrent que cette expression est sans doute coordonnée. Pour cette raison il nous a semblé intéressant de comparer les régions promotrices des gènes codant pour les protéines T et H (Macherel *et al.*, 1992) afin d'identifier d'éventuelles séquences consensus susceptibles d'être présentes dans les région 5'des deux gènes. Nous avons également comparé la séquence nucléotidique de la région promotrice du gène codant pour la protéine T avec celle d'autres gènes régulés à la lumière. Ainsi, toutes ces comparaisons ont permis de localiser dans la région promotrice du gène *gdcT* trois séquences consensus.

A - IDENTIFICATION DE TROIS POTENTIELS SEQUENCES CONSENSUS DE REGULATION DANS LA REGION PROMOTRICE DE gdcT

1 - Caractérisation d'une région riche en nucléotides AT

La comparaison des régions promotrices des gènes codant pour les protéines H et T révèle la présence d'une séquence riche en nucléotides AT au sein des deux promoteurs (figure 41). La comparaison de cette séquence avec celles d'autres régions promotrices présentes dans les banques de données, a permis de déceler de fortes homologies avec certaines régions riches en nucléotides AT présents dans la zone promotrice d'autres gènes. Ainsi, la plus forte homologie a été déterminée avec la séquence consensus (5'-AATATTTTTATT-3') fixant le

facteur de transcription AT-1 (figure 41). Cette séquence a été caractérisée dans la région promotrice du gène codant pour la petite sous-unité de la rubisco rbcS 3,6 de pois (Datta et Cashmore, 1989). On trouve également cette séquence consensus dans d'autres régions promotrices de gènes comme la petite sous-unité de la rubisco rbcS3A et la protéine liant la chlorophylle a/b (cab) (Datta et Cashmore, 1989). D'autres homologies ont pu être également décelées avec les régions promotrices des gènes codant pour la protéine leghemoglobine (impliquée dans le transport de l'oxygène au sein des nodules) chez la graine de soja « -239 AATATTTTAAT -229 » et chez un légume tropical « -223 ATAATTTTTAA -213 » (*Sesbania rostrata*) (Jensen *et al.*, 1988; Metz *et al.*, 1988), pour la β-phaséoline de haricot « -650 TTAATTTTAAG -640» (Bustos *et al.*, 1989) et pour l'hélianthinine, protéine majeure de la graine de tournesol « -696 TTTATTTTGAT -686 » (Jordano *et al.*, 1989).

2 - Présence d'une boîte GT-1 dans la région promotrice du gène *gdcT*

Parce que le niveau de transcription des gènes codant pour les protéines de la GDC subit une forte induction à la lumière, nous avons comparé la région promotrice du gène *gdcT* de pois avec plusieurs éléments différents agissant en cis (cis-acting-element) connus pour être impliqués dans la régulation de gènes sensibles à la lumière. Cette analyse de la région promotrice du gène *gdcT,* nous a révèlé la présence de motifs connus (figure 41). Notamment, une séquence localisée entre les nucléotides -126 et -114 possède une forte homologie avec la boîte II qui agit comme un élément de régulation positif sur la transcription des gènes dépendant de la lumière, en particulier sur le gène de la petite sous-unité de la Rubisco rbcS 3A (Green *et al.*, 1987; Green *et al.*, 1988). De même, une autre séquence située dans le gène *gdcT* « -225 GTGAGGTAATAT -214 » présente une certaine homologie avec une autre boîte située en amont de la boîte II nommée boîte II* dans le cas de rbcS3A « -210 AAGAGGTAATGG -201» (Green *et al.*, 1987). Récemment, une étude de la région promotrice du gène codant pour la protéine H du complexe de la GDC chez *Arabidopsis thaliana*, a révélé qu'une zone située entre -376 et -117 pb et contrôlant l'expression à la lumière du gène, contenait une série de boîtes GT (Srinivasan et Oliver, 1995). De plus, de tels motifs se retrouvent également dans la région promotrice du gène codant pour une protéine péroxisomale impliquée dans le cycle photorespiratoire: l'hydroxypyruvate réductase dont le niveau de transcription serait également sous l'influence de la lumière (Schwartz *et al.*, 1991).

3 - La région 5' en amont de *gdcT* possède l'élément GATA

La littérature concernant la spécificité tissulaire des gènes est dense; elle présente de nombreuses études sur différentes familles de gènes impliquées dans plusieurs espèces. En effet, les recherches entreprises sur la spécificité tissulaire de certains gènes de plantes ont eu un impact important sur les biotechnologies en agriculture. La manipulation récente des promoteurs, à l'aide des plantes transgéniques, montre la présence de certains éléments de régulation en cis (cis-acting element) capables de conférer à certains gènes une régulation à la lumière ou une spécificité tissulaire. Ainsi, la comparaison entre la région promotrice de *gdcT* et d'autres régions en 5', nous a permis d'identifier (figure 41) un motif GATA en tandem (-366 à -350) présent au niveau des promoteurs des gènes cab de la région 5' du promoteur 35S du virus de la mosaïque du chou-fleur (CaMV 35S) et d'une des petites sous-unités de la rubisco rbcS 3C (Castresana *et al.*, 1987; Lam et Chua, 1990; Manzara *et al.*, 1991). Cependant, la distance à laquelle se situe cet élément GATA en tandem de *gdcT* est importante alors que celui principalement étudié dans la littérature se situe entre la séquence CAAT et la boîte TATA. De plus, selon Benfey *et al.*, (1990), le mécanisme régulant la spécificité tissulaire est beaucoup trop complexe pour être dû à la seule présence d'une séquence consensus. En effet, ces auteurs ont montré que la région promotrice de CaMV 35S est constituée de deux domaines principaux, A et B, chacun subdivisé en plusieurs sous-domaines et ce serait la combinaison de chacun des sous-domaines qui permettrait de définir la spécificité tissulaire au cours du développement. Par conséquent, il apparaît prématuré d'assigner à l'élément GATA (-366 à -350) du gène *gdcT* un rôle certain dans l'expression tissulaire de la protéine T.

A

```
-905 attttgcataattgaaacgaatgctgagagaagtggacgaagggaaggagagagaagagagactaaaatgaaagaaagaactt
                                                                      IR-1
-820 atcttttagttatcttatagagtgaagaaataaaatgtagttatgaaggagccaaggaaagaatgagacataatgatgatttaag
              IR-1
-735 gagattgagataagatgatggacgagattctggtgtggaatagatgatcttgcacgtggaggaaagctcctgtagtaaaaaaatt

-650 tgatttgaccaagaatgatgaaaaattagcgtgaagaagagatatatgattagaatttaattcttctaataccatgttaaattta
                                                          IR-2          IR-2
-565 gtctattggacaccttaaccaatatataatttttataacaaaaaaagaagcaaaaaacaattaaaattgaaataattttttcttct
        IR-3          IR-3          AT-1
-479 tattattatatttactacttagtgggcttatttatatatccatgttgttaacaaaaagcccacgatatggatagattgttacaac
                                                IR-4 <---p--->        IR-4 GATA    IR-5
-394 aaacaccatagagaagtagcaattaattaaaattaaagaatacactaataatgatgtgtaacaaatgcattacatagcacttcat
                      IR-5                                                          DR-1
-309 atctacttcataagaaaaaaaaattatttttaaatttattaaagtttattaaataaacttggttggtggagaaaagaaaaagaa
            DR-1                                      IR-7   IR-7
-224 aaaaaagaatgggaaagaggtaatggtaatgttgctaccactcactaaggatggggtatcataggctctcttcatttggggcact
                boîte II* DR-2  DR-2
-139 cactcaatcaagatgaggttaagttgttctcaatctcacatccaccaactccctctttttatatttttatttaaattaaataaat
                 boîte II IR-6                  IR-6
                                                      +1 ↱                          +31 ↱
-54  atattaatatattattttatttatttattcatttatgtccaacaaattcatttgataaacgaggttctctcttctcgtttctt
       boîte TATA
+33  gtctctcgcagtgtttcataaaaacaaaaacaagttacctaagcttctccattgttttgttcct
```

B

Boîte AT-1

gdc T de pois	- 541	ATATA**ATTTTTA**TAACAAA	- 525	
gdc H de pois	- 672	TTGAT**ATTTTTA**ATTCTTT	- 653	(1)
rbcS 3A de pois	- 258	TCGAT**ATTTTTA**TTTTGTA	- 288	(2)
rbcS3.6 de pois	- 553	TAATT**ATTTTTA**TTCTCTT	- 547	(3)
cab-E de tabac	-1107	TAAAT**ATTTTTA**TTTTCAT	-1137	(4)
rbcS 3A de tom.	-1014	TTAAT**ATTTTTA**TTTAGTA	- 996	(5)
		. ..*******.. .		

Boîte II du site de fixation de GT - 1

gdc T de pois	-127	ATGA**GGTTAA**GTTG	-114	
rbcS E9 de pois	-149	GTGT**GGTTAA**TATG	-136	(2)
rbcS 3A de pois	-151	GTGT**GGTTAA**TATG	-137	(2)
rbcS 3C de pois	-108	GTGT**GGTTAA**TATG	- 95	(2)
		.****..**		
gdc H d'A.thalian.	-138	AAAA**GGTTA**CAAAA	- 151	(6)
		. .***** ..		

Boîte GATA

gdc T de pois	-366	CAC**GATA**TG**GATA**GATT	-350	
cab 22R de Pet.	- 80	ATA**GATA**GT**GATA**TTCA	- 64	(7)
cab-E Nic.P.	- 50	GTA**GATA**TA**GATA**CTCA	- 35	(8)
rbcS 3C de pois	-283	TAT**GATA**TA**GATA**TTTA	-267	(2)
cab 3C de Tom.	- 49	GCA**GATA**AT**GATA**TTCT	- 38	(9)
		**** ****		
CaMV 35S	-101	ATT**GATG**T-**GATA**TCTC	- 86	(10)
		*** ****		

Figure 41: Principales caractéristiques de la région promotrice du gène codant pour la protéine T du complexe de la glycine décarboxylase. (A) Séquence de la région promotrice avec les deux sites d'initiations de transcriptions indiquée par une flèche au dessus des nucléotides +1 et +31. La séquence soulignée par un trait plein et en caractères gras correspond à la région contenant la boîte TATA. Les séquences soulignées correspondent aux séquences directement (DR-1 à -2) et inversement (IR-1 à -6) répétées. Les motifs « GATA », « AT-1 » et « boîte II » sont encadrés. p indique un palindrome. **(B)** Caractérisation des motifs « GATA », « AT-1 » et « GT-1 » par comparaison avec la région promotrice de différents gènes (1, Macherel *et al.*, 1992; 2, Fluhr *et al.*, 1986a, 1986b; 3, Herrera-Estrella *et al.*, 1984; 9, 4, Castresana *et al.*, 1987, 1988; 5, Ueda *et al.*, 1989; 6, Srinivasan et Oliver, 1995; 7, Gidoni *et al.*, 1989; 8, Pickersky *et al.*, 1985; 10, Odell *et al.*, 1985).

Les comparaisons de séquences nucléotidiques du gène codant pour la protéine T avec celles de nombreux gènes régulés par la lumière, ont permis de localiser certains éléments consensus susceptibles de conférer une expression à la lumière. A ce stade du travail, il nous a semblé intéressant, sans nous engager dans des études trop longues d'expression de promoteur chimérique à l'aide de plantes transgéniques, d'analyser les interactions Protéines/ADN sur les régions consensus déterminées dans ce paragraphe. Ainsi, nous avons essayé, par des expériences de gel retard avec ou sans compétition, de mettre en évidence d'éventuels facteurs de transcription capables d'interagir avec les sites précédemment définis.

B - EXPERIENCE D'INTERACTION PROTEINES/ADN

Notre approche expérimentale a consisté à rechercher si les éléments cis caractérisés sont susceptibles de fixer des protéines nucléaires. Pour cela, un extrait brut de protéines nucléaires provenant de noyaux de tissus chlorophylliens de pois a été incubé avec un duplex d'oligonucléotide marqué au phosphore 32 dans lequel se trouve la séquence consensus précédemment caractérisée. La formation de complexes ADN-protéine a été mise en évidence sur un gel de polyacrylamide en condition non dénaturante par la diminution de la mobilité électrophorétique du fragment d'ADN complexé. Ainsi, trois complexes nucléoprotéiques dont deux importants, ont pu être mis en évidence lorsque la région promotrice de *gdcT* contenant la boîte II a été incubée avec un extrait nucléaire de feuille de pois (figure 42). Par contre, un seul retard est visible au niveau de la région riche en nucléotides AT située dans la région promotrice de *gdcT*. Les expériences de compétition avec un excès de sonde froide montrent que les protéines impliquées dans la formation des complexes interagissent spécifiquement avec les séquences nucléotidiques contenant les régions consensus. Finalement, nous avons voulu déterminer si ces séquences consensus capables de lier des

protéines nucléaires pouvaient interagir avec les séquences de l'élément AT et de la boîte II consensus des gènes rbcS3,6 et 3A.

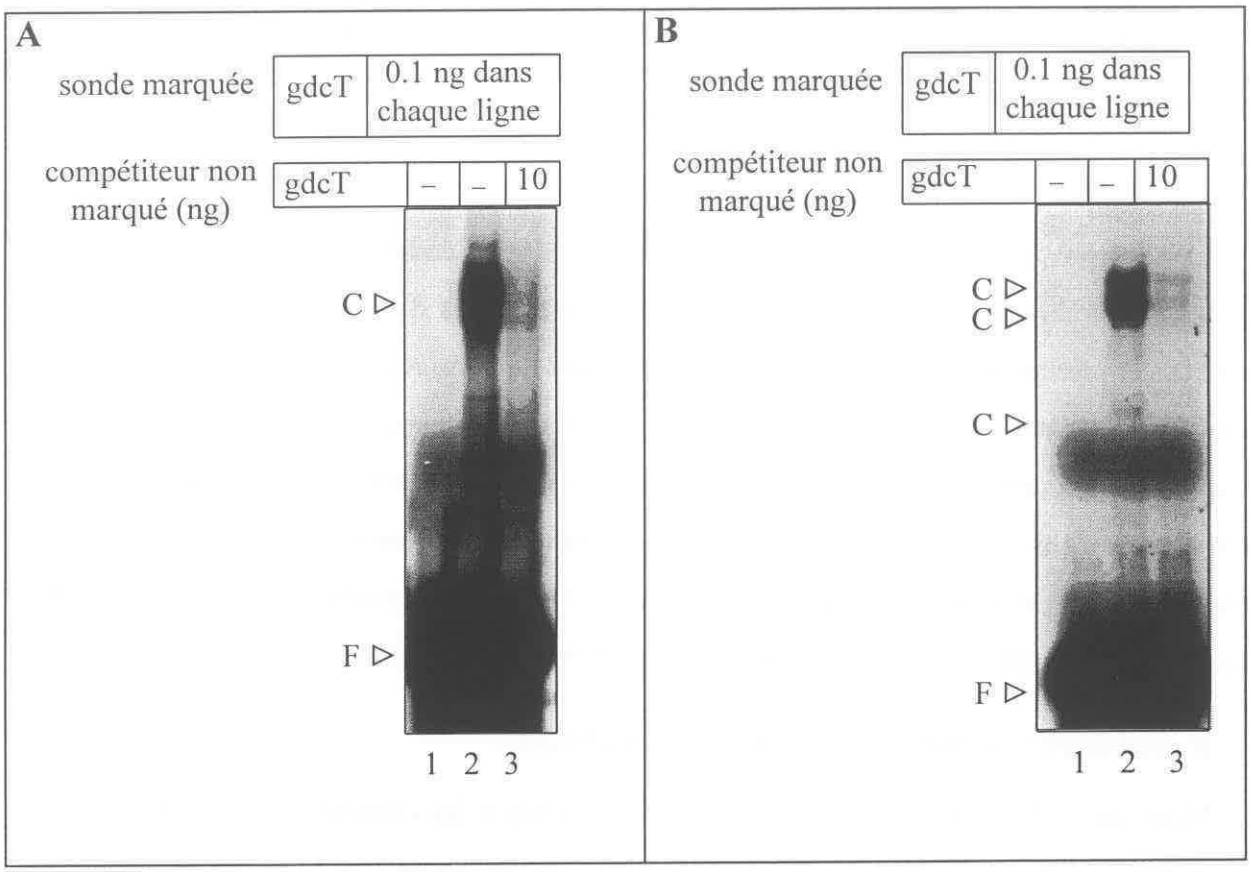


Figure 42: Formation de complexes nucléoprotéiques avec un fragement d'ADN contenant la région riche en nucléotides AT et la boîte GT-1.

A) *gdcT*: Fragment d'ADN de *gdcT* contenant une région riche en nucléotides AT: TTTACACCTTAACCA ATATATAATTTTTATAACAAAAAAAAAG.

B) *gdcT*: Fragment d'ADN de *gdcT* contenant un site équivalent à la boîte II présent dans la zone promotrice de rbcS 3A: TTTAATCAAGATGAGGTTAAGTTGCTCTCAATC. Ligne 1, sonde seule; lignes 2 et 3 sonde en présence de 5 µg d'extrait nucléaire de feuille de pois âgés de 7 jours. C: complexe nucléoprotéique; F: sonde libre.

1 - Relation entre la région riche en nucléotides AT du promoteur de gdcT et l'élément cis, riche en nucléotides AT, du promoteur de rbcS 3,6

L'incubation de l'oligonucléotide (duplex) contenant la région riche en nucléotides AT de *gdcT* marquée au phosphore 32 avec de l'extrait nucléaire de feuille de pois, permet la formation d'un retard important spécifique qui disparaît en présence d'un excès de dimère contenant la région riche en nucléotides AT de rbcS 3,6 mais qui ne semble pas être influencé par la présence d'un excès de 100 fois (en masse) du plasmide pBS (figure 43B). Par ailleurs, si ce compétiteur est utilisé comme sonde marquée, il entraîne un retard similaire à celui de la

boîte riche en nucléotides AT du gène *gdcT* (figure 43A). Il semble donc que l'élément AT du gène gdcT soit reconnu par une ou des protéine(s) nucléaire(s), et il est fort probable que ce soient les mêmes protéines qui interagissent avec l'élément AT du gène rbcS. En effet, les 2 types de séquences entraînent les mêmes variations de mobilité dans l'essai, et agissent comme des compétiteurs spécifiques vis à vis l'une de l'autre. Cependant, il est à noter que les régions promotrices des gènes des plantes au même titre que celles des levures, contiennent souvent des éléments riches en nucléotides AT. Ainsi, il a été montré que chez les levures, cette région stimule la transcription alors que chez les plantes, ses fonctions sont moins connues et semblent montrer des effets contradictoires (Russell *et al.*, 1983; Struhl, 1985; Chen *et al.*, 1987). En effet, Castresana *et al.*, 1988 et Datta et Cashmore, 1989 ont montré que l'élément riche en nucléotides AT, situé dans la région promotrice de rbcS et cab, aurait un effet négatif sur la transcription. Ces auteurs ont caractérisé un facteur de transcription AT-1 capable de se lier à la région riche en nucléotides AT mais dont la capacité de liaison était modulée par phosphorylation. Ainsi, ce facteur AT-1, actif sous une forme non phosphorylée perdrait toute capacité de liaison avec l'ADN sous une forme phosphorylée. A l'opposé, Bustos *et al.*, (1989) ont caractérisé un élément riche en nucléotides AT, situé dans le promoteur d'un gène codant pour la β-phaséoline de haricot, qui augmente l'expression de la β-glucuronidase (GUS) chez les plantes de tabac transformées avec un gène reporteur. La même observation a été effectuée au niveau d'une région riche en nucléotides AT dans la région 5' du gène codant pour la leghemoglobine chez le haricot et S. rostrata (Laursen *et al.*, 1994; Szczyglowski *et al.*, 1994). Il semblerait donc que les régions riches en nucléotides AT (dont la plupart sont éloignées de la boîte TATA) soient l'objet de nombreuses interactions avec des protéines nucléaires capables de jouer un rôle d'activateur ou de répresseur de gène. Cependant, de nombreuses questions restent en suspens quant au rôle exact de ces zones consensus.

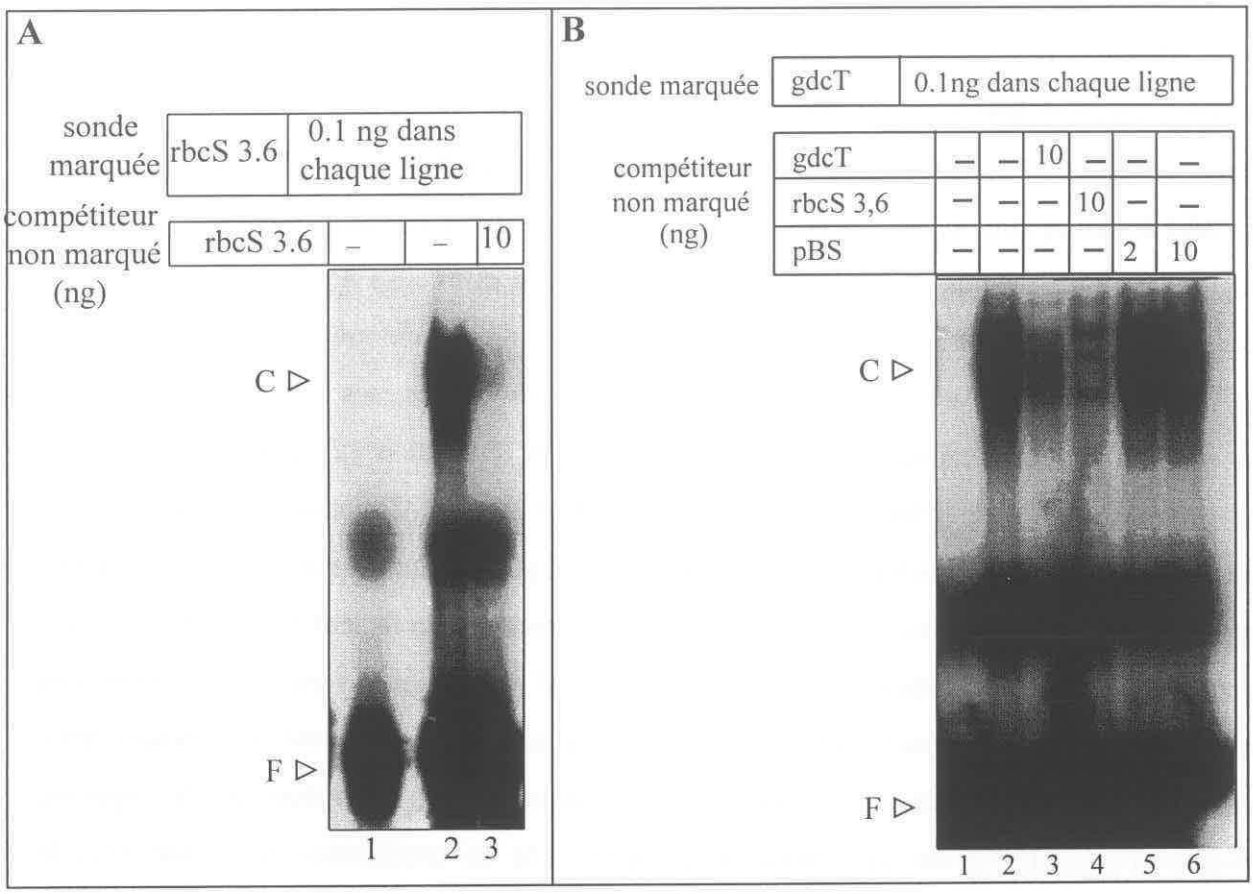


Figure 43:

A) Formation de complexes nucléoprotéines avec le fragment d'ADN de rbcS 3.6 contenant la région riche en nucléotides AT. rbcS 3.6, fragment d'ADN de rbcS 3.6 contenant la région riche en nucléotides AT: TTTATATATTTTTAATTATTTTTATTCTCTTAA. Ligne 1, sonde seule; lignes 2 et 3 sonde en présence de 5 µg d'extrait nucléaire de feuille de pois et en présence ou non de compétiteur. C: complexe nucléoprotéique; F: sonde libre.

B) Expérience de compétition entre les éléments riches en nucléotides AT de *gdcT* et rbcS 3.6. *gdcT*, Fragment d'ADN de *gdc*T contenant une région riche en nucléotides AT: TTTACACCTTAACCAATATATA ATTTTTATAACAAAAAAAAAG. rbcS 3.6, fragment d'ADN contenant la région riche en nucléotides AT (Cf. A). pBS: plasmide linéarisé à l'aide de l'enzyme de restriction *Eco* RI. Ligne 1, sonde seule; lignes 2 à 6 sonde en présence de 5 µg d'extrait nucléaire de feuille de pois et en présence ou non de compétiteur. C: complexe nucléoprotéique; F: sonde libre.

2 - Relation entre la boîte II située dans la région promotrice de *gdcT* et celle située dans la région 5' du gène codant pour rbcS 3A.

Les expériences d'interactions protéines/ADN nous ont permis de déterminer 3 retards spécifiques au niveau de l'oligonucléotide contenant la boîte II de *gdcT*. Ce même profil est observé lorsque l'expérience est effectuée avec la sonde contenant la région consensus de rbcS 3A (figure 44A). On attribue ce nombre élevé de retards au fait que nous avons utilisé une séquence nucléotidique encadrant la boîte II plutôt qu'un dimère ou un tétramère de la

séquence consensus. De ce fait, cela laisse la possibilité à d'autres protéines nucléaires d'interagir avec les séquences se situant de part et d'autre de la séquence consensus. En outre, les retards observés diminuent fortement lorsque la boîte II de *gdcT* marquée au phosphore 32 est mise en présence d'un excès de la région consensus de rbcS 3A (figure 44B). Cependant, la présence d'un excès de 100 fois (masse) du plasmide pBS ne modifie en rien le profil du gel retard. Ainsi, les résultats obtenus tendent à montrer que les régions consensus contenant la boîte II situées dans les régions promotrices des gènes *gdcT* et rbcS 3A sont apparentées. Il n'est donc pas exclu que le gène *gdcT* soit régulé par des facteurs nucléaires similaires à ceux intervenant dans la régulation du gène rbcS 3A. Green *et al.* (1987) ont caractérisé une protéine nucléaire GT-1 capable de se lier à la boîte II de rbcS dont la fonction serait d'être impliquée dans la réponse à la lumière. Ainsi, Lam et Chua (1990) ont démontré par l'intermédiaire de plantes transgéniques, qu'un seul tétramère synthètique (correspondant à la séquence de la boîte II) fusionné avec une partie du promoteur 35S du virus de la mosaïque du chou-fleur, suffisait à activer la transcription à la lumière et à obtenir une spécificité tissulaire. Récemment, l'ADNc correspondant au facteur de transcription GT-1 (trans-acting-factor) a été caractérisé. Cette protéine nommée GT-1a par Gilmartin *et al.* (1992) et B2F (pour facteur de la boîte II) par Perisic et Lam (1992) interagit avec le coeur de la boîte GT (pour une revue voir Gilmartin *et al.*, 1990). Il semble que ce facteur nucléaire soit exprimé dans tous les tissus: tissus foliaires de plantes cultivées à la lumière, tissus de feuilles qui ont subi trois jours d'obscurité, de tiges, de racines et de graines (Gilmartin *et al.*, 1992; Perisic et Lam, 1992). Afin d'expliquer de tels résultats, Lam et Chua (1990) et Gilmartin *et al.* (1992) ont proposé trois modèles. Deux font intervenir une protéine supplémentaire qui agirait, soit comme un répresseur à l'obscurité, soit comme un activateur à la lumière. Le troisième modèle envisage une modification photoréversible du facteur de transcription GT-1a qui permettrait une fonction activatrice à la lumière. En fait, le mécanisme exact qui confère à la boîte II une activité spécifique dans les cellules chlorophylliennes, reste encore une énigme. Enfin, l'analyse de la région promotrice de la protéine H du complexe de la GDC chez *Arabidopsis thaliana* par Srinivasan et Oliver, (1995), a montré que la délétion d'un fragment allant de -376 et -117 pb entraînait une chute dramatique de l'activité du gène reporteur GUS chez les plantes exposées à la lumière. Cette région contiendrait les séquences qui induisent l'expression à la lumière. En effet, l'analyse de la séquence nucléotidique de cette région a révélé la présence de nombreuses boîtes GT dont la séquence consensus présente une certaine

homologie avec celle de la boîte II. Ainsi, des régions consensus similaires existeraient entre les gènes codant pour les protéines H et T. Il n'est donc pas exclu qu'un mécanisme de régulation comparable à celui de la rbcS soit présent au niveau de l'expression de ces 2 gènes induits à la lumière.

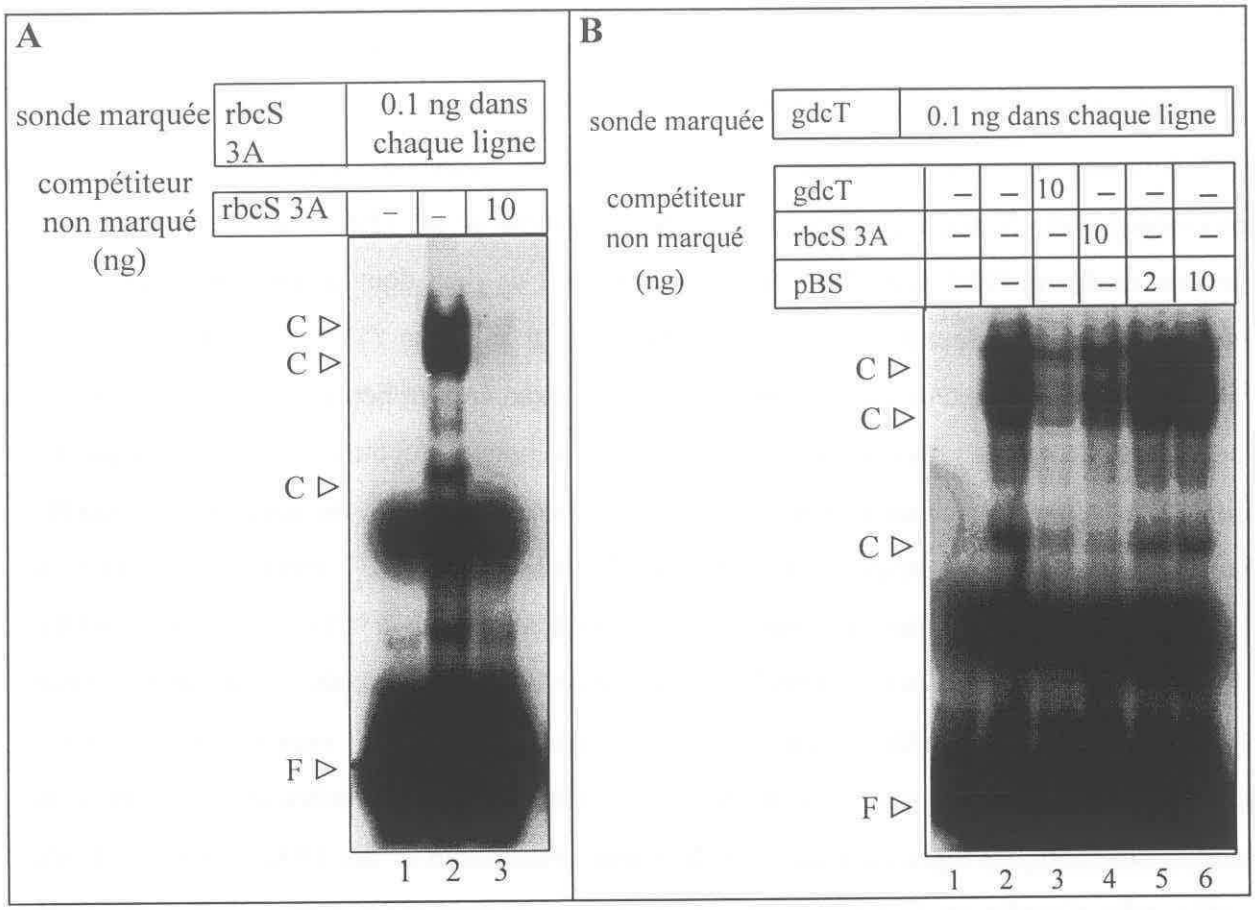


Figure 44:

A) Formation de complexes nucléoprotéines avec le fragment d'ADN contenant la boîte II de rbcS 3A. rbcS 3A, fragment d'ADN contenant la boîte II de rbcS 3A: TTTAATCAAGATGAGGT TAAGTTGCTCTCAATC. Ligne 1, sonde seule; lignes 2 et 3 sonde en présence de 5 µg d'extrait nucléaire de feuille de pois et en présence ou non de compétiteur. C: complexe nucléoprotéique. F: sonde libre.

B) Expérience de compétition entre les boîtes II de *gdcT* et rbcS 3A. *gdcT*, fragment d'ADN de *gdc*T contenant un site équivalent à la boîte II présent dans la zone promotrice de rbcS 3A: TTTAATCAAGATGAGGTTAAGTTGCTCTCAATC. rbcS 3A, fragment d'ADN contenant la boîte II qui fixe la facteur de transcription GT-1: TTTCAATCTTGTGTGGTTAATATGACTGCAAAG. pBS: plasmide linéarisé à l'aide de l'enzyme de restriction *Eco* RI. Ligne 1, sonde seule; lignes 2 à 6 sonde en présence de 5 µg d'extrait nucléaire de feuille de pois et en présence ou non de compétiteur. C: complexe nucléoprotéique. F: sonde libre.

IV - CONCLUSION

L'isolement et le séquençage d'un fragment de 3740 pb dans lequel se trouve le gène codant pour la protéine T de pois, ont révélé que ce gène présent en une seule copie dans le génome du pois est constitué de 4 exons et de 3 introns dont l'un se situe dans la région 5' non

codante. Nous avons découvert dans la région 3' non codante du gène le site consensus AATAAA de polyadénylation absent dans l'ADNc correspondant à la protéine T (Cf. chapitre I). De plus, l'étude du gène codant pour la protéine T a permis de montrer certaines homologies avec le gène codant pour la protéine H. En effet, ces 2 gènes possèdent 2 sites d'initiations dont un se situe dans une région riche en pyrimidine. Ce site, éloigné de la boîte TATA ressemble à l'élément Inr capable de conférer une transcription basale (Smale et Baltimor, 1989). Cependant, le rôle exact de ces deux sites d'initiation dans la transcription des gènes de la GDC reste encore à élucider. En effet, il serait intéressant de suivre au cours du développement l'évolution de ces deux populations de transcrits correspondant aux protéines T et H. Par ailleurs, l'étude de la région promotrice du gène codant la protéine T a révélé la présence de deux séquences consensus (régions riches en nucléotides AT et boîtes GT-1) capables de fixer certains facteurs protéiques nucléaires. En effet, les expériences de compétition ont montré que ces éléments seraient apparentés à ceux présents chez les gènes codant pour rbcS 3,6 et rbcS 3A. En outre, ces éléments de régulation en cis ont pu être retrouvés au niveau de la région promotrice du gène codant pour la protéine H de pois et d'*Arabidopsis thaliana*. Tous ces motifs possèdent des propriétés qui les impliquent plus ou moins indirectement dans la régulation à la lumière des gènes.

Tous ces résultats ne sont qu'une première approche dans l'étude des phénomènes qui régulent l'expression des gènes de la GDC. En effet, il serait indispensable d'entreprendre (à l'aide de gènes reporteurs tels que la β-glucuronidase: GUS) des études de délétion de promoteur ou de mutations ponctuelles par l'obtention de plantes transgéniques. Nous pourrions ainsi, d'une part déterminer la fonction de ces régions consensus et, d'autre part caractériser d'autres éléments de régulation en cis au niveau de la région promotrice du gène codant pour la protéine T. Enfin, il serait tout à fait intéressant de savoir à quel stade du développement de la plante les différentes régions impliquées dans la régulation des gènes de la GDC interviennent, dans le but de mieux comprendre comment la mise en place des différentes protéines du complexe de la GDC est coordonnée.

CHAPITRE IV

CHAPITRE IV
EXPRESSION DE LA PROTEINE T DU COMPLEXE DE LA GLYCINE DECARBOXYLASE CHEZ *ESCHERICHIA COLI*

Afin de réaliser une étude structurale détaillée de la protéine T, il était nécessaire dans un premier temps de surexprimer la protéine T dans un organisme procaryote. En effet, l'obtention de cristaux de protéine dans le but d'étudier la structure tridimentionnelle de la protéine ainsi que les expériences de mutagénèse dirigée (nécessaires pour l'étude des sites catalytiques) ne sont possibles qu'après obtention de quantités relativement importantes d'enzymes purifiées.

I - ELABORATION D'UN GENE SYNTHETIQUE CORRESPONDANT A LA PROTEINE T MATURE EN VUE DE SON EXPRESSION DANS *ESCHERICHIA COLI*

A - RAPPEL SUR LE MECANISME DE SUREXPRESSION DE PROTEINES DANS UN ORGANISME PROCARYOTE

Le principe de base d'expression de protéines eucaryotes dans un système procaryote est fondé sur l'universalité du code génétique. En effet, dans un système d'expression procaryotique, le plasmide est un élément indispensable car il permet, en réquisitionnant la machinerie enzymatique de la cellule hôte (bactérie), de traduire un fragment d'ADN qu'on lui aura inséré. Ainsi, les vecteurs d'expression procaryotique utilisés pour la surproduction de protéines eucaryotes, contiennent généralement un promoteur ayant une forte affinité pour l'ARN polymérase bactérienne. Un gène étranger, inséré en aval d'un tel promoteur, sera transcrit en ARNm et traduit en une protéine qui correspondra à la séquence introduite dans le plasmide. La séquence d'ADN insérée doit se situer en aval de la séquence de Shine & Dalgarno (AGGA) et à 8-15 nucléotides en amont du codon initiateur de la traduction (AUG)

de la protéine à produire. En effet, la séquence de Shine & Dalgarno composée de 4 nucléotides, permet l'appariement spécifique de l'ARNm avec le ribosome procaryote et l'initiation de la synthèse protéique (voir Mandrup *et al.*, 1991).

Nous avons, dans un premier temps, entrepris d'essayer de produire chez *E. coli*, la protéine T mature de pois.

B - CONSTRUCTION ET CLONAGE DE LA FORME MATURE DE LA PROTEINE T

Afin de disposer d'une forme active de la protéine T, il est nécessaire d'exprimer uniquement la protéine mature. En effet, les procaryotes n'ont pas la possibilité d'exciser le peptide transit des protéines eucaryotes. La stratégie employée est décrite ci-dessous:

- Il s'agit de créer un insert dépourvu de sa séquence nucléotidique codant pour le peptide de transit.
- Il faut introduire, en amont de la séquence codant pour la protéine mature, un codon initiateur de la traduction dans le même cadre de lecture.
- Afin de pouvoir insérer le fragment dans les vecteurs d'expression procaryotique choisis, nous avons placé près du codon d'initiation, des sites de restriction.

La construction de l'insert correspondant à la protéine T mature à partir de l'ADNc TB2 a été réalisée à l'aide de la technique de PCR. La stratégie de clonage est schématisée dans la figure 45. La technique employée consiste à amplifier l'insert *in vitro*, avec la taq polymérase et deux oligonucléotides de synthèse spécifiques. Le premier oligonucléotide se situe dans la région 5' codant de l'ADNc et contient le codon d'initiation de la traduction et deux sites de sous-clonages *Eco* RI et *Nco* I. Le deuxième correspondant à la région 3' codant de l'ADNc possède un site de restriction *Bam* HI. L'insert ainsi amplifié de 1137 pb est digéré à l'aide des enzymes de restriction *Eco* RI et *Bam* HI puis sous-cloné dans le plasmide pUC 19 (voir figure 45). Le séquençage complet du fragment à l'aide d'amorces spécifiques nous a permis de détecter une erreur induite par la taq polymérase qui a engendré une mutation sur la première base d'un triplet codant pour une lysine. Cette mutation se situe à 85 pb en amont du codon stop TAG et remplace une adénine par une thymidine formant ainsi un codon stop (**A**AG→**T**AG). Cette mutation aboutirait donc à la formation d'une protéine incomplète. Le remplacement de la thymidine par une adénine à l'aide de la mutagenèse dirigée nous a permis de corriger l'erreur commise par la taq polymérase (voir figure 45 et Matériel et méthodes).

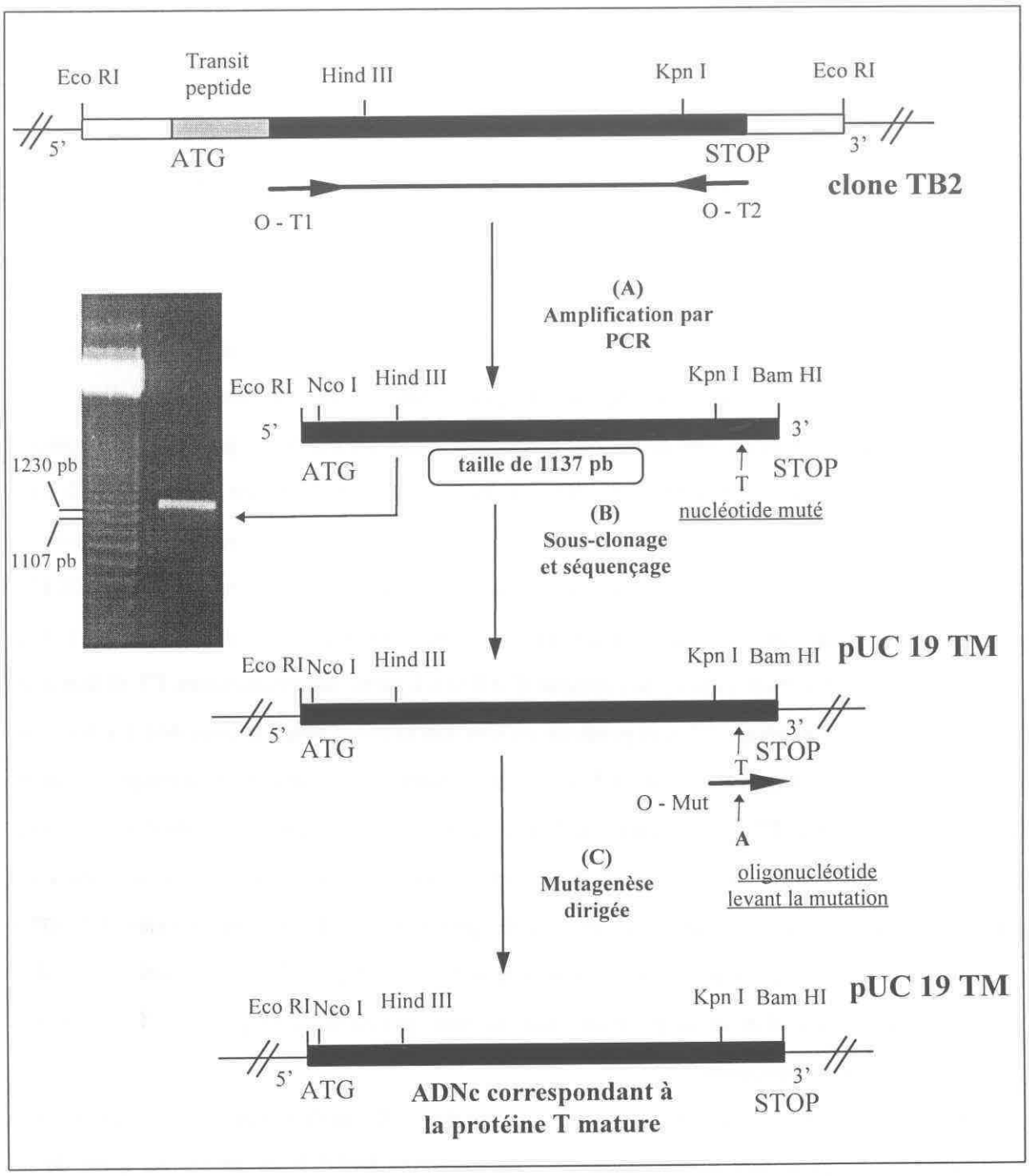


Figure 45: Construction de l'insert codant pour la protéine T mature: (A) Amplification par la *taq* polymérase, du fragment codant pour la protéine T mature à l'aide de l'oligonucléotide O-T1 contenant les sites *Eco*RI et *Nco*I (pGGGGGAATTCCATGGCTACAGAATCTGAGCTGA) et à l'aide de l'oligonucléotide O-T2 complémentaire contenant le site *Bam*H1 (pGGGGGGATCCCTATGAAGGCTTATAGTA). (B) Sous-clonage de ce fragment *Eco*RI-*Bam*H1dans le plasmide pUC 19 et séquençage de l'insert. (C) Mutagénèse dirigée du fragment muté (T correspond à la base mutée: une thymidine à la place d'une adénine) à l'aide de l'oligonucléotide O-Mut qui remplace la base mutée thymidine (T) par une adénine (**A**) (pGCAGGTACCAAAGTA**A**AGATCATTATTCGAGG). L'insert ainsi réparé code pour la forme mature de la protéine T.

Ayant en notre possession un plasmide nommé pUC 19 TM renfermant un insert codant pour la protéine T mature de pois, nous avons sous-cloné cet insert dans le vecteur d'expression procaryotique pET-3d.

II - UTILISATION DU SYSTEME D'EXPRESSION pET (Studier *et al.*, 1990)

A - PRINCIPE DE LA SUREXPRESSION DANS Escherichia coli

Ce système d 'expression est basé sur l'utilisation de l'ARN polymérase du phage T7 (Studier *et al.*, 1990). Cette polymérase de petite taille (elle est composée d'un seul polypeptide) est une enzyme très active qui possède une très forte affinité pour son promoteur.

La vitesse d'élongation des brins d'ARN par la polymérase du phage T7 est 5 à 10 fois supérieure à celle de l'ARN polymérase bactérienne. Les promoteurs, T7 lac et T7Φ10 (pΦ10: promoteurs très puissants de classe III), ont été introduits dans un dérivé du plasmide pBR322 pour créer toute une série de vecteurs d'expression procaryotique, dont pET-3d et pET-11d (Studier *et al.*, 1990). Pour induire la synthèse d'ARNm à partir des promoteurs T7, il faut que l'ARN polymérase du phage T7 soit produite au sein des cellules hôtes (figure 46). La série de vecteurs pET utilise une souche hôte d'*E. coli* qui contient une copie chromosomique du gène de l'ARN polymérase T7 sous contrôle de l'opérateur *lac*. Ainsi, dans ce système la présence d'un inducteur artificiel (IPTG) induit la production de la polymérase qui, à son tour, va transcrire spécifiquement l'insert situé en aval du promoteur de l'ARN polymérase T7 (Φ10 dans pET-3d et T7*lac* dans pET-11d). La souche d'*E. coli* BL21(DE3) (F^-, ompT, r_B-m_B-) a été choisie tout d'abord pour un premier essai de surexpression de la protéine T mature de pois pour les raisons suivantes:

- Elle possède une copie chromosomique du gène de l'ARN polymérase T7 et du gène *lac*I. Le produit de ce gène *lac*I bloque le promoteur bactérien lacUV5 en amont du gène de la polymérase. L'IPTG est en fait un inhibiteur de ce répresseur, et sa présence induit le promoteur lacUV5 et déclenche la synthèse de la polymérase (figure 46). Cependant, le blocage par ce système lac n'est pas total, assurant ainsi un niveau basal d'expression de la protéine cible même en l'absence d'IPTG.

- Elle est déficiente dans la synthèse de deux protéases (l'une est soluble et l'autre membranaire) codées par les gènes *lon* et *omp* respectivement. Ces caractéristiques sont un atout supplémentaire important pour l'expression de la protéine.

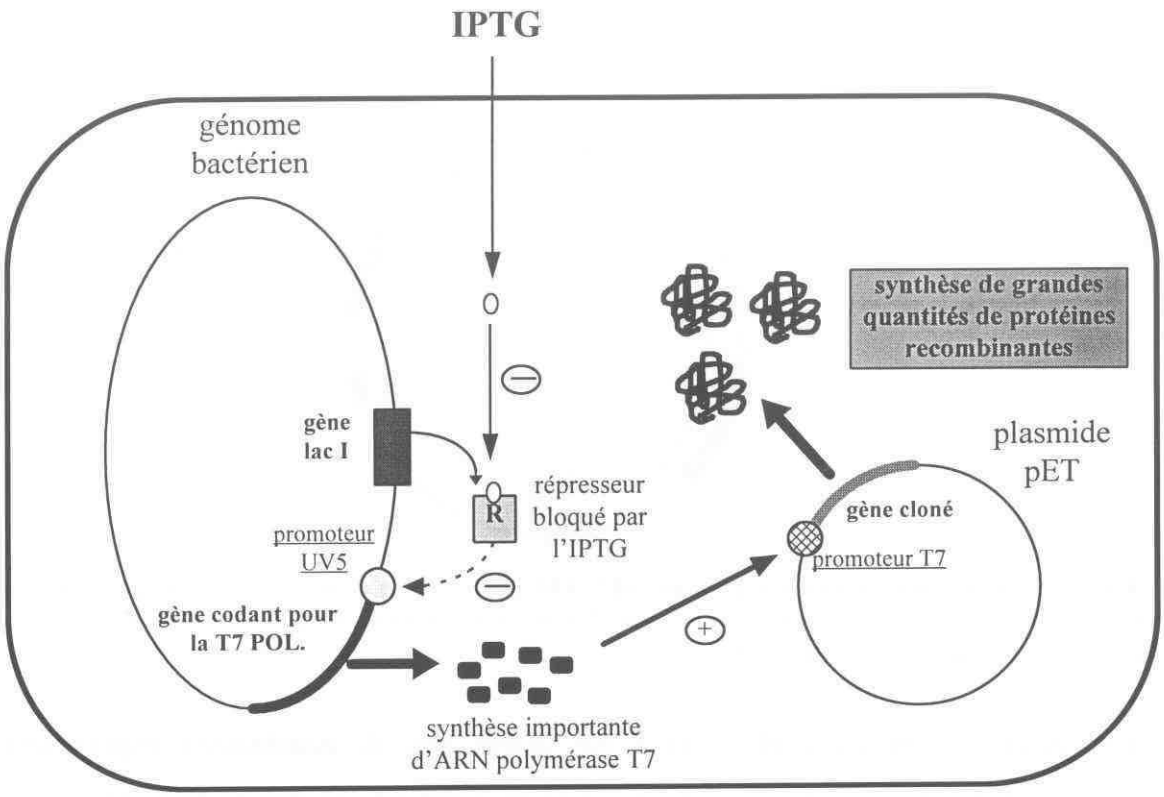


Figure 46: Système d'expression procaryotique (Studier *et al.*, 1990). T7 POL: ARN polymérase du phage T7. IPTG: inducteur artificiel. En absence d'inducteur (IPTG), le répresseur codé par le gène lac I bloque le promoteur Lac UV5 et l'ARN polymérase est alors faiblement produite (son activité peut être inactivée par le lysosozyme T7 qui est présent uniquement dans certaines souches). En présence d'IPTG (inhibiteur du répresseur), le promoteur lacUV5, totalement déréprimé, permet la synthèse d'importantes quantités d'ARN polymérases T7, qui transcrivent l'ADNc codant pour la protéine recombinante qui s'accumule alors dans la cellule bactérienne.

B - TENTATIVE D'EXPRESSION DE LA PROTEINE T MATURE DANS E.COLI EN UTILISANT LE VECTEUR D'EXPRESSION pET

Notre choix s'est porté sur le vecteur pET-3d (STRATAGENE). Le fragment (*Nco*I-*Bam*HI) de 1137 pb codant pour la protéine T mature a été extrait après digestion du plasmide pUC 19 TM et inséré dans le plamide pET (préalablement linéarisé par des digestions *Nco*I et *Bam*HI), afin de créer pET-TM (figure 47). Cette construction a permis de placer le codon d'initiation de la traduction (AUG), 8 nucléotides en aval de la séquence Shine-Dalgarno (SD) du vecteur:

SD 8 pb Met Ala Thr Glu Ser Glu Leu Lys

GAAGGAGATAACC ATG GCT ACA GAA TCT GAG CTG AAG

*Nco*I

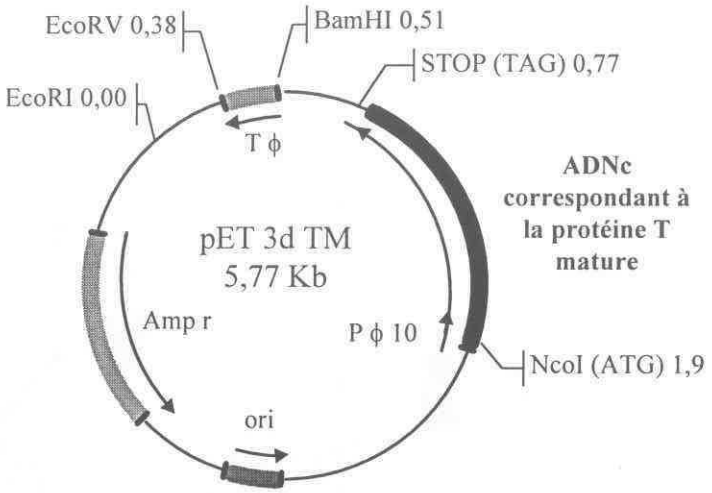


Figure 47: Carte du vecteur d'expression pET-TM. Amp r; gène de résistance à l'ampicilline. P ϕ 10: promoteur sous contrôle du répresseur *lac* (si placé dans une souche bactérienne lac Iq), inductible par addition d'IPTG. T ϕ: terminaison de la transcription.

Le vecteur d'expression pET-TM a été utilisé lors de nombreuses expériences de surexpression. Malheureusement, toutes nos tentatives ont échoué malgré de nombreux essais qui ont consisté à tester, dans différentes conditions de cultures, la souche d'*E. coli* BL21(DE 3) puis par la suite la souche B834(DE3)pLysS (Cf. Matériel et méthodes). Dans toutes ces expériences, les analyses électrophorétiques des protéines bactériennes ainsi que les analyses par Western-blot n'ont jamais révélé une accumulation de protéine T. A ce stade de nos recherches, nous avons pensé que la protéine T ne pouvait pas être produite car elle avait besoin d'être stabilisée par les autres protéines du complexe de la GDC et en particulier par la protéine H. Comme il a été montré précédemment, la protéine T a tendance à s'agréger lorsqu'elle se trouve à des concentrations élevées. L'addition de protéine H empêche ce phénomène d'agrégation. la protéine H semble jouer un rôle de « chaperon » pour la protéine T. La protéine H ayant déjà été surexprimée dans *E. coli* (Macherel *et al.*, 1996), nous avons pensé que la coexpression dans *E. coli* des deux protéines T et H, pouvait permettre de stabiliser la protéine T ou de favoriser sa synthèse. Nous avons alors entrepris la constructuion d'un vecteur comportant les deux séquences: HM (protéine H mature) et TM en tandem afin de coexprimer les deux protéines.

C - TENTATIVE DE COEXPRESSION DES PROTEINES T ET H MATURES

1 - Construction du vecteur d'expression procaryote pET-HTM contenant en tandem les séquences codant pour les protéines T et H matures

Un insert contenant la séquence de Shine et Dalgarno suivie de la séquence codant pour la protéine T mature a été excisé du plasmide pET-TM à l'aide des enzymes de restriction *Xba* I et *Bam* H1. Ce fragment a été inséré en aval de la séquence codant pour la protéine H mature, par clonage forcé dans le site *Bam* H1 ou *BPU* 1102 (figure 48) du plasmide pET-HM (Macherel *et al.*, 1996).

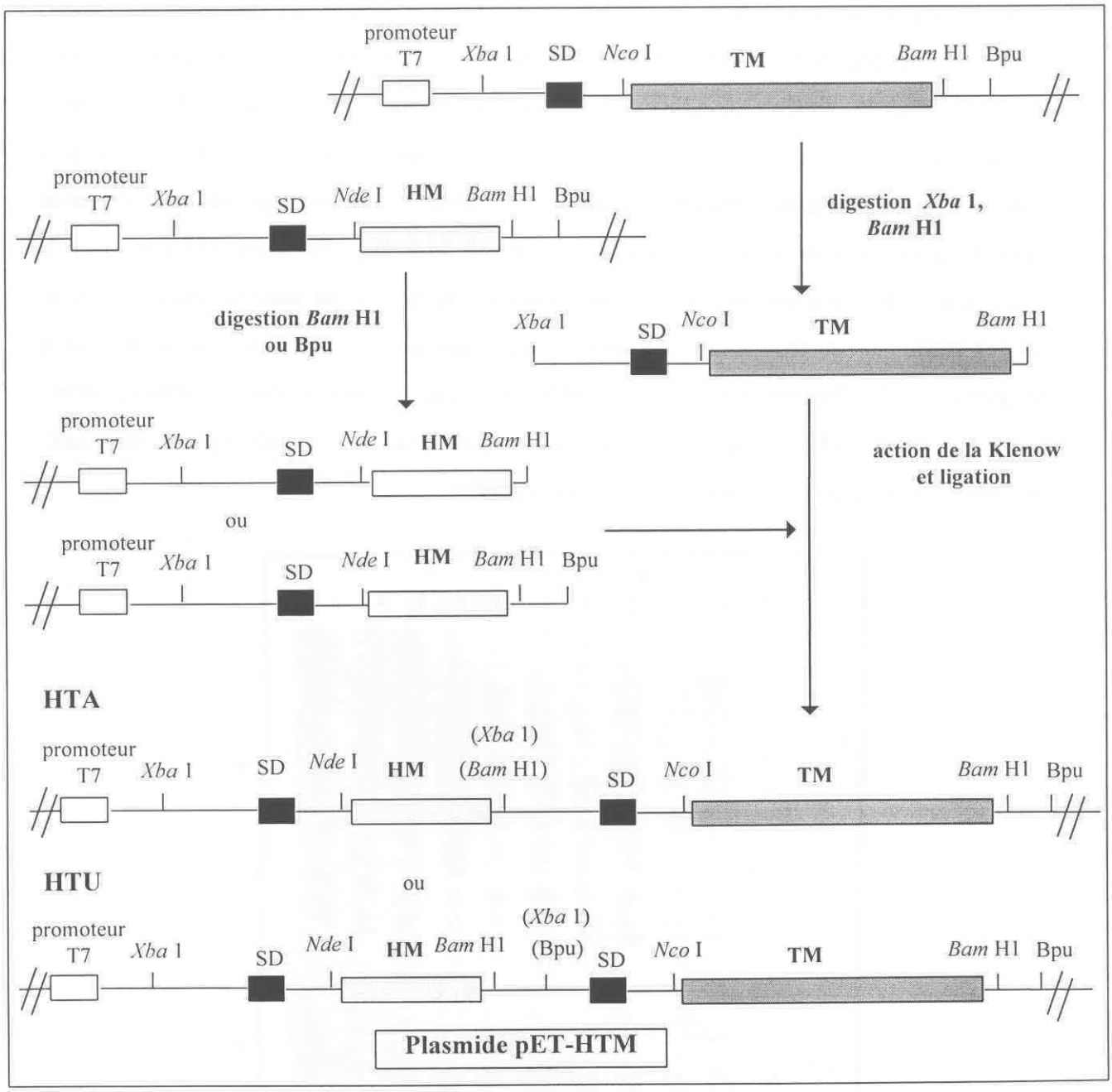


figure 48: Schéma représentant la stratégie de construction du plasmide en tandem pET-3d HTM. SD: séquence de Shine et Dalgarno. T7: Promoteur de la T7 polymérase. L'insert correspondant à la protéine T mature, digéré avec *Xba* I-*Bam* HI est ligué dans la construction pET-HM digérée soit avec *Bam* HI (construction HTA) soit avec *Bpu* (construction HTU).

Les vecteurs pET-HTA et pET-HTU ainsi formés peuvent permettre la transcription d'un seul ARNm codant pour la protéine H puis pour la protéine T. Les 2 constructions diffèrent légèrement par la longueur du fragment séparant le codon stop de la protéine H et le codon d'initiation de la protéine T.

2 - Essai de coexpression de la protéine H et T dans Escherichia coli à l'aide des vecteurs pET- HTA et pET-HTU

Nous avons utilisé les deux souches BL21(DE3) et B834(DE3)pLysS dans les mêmes conditions d'expression que précédemment. Malheureusement, les résultats de la coexpression de la protéine H et de la protéine T n'ont pas révélé une expression de la protéine T recombinante. En effet, la protéine H est rapidement exprimée (avec semble-t-il un meilleur rendement pour la souche transformée avec la construction HTU), alors qu'aucune protéine de la taille de la protéine T ne s'accumule, que ce soit avec la construction HTA ou avec la construction HTU (figure 49). Nous avons, toutefois, pu révéler par analyse immunologique, un très faible taux de production d'un polypeptide réagissant avec les anticorps dirigés contre la protéine T. Malheureusement, ce très faible niveau de production était systématiquement associé à une dégradation importante, matérialisée par la présence de polypeptides de masses moléculaires inférieures réagissant avec les anticorps.

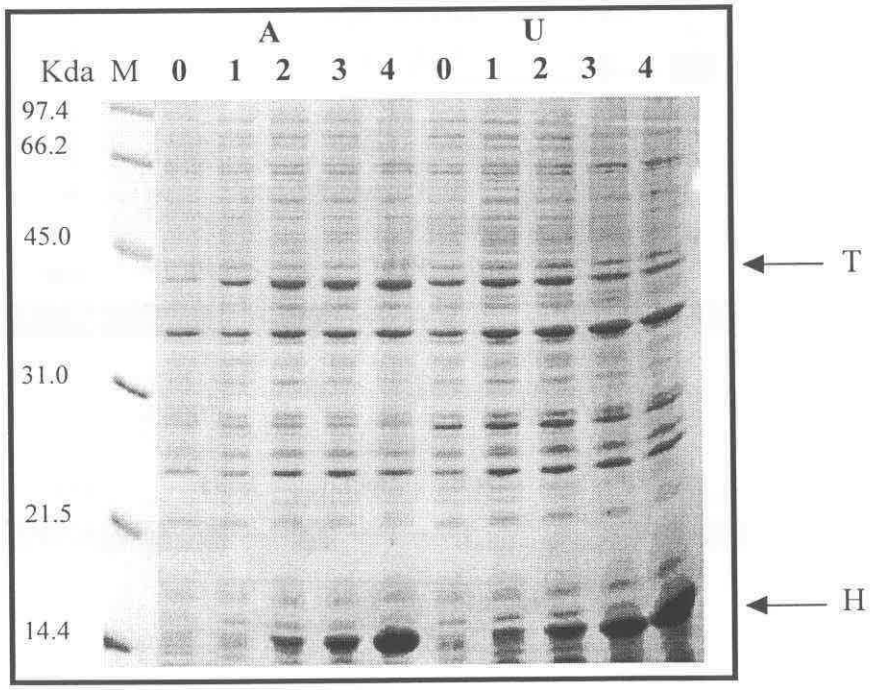


Figure 49: Protéines totales de l'extrait bactérien séparées sur gel de polyacrylamide (12%). Les Pistes 0, 1, 2, 3, 4 correspondent respectivement aux protéines de B834(DE3)pLysS transformée par pET-HT, avant l'induction, après, 30min, 1 h, 2 h et 3h d'induction. A: construction HTA, U: construction HTU (voir figure 48). M: marqueurs de masse moléculaire.

Toutes les tentatives d'expression de la protéine T dans *E. coli* ayant échoué, on peut supposer que l'absence d'expression de la protéine T puisse provenir d'un problème de traduction dans la bactérie. L'examen de la séquence du gène codant pour la protéine T de pois montre qu'elle contient 15 arginines dont 10 sont codées par les triplets AGG et AGA (voir figure 50). Ces codons sont très peu utilisés par la machinerie traductionnelle de *E. coli* et sont donc considérés comme rares. On les retrouve essentiellement dans des gènes faiblement exprimés comme le represseur lac (Farabaugh, 1978). La présence de tels codons ne pose généralement pas de problèmes s'ils sont peu abondants, par contre si leur nombre augmente et en particulier s'ils sont présents les uns à côté des autres, ceci peut réduire d'une manière significative l'expression de la protéine dans *E. coli* en provoquant des pauses au cours de la lecture du messager par les Ribosomes (Robinson *et al.*, 1984; Rosenberg *et al.*, 1993). De telles interruptions de lecture empêchent un couplage normal de la transcription et de la traduction, de ce fait les messagers peuvent être exposés à des dégradations par des ribonucléases ce qui rendrait impossible la traduction complète de la protéine. Un tel phénomène peut expliquer le fait que nous ayons vu au cours de nos expériences d'expression précédentes des produits de « dégradation » de la protéine T.

```
     GCT ACA GAA TCT GAG CTG AAG AAG ACA GTT CTG TAT GAC TTC CAT GTT GCT CAT GGT GGA AAG ATG
     Ala Thr Glu Ser Glu Leu Lys Lys Thr Val Leu Tyr Asp Phe His Val Ala His Gly Gly Lys MET

  67 GTT CCT TTT GCT GGT TGG AGT ATG CCT ATT CAA TAC AAA GAC TCA ATC ATG GAC TCA ACG TTA AAC
  23 Val Pro Phe Ala Gly Trp Ser MET Pro Ile Gln Tyr Lys Asp Ser Ile MET Asp Ser Thr Leu Asn

 133 TGT AGA CAG AAC GGT AGT CTT TTC GAT GTT TCT CAT ATG TGT GGT CTT AGT CTC AAG GGA AAG GAT
  45 Cys Arg Gln Asn Gly Ser Leu Phe Asp Val Ser His MET Cys Gly Leu Ser Leu Lys Gly Lys Asp

 199 GTT ATT TCG TTC CTT GAG AAG CTT GTC ATT GCG GAT GTT GCT GCT CTT GCC AAC GGA ACC GGG ACT
  67 Val Ile Ser Phe Leu Glu Lys Leu Val Ile Ala Asp Val Ala Ala Leu Ala Asn Gly Thr Gly Thr

 265 TTG ACT GTT TTC ACT AAT GAA AAG GGA GGG GCG ATC GAT GAT TCT GTG ATT ACT AAG GTG ACA GAG
  89 Leu Thr Val Phe Thr Asn Glu Lys Gly Gly Ala Ile Asp Asp Ser Val Ile Thr Lys Val Thr Glu

 331 GAT CAT CTT TAC TTG GTT GTC AAT GCT GGG TGC AGA GAT AAA GAT TTG GCT CAT ATT GAG GAG CAT
 111 Asp His Leu Tyr Leu Val Val Asn Ala Gly Cys Arg Asp Lys Asp Leu Ala His Ile Glu Glu His

 397 ATG AAG GCA TTC AAG GCC AAA GGC GGT GAT GTG TCG TGG CAC ATT CAC GAT GAG AGA TCT CTA CTC
 133 MET Lys Ala Phe Lys Ala Lys Gly Gly Asp Val Ser Trp His Ile His Asp Glu Arg Ser Leu Leu

 463 GCT CTT CAG GGT CCT CTT GCT GCT CCT GTT CTT CAA CAT CTG ACA AAA GAG GAT TTG AGC AAG CTG
 155 Ala Leu Gln Gly Pro Leu Ala Ala Pro Val Leu Gln His Leu Thr Lys Glu Asp Leu Ser Lys Leu

 529 TAC TTT GGG GAG TTC CGC GTG TTG GAC ATC AAT GGC TCG CAG TGC TTT CTC ACA CGG ACA GGG TAC
 177 Tyr Phe Gly Glu Phe Arg Val Leu Asp Ile Asn Gly Ser Gln Cys Phe Leu Thr Arg Thr Gly Tyr
```

```
595  ACT GGT GAA GAT GGA TTT GAG ATC TCA GTT CCT TCA GAG CAT GGA GTT GAA CTT GCC AAG GCA CTG
199  Thr Gly Glu Asp Gly Phe Glu Ile Ser Val Pro Ser Glu His Gly Val Glu Leu Ala Lys Ala Leu

661  CTG GAA AAA TCT GAA GGG AAA ATA AGA TTG ACA GGA CTC GGT GCT AGA GAT AGT CTA CGT CTC GAA
221  Leu Glu Lys Ser Glu Gly Lys Ile Arg Leu Thr Gly Leu Gly Ala Arg Asp Ser Leu Arg Leu Glu

727  GCT GGA CTG TGC TTA TAC GGA AAT GAC CTG GAA CAG CAC ATT ACA CCT ATT GAA GCA GGA CTG ACA
243  Ala Gly Leu Cys Leu Tyr Gly Asn Asp Leu Glu Gln His Ile Thr Pro Ile Glu Ala Gly Leu Thr

793  TGG GCT ATA GGT AAG AGA AGG AGA GCA GAA GGT GGT TTT CTA GGA GCT GAT GTT ATC CTG AAA CAG
265  Trp Ala Ile Gly Lys Arg Arg Arg Ala Glu Gly Gly Phe Leu Gly Ala Asp Val Ile Leu Lys Gln

859  CTC GCA GAC GGT CCT TCC ATT AGG CGT GTC GGT TTC ATT TCT TCT GGT CCA CCT CCA AGA AGC CAC
287  Leu Ala Asp Gly Pro Ser Ile Arg Arg Val Gly Phe Ile Ser Ser Gly Pro Pro Pro Arg Ser His

925  AGC GAG ATT CAA GAT GAA GGA GGA AAC AAC ATT GGT GAA GTG ACC AGT GGT GGA TTC AGT CCT TGT
309  Ser Glu Ile Gln Asp Glu Gly Gly Asn Asn Ile Gly Glu Val Thr Ser Gly Gly Phe Ser Pro Cys

991  CTC AAG AAG AAC ATA GCT ATT GGA TAT GTC AAA TCT GGA TTG CAC AAA GCA GGT ACC AAA GTA AAG
331  Leu Lys Lys Asn Ile Ala Ile Gly Tyr Val Lys Ser Gly Leu His Lys Ala Gly Thr Lys Val Lys

1057 ATC ATT ATT CGA GGT AAA CAG AAT GAA GGA GTT GTC ACG AAA ATG CCG TTC GTA CCC ACA AAA TAC
353  Ile Ile Ile Arg Gly Lys Gln Asn Glu Gly Val Val Thr Lys MET Pro Phe Val Pro Thr Lys Tyr

1123 TAT AAG CCT TCA TAG
375  Tyr Lys Pro Ser ---
```

Figure 50: Séquence nucléotidique correspondant à la protéine T mature et sa séquence en aminoacides déduite de la séquence nucléotidique. Les cadres en grisés indiquent les codons rares.

Une approche possible pour remédier à la présence des codons rares, pourrait s'effectuer par mutagénèse dirigée, de façon à les remplacer par des codons normalement utilisés dans les bactéries. Cela représente une expérimentation très lourde, nous avons alors préféré adapter une approche inspirée par les travaux de Shenk *et al.* (1995) qui consiste à fournir à la bactérie l'ARN de transfert codant pour le codon le plus rare. D'autre part, le produit du gène *arg*U (ou DNA Y) a un effet positif sur la réplication et permet d'augmenter le nombre de copies du plasmide (Garcia *et al.*, 1986). Le codon AGA est traduit chez E. coli par l'ARNt^{Arg4} codé par le gène *arg*U (figure 51) (Garcia *et al.*, 1986). Schenk *et al.*, (1995) ont réussi à surexprimer des protéines dont le gène contient de nombreux codons rares, en insérant une cassette d'expression extraite de vecteur pET dans un plasmide comportant le gène *arg*U. Notre approche inverse a consisté à insérer dans nos vecteurs d'expression, un gène synthétique codant pour l'ARNt^{Arg4}.

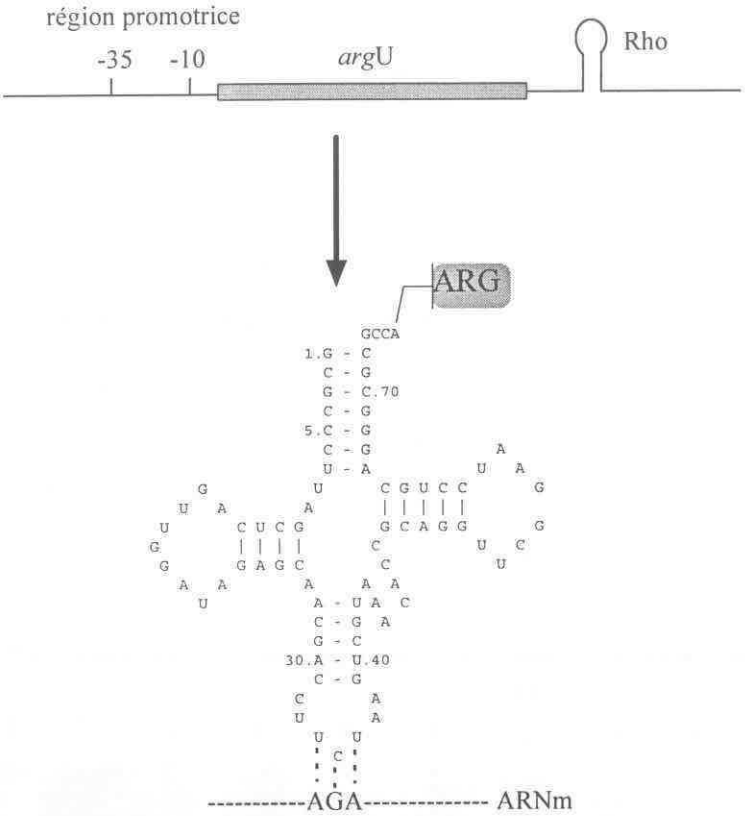


figure 51: Structure secondaire potentielle de l'ARNt$_{AGA/AGG}$ rare codé par le gène *arg*U.

D - MODIFICATION DU SYSTEME D'EXPRESSION DANS E. COLI PAR INSERTION DANS LES VECTEURS pET-TM ET pET-HTA ET pET-HTU DU GENE Arg U CODANT POUR UN ARNt^{ARG} RARE

1 - Insertion du gène argU dans les vecteurs pET-TM, pET-HTA et pET-HTU

Par l'intermédiaire du Genetic Stock Research Center (National Institute of genetics, Mishima, Japan), nous avons obtenu une souche d'*E. coli* possédant un plasmide PLC 22-1 porteur d'un fragment génomique d'*E. coli* d'environ 7 kpb contenant le gène *arg*U (Nishimura *et al.*, 1992). Ce plasmide a été amplifié au chloramphénicol et purifié puis, a servi de matrice pour amplifier par PCR un fragment de 276 pb contenant le promoteur, la séquence codante et le terminateur Rho du gène *arg*U caractérisé par Garcia *et al.* (1986). Des sites *Eco* RI ont été incorporés dans la région 5' des oligonucléotides et ont donc permis

d'insérer ce fragment dans le site *Eco* RI unique des vecteurs pET-TM, pET-HTA et pET-HTU dans 2 orientations possibles, sens et anti-sens par rapport à la transcription du gène T. Nous avons ainsi obtenu les plasmides pETY-TM1 (sens), pETY-HTA1, (sens), pETY-HTU1 (sens) et pETY-TM1 (antisens), pETY-HTA1, (antisens), pETY-HTU1 (antisens).

2 - Expression et coexpression des protéines H et T dans Escherichia coli à l'aide des vecteurs pETY-TM, HTA, HTU

Les souches de *E. coli* B834(DE3)pLysS transformées par les 6 types de constructions ont été cultivées dans les mêmes conditions que celles décrites dans le chapitre Maériel et méthodes. Ainsi, l'électrophorèse, en conditions dénaturantes, des échantillons prélevés à différents temps d'induction, révèle une rapide accumulation d'une protéine dont la masse relative (43 kDa) est identique à celle de la protéine T de pois et ceci, dans les 6 constructions (figure 52). De plus, nous remarquons que cette protéine est coexprimée en parallèle avec la protéine H dans les constructions en tandem.

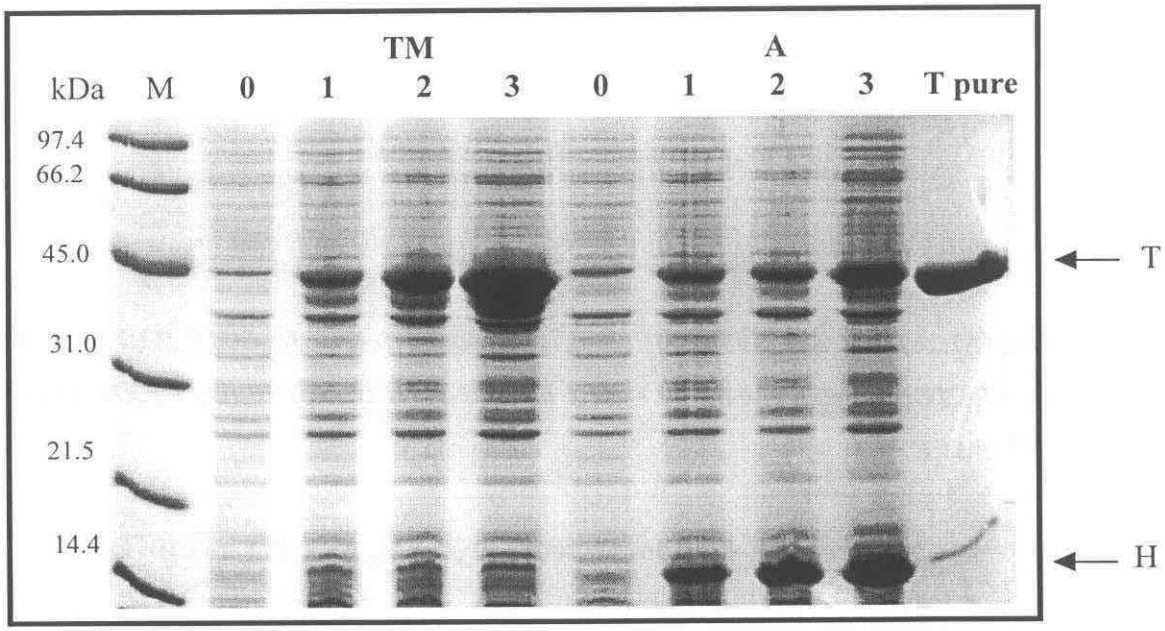


Figure 52: Expression de la protéine T mature de pois chez *E. coli*. Les protéines des bactéries B834(DE3)pLysS transformées ont été séparées sur un gel de polyacrylamide (12%) en conditions dénaturantes (SDS). Les pistes 0, 1, 2 et 3 correspondent respectivement aux protéines de E coli avant, et après 1h, 2h et 4h d'induction. TM: construction pETY-TM1; A: construction pET-HTA1; M: marqueur de masse moléculaire.

III - DISCUSSION-CONCLUSION

Suite aux échecs répétitifs survenus lors des différents essais d'expression de la protéine T dans un système procaryote classique, nous avons réussi à développer au laboratoire une technique de production de la protéine T recombinante chez *E. coli*. En effet, l'étude de la séquence nucléotidique de l'ADNc correspondant à la protéine T mature nous a révélé une forte teneur en codons AGA et AGG qui sont traduits chez *E coli* par des ARNt rares, empêchant toute traduction complète de l'ADNc. De ce fait, pour produire la protéine T en grande quantité, il a fallu adapter nos vecteurs classiques d'expression au système procaryotique, en leur incorporant le gène *arg*U codant pour l'ARNt limitant. Les premiers essais d'activité de la protéine T réalisés à partir d'un extrait brut d'*E. coli*, montrent que malheureusement la protéine T n'est pas active. Les analyses effectuées récemment afin de localiser la protéine T dans *E. coli* révèlent que cette protéine surexprimée contrairement à la protéine H de pois (Macherel *et al.*, 1996) ou à la protéine T de poulet (Okamura-Ikeda *et al.*, 1992), n'est pas soluble et s'accumule sous forme de corps d'inclusion.

Des études préliminaires effectuées récemment au Laboratoire de Spectrométrie de Masse ont montré que la protéine T possèderait 3 ponts disulfures (résultats à confirmer). Les ponts disulfures ont une grande importance dans la structure des protéines et l'absence de ceux-ci entraîne un mauvais repliement de la protéine pouvant rendre celle-ci inactive. Il est possible que le milieu très réducteur des bactéries ne permette pas l'oxydation des groupements SH des cystéines de la protéine T provoquant de ce fait l'agrégation de la protéine incorrectement repliée. Devant ce problème, une des solutions envisageables serait alors d'exporter la protéine T dans le périplasme, afin de favoriser la formation des ponts disulfures. La formation des ponts disulfures est catalysée chez *E. coli* par 2 protéines principalement: DsbA, et DsbB. La protéine DsbA possède des ponts disulfures très oxydatifs et DsbB est nécessaire à la réoxydation de DsbA. Ainsi, les groupements SH libres de la protéine seraient oxydés lors de leur passage dans le périplasme (pour une revue Bardwell, 1994). L'autre solution serait d'utiliser le mutant *trxB* sélectionné par Derman *et al.* (1993). En effet, cette souche possède une mutation qui touche le gène codant pour la thioredoxine reductase, réduisant, voire éliminant l'activité de cette enzyme. De ce fait, ce mutant permettrait la formation de ponts disulfures au sein du cytoplasme de *E. coli*. Grâce à cette souche, il serait donc possible d'exprimer des protéines à pont(s) disulfure(s) au sein du cytoplasme

(Bardwell, 1994). Ainsi, l'utilisation de la souche mutante *trxB* pourrait être une solution à l'éventuel problème de conformation que pourrait rencontrer la protéine T recombinante. Cette souche serait un outil précieux pour purifier la protéine T sous sa forme active. Une autre alternance va consister à purifier la protéine T à partir de corps d'inclusion puis à la solubiliser par l'utilisation d'agents dénaturant tels que l'urée ou le chlorure de guanidine et enfin à la renaturer dans des conditions permettant un repliement correct. Cette stratégie biochimique s'est révélée fructueuse dans de nombreux cas.

CONCLUSION GENERALE ET PERSPECTIVES

CONCLUSION GENERALE ET PERSPECTIVES

L'étude de la structure et de la régulation du complexe mitochondrial de la glycine décarboxylase est depuis une dizaine d'années un thème de recherche majeur du laboratoire. Ainsi, lorsque nous avons commencé ces travaux, les ADNc correspondant aux protéines H, L et P du complexe de la GDC et le gène codant pour la protéine H avaient été isolés et caractérisés par Macherel *et al.* (1990, 1992), Bourguignon *et al.* (1992) et Turner *et al.* (1992). Nous avons commencé par le criblage d'une banque d'ADNc, réalisée à partir d'ARN messagers de feuilles de pois, avec des anticorps polyclonaux dirigés spécifiquement contre la protéine T. Nous avons ainsi isolé et séquencé des clones qui ont permis de déduire la structure primaire du précurseur de la protéine T (408 acides aminés) dont l'identité ainsi que l'exactitude de la séquence peptidique ont été confirmées grâce à plusieurs éléments:

- La séquence de certains peptides obtenue après digestion chimique de la protéine T et microséquençage (35% de la protéine totale), est strictement identique à la séquence primaire déduite de la séquence nucléotidique des ADNc isolés. La connaissance de la séquence amino-terminale de la protéine T nous a permis de distinguer le peptide de transit composé de 30 acides aminés.

- La masse de la protéine T purifiée, analysée par spectrométrie de masse correspond à la valeur déduite de la séquence nucléotidique de l'ADNc. De même, les masses des peptides séparés par HPLC et provenant du clivage chimique de la protéine T, mesurées par spectrométrie de masse ont des valeurs similaires à celles calculées à partir de la structure primaire déduite de l'ADNc. De plus, grâce à cette technique, nous avons pu montrer que le tétrahydrofolate n'était pas lié de façon covalente à la protéine T.

Les analyses de comparaison de séquences ont révélé que la séquence primaire de la protéine T de pois présente de fortes homologies avec celles des autres protéines T (homme, poulet, etc...). Ainsi, plusieurs régions fortement conservées ont pu être caractérisées, notamment une région riche en acides aminés chargée positivement qui pourrait être impliquée dans les interactions avec les résidus polyglutamate du tétrahydrofolate. Cependant, des expériences de mutagénèse dirigée seront nécessaires pour définir les acides aminés qui jouent un rôle clef dans la réaction enzymatique et en particulier, ceux qui interagissent dans

la reconnaissance de la protéine H et du tétrahydrofolate. Pour commencer cette étude, nous avons utilisé l'ADNc correspondant à la protéine T de pois afin de produire en grande quantité cette enzyme dans un hôte procaryote. Cependant, contrairement à la protéine H et malgré de nombreuses tentatives, nous n'avons pas réussi à surexprimer la protéine T en utilisant des vecteurs classiques d'expression. L'analyse de la séquence nucléotidique de l'ADNc a révélé un nombre élevé de codons rares qui empêchaient l'expression de la protéine recombinante. Grâce à l'insertion du gène *arg*U (codant pour l'ARNt^{arg}) dans le vecteur d'expression, la protéine T a pu être surexprimée. Cependant, les dernières expériences effectuées montrent que la protéine T s'accumule en très grandes quantités (200 mg de protéine T produits par litre de culture) dans *E. coli* sous forme de corps d'inclusion. Le milieu fortement réducteur des bactéries pourrait être une entrave à la formation d'éventuels ponts disulfures de la protéine T et de ce fait la rendre insoluble. Pour résoudre ce problème, plusieurs solutions sont envisagées: exporter la protéine recombinante dans le périplasme, utiliser une souche mutante *trxB* qui permettrait la formation des ponts disulfures au sein du cytoplasme et enfin, solubiliser la protéine T par des agents dénaturants puis essayer de la renaturer en trouvant les conditions qui permettent un repliement correct de la protéine. Les premières tentatives de solubilisation et de renaturation effectuées au laboratoire sont tout à fait prometteuses. En effet, la solubilisation de la protéine T présente dans les corps d'inclusion (par de l'urée 6 M) suivie d'une renaturation lente par dialyse en présence de protéine H a permis d'obtenir une protéine T active. Le pourcentage de renaturation est pour l'instant faible (2-3 %) mais de nombreuses conditions restent à tester pour améliorer ce rendement. Ces expériences confirment l'existence de fortes interactions entre les protéines T et H et montrent que la protéine H joue, en plus de son rôle catalytique, un rôle dans la stabilisation de la protéine T.

En collaboration avec C. Cohen Addad du Laboratoire de Cristallographie Macromoléculaire de l'Institut de Biologie Structurale de Grenoble, des premières expériences de cristallisation de la protéine T en présence de différentes formes de la protéine H (Oxydée et chargée en méthylamine) ont été entreprises. En effet, il est fort probable pour les raisons évoquées précédemment que l'étude structurale de la protéine T passe par la co-cristallisation de la protéine T avec la protéine H. Cette étude permettra de déterminer les régions d'interaction entre ces deux protéines. Ce travail est mené en parallèle avec des expériences de « cross-linking » qui ont pour but également de déterminer les sites d'interactions entre les protéines H et T.

A l'aide de tous les outils moléculaires (ADNc codant pour les protéines H, P, L et T) ainsi que des anticorps dirigés contre toutes les protéines, nous avons entrepris d'étudier la régulation et la biogénèse de ce complexe. Il est clairement établi que seules les mitochondries issues de feuilles chlorophylliennes sont capables d'oxyder la glycine. En effet, l'analyse des protéines H, P et T et de leurs transcrits, a révélé une très forte expression des gènes correspondant et une accumulation des protéines au sein de la matrice mitochondriale des feuilles de plantes cultivées à la lumière. A l'opposé, la faible capacité d'oxydation des mitochondries provenant de feuilles exposées à l'obscurité correspond à une nette diminution des protéines P, H et T, et à un niveau négligeable des transcrits. Quant à la protéine L, son implication dans d'autres complexes multienzymatiques comme la pyruvate dehydrogénase (Bourguignon *et al.*, 1995) lui confère une régulation différente, à savoir une expression dans tous les tissus et une faible influence de la lumière. Toutes ces observations ayant été réalisées sur des feuilles matures, nous avons alors cherché à suivre la mise en place de ce complexe au cours du développement du pois. Les résultats obtenus ont clairement montré que la capacité d'oxydation de la glycine des mitochondries isolées aux jeunes stades de développement est quasiment nulle, alors qu'elle subit une induction importante lorsque les feuilles deviennent des capteurs solaires efficaces. En effet, à ce stade, l'accumulation au sein de la matrice mitochondriale du complexe de la GDC est telle que les mitochondries voient leur densité augmenter. Ainsi, contrairement à la ribulose 1,5 bisphosphate carboxylase/oxygénase (ou Rubisco) et à la protéine L, qui sont présentes dès les tout premiers stades de développement (4 jours), les protéines spécifiques de la GDC apparaissent beaucoup plus tardivement (9 jours) à un stade où les feuilles, totalement déployées, reçoivent une intensité lumineuse maximale. Il semble donc que la photorespiration apparaît lorsque la photosynthèse est bien engagée.

L'analyse des transcrits correspondant aux protéines P, H et T provenant de feuilles de pois exposées à la lumière a révélé que le taux de transcription des gènes augmente de façon coordonnée avec un pic au 7ième jour de croissance. Ce profil d'évolution, similaire à celui de la petite sous-unité de la Rubisco (rbcS) et à celui de la protéine liant les chlorophylles a,b (cab), suggère ainsi l'existence de mécanismes de régulation proches au niveau de la transcription de leurs gènes. Cependant, alors que les transcrits correspondant à la Rubisco sont traduits dès les premiers stades du développement, ceux de la GDC apparaissent

beaucoup plus tardivement. Il semblerait donc qu'il existe un contrôle post-transcriptionnel de l'expression des gènes de la GDC. A ce titre, il serait intéressant de déterminer les facteurs responsables de cette régulation. En effet, des expériences préliminaires laissent à penser qu'une régulation de type métabolique serait envisageable.

Suite à ces observations et à la caractérisation par Macherel *et al.* (1992) du gène codant pour la protéine H, nous nous sommes dirigés vers l'étude structurale du gène codant pour la protéine T dans le but de découvrir d'éventuels éléments communs de régulation. Ainsi, à l'aide de l'ADNc correspondant à la protéine T, nous avons criblé une banque génomique de pois et isolé le gène codant pour la protéine T. Ce dernier, d'une taille approximative de 3 kpb, comporte quatre exons interrompus par trois introns et deux sites d'initiation de transcription dont l'un ressemble à l'élément Inr décrit par Smale et Baltimore (1989). En effet, cette séquence riche en pyrimidine que l'on retrouve chez le gène codant pour la protéine H est capable d'initier la transcription de façon indépendante sans être sous le contrôle d'une boîte TATA. Nous avons pu, également, caractériser plusieurs boîtes de régulation communes grâce à l'analyse de la région promotrice par comparaison de séquences avec des gènes subissant une induction à la lumière comme ceux de la rbcS, de cab et de la protéine H. Notamment, nous avons observé la présence d'une boîte GT, impliquée dans l'induction à la lumière, et d'une région riche en nucléotides AT dont la fonction semble être soit activatrice, soit répressive (Datta et Cashmore, 1989; Bustos *et al.*, 1989). L'analyse de ces séquences, par des expériences de gel retard, a montré que des protéines nucléaires étaient capables d'interagir avec ces séquences. En effet, des analyses par compétitions ont permis de déterminer que ces éléments de régulation en cis sont probablement apparentés à ceux présents dans la région promotrice des gènes codant pour rbcS. Il est donc fort probable que les gènes codant pour les protéines du complexe de la GDC possèdent certains mécanismes de régulation identiques à ceux de la petite sous-unité de la Rubisco. Cependant, une analyse fonctionnelle plus détaillée de la zone promotrice du gène de la protéine T par l'utilisation de plantes transgéniques ou par expression transitoire dans des protoplastes, permettra de caractériser d'autres régions impliquées dans la régulation du gène codant pour la protéine T et de confirmer celles déjà mises en évidence.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

Albert B., Bray D., Lewis J., Raff M., Roberts K. & Watson J. D. (1983). Molecular Biology of the Cell. New York: Garland.

Alberth W., Straub K. M. & Burlinghame A. L. (1982). Secondary ion mass spectrometry with cesium ion primary beam and liquid target matrix for analysis of bioorganic compounds. *Anal. Chem.* **54**, 2029-2034.

Arneson D., Chien L. T., Chance P. & Wilroy R. S. (1979). *Pediatrics* **63**, 369-373.

Arron G. P., Spalding M. H. & Edwards G. E. (1979). Isolation and oxydative properties of intact mitochondria from the leaves of *Sedum praealtum*. *Plant Physiol.* **64**, 182-186.

Ayer D. A. & Dynan W. S. (1988). Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. *Mol. Cell. Biol.* **8**, 2021-2033.

Badger M. R. & Dean Price G. (1994). The role of carbonic anhydrase in photosynthesis. *Ann. Rev. Plant Physiol.* **45**, 369-392.

Badger M. R., Kaplan A. & Berry J. A. (1977). Carnegie Inst. Washinton, Yearb. **76**, 362-366.

Badger M. R., Kaplan A. & Berry J. A. (1978). Carnegie Inst. Washinton, Yearb. **77**, 251-261.

Banno K. & Horimoto S. (1991). Separation and quantitation of nicergoline an drelated substances by high-performance liquid chromatography/atmospheric pressure ionization mass spectrometry. *Chromatographia* **31**, n° 1/2, 50-54.

Barber M., Bordoli R., Sedgwick R. & Tyler A. (1981). Fast atom bombardments of solids as an ion source in mass spectrometry. *Nature* **293**, 270-275.

Bardwell J. C. A. (1994). Building bridges: disulphide bond formation in the cell. *Mol. Microbiol.* **14**, 199-205.

Beckey H. D. & Schulten H. R. (1975). Field desorption mass spectrometry. *Angew. Chem. Int. Ed. Engl.* **14**, 403-415.

Benfey P. N., Ren L. & Chua N. -H. (1990). Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *EMBO J.* **9**, 1677-1684.

Benson A. A. et Calvin M. (1950). The path of carbon in photosynthesis: respiration and photosynthesis. *J. Exp. Bot.***1**, 63-68

Berry J., Boynton J., Kaplan A. & Badger M. (1976). Growth and photosynthesis of *Chlamydomonas reinhardtii* as a function of CO_2 concentration. *Carnegie Inst. Washinton, Yearb.* **75**, 423-432.

Besson V., Rebeille F., Neuburger M., Douce R. & Coosins E. A. (1993). Effects of tetrahydrofolate polyglutamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. *Biochem. J.* **292**, 425-430.

Blakley R. L. (1969). The biochemistry of folic acid and related pteridines. In Frontiers of biology, Vil. XIII (A. Neuburger and E. L. Tatum, Eds). Elsiever, New York.

Blakley R. L. (1984). In Folates and pterins (Blakley R. L. & Benokovic S. J., eds) vol. 1, pp. 191-253, Wiley Interscience, New York.

Bligny R. & Douce R. (1978). Calcium-dependent lipolytic acyl-hydrolase activity in purified plant mitochondria. *Biochem. Biophys. Acta* **529**, 419-428.

Boowes G., Ogren W. L. & Hageman R. H. (1971). Phosphoglycolate production catalysed by Ribulose diphosphate carboxylase. *Biochem. Biophys. Res. Commun.* **45**, 716-722.

Bourguignon J. (1988). Etude du complexe enzymatique impliqué dans l'oxydation de la glycine par les mitochondries de tissus foliaires. Doctorat de l'université Joseph Fourier de Grenoble.

Bourguignon J., Neuburger M. & Douce R. (1988). Resolution and characterization of the glycine-cleavage reaction in pea leaf mitochondria. *Biochem. J.* **255**, 169-178.

Bourguignon J., Macherel D., Neuburger M. & Douce R. (1992). Isolation, characterization, and sequence analysis of a cDNA clone encoding L-protein, the dihydrolipoamide dehydrogenase component of the glycine cleavage system from pea-leaf mitochondria. *Eur. J. Biochem.* **204**, 865-873.

Bourguignon J., Vauclare P., Merand V., Forest E., Neuburger M. & Douce R. (1993). Glycine decarboxylase complex from higher plants: molecular cloning, tissue distribution and mass spectrometry analyses of the T protein. *Eur. J. Biochem.* **217**, 377-386.

Bourguignon J., Merand V., Rawsthorne S., Forest E. & Douce R. (1996). Glycine decarboxylase and pyruvate dehydrogenase complexes share the same dihydrolipoamide dehydrogenase in pea leaf mitochondria: evidence from mass spectrometry and primary-structure analysis. *Biochem. J.* **313**, 229-234

Bowes G. (1985). Pathways of CO2 fixation by aquatic organism, in *Inorganic Carbon Uptake by aquatic Photosynthetic Organisms*, Lucas, W. J. and Berry, J. A., Eds., American Society of Plant Physiol., Rockville. Md., 187.

Bradford M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein, utilizing the principle of protein-dye binding. *Anal Biochem.* **72**, 248-254.

Breathnach R. & Chambon P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**, 349-383.

Brody T., Shin Y. S. & Stokstad E. L. R. (1976). Rat brain folate identification. *J. Neurochem.* **27**, 409-413.

Bucher P. (1990). Weight matrix descriptions of four eucaryotic RNA polymérase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Cell. Biol.* **212**, 563-578.

Buratowski S. (1994). The basics of basal transcription by RNA polymerase II. *Cell.* **77**, 1-3.

Bustos M. M., Guiltinan M. J., Jordano J., Begum D., Kalkan F. A. & Hall T. C. (1989). Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a french bean β-phaseolin gene. *Plant Cell.* **1**, 839-853.

Caprioli R. M. & Suter M. J. F. (1992). Continuous-flow fast atom bombardment: recent advances and applications. *Int. J. Mass Spectrom. and Ion Processes* 118-119, 449-476.

Caprioli R. M. (1988a). Continuous-flow fast atom bombardment mass spectrometry. *Trends in Analyt. Chem.* **7**, 328-333.

Caprioli R. M. (1988b). One-line Fast Atom Bombardment Analysis of Dynamic Biological Systems. Biomed. Environ. Mass Spectrom. **16**, 35-39.

Carcamo J., Buckbinder L. & Reinberg D. (1991). The initiator directs assembly of a transcription factor IID-dependent transcription complex. *Proc. Natl. Acad. Sci.* USA. **88**, 8052-8056.

Castresana C., Staneloni R., Malik V. S. & Cashmore A. R. (1987). Molecular characterization of two clusters of genes encoding the type I CAB polypeptides of PSII in *Nicotiana plumbaginifolia*. *Plant Mol. Biol.* **10**, 117-126.

Castresana C., Garcia-Luque I., Alonso E., Malik V. S. & Cashmore A. R. (1988). Both positive and negative elements mediate expression of a photoregulated CAB gene from *Nicotiana plumbaginifolia*. *EMBO J.* **7**, 1929-1936.

Chen W., Tabor S. & Struhl K. (1987). Distinguishing between mechanisms of eucaryotic transcriptional activation with bacteriophage T7 polymerase. *Cell.* **50**, 1047-1055.

Cichowicz D. J. & Shane B. (1987). Mammalian folylpoly-γ-glutamate synthetase: purification and general properties of hog liver enzyme. *Biochemistry* **26**, 504-512.

Clandinin M. T. & Cossins E. A. (1972). Localisation and interconversion of tetrahydropteroylglutamates in isolated pea mitochondria. *Biochem. J.* **128**, 29-40.

Cohen-Addad C., Pares S., Sieker L. Neuburger M. & Douce R. (1995). The lipoamide arm in the glycine decarboxylase complex is not freely swinging. *Nature Struct. Biol.* **2**, 63-68.

Colman B. & Cook C. M. (1985). Photosynthetic characteristics of the marine macrophytic red alga *Rhodymenia palmata*: evidence for bicarbonate transport, in *Inorganic Carbon Uptake by aquatic Photosynthetic Organisms*, Lucas, W. J. and Berry, J. A., Eds., American Society of Plant Physiol., Rockville. Md., 97.

Connelly S. & Manley J. L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2**, 440-452.

Cossins E. A. & Shah S. P. J. (1972). Pteroylglutamates of higher plant tissues. *Phytochemestry* **11**, 587-593.

Cossins E. A. (1980). One-carbone metabolism. In the biochemestry of plants, vol. 2 Metabolism and respiration (D.D.Davies, Ed), pp 365-418. Academic press, NewYork.

Cossins E. A. (1984). Folates in biological materials. In Folates and Pterins (R. L. Blakley et S. J. Benkovic, Eds), vol.1, pp 1-59.

Covey T. R., Bonner R. F., Shushan B. I. & Henion J. D. (1988). The determination of protein, oligonucleotide and peptide molecular weights by ionspray mass spectrometry. *Rapid Commun. Mass Spectrom.* **2**, 249-256.

Covey T. R., Huang E. C. & Henion J. D. (1991). Structural characterization of protein tryptic peptides via liquid chromatography/mass spectrometry and collision-induced dissociation of their doubly charged molecular ions. *Anal. Chem.* **63**, 1193-1200.

Curtis D., Lehmann R. & Zamore P. D. (1995). Translational Regulation in Development. *Cell* **81**, 171-178.

Cybulski R. L. & Fisher R. R. (1976). Intramitochondrial localization and proposed metabolic significance of serine transhydroxymethylase. *Biochemistry*, **15**, 3183-3187.

Dai Z., Ku M. S. B. & Edwards G. E. (1993). The CO_2-concentrating mechanism and photorespiration. *Plant Physiol.* **103**, 83-90.

Datta N. & Cashmore A. R. (1989). Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. *Plant Cell.* **1**, 1069-1077.

Day D. A. & Wiskich J. T. (1980). Glycine transport by pea leaf mitochondria. *FEBS Letters* **112**, 191-194.

Day D. A., Neuburger M., & Douce R. (1985). Biochemical characterization of chlorophyll-free mitochondria from pea leaves. *Aust. J. Plant Physiol.* **12**, 219-228.

de Groot C. J., Boelleverts V., Touwen B. C. L. & Hommes F. A. (1978). *Prog. Brain Res.* **48**, 199-207.

Delieu T. & Walker D. A. (1972). An improved cathode for the measurement of photosynthetic oxygen evolution by isolated chloroplasts. *New Phytol.* **71**, 201-225.

Derman A. I., William A. P. Belin D. & Beckwith J. (1993). Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science*. **262**, 1744-1747.

Doherty A. J., Ashford S. R., Brannigan J. A. & Wigley D. B. (1995). A superior host strain for the over-expression of cloned genes using the T7 promoter based vectors. *Nucleic Acids Res*. **23**, 2074-2075.

Dole M., Mach L. L., Hines R. L., Mobley R. C., Fergusson L. D. & Alice M. B. (1968). Molecular beams of macroions. *J. Chem. Phys*. **49**, 2240-2249.

Douce R. (1985). Mitochondria in Higher Plants: Structure, Function and Biogenesis. Academic Press, London.

Douce R., Bourguignon J., Brouquisse R. & Neuburger M. (1987). Isolation of plant mitochondria: general principles and criteria of integrity. *Methods Enzymol*. **148**, 403-415.

Douce R., Bourguignon J., Macherel D. & Neuburger M. (1994). The glycine decarboxylase system in higher plant mitochondria-structure, function and biogenesis. *Biochem. Soc. Trans*. **22**, 184-188.

DuBell A. N. & Mullet J. E. (1995). Differential transcription of pea chloroplast genes during light-induced leaf development. *Plant Physiol*. **109**, 105-112.

Dürre P. & Andreesen J. R. (1983). Purine and glycine metabolism by purinolytic clostridia. *J. Bacteriol*. **154**, 192-199.

Edwards G. E. & Walker D. A. (1983). C3, C4: Mechanisms, and Cellular and Environmental Regulation of Photosynthesis. Blackwell Scientific, Oxford, UK.

Ehleringer J. R., Sage R. F., Flanagan L. B. & Pearcy R. W. (1991). Climate Change and the Evolution of C4 Photosynthesis. *Trends in Ecology and Evolution*, vol. **6**. Issue 3, 95-99.

Eisenberg D., Weiss R. M. & Terwilliger T. C. (1984). The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc. Natl. Acad. Sci*. USA **81**, 140-144.

Estabrook R. W. (1967). Mitochondrial respiratory control and the polarographic measurement of ADP/O ratios. *Methods Enzymol*. **10**, 41-47.

Farabaugh P. J. (1978). Sequence of the lac I gene. *Nature*. **174**, 765-769.

Fenn J. B., Mann M., Meng C. K., Wong S. F. & Whitehouse C. M. (1990). Electrospray ionization-principles and practice. *Mass Spectrom. Rev*. **9**, 37-70.

Findinegg, G. R. (1976). Correlations between accessibility of carbonic anhydrase for external substrate and regulation of photosynthetic use of CO2 and HCO3- by *Scenedesmus obliquus*. *Z. Pflanzenphysiol*. **79**, 428-437.

Findinegg, G. R. (1977). Interactions of glycolate-, HCO_3^- , Cl^--, and H^+- balance of *Scenedesmus obliquus*. *Planta* **135**, 33-38.

Fluhr R., Kuhlemeier C., Nagy F. & Chua N. -H. (1986a). Organ-specific and light-induced expression of plant genes. *Science* **232**, 1106-1112.

Fluhr R., Moses P., Morelli G., Coruzzi G. & Chua N. -H. (1986b). Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts. *EMBO J.* **5**, 2063-2071.

Freudenberg W. & Andreesen J. R. (1989). Purification and partial characterization of the glycine decarboxylase multienzyme complex from *Eubacterium acidaminophilum. J. Bacterio.* **171**, 2209-2215.

Gallagher T. F. & Ellis R. J. (1982). Light-stimulated transcription of genes for two chloroplast polypeptides in isolated pea leaf nuclei. *EMBO J.* **1**, 1493-1498.

Garcia G. N., Mar P. K., Mullin D. A., Walker J. R. & Prather N. E. (1986). The *E. coli* dnaY gene encodes an arginine transfer RNA. *Cell.* **45**, 453-459.

Gardeström P., Bergman., A. & Ericson I. (1980). Oxidation of glycine via the respiratory chain of mitochondria prepared from different parts of spinach. *Plant Physiol.* **65**, 389-391.

Gariboldi R. T. & Drake H. L. (1984). Glycine synthase of the purinolytic bacterium, *Clostridium acidiurici*: Purification of the glycine-CO_2 exchange system. *J. Biol. Chem.* **259**, 6085-6089.

Gidoni D., Brosio P., Bond-Nutter D., Bedbrook J. & Dunsmuir P. (1989). Novel cis-acting elements in petunia cab gene promoters. *Mol. Gen. genet.* **215**, 337-344.

Gietl C., Lehnerer M. & Olsen O. (1990). Mitochondrial malate dehydrogenase by isolated watermelon mitochondria. *Z. Pflanzenphysiol. Bd.* **114**, 393-401.

Gilmartin P. M., Memelink J., Hiratsuka K., Kay S. A. & Chua N. -H. (1992). Characterization of a gene encoding a DNA binding protein with specificity for a light-responsive element. *Plant Cell.* **4**, 839-849.

Gilmartin P. M., Sarokin L., Memelink J. & Chua N. -H. (1990). Molecular light switches for plant genes. *Plant Cell.* **2**, 369-378.

Gitzelmann R., Steinmann B. & Cuénod M. (1978). *N. Engl. J. Med.* **298**, 1424.

Givan C. V., Joy K. W., Kleczkowski L. A. (1988). A decade of photorespiratory nitrogen cycling. *Trends in Biochemical Sciences.* **13**, 433-437.

Green P. J., Kay S. A. & Chua N. -H. (1987). Sequence-specific interactions of pea nuclear factor with light-responsive elements upstream of the rbcS-3A gene. *EMBO J.* **6**, 2543-2549.

Green P. J., Kay S. A., Lam E. & Chua N. -H. (1989). In vitro DNA footprinting. Plant Mol. Biol. Manual B11, Kluwer Academic Publishers, Dordrecht-Belgium, pp. 5-10.

Green P. J., Yong M. -H., Cuozzo M., Kano-Murakami Y., Silverstein P. & Chua N. -H. (1988). Binding site requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-depedent transcriptional activation of the rbcS-3A gene; *EMBO J.* **7**, 4035-4044.

Gross E. (1967). The cyanogen bromide reaction. *Methods Enzymol.* **193**, 131-153.

Hardy L. W., Finer-Moore J. S., Montfort W. R., Jones M. O., Santi D. V. & Stroud R. M. (1986). Atomic Structure of Thymidilate synthase: Target for Rational Drug Design. *Science.* **235**, 448-455.

Hatch M. D. (1987). C4 photosynthesis: a unique blend of modified biochemestry, anatomy and ultrastructure. *Biochim. Biophys. Acta* **895**, 81-106.

Hayasaka K. Nanao K. Takada K. Okamura-Ikeda K. & Motokawa Y. (1993). Isolation and sequence determination of cDNA encoding human T-protein of the glycine cleavage system. *Biochem. Biophys. Res. Commun.* **192**, 766-771.

Hemling M. E., Roberts G. D., Johnson W., Carr S. A. & Covey T. R. (1990). Analysis of proteins and glycoproteins at the picomole level by on-line coupling of microbore high-performance liquid chromatography with continuous flow fast atom bombardment and electrospray mass spectrometry: a comparative evaluation. *Biom. Env. Mass Spectrom.* **19**, 677-691.

Henderson S. A., von Caemmerer S. & Farqyhar G. D. (1992). Short-term measurements of carbon isotope discrimination in several C4 species. *Aust. J. Plant Physiol.* **19**, 263-285.

Herrera-Estrella L., van den Broeck G., Maenhaut R., van Montagu M., Schell J., Timko M. & Cashmore A. (1984). Light-inducible and chloroplast-associated expression of a chimaeric gene introduced into *Nicotiana tabacum* using a Ti plasmid vector. *Nature* **310**, 115-120.

Higgins D.G. & Sharp P. M. (1988). CLUSTAL: a package for performing multiple sequence alignement on a microcomputer. *Gene.* **73**, 237-244.

Hiraga K. & Kikuchi K. (1980). The mitochondrial glycine cleavage system: purification and properties of glycine decarboxylase from chicken liver mitochondria. *J. Biol. Chem.* **255**, 11664-11670.

Hopp T. P. & Woods K. R. (1981). Prediction of protein antigenic determinants from aminoacid sequences. *Proc. Natl. Acad. Sci.* USA **78**, 3824-3828.

Hurkman W. J. & Tanaka C. K. (1986). Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol.* **81**, 802-806.

Husic D. W., Husic H. D. & Tolbert N. E. (1987). The oxidative photosynthetic carbon cycle or C2 cycle. Critical Reviews in Plant Sciences, vol. **5**. Issue 1, 45-100.

Hylton C. M., Rawsthorn S., Smith A. M., Jones D. A. & Woolhouse H. W. (1988). Glycine decarboxylase is confined to the bundle-sheath cells of leaves of C3-C4 intermediate species. *Planta* **175**, 452-459.

Imeson H. C., Zheng L. & Cossins E. A. (1990). Folylpolyglutamate derivaties of *Pisum sativum* L. Determination of polyglutamate chain lengths by high performance liquid chromatography following conversion to p-aminobenzoylpolyglutamates. *Plant Cell Physiol.* **31**, 223-231.

Ingle R. K. & Colman B. (1976). The relationship between carbonic anhydrase activity and glycolate excretion in the blue-green alga *Coccochloris peniocystis*. *Planta* **128**, 217-223.

Jenkins C. L. D., Furbank R. T. & Hatch M. D. (1989). Mechanism of C4 photosynthesis. *Plant Physiol.* **91**, 1372-1381.

Jennings I. & Cotton R. (1990). Structural similarities among enzyme pterin binding sites as demonstrated by monoclonal anti-idiotypic antibody. *J. Biol. Chem.* **265**, 1885-1889.

Jensen E. O., Marker K. A., Schell J. & deBruijn F. J. (1988). Interaction of a nodule specific, trans-acting factor with distinct DNA elements in the soybean leghaemoglobin *Ibc*$_3$ 5' upstream region. *EMBO J.* **7**, 1265-1271.

Jordano J., Almoguera C. & Thomas T. L. (1989). A sunflower helianthinin gene upstream sequence ensemble contains an enhacer and sites of nuclear protein interaction. *Plant Cell* **1**, 855-866.

Karas M. & Hillenkamp F. (1988). Laser desorption ionisation of proteins with molecular masses exceeding 10 000 daltons. *Anal. Chem.* **60**, 2299-2301.

Karlsson J. L. & Barker H. A. (1949). Tracer experiments on the mechanism of uric acid decomposition and acetic acid synthesis by *Clostridium acidi-urici. J. Biol. Chem.* **178**, 891-902.

Keys A. J. (1980). Synthesis and interconversion of glycine and serine. In The biochemestry of Plants, Vol. 5 (J. Miflin Ed), pp 359-374. Academic Press, NewYork.

Keys A. J., Bird I. F., Cornelius M. J., Lea P. J., Wallsgrove R. M. & Miflin B. J. (1978). The photorespiratory nitrogen cycle. *Nature.* **275**, 741-743.

Kikuchi G. (1973). The glycine cleavage system. Composition, reaction mechanism, and physiological significance. *Mol. Cell. Biochem.* **1**, 169-187.

Kim W. K. (1970). Effect of excision and benzamidazole treatment on folate content of wheat leaves and wheat leaf chloroplasts. *Can. J. Biochem.* **48**, 1091-1095.

Kim Y. & Oliver D. J. (1990). Molecular cloning, transcriptional characterization, and sequencing of the cDNA encoding the H-protein of the mitochondrial glycine decarboxylase complex in peas. *J. Biol. Chem.* **265**, 848-853.

Kim Y., Shah K. & Oliver D. J. (1991). Cloning and light dependent expression of the gene coding for the p-protein of the glycine decarboxylase complex from peas. *Physiol. Plant.* **81**, 501-506.

Kimura T., Takeda S., Asahi T. & Nakamura K. (1990). Primary structure of a precursor for the δ-subunit of sweet potato mitochondrial F1-ATPase deduced from full-length cDNA. *J. Biol. Chem.* **265**, 6079-6085.

Kisaki T. & Tolbert N. E. (1969). Glycolate and glyoxylate métabolisme by isolated peroxisomes or chloroplasts. *Plant Physiol.* **44**, 242-250

Kisaki T. & Tobert N. E. (1970). Glycine as a substrate for photorespiration. *Plant Cell. Physiol.* **11**, 247-258.

Kisaki T., Imai A. & Tolbert N. E. (1971). Intracellular localisation of enzymes related to photorespiration in green leaves. *Plant Cell. Physiol.* **12**, 267-273.

Klein S. M. & Sagers R. D. (1966). Glycine metabolism. II. Kinetic and optical studies on the glycine decarboxylase system from *Peptococcus glycinophilus*. *J. Biol. Chem.* **241**, 206-209.

Klein S. M. & Sagers R. D. (1967 a). Glycine metabolism. A flavin-linked dehydrogenase associated with the glycine cleavage system in *Peptococcus glycinophilus*. *J. Biol. Chem.* **242**, 297 - 300.

Klein S. M. & Sagers R. D. (1967 b). Effect of borohydride reduction on the pyridoxal phosphate-containing glycine decarboxylase from *Peptococcus glycinophilus*. *J. Biol. Chem.* **242**, 301 - 305.

Kochi H. & Kikuchi G. (1969). Reaction of glycine synthesis and glycine cleavage catalysed by extracts of *Arthrobacter globiformis*. *Arch. Biochem. Biophys.* **132**, 359 - 369.

Kopriva S. & Bauwe H. (1994). T-protein of Glycine Decarboxylase from *Solanum tuberosum*. *Plant physiol.* **104**, 1079-1080.

Kozak M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research.* **15**, 8125-8148.

Krieger I., Winbaum E. S. & Eisenbrey A. B. (1977). *Metabolism* **26**, 517-524.

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

Laemmli U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. *Nature* **227**, 680-683.

Lam E. & Chua N. -H. (1990). GT-1 binding site confers light-responsive expression in transgenic tobacco. *Science*. **248**, 471-474.

Lasaga A. C., Berner R. A. & Garrels R. M. (1985). In the Carbon Cycle and Atmospheric CO_2: atural Variations Archean to present (Sundquist, E. T. and Broecker W. S., eds), pp 397-411, American Geophysical Union.

Laursen N. B., Larsen K., Knudsen J. Y., Hoffman H. J., Poulsen C., Marcker K. A. & Jensen E. Ø. (1994). A protein binding AT-rich sequence in the soybean leghemoglobin *c3* promoter is a general cis element that requires proximal DNA elements to stimulate transcription. *Plant Cell* **6**, 659-668.

Lenne C., Block M. A., Garin J. & Douce R. (1995). Sequence and expression of the mRNA encoding HSP22, the mitochondrial small heat-shock protein in pea leaves. *Biochem. J.* **311**, 805-813.

Logan J. E., Falk-Pederson J. E., Darnell J. E. & Shenk T. (1987). A poly(A) addition site and a dowstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse β-globin gene. *Proc. Natl. Acad. Sci.* USA **84**, 8306-8310.

Lord, J. M. & Merrett, M. J. (1970). The pathway of glycollate utilization in *Chlorella pyrenoidosa*. *Biochem. J.* **117**, 929-937.

Lorimer G. H. & Andrews T. J. (1981). The C2 chemo- and photorespiratory carbon oxydation cycle. In The biochemestry of Plants, vol. 8. Photosynthesis (Hatch, M. D. and Boardman N. K. eds), pp 329-374. Academic Press, New-York.

Lorimer G. H. (1982). Activities of RuBP carboxylase-oxygenase. In the Methods in chloroplast molecular Biology (Edelman, M., Hallick, R. B. & Chua N. -H., eds) pp. 803-808, Essevier Biomedical Press.

Lowry O. H., Rosenbrough N. J., Farr A. L. & Randall R. K. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275.

Lugert T. & Werr W. (1994). A novel DNA-binding domain in the Shrunken initiator-binding protein (IBP1). *Plant Mol. Biol.* **25**, 493-506.

Lûtcke H. A., Chow K. C., Mickel F. S., Moss K. A., Kern H. F. & Scheele G. A. (1987). Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6**, 43-48.

Maas C., Schaal S. & Weer W. (1990). A feedback control element near the transcription start site of the maize Shrunken gene determines promoter activity. *EMBO J.* **9**, 3447-3452.

Mac Dermot K. D., Nelson W., Reichert C. M. & Schulman J. D. (1980). *Pediatrics* **65**, 61-64.

Mac Guire J. J. & Bertino J. R. (1981). Enzymatic synthesis and function of folylpolyglutamates. *Mol. Cell. Biochem.* **38**, 19-48.

Mac Guire J. J. & Coward J. K. (1984). Pteroylpolyglutamates: biosynthesis, degradation and function. In Folates and Pterins, vol. 1 (R. L. Blakley and S. K. Benkovic, Eds), pp 135-190. Wiley (interscience), New York.

Mac Kenzie R. E. (1984). In Folate and Pterins (Blakley R. L. and Benkovic S. J., eds), vol. **1**, pp. 255-306, John Wiley and Sons, New York.

Macherel D., Lebrun M., Gagnon J., Neuburger M. & Douce R. (1990). Primary structure and expression of H-protein, a component of the glycine cleavage suystem of pea leaf mitochondria. *Biochem. J.* **268**, 783-789.

Macherel D., Bourguignon J. & Douce R. (1992). Cloning of the (*gdcH*) encoding H-protein, a component of the glycine decarboxylase complex of pea (*Pisum sativum* L.). *Biochem. J.* **286**, 627-630.

Macherel D., Bourguignon J., Forest E., Faure M., Cohen-addad C. & Douce R. (1996). Expression, lipoylation and structure determination of recombinant pea H-protein in Escherichia coli. *Eur. J. Biochem.* In press.

Maley G. F., Bellisario R. L., Guarino D. U. & Maley F. (1979). The primary structure of *Lactobacillus casei* thymidylate synthetase. III. The use of 2-(2-Nitrophenylsulfenyl)-3-methyl-3-bromoindolenine and limited tryptic peptides to establish the complete amino acid sequence of the enzyme. *J. Biol. Chem.* **254**, 1301-1304.

Maley G. F., Maley F. & Baugh C. M. (1982). Studies on identifying the folylpolyglutamate binding sites of *Lactobacillus casei* thymidilate synthase. *Arch. Biochem. Biophys.* **216**, 551-558.

Mandrup S., Hojrup P., Kristiansen K. & Knudsen J. (1991). Gene synthesis, expression in *Escherichia coli*, purification and characterization of the recombinant bovine acyl-Coa-binding protein. *Biochem. J.* **276**, 817-823.

Maniatis T., Fritsh E. F. & Sambrook J. (1982). Molecular cloning: a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mann M., Meng C. K. & Fenn J. B. (1989). Interpreting mass spectra of multiply charged ions. *Anal. Chem.* **61**, 1702-1708.

Manzara T., Carrasco P. & Gruissem W. (1991). Developmental and organ-specific changes in promoter DNA-protein interactions in the tomato *rbcS* gene family. *Plant Cell* **3**, 1305-1316.

Martini F., Maras B., Tanci P., Angelaccio S., Pascarella S., Barra D., Bossa F. & Schirch V. (1989). The primary structure of rabbit liver mitochondrial serine hydroxymethyltransferase *J. Biol. Chem.* **264**, 8509-8519.

Meng C. K., Mann M. & Fenn J. B. (1988). *Z. Phys. D.* **10**, 361-368.

Mérand V. (1994). Développement du couplage de la chromatographie liquide avec la spectrométrie de masse par l'interface FAB à flux continu. Caractérisation des protéines H, T, et L du complexe de la glycine décarboxylase des mitochondries des feuilles de pois. Thèse de 3ème cycle, Université de Grenoble.

Mérand V., Forest E., Gagnon J., Monnet C., Thibault P., Neuburger M. & Douce R. (1993). *Biol. Mass. Spectrometry.***22**, 447-453.

Metz B. A., Welters P., Hoffman H. J. Jensen Ø., Schell J. & de Bruijn F. J. (1988). Primary structure and promoter analysis of leghemoglobin genes of the stem-nodulated tropical legume *Sesbania rostrata*: Conserved coding sequences, cis-element and *trans*-acting factors. *Mol. Gen. Genet.* **214**, 181-191.

Mitsuda, H., Suzuki, Y., Tadera, K. & Kawai, F. (1966). Biochemical studies on pteridines in plants; biogenesis of folic acid in green leaves: enzymatic synthesis of dihydropteroic acid from guanosine compounds and mechanism of its synthetic pathway. *J. Vitaminol.*, **12**, 192-204.

Moore A. L., Wood C. K. & Watts F. Z. (1994). Protein import into plant mitochondria. Annu. Rev. Plant Physiol. *Plant Mol. Biol.* **45**, 545-575.

Mourioux G. & Douce R. (1981). Slow passive diffusion of orthophosphate between intact isolated chloroplasts and suspending medium. *Plant Physiol.* **67**, 470-473.

Nakajima N., Horikoshi M. & Roeder R. G. (1988). Factors involved in specific transcription by mammalian RNA polymerase II: Purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* **8**: 4028-4040.

Navarre D. A. & Wolpert T. J. (1995). Inhibition of the glycine decarboxylaes multienzyme complex by the host-selective toxin victorin. *Plant Cell* **7**, 463-471.

Neuburger M. & Douce R. (1977). Oxydation du malate, du NADH et de la glycine par les mitochondries de plantes en C3 et C4. C. R. Hebd. *Sceance Acad. Sci. Ser* **285**, 881-884.

Neuburger M., Bourguignon J. & Douce R. (1986) Isolation of a large complex from the matrix of pea leaf mitochondria involved in the rapid transformation of glycine into serine. *FEBS Lett.* **207**, 18-22.

Neuburger M., Journet E. P., Bligny R., Carde J. P. & Douce R. (1982). Purification of plant mitochondria by isopicnic centrifugation in density gredients of percoll. *Arch. Biochem. Biophys.* **217**, 312-323.

Nishimura A., Akiyama K., Kohara Y. & Horiuchi K. (1992). Correlation of a subset of the pLC plasmids to the physical map of *Escherichia coli* K-12. *Microbiological Reviews.* **56**, 137-151.

Nyhan W. L. (1978). In the metabolic basis of inherited disease (Stanbury J. B., Wyngaarden J. B. & Fredrickson D. S. eds.) 4th edn., pp. 518-527., Mc Graw-Hill, New York.

O'Shea-Greenfield A. & Smale S. (1992). Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *J. Biol. Chem.* **267**, 1391-1409.

Odell J. T., Nagy F. & Chua N. H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **313**, 810-812.

Ohnishi J. & Kanai R. (1983). Differentiation of photorespiratory activity between mesophyll and bundle sheath cells of C4 plants. I. Glycine oxidation by mitochondria. *Plant Cell Physiol.* **24**, 1411-1420.

Ohnishi J., Yamazaki M & Kanai R. (1985). Differentiation of photorespiratory activity between mesophyll and bundle sheath cells of C4 plants. II. Peroxisomes of *Panicum miliaceum* L. *Plant Cell Physiol.* **26**, 797-803.

Okamura-Ikeda K., Fujiwara K. & Motokawa Y. (1982). Purification and characterization of chicken liver T-protein, a component of the glycine cleavage system. *J. Biol. Chem.* **257**, 135-139.

Okamura-Ikeda K., Fujiwara K., Yamamoto M., Hiraga K. & Motokawa Y. (1991). Isolation and sequence determination of cDNA encoding T-protein of the glycine cleavage system. *J. Biol. Chem.* **266**, 4917-4921.

Okamura-Ikeda K., Fujiwara K. & Motokawa Y. (1992). Molecular cloning of a cDNA encoding chicken T-protein of the glycine cleavage system and expression of the functional protein in Escherichia coli: Effect of mRNA secondary structure in the translational initiation region on expression. *J. Biol. Chem.* **267**, 18284-18290.

Okamura-Ikeda K., Ohmura Y., Fujiwara K. & Motokawa Y. (1993). Cloning an dinucleotide sequence of the *gcv* operon encoding the *Escherichia coli* glycine-cleavage system. *Eur. J. Biochem.* **216**, 539-548.

Oliver D. J. & Sarojini G. (1987). Regulation of glycine decarboxylase by serine. In progress in Photosynthesis Research, vol. III, (Biggins, J., ed), pp. 219-226, Martinus Nijhoff Publishers, The Hague.

Oliver D. J., Neuburger M., Bourguignon J. & Douce R. (1990a). Glycine metabolism by plant mitochondria. *Physiol. Plant.* **80**, 487- 491.

Oliver D. J., Neuburger M., Bourguignon J. & Douce R. (1990b). Interaction between the component enzymes of the the glycine decarboxylase multienzyme complex. *Plant Physiol.* **94**, 833-839.

Perisic O. & Lam E. (1992). A tobacco DNA binding protein that interacts with a light-responsive Box II element. *Plant Cell.* **4**, 831-838.

Pichersky E., Bernatzky R., Tanksley S. D., Breidenbach R. B., Kausch A. P. & Cashmore A. R. (1985). Molecular characterization and genetic mapping of two clusters of genes encoding chlorophyll a/b-binding proteins in *Lycopersicon esculentum* (tomato). *Gene* **40**, 247-258.

Prasad T. K. & Stewart C. R. (1992). cDNA clones encoding *Arabidopsis thaliana* and Zea mays mitochondrial chaperonin HSP60 and gene expression during seed germination and heat shock. *Plant Mol. Biol.* **18**, 873-885.

Priest D. G., Happel K. K., Mangum M., Bednarek J. M., Doig M. T. & Baugh C. M. (1981). Tissue folylpolyglutamate chain-length characterization by electrophoresis as thymidilate synthetase-fluorodeoxyuridylate ternary complexes. *Anal. Bioch.* **115**, 163-169.

Pritchard G. G., Griffin W. H. & Whittingham C. P. (1962). The effect of isonicotinyl hydrazide on the photosynthetic incorporation of radioactive CO_2 into ethanol-soluble compounds of Chlorella. *J. Exp. Bot.* **14**, 281-289.

Proodfoot N. J. & Brownlee G. G. (1976). 3' non-coding region sequences in eucaryotic messenger RNA. *Nature* **263**, 211-214.

Pugh B. F. & Tjian R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.* **5**, 1935-1945.

Rathman C. K. M. (1978). Studies with isolated bundle sheath mitochondria. Evidence for NAD-malic enzyme-catalysed decarboxylation of C4 acids in species representing the three C4 metabolic subtypes. *FEBS Lett.* **96**, 367-372.

Ravanel P., Tissut M. & Douce R. (1981). Effects of flavone on the oxidative properties of intact plant mitochondria. *Phytochem.* **20**, 2101-2103.

Ravanel P., Tissut M. & Douce R. (1982). Effects of kaempferol on the oxidative properties of intact plant mitochondria. *Plant physiol.* **69**, 375-378.

Ravanel P., Tissut M. & Douce R. (1984). Effects of rotenoids on isolated plant mitochondria. *Plant physiol.* **75**, 414-420.

Raven J. A., Osborne B. A. & Johnston A. M. (1985). Uptake of CO_2 by aquatic vegetation. *Plant Cell Environ.* **8**, 417.

Rawsthorne S. (1992). C3-C4 intermediate photosynthesis: linking physiology to gene expression. *Plant J.* **2**, 267-274.

Rawsthorne S., Hylton C. M., Smith. A. M. & Woolhouse H. W. (1988a). Photorespiratory metabolism and immunogold localization of photorespiratory enzymes in leaves of C3 and C3-C4 intermediate species of Moricandia. *Planta* **173**, 298-308;

Rawsthorne S., Hylton C. M., Smith. A. M. & Woolhouse H. W. (1988b). Distribution of photorespiratory enzymes between bundle-sheath and mesophyll cells in leaves of the C3-C4 intermediate species *Moricandia arvensis* (L.) DC. *Planta* **176**, 527-532.

Rebeille F., Neuburger M. & Douce R. (1994). Interaction between glycine decarboxylase, serine hydroxymethyltransferase and tetrahydrofolate polyglutamates in pea leaf mitochondria. *Biochem. J.* **302**, 223-228.

Reed M. L. & Graham D. (1977). *Aust. J. Plant Physiol.* **4**, 87-98.

Robinson M., Lilley R., Little S., Emtage J. S., Yarranton G., Stephens P., Millican A., Eaton M. & Humphreys G. (1984). Codon usage can affect efficienty of translation of genes in Escherichia coli. *Nucleic Acids Res.* **12**, 6663-6671.

Rosenberg A. H., Goldman E., Dunn J. J., Studier F. W. & Zubay G. (1993). Effects of consecutive AGG codons on translation in Escherichia coli, demonstrated with a versatile codon test system. *J. Bacteriol.* **175**, 716-722.

Roy A. L., Meisterernst M., Pognonec P. & Roeder R. (1991). Cooperative interaction of an initiator-binding transcription factor and the helix-loop-helix activator USF. *Nature* (london). **354**, 245-248.

Russel D. W., Smith M., Cox D., Williamson V. M. & Young E. T. (1983). DNA sequences of two yeast promoter-up mutants. *Nature* **304**, 652-654.

Sagers R. D. & Gunsalus I. C. (1961). *J. Bacteriol.* **81**, 541-549.

Sambrook J., Fritsch E. & Maniatis T. (1989). Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger F., Nicklen S. & Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* USA. **74**, 5463-5466.

Sarojini G. & Oliver D. J. (1983). Isolation and partial characterization of glycine decarboxylase from pea leaf mitochondria, *Plant physiol.* **72**, 194-199.

Sarojini G. & Oliver D. J. (1985). Inhibition of glycine oxidation by carboxymethoxylamine, methoxylamine, and acethydrazide. *Plant Physiol.* **77**, 786-789.

Schenk P. M., Baumann S., Mattes R. & Steinbiβ H. H. (1995). Improved high-level expression system for eucaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare ^{Arg}tRNAs. *BioTechniques.* **19**, 196-200.

Schwartz B. W., Sloan J. S. & Becker W. M. (1991). Characterization of genes encoding hydroxypyruvate reductase in cucumber. *Plant Mol. Biol.* **17**, 941-947.

Seto E., Lewis B. & Shenk T. (1993). Interaction between transcription factors Sp1 and YY1. *Nature J.* **365**, 462-464.

Shane B. (1982). High performance liquid chromatography of folates: identification of poly-γ -glutamate chain lengths of labeled and unlabeled folates. *Am. J. Clin. Nutr.* **35**, 599-608.

Shatkin A. J. (1985). mRNA cap binding proteins : essential factors for initiation translation. *Cell* **40**, 223-224.

Shingles R., Woodrow L. & Grodzinski B. (1984). Effects of glycolate pathway intermediates on glycine decarboxylation and serine synthesis in pea (*Pisum sativum* L.). *Plant Physiol.* **74**, 705-710.

Shirch V. & Strong W. B. (1989). Interaction of folylpolyglutamates with enzymes in one-carbon metabolism. *Arch. Biochem. Biophys.* **269**, 371-380.

Smale S. T. & Baltimor D. (1989). The initiator as a transcription control element. *Cell.* **57**,103-113.

Smale S. T., Schimidt M. C., Berk A. J. & Baltimor D. (1990). Transcriptional activation by SP1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc. Natl. Acad. Sci.* USA. **87**, 4509-4513.

Smith B. N. (1986). Evolution of C4 photosynthesis in response to changes in carbon and oxygen concentrations in the atmosphere through time. *Biosystems* **8**, 24-32.

Smith R. D., Loo J. A., Edmonds C. G., Barinaga C. J. & Udseth H. R. (1990). New developments in biochemical mass spectrometry: electrospray ionization. *Anal. Chem.* **62**, 882-899.

Smith S. M. & Ellis R. J. (1979). Processing of small subunit precursor of ribulose bisphosphate carboxylase and its assembly into whole enzyme are stromal events. *Nature* **278**, 662-664.

Srinivasan R. & Olicer D. J. (1995). Light-dependent and tissue-specific expression of the H-protein of the glycine decarboxylase complex. *Plant Physiol.* **109**, 161-168.

Srinivasan R., Kraus C. & Oliver D. J. (1992). Developemental expression of the glycine decarboxylase multienzyme complex in greening pea leaves. In molecular, Biochemical, and physiological Aspects of plant Respiration, ed. H Lambers, LHW van der plas, pp. 323-334. The Hague: SPB Academic.

Srinivasan R., Berndt W. A. & Oliver D. J. (1993). Coordinated expression of photosynthetic and photorespiratory genes. In plant Mitochondria, ed. A. Brennicke, U; Kuck, pp. 160-169. Weinheim, Germany: VCH.

Staben C. & Rabinowitz J. C. (1986). Nucleotide sequence of the Saccharomyces cerevisiae ADE3 gene encoding C1-Tetrahydrofolate synthase. *J. Biol. Chem.* **261**, 4629-4637.

Stenlid G. (1970). Flavonoids as inhibitors of the formation of adenosin triphosphate in plant mitochondria. *Phytochem.* **9**, 2251-2256.

Struhl K. (1985). Naturally occuring poly (dA-dT) sequences are upstream promoter elements for constitutive transcription on yeast. *Proc. Natl. Acad. Sci.* USA. **82**, 8419-8423.

Studier F. W. & Moffat B. A. (1986). Use of a Bacteriophage T7 RNA polymerase to direct selective expression of cloned genes. J. *Mol. Biol.* **189**, 113-130.

Studier F. W., Rosenberg A. H., Dunn J. J. & Dubendorff J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60-89.

Suter M. J. F., DaGue B. B., Moore W. T., Shen-Nan Lin & Caprioli R. M. (1991). Recent advances in liquid chromatography-mass spectrometry and capillary zone electrophoresis-mass spectrometry for protein analysis. *J. Chromatogr.* **553**, 101-116.

Szczyglowski K., Szabados L., Fujimoto S. Y., Silver D. & de Bruijn F. J. (1994). Site-specific mutagenesis of the nodule-infected cell expression (NICE) element and the AT-rich elmeent ATRE-BS2 of the *Sesbania rostrata* leghemoglobin *glb3* promoter. *Plant Cell* **6**, 317-332.

Thomas P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci.* **77**, 5201-5205.

Thompson P. D. & Freisheim J. H. (1991). Conversion of arginine to lysine at position 70 of human dihydrofolate reductase: generation of a methotrexate-Insensitive mutant enzyme. *Biochemistry* **30**, 8124-8130.

Thomson S. A. & Irbane J. V. (1979). Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* **71**, 4451-4463.

Tillberg J. E., Giersch C. & Heber U. (1977). CO_2 reduction by intact chloroplasts under a diminished proton gradient. *Biochim. Biophys. Acta* **461**, 31-47.

Tolbert N. E. (1971a). Photorespiration. In " Photosynthesis and Photorespiration " (Hatch M. D., Osmond C. B. and Slatyer R. O., eds.), pp. 458-471. Wiley (interscience), New York.

Tolbert N. E. (1971 b). In " Photosynthesis and Photorespiration " (Hatch M. D., Osmond C. B. and Slatyer R. O., eds.), pp. 506-513. Wiley (interscience), New York.

Tolbert N. E. (1974). In " algal Physiology and Biochemistry " (W. D. P. Stewart, ed.), pp. 474-504. Blackwell, Oxford.

Tolbert N. E. (1980). Photorespiration. In The biochemistry of plants (D. D. Davies, Ed), Vol. 2, pp 487-523. Academic Press, New York.

Tolbert N. E. & Zill L. P. (1956). Excretion of glycolic acid by algae during photosynthesis. *J. Biol. Chem.* **22**, 895-906.

Tomer K. B. & Parker C. E. (1989). Biochemical applications of liquid chromatography-mass spectrometry. *J. Chromatogr.* **492**, 189-221.

Turner S. R., Hellen R., Ireland R., Ellis N. & rawsthorne S. (1993). The organisation and expression of the gene encoding the mitochondrial glycine decarboxylase complex and serine hydroxymethyltransferase in pea (Pisum sativum). *Mol. Gen. Genet.* **236**, 402-408.

Turner S. R., Ireland R. & Rawsthorne S. (1992a). Cloning and characterization of the p subunit of glycine decarboxylase from pea (Pisum sativum). *J. Biol. Chem.* **267**, 5355-5360.

Turner S. R., Ireland R., Morgan C. & Rawxthorne S. (1992c). Identification and localization of multiple forms of serine hydroxymethyltransferase in pea (Pisum sativum) and characterization of a cDNA encoding a mitochondrial isoform. *J. Biol. chem.* **267**, 13528-13534.

Turner S., Ireland R. & Rawsthorne S. (1992b). Purification and primary amino acid sequence of the L subunit of glycine decarboxylase. Evidence for a single lipoamide dehydrogénase in plant mitochondria. *J. Biol. Chem.* **267**, 7745-7750.

Ueda T., Pichersky E., Malik V. S. & Cashmore A. R. (1989). Level of expression of the tomato *rbcS-3A* gene is modulated by a far upstream promoter element in a developmentally regulated manner. *Plant Cell* **1**, 217-227.

Usha R., Savithri H. S. & Rao N. A. (1992). Arginine residues involved in binding of tetrahydrofolate to sheep liver serine hydroxymethyltransferase. *J. Biol. Chem.* **267**, 9289-9293.

Usheva A. & Shenk T. (1994). TATA-binding protein-independant initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA. *Cell.* **76**, 1115-1121.

Van Dorsselear A., Bitsch F., Green B., Jarvis S., Lepage P., Bischoff R., Kolbe H. V. J. & Roitsch C. (1990). Application of electrospray mass spectrometry to the characterization of recombinant proteins up to 44 kDa. *Biomed. Env. Mass Spectrom.* **19**, 692-704.

Vanden Boom T. J., Reed K; E. & Croman J. E. (1991). Lipoic acid metabolism in *Escherichia coli*: isolation of null mutants defective in lipoic acid biosynthesis, molecular cloning and characterization of the E. coli lip locus, and identification of the lipoylated protein of the glycine cleavage system. *J. Bacteriol.* **173**, 6411-6420.

Veira J. & Messing J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-268.

von Heijne G., Steppuhn J. & Herrmann R. G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* **180**, 535-545.

Walker G. H., Oliver D. J., Sarojini G. (1982). Simultaneous oxidation of glycine and malate by pea leaf mitochondria. *Plant Physiol.* **70**, 1465-1469.

Walker J. L. & Oliver D. J. (1986a). Glycine décarboxylase multienzyme complex/ purification and partial characterization from pea leaf mitochondria. *J. Biol. Chem.* **261**, 2214-2221.

Walker J. L. & Oliver D. J. (1982). Identification of a glycine transporter from pea leaf mitochondria. *Biochem. Biophys. Res. Commun.* **107**, 856-861.

Walker J. L. & Oliver D. J. (1986b). Light-induced increases in the glycine decarboxylase multienzyme complex from pea leaf mitochondria. *Arch. Biochem. Biophys.* **248**, 626-638.

Warburton D., Boyle R. J., Keats J. P., Vohr B., Peuschel S. & Oh W. (1980). *Am. J. Dis. Child.* **134**, 273-275.

Watson J. C. & Thompson W. F. (1986). Purification and restriction endonuclease analysis of plant nuclear DNA. *Methods in enzymology* **118**, 57-75.

Watts F. Z., Walter A. J. & Moore A. L. (1992). Characterization of PHSP1, a cDNA encoding a mitochondrial HSP70 from *Pisum sativum*. *Plant mol. Biol.* **18**, 23-32.

Whitelaw E. & Proudfoot N. (1986). α-Thallassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human α-2 globin gene. *EMBO J.* **5**, 2915-2922.

Wolf S. M., Annan R. S., Vouros P. & Giese R. W. (1992). Characterization of amino polyaromatic hydrocarbon-DNA adducts using continuous flow fast atom bombardment and collision-induced dissociation: positive and negative ion spectra. *Biol. Mass Spectrom.* **21**, 647-654.

Wolpert T. J. & Macko V. (1989). Specific binding of victorin to a 100-kD protein from oats. *Proc. Natl. Acad. Sci.* USA **86**, 4092-4096.

Wolpert T. J., Navarre D. A., Moore D. L. & Macko V. (1994). Identification of the 100-kD victorin binding protein from oats. *Plant Cell* **6**, 1145-1155.

Woo K. C., Berry J. A. & Turner G. L. (1978). Release and refixation of ammonia during photorespiration. Carnegie Institution of Washinton, Yearbook. **77**, 240-245.

Woo K. C., Morot-Gaudry J. F., Summons R. E. & Osmond C. B. (1982). Evidence for the glutamine synthetase/glutamate synthase pathway during the potorespiratory nitrogen cycle in spinach leaves. *Plant physiol.* **70**, 1514-1517.

Yoshida T. & Kikuchi (1973). Major pathways of serine and glycine catabolism in various organs of the rat and cock. *J. Biochem.* (Tokyo). **73**, 1013-1022.

Zelitch I. (1975). Pathways of carbon fixation in green plants. *Annu. Rev. Biochem.* **44** , 132.

Zenzy-Gregory B., Khachi A., Garraway I. P. & Smale S. T. (1993). Mechanism of initiator-mediated transcription: evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence. *Mol. Cell. Biol.* **13**, 3841-3849.

Zenzy-Gregory B., O'Shea-Greenfield A. & Smale S. T. (1992). Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. *J. Biol. Chem.* **267**, 2823-2830.

Zindel U., Freudenberg W., Rieth M., Andreesen J. R., Schnell J. & Widdel F. (1988). *Eubacterium acidaminophilum* sp. nov., a versatil amino acid-degradating anaerobe producing or utilizing H_2 or formate. Description and enzymatic studies. *Arch. Microbiol.* **150**, 254 - 266.

Zutter M. M., Santoro S. A. Audrey S. P., Ying li Tsung & Gafford A. (1994). The human $\alpha 2$ Integrin gene promoter. *J. Biol. Chem.* **269**, 463-469.

PUBLICATIONS ET PARTICIPATION A DES CONGRES

PUBLICATIONS ET PARTICIPATION A DES CONGRES

I - PUBLICATIONS

A - ARTICLE PUBLIE DANS UNE REVUE A COMITE DE LECTURE

• BOURGUIGNON J., **VAUCLARE P.**, MERAND V., FOREST E., NEUBURGER M., & DOUCE R. (1993). Glycine decarboxylase complex from higher plants. Molecular cloning, tissue distribution and mass spectrometry analyses of the T protein - *Eur. J. Biochem.*, **217**, 377-386.

B - ARTICLES EN PREPARATION

• **VAUCLARE P.,** DIALLO N., BOURGUIGNON J., MACHEREL D. & DOUCE R. (1996). Regulation of the expression of the glycine decarboxylase complex during pea leaf development: soumis à publication.

• **VAUCLARE P**., BOURGUIGNON J. & DOUCE R. (1996). Isolation, characterization and promoter studies of the gene encoding the T protein of the glycine decarboxylase complex involved in the mitochondrial step of the photorespiratory pathway in plants: en préparation.

II - PARTICIPATION A DES CONGRES

A - SEMINAIRE

• XXIe FORUM DES JEUNES CHERCHEURS DE LA SOCIETE FRANÇAISE DE BIOCHIMIE ET BIOLOGIE MOLECULAIRE.
REIMS (France).
VAUCLARE P., BOURGUIGNON J., NEUBURGER M., & DOUCE R. (1994). Etude du complexe de la glycine décarboxylase des mitochondries de tissus foliaires des plantes supérieures.

B - PRESENTATION DE POSTERS

• HSFP-WORKSHOP « RNA EDITING-PLANT MITOCHONDRIA » & DFG-WORKSHOP « INTERACTION OF THREE GENOMES ».
BERLIN (Allemagne).
BOURGUIGNON J., MACHEREL D., **VAUCLARE P.**, NEUBURGER M. & DOUCE R. (1992). Glycine oxydation by green leaf mitochondria.

• XTH INTERNATIONAL CONGRESS ON PHOTOSYNTHESIS,
MONTPELLIER (France).
BOURGUIGNON J., **VAUCLARE P.**, DIALLO N., NEUBURGER M. & DOUCE R. (1995). The glycine decarboxylase system from higher plants.

• HUITIEME ENTRETIEN DU CENTRE JACQUES CARTIER.
LYON-Ecole Normale Supérieure (France).
VAUCLARE P., BOURGUIGNON J. & DOUCE R. (1995). Glycine decarboxylase complex from higher plants.
DOUCE R., **VAUCLARE P.**, BOURGUIGNON J., MACHEREL J. & NEUBURGER M. (1995). The glycine decarboxylase system from higher plants: developmental control.

• 3 ème RENCONTRE DU DEPARTEMENT DE BIOLOGIE MOLECULAIRE ET STRUCTURALE DE GRENOBLE-3RD DBMS WORKSHOP.
VILLARD DE LANS (France).
VAUCLARE P., BOURGUIGNON J., MACHEREL D. & DOUCE R. (1996). The glycine decarboxylase system from higher plants: Molecular cloning of T protein and regulation analyses.

RESUME

Nous avons utilisé des anticorps polyclonaux dirigés contre la protéine T du complexe de la glycine décarboxylase (GDC) pour isoler un ADNc de 1430 pb codant pour la protéine T de pois. L'analyse de la séquence nucléotidique a montré que la protéine T était synthétisée sous forme d'un précurseur comportant un peptide de transit mitochondriale de 30 aminoacides précédant la protéine T mature qui comporte 378 aminoacides. La masse déduite de la séquence primaire est de 40 961 Da et correspond exactement à la valeur mesurée au spectromètre de masse, ce qui indique que le cofacteur de la protéine T, le tétrahydrofolate n'est pas fixé de manière covalente sur la protéine. La structure primaire de la protéine T a pu également être confirmée par analyse en HPLC-ESI/MS et par microséquençage des fragments générés par protéolyse chimique de la protéine T purifiée. Des homologies de séquences avec les protéines T de foie de poulet, de boeuf et de pomme de terre, nous ont permis de localiser des régions hautement conservées comportant des résidus chargés positivement qui pourrait être impliquées dans les intéractions avec le tétrahydrofolate. L'étude de l'expression de la protéine T par Northern et Western-blot a montré que la protéine T et les messagers correspondants était principalement présents dans les tissus foliaires et subissaient une forte induction (8 à 10 fois) à la lumière.

Disposant de tous les outils moléculaires nécessaires à l'étude de la biogénèse de la GDC, nous avons étudié les modalités de mise en place du complexe au cours du développement de la feuille de pois. Nous avons observé que la capacité d'oxydation des mitochondries de tissus foliaires durant les premiers stades était négligeable pour augmenter de façon dramatique lorsque la feuille était pleinement ouverte (9 jours). En effet, contrairement à la Rubisco qui est présente dans les chloroplastes dès les tout premiers stades du développement, les protéines de la GDC se mettent en place beaucoup plus tardivement au sein de la matrice mitochondriale. Ce chargement en GDC des mitochondries est tellement important qu'il se traduit par une augmentation de leur densité. L'analyse des transcrits des gènes codant pour les protéines de la GDC (P, H et T) et la petite sous unité de la Rubisco révèle qu'ils sont exprimés dès le 4ième jour de croissance, stade où les mitochondries sont incapables d'oxyder la glycine. Cela suggère l'existence d'un contrôle post-transcriptionnel de l'expression des gènes codant pour les protéines P, H et T de la GDC.

Nous avons isolé le gène codant pour la protéine T qui est composé de 4 exons et dont la taille approximative est de 3 kpb. Nous avons détecté deux démarrages de transcription dont l'un présente une séquence riche en pyrimidine proche de la séquence Inr (Initiator element). L'analyse de la région promotrice du gène révèle essentiellement trois régions consensus, découvertes chez les gènes nucléaires codant pour la petite sous-unité de la Rubisco et pour la protéine liant les chlorophylles a/b : une région riche en nucléotide AT, un motif GATA en tandem et la boîte II (GGTTAA). Ces séquences semblent être impliquées dans la réponse à la lumière et dans la spécificité tissulaire. L'incubation de la région riche en nucléotides AT et de la boîte II avec un extrait nucléaire provenant de feuilles vertes a permis de caractériser certains facteurs de transcription similaires à ceux présents chez rbcS3,6 et rbcS3A.

Dans la perspective de l'étude structurale de la protéine T, nous avons réussi à surexprimer chez *E. coli* la protéine T, seule ou en tandem avec la protéine H. La présence de nombreux codons rares dans la séquence nucléotidique, nous a amené à construire des vecteurs co-exprimant l'ARNt codant pour l'arginine, de façon à permettre la traduction de la protéine dans la bactérie.

CEA/GRENOBLE/UST/EC/MCO/REPROGAPHIE /1996