

**UNIVERSITE LOUIS PASTEUR
STRASBOURG I
2006**



THESE

présentée à la

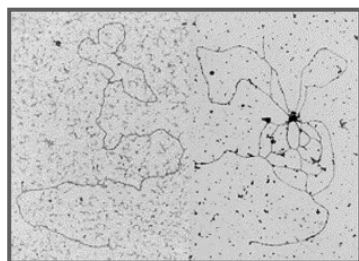
FACULTE DES SCIENCES

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR

Discipline : SCIENCES DU VIVANT

Spécialité : ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE



par

Vincent ZAEGEL

*Etude d'une famille de protéines spécifique des plantes impliquée
dans la maintenance des génomes des organites*

Soutenue le 10 février 2006 devant la Commission d'Examen :

**Mario KELLER
Thomas BÖRNER
Françoise FOURY
Françoise BUDAR
Patrice IMBAULT**

**RAPPORTEUR INTERNE
RAPPORTEUR EXTERNE
RAPPORTEUR EXTERNE
EXAMINATEUR
DIRECTEUR DE THESE**



Institut de Biologie Moléculaire des Plantes (IBMP) UPR-CNRS 2357

SOMMAIRE

SOMMAIRE	1
Abréviations	4
Introduction	5
Chapitre I: Mécanismes de la maintenance des génomes	6
1 La réplication.....	6
1.1 L'initiation.....	7
1.2 L'élongation.....	8
1.3 Remarques et conclusion.....	9
2 La recombinaison.....	10
2.1 L'échange de brins, cœur du mécanisme.....	10
2.2 Les événements produisant des cassures de l'ADN.....	12
2.3 Le modèle DSBR chez <i>E.coli</i>	12
2.4 Les données concernant les eucaryotes.....	15
3 Conclusion.....	16
Chapitre II: les protéines liant l'ADN simple brin (SSB)	18
1 Propriétés générales.....	18
1.1 Affinité pour l'ADNsb et abondance dans les cellules.....	19
1.2 Organisation multimérique et modes de fixation.....	19
1.3 La coopérativité.....	20
2 L'« OB-fold » : un motif commun pour la fixation de l'ADNsb.....	21
3 Conclusion.....	23
Chapitre III: les génomes des organites	24
1 Situation générale.....	24
2 Composition en séquences des chondriomes embryophytes.....	25
3 Conséquences des séquences répétées sur les chondriomes végétaux.....	27
4 L'organisation in vivo et les modes de réplication des organites.....	29
5 Mutations et « substoichiometric shifting » chez les mitochondries des plantes supérieures.....	30
5.1 Les recombinaisons ectopiques.....	30
5.2 « Substoichiometric shifting ».....	31
5.3 Exemples de phénotypes liés aux mutations sur l'ADNmt.....	33
5.4 Les tissus transmetteurs.....	36
MATériels et methodes	38
Chapitre I : Matériels et Outils Informatiques	39
1 Matériel végétal.....	39
1.1 Plantes entières.....	39
1.2 Cultures in vitro.....	39
1.3 Cultures cellulaires.....	39
2 Souches bactériennes.....	39
2.1 <i>E. coli</i>	39
2.2 <i>Agrobacterium tumefaciens</i>	40
3 Plasmides.....	41
3.1 pSK (-/+).....	41
3.2 pRSETc.....	41
3.3 PGEX.....	41
3.4 pUCAP, pUCAP35S et pUCAP35S-TAP.....	42
3.5 pBINPLUS.....	42
3.6 pER8.....	43
3.7 PBI 101.....	43
3.8 pCK-GFP3.....	43
3.9 pCR [®] 2,1-TOPO.....	44

4	Oligodésoxyribonucléotides.....	44
5	Outils informatiques.....	44
5.1	Banques de données	44
5.2	Prédictions informatiques d'adressage vers les organites	44
Chapitre II : METHODES.....		45
1	Purification de mitochondries	45
1.1	Mitochondries de pommes de terre et de chou-fleur.	45
1.2	Fractionnement subcellulaire et purification de mitochondries d'Arabidopsis	46
2	Préparation des bactéries compétentes.....	46
3	Techniques communes aux différents acides nucléiques	47
3.1	Quantification des acides nucléiques.....	47
3.2	Séparation par électrophorèse.....	47
4	Techniques relatives à l'ADN	49
4.1	Techniques de clonage	49
4.2	Préparation de l'ADN plasmidique.....	50
4.3	Mutagenèse dirigée.....	51
4.4	Séquençage de l'ADN.....	52
4.5	Amplification par PCR ("Polymerase Chain Reaction")	52
4.6	Marquage d'oligonucléotides.....	52
4.7	Marquage d'ADNdb.....	53
5	Techniques relatives à l'ARN.....	53
5.1	Extraction des ARN totaux.....	53
5.2	Traitement ADNase « RNase free »	54
5.3	Réaction de transcription inverse	54
5.4	Transfert d'ARN sur membrane : « northern blot »	55
6	Techniques relatives aux protéines	56
6.1	Quantification.....	56
6.2	Techniques de concentration	57
6.3	Fractionnement par électrophorèse en conditions dénaturantes	58
6.4	Détection	60
6.5	Surexpression	61
6.6	Purification des protéines surexprimées	62
6.7	Purification des corps d'inclusion	64
6.8	Couplage de protéines à une matrice de Sépharose-CNBr activé (système « CNBr-activated Sepharose 4B » ; Amersham biosciences).....	64
6.9	Séquençage N-terminal des protéines.....	65
7	Techniques d'immunologie	65
7.1	Production d'anticorps polyclonaux.....	65
7.2	Purification d'anticorps.....	65
7.3	Western blot	66
7.4	Immunoprécipitation	67
8	Techniques relatives au matériel végétal	68
8.1	Transformation d'Arabidopsis par Agrobacterium	68
8.2	Croisement entre deux lignées d'Arabidopsis.....	69
8.3	Transfection par bombardement.....	69
8.4	Observation au microscope	70
8.5	Analyse des fusions-GUS.....	71
8.6	Etablissement de cultures de cellules transgéniques.....	71
9	Autres méthodes.....	73
9.1	Extraction simultanée d'ARN, d'ADN et de protéines au trizol	73
9.2	Test d'importation in vitro de protéines dans les mitochondries.....	74
9.3	Southwestern	75
9.4	Le gel retard ou EMSA (« Electromobility Shift Assay »).....	76
9.5	Test de recombinaison <i>in vitro</i>	76
Resultats.....		78
Chapitre I : Identification de protéines mitochondriales liant l'Adn simple brin chez les plantes supérieures.....		79

1	Identification de SSB chez <i>Solanum tuberosum</i> par chromatographie d'affinité sur matrice d'ADNsb	79
2	Confirmation de la capacité de liaison à l'ADN :	80
3	Les orthologues de St-SBP40 :	80
3.1	Le cas d' <i>Arabidopsis thaliana</i> .	80
3.2	Les autres espèces	81
Chapitre II: Developpement d'outils en vue de l'étude fonctionnelle		83
1	Expression de protéines de fusion dans différents systèmes :	83
1.1	Dans <i>Escherichia coli</i>	83
1.2	Dans <i>Pichia pastoris</i>	85
1.3	Dans <i>Nicotiana benthamiana</i>	86
1.4	Bilan	86
2	Production d'anticorps polyclonaux et test des plantes transgéniques:	87
2.1	Production et purification des anticorps	87
2.2	Plantes surexprimant les protéines At-SBP40	88
2.3	Mutants d'insertion	89
2.4	Conclusion	89
3	Cultures cellulaires transgéniques :	89
Chapitre III: etude fonctionnelle		91
1	Confirmation de l'adressage aux organites :	91
1.1	Immunodétection	91
1.2	Importation in vitro	92
1.3	Expression transitoire dans des feuilles de tabac	93
1.4	Conclusion	93
2	Recherche de partenaires nucléiques :	93
2.1	Détermination in vitro du type d'acide nucléique fixé	94
2.2	Essai de caractérisation de l'activité des At-SBP40	99
2.3	Recherche de partenaires in vivo	103
2.4	Conclusion	104
3	Etude du domaine de fixation aux acides nucléiques :	105
4	Mise en évidence de la participation d'At-SBP40a à la maintenance du génome mitochondrial :	106
5	Etude de l'expression dans les organes :	107
5.1	Le cas d'At-SBP40a	108
5.2	Le cas d'At-SBP40b	108
5.3	Le cas d'At-SBP40c	109
Discussion et perspectives		110
1	Le motif PDF	111
2	Points communs et divergences entre les At-SBP40 et les SSB	113
3	Quelle fonction pour At-SBP40a ?	115
4	Conclusion générale	117
Bibliographie		118
ANNEXE		130
(LISTE DES Oligonucleotides UTILISES)		130
1	Sondes	130
2	Clonages	131
2.1	Clonage des séquences codant pour les précurseurs des protéines de la famille SBP40	131
2.2	Fusions avec l'étiquette TAP	132
2.3	Fusions avec l'eGFP	133
2.4	Fusions avec le marqueur GUS	133

ABREVIATIONS

ADN	Acide désoxyribonucléique	kb	Kilobase
ADNdb	ADN double brin	kDa	Kilodalton
ADNc	ADN complémentaire	LB	Luria Broth
ADNmt	ADN mitochondrial	MES	Acide 2-(N-morpholino) éthansulfonique
ADNsb	ADN simple brin	nt	Nucléotide
APS	Persulfate d'ammonium	ORF	Cadre ouvert de lecture
ARN	Acide ribonucléique	PAGE	« polyacrylamide gel electrophoresis »
ARNm	ARN messenger	PCR	Réaction de polymérisation en chaîne
ARNr	ARN ribosomique	pI	point isoélectrique
ARNt	ARN de transfert	PVDF	Polyvinylidène difluorure
BET	Bromure d'éthidium	RT	Transcription inverse
BSA	Sérum albumine bovine	SDS	Dodécyl sulfate de sodium
CBP	Calmoduline binding peptide	SSC	« Standard sodium citrate »
CMS	Stérilité mâle cytoplasmique	TEB	Tris-EDTA-borate
Ci	Curie	TEMED	N,N,N',N'- tetraméthyléthylène diamine
cpm	Coup par minute	Tris	N-tris(hydroxyméthyl) aminométhane
dNTP	Désoxyribonucléotide-5'-triphosphate	U	unité enzymatique
DO	Densité optique	X-gal	5-bromo 4-chloro 3-indole β- D-galactopyranoside
DTT	Dithiotréitol		
EDTA	Acide étylène diamine tétra- acétique		
EMSA	Gel retard		
EST	« Expressed sequence tag »		
GFP	« Green fluorescent protein »		
GST	Glutathion S-transférase		
IPTG	Isopropyl-β-D- thiogalactopyranoside		

INTRODUCTION

Ce travail de thèse porte sur l'étude d'At-SBP40, une famille de protéines d'*Arabidopsis thaliana* participant au métabolisme de l'ADN dans la mitochondrie et le chloroplaste, vraisemblablement dans le cadre d'un processus de réplication ou de recombinaison. Ces mécanismes sont à l'heure actuelle encore peu connus dans les organites des plantes supérieures. Le premier chapitre de l'introduction sera donc consacré à une synthèse des informations disponibles pour les systèmes les plus étudiés. Nous verrons que les bases du fonctionnement de la réplication et de la recombinaison sont ubiquitaires, ce qui permet d'émettre l'hypothèse que ceux-ci soient également en charge de la gestion des génomes des organites. Le second chapitre sera consacré à la description d'une classe de protéines impliquées dans ces étapes, les protéines fixant l'ADN simple brin, ou SSB (Single Stranded DNA Binding proteins), à laquelle appartiennent les protéines At-SBP40. Enfin, le troisième chapitre évoquera la situation dans les organites et le cas particulier des mitochondries des plantes supérieures.

	Organisme	<i>E.coli</i>	Phage λ	Phage T4	SV40/ <i>H.sapiens</i>	Levures
Initiation	Prot. d'initiation	DnaA	λ O	aucune	antigène T	ORC
	Prot. de chargement/ remodelage	DnaC	λ P, DnaJ, DnaK	gp59	chaperone cellulaire?	Cdc6
	Hélicase	DnaB		gp41	AgT	prot. MCM?
Fourche de réplication	Primase	DnaG		gp61	sous-unité primase de pol α -primase	sous-unité primase de pol α - primase
	Polymérase	sous-unité α de pol III HE		gp43	pol δ	pol δ et pol ϵ
	Exonucléase de correction	sous-unité ϵ de pol III HE		partie de gp43	partie de pol δ	parties de pol δ et pol ϵ
	Pince de processivité	sous-unité β		gp45	PCNA	PCNA
	« chargeur » de la pince	complexe γ		gp44/62	RF-C	RF-C
	SSB	SSB		gp32	RPA	RPA

Tableau i 1.1: Protéines impliquées dans des fonctions analogues lors de la réplication

d'après Baker TA et Bell SP, 1998

CHAPITRE I: MECANISMES DE LA MAINTENANCE DES GENOMES

La conservation d'une organisation efficace du patrimoine génétique au cours de la vie d'un organisme, et particulièrement pendant la division cellulaire, requiert de très nombreux processus enzymatiques qui sont usuellement regroupés en trois types distincts : la réplication, la recombinaison et la réparation. Au cours des vingt dernières années, des données de plus en plus nombreuses ont démontré leur étroite interconnexion. En particulier, la recombinaison peut à présent être vue comme une étape essentielle au bon déroulement de la réplication. Elle permet en effet la ré-initiation, indépendamment d'une origine, d'une fourche de réplication bloquée en raison d'un dommage sur la matrice. Nous allons toutefois, pour des raisons de clarté, évoquer ces deux processus séparément, en commençant par la réplication.

1 La réplication

La réplication de l'ADN peut être décomposée en trois étapes : l'initiation, l'élongation et la terminaison. L'étape d'initiation nécessite la reconnaissance d'une séquence nucléique appelée origine de réplication par un complexe protéique, l'ouverture localisée de la double hélice par une hélicase puis la synthèse d'une amorce ARN. Il s'agit vraisemblablement du lieu du contrôle, c'est à dire du moment auquel les protéines diffusant les signaux représentatifs de l'état d'avancement du cycle cellulaire, comme les cyclines, sont en charge de la coordination de ce mécanisme avec l'ensemble des autres processus cellulaires. L'étape d'élongation consiste en la progression bidirectionnelle de la fourche de réplication, qui est à la fois le lieu d'ouverture du duplexe d'ADN et de la production des molécules néo-synthétisées. Enfin, l'étape de terminaison implique le blocage de la fourche de réplication mais également la non réutilisation des molécules qui viennent d'être synthétisées. Il s'agit à l'heure actuelle de la partie la moins bien caractérisée de la réplication. Nous allons voir au cours de ce paragraphe que la base de ces différentes étapes est conservée au cours de l'évolution (voir **tableau i1.1**), mais que l'augmentation de la taille des génomes a rendu nécessaire l'intervention d'un nombre croissant de protéines.

1.1 L'initiation

1.1.1 Chez *E. coli*

Chez *E. coli*, l'origine de réplication *oriC* est présente à un exemplaire par chromosome bactérien (Kaguni et Kornberg 1984). Cette région d'environ 250 pb contient plusieurs séquences répétées, contenant un élément consensus de 9 pb appelé boîte DnaA (Fuller *et al.* 1984). Chacune d'elles est la cible d'un monomère de la protéine DnaA ; ces monomères interagissent entre eux par le biais d'interactions protéines-protéines et provoquent une distorsion de la molécule d'ADN qui s'ouvre localement en une région proche riche en paires de bases A-T (Speck et Messer 2001). Cette ouverture conduit au recrutement de l'hélicase DnaB, dont l'incorporation nécessite l'action d'une autre protéine, DnaC, et d'ATP (Fang *et al.* 1999). Le déplacement de DnaB de quelques nucléotides permet la création d'une région d'ADNsb suffisamment longue pour l'arrivée de la primase, DnaG, qui synthétise des amorces ARN de 8-12 nt au niveau de séquences 5'-CTG-3' (Yoda *et al.* 1988).

1.1.2 Cas des eucaryotes

L'identification de séquences d'initiation eucaryotes est beaucoup plus ardue, en raison de leur nombre (jusqu'à plusieurs milliers) et de leur taux de conservation plus faible que chez les procaryotes. A cela s'ajoute le fait qu'elles soient asynchrones (utilisées différemment en fonction des conditions de culture des organismes ou des stades de développement) et que des facteurs épi génétiques gouvernent leur activation (Antequera 2004). Nous allons donc seulement évoquer ici le cas du système le plus simple, celui retrouvé chez *S. cerevisiae*, le seul pour lequel des séquences conservées ont été retrouvées aux origines de réplication. Ces séquences, nommées ARS (« Auto Replicating Sequences »), constituées d'ADN non codant et longues en moyenne de 100 à 200 pb, sont composées des éléments A (très conservés) et B (moins conservés) (Bell et Dutta 2002). Les éléments A sont reconnus par le complexe ORC (« Origin Recognition Complex ») constitué de six protéines (Orc 1-6), et dont la fonction est analogue à celle de la protéine bactérienne DnaA (Stillman 2005). Les différences fondamentales avec le système bactérien résident dans le fait que le complexe ARS-ORC est stable durant tout le cycle cellulaire et qu'il ne semble pas capable à lui seul d'ouvrir la double hélice (Geraghty *et al.* 2000). Le meilleur candidat pour remplir le rôle d'hélicase est le complexe hexamérique MCM (Minichromosomal Maintenance complex), composé lui aussi de six protéines (Mcm 2-7) (Stillman 2005). Comme chez *E. coli*, son recrutement nécessite des protéines supplémentaires, parmi lesquelles

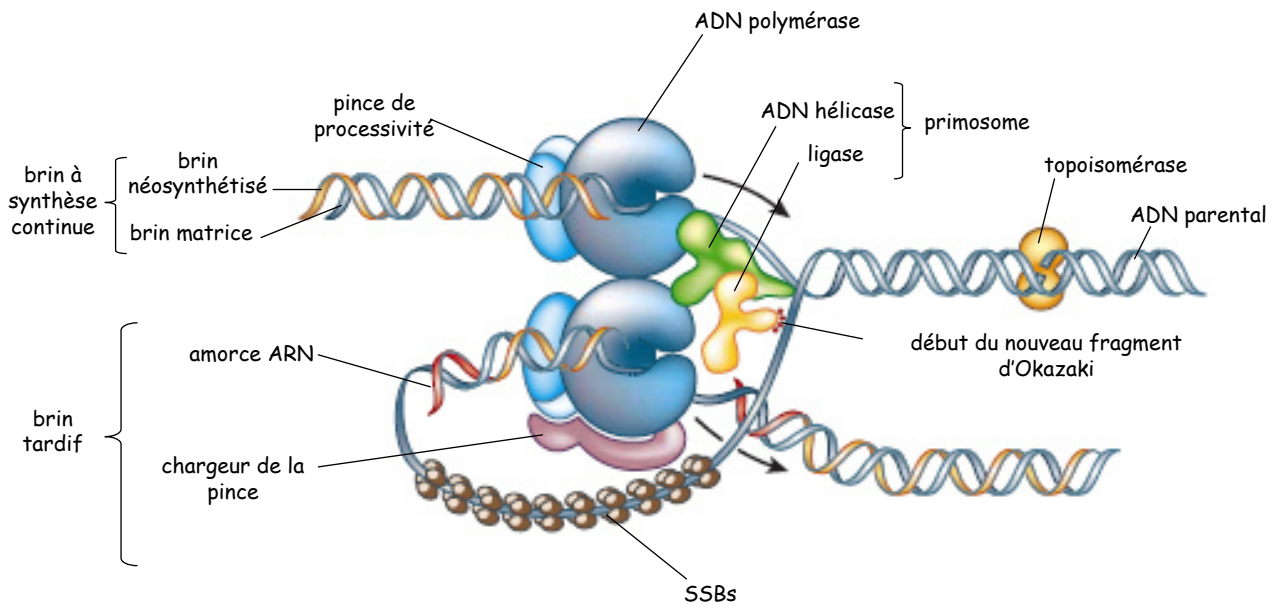


Figure i1.1: Représentation schématique d'une fourche de réplication

Deux molécules d'ADN polymérase sont actives en permanence. L'une est en charge du brin à synthèse continue alors que l'autre produit une série de courts « fragments d'Okazaki » sur le brin tardif. Toutes deux sont maintenues sur leur matrice par des protéines accessoires, la pince (« clamp ») et le chargeur de pince (« clamp loader »), afin d'augmenter leur processivité. Une ADN hélicase, hydrolysant l'ATP, est nécessaire à la séparation des deux brins de la double hélice et une topoisomérase élimine les super-tours créés lors de l'ouverture. En plus de la matrice, la polymérase nécessite l'extrémité 3' d'une amorce sur laquelle ajouter les nucléotides. Une ADN primase ajoute donc une courte amorce d'ARN en 5' de chaque fragment d'Okazaki, sur laquelle la polymérase du brin tardif continue la synthèse. Les régions d'ADNsb sont protégées par une SSB (« Single-Strand DNA-Binding protein »), qui a également pour fonction d'empêcher la ré-hybridation de la matrice. Les deux polymérases sont maintenues proches l'une de l'autre par des interactions protéine-protéine, ce qui permet à celle du brin tardif de se maintenir au contact de la fourche, afin de pouvoir servir à la synthèse de tous les fragments d'Okazaki. Cela a pour conséquence de nécessiter le repliement du brin tardif. En plus des protéines représentées, la réplication utilise un groupe de protéines en charge de la reconnaissance de l'origine de réplication. Enfin, les amorces ARN sont éliminées par une activité exonucléase 5'→3' (portée par une sous-unité de la polymérase ou par une protéine particulière dans certains cas). Les trous ainsi formés sont ensuite comblés par la polymérase et les fragments d'Okazaki reliés par une ADN ligase.

d'après Alberts B, 2003

Tableau i1.2 : voir page suivante ...□□

..

Cdc6 et Cdt1 et ORC lui-même (Bowers *et al.* 2004). Enfin, la position des nucléosomes et le niveau d'acétylation de la chromatine semblent avoir un rôle dans la formation du complexe d'initiation (Vogelauer *et al.* 2002). Ces étapes de reconnaissance de l'origine et d'ouverture de la double hélice permettent ensuite l'assemblage, sur un brin d'ADNsb protégé par la protéine RPA (équivalent eucaryotique des SSB), du complexe polymérase α -primase, en charge de la synthèse d'une amorce hybride ARN-ADN (Murakami *et al.* 1992).

1.2 L'élongation

Une fois le complexe d'initiation formé, la réplication proprement dite peut commencer. La **Figure i1.1** explique le mode de progression d'une fourche de réplication. Celle-ci nécessite l'emploi d'un jeu d'activités enzymatiques commun à tous les règnes du vivant. Au fur et à mesure de l'évolution, le nombre de sous unités employées croît considérablement, sans bouleversement des mécanismes de bases (voir **tableau i1.1**). Par exemple, le noyau enzymatique (« core enzyme ») comprend quatre protéines chez le phage T7, 14 chez *E. coli* (Benkovic *et al.* 2001) et au moins 27 chez *S. cerevisiae* (Waga et Stillman 1998). Nous allons à présent détailler les caractéristiques de quelques unes des activités majeures de la fourche d'*E. coli* et de celles rencontrées chez les eucaryotes (levures et SV40/*H.sapiens*).

1.2.1 Chez *E. coli*

L'hélicase DnaB interagit avec les 16 résidus C-terminaux de la primase DnaG (Tougu et Marians 1996), dont l'activité se trouve augmentée d'un facteur 10^3 (Johnson *et al.* 2000) et dont l'action est dirigée vers la fourche par un élargissement de son domaine de reconnaissance à la séquence 5'-CAG-3' (Bhattacharyya et Griep 2000). L'allongement des amorces d'ARN est ensuite prise en charge par l'enzyme répllicative, l'ADN polymérase III, composée de 10 sous-unités enzymatiques (voir **tableau i1.2**) (Herendeen et Kelly 1996). La polymérase constituée de $(\alpha\epsilon\theta)_2\tau_2\gamma_2\delta\delta'\chi\psi$ est peu processive (environ 10 nt à 20 nt.s⁻¹) (Fay *et al.* 1981), mais le devient (50000 nt à 750 nt.s⁻¹) (Studwell et O'Donnell 1990) en présence de deux sous-unités β , jouant le rôle de pince en formant un anneau de 35 Å qui glisse librement autour de la double hélice d'ADN (Stukenberg et O'Donnell 1995). L'assemblage du complexe enzymatique processif implique la fixation (sans hydrolyse) de l'ATP par γ , ce qui induit un changement conformationnel exposant δ à la fixation de β (Naktinis *et al.* 1995). Cette fixation ouvre la « pince » β , son positionnement sur le duplexe consomme ensuite une molécule d'ATP dont une autre est utilisée pour le déplacement de γ . Les segments d'ADNsb sont protégés par la

Sous-unité	Masse moléculaire (kDa)	Fonction	Contacts
α	130	ADN polymérase	} α - ϵ - θ
ϵ	27,5	3'-5' exonucléase	
θ	8,6	Stimule ϵ	
τ	71	Stimule l'hélicase et dimérise le « core »	θ ϵ α - τ - τ - α ϵ θ γ - γ
γ	47,5	« chargeur » de la pince, nécessite de l'ATP	} γ - γ ψ δ' χ - δ
δ	38,7	protéine accessoire, se lie à β	
δ'	36,9	protéine accessoire, stimule γ	
χ	16,6	protéine accessoire, se lie à la protéine SSB	
ψ	15,2	protéine accessoire, augmente l'affinité pour γ	
β	40,6	pince	

Tableau i 1.2: Composition de l'ADN polymérase III holoenzyme de *E.coli*

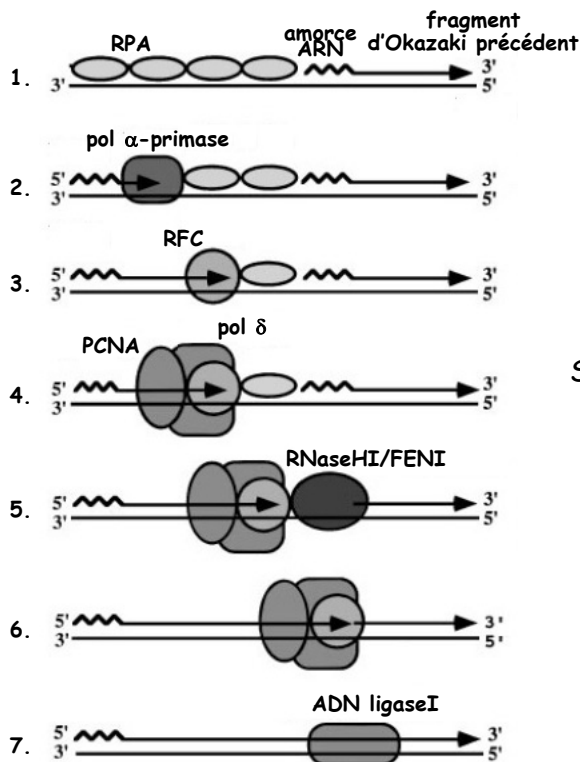


Figure i 1.2:
Synthèse du brin tardif chez les eucaryotes
(voir texte pour les détails)

D'après Waga et al. 1998

protéine SSB (Lohman et Ferrari 1994), présente sous forme d'homo tétramère et interagissant également avec la sous-unité χ de la polymérase III et DnaG. SSB est également essentielle au couplage de la synthèse des brins précoce et tardif (Meyer *et al.* 1979).

1.2.2 Cas des eucaryotes

Les différentes étapes de la synthèse du brin tardif sont schématisées **Figure i1.2**. Les parties de la matrice non encore copiée sont protégées par l'hétéro trimère RPA. Plusieurs études, chez la levure et l'humain, ont rapporté que différentes sous-unités de cette protéine peuvent être phosphorylées (Din *et al.* 1990) par des kinases cyclines- ou DNA-dépendantes, mais la signification fonctionnelle de ce phénomène reste pour l'heure inconnue. L'hypothèse la plus couramment avancée est celle d'un rôle de la phosphorylation dans le contrôle de l'état d'avancement de la phase S (Henricksen *et al.* 1996). RPA joue également un rôle dans le recrutement de la polymérase α -primase (Tsurimoto et Stillman 1991), qui synthétise les amorces. Lorsque celles-ci atteignent environ 30 nt, RFC, qui reconnaît spécifiquement l'assemblage amorce-matrice, se fixe en 3' de manière ATP-dépendante, déplaçant la primase. Le rôle principal de RFC est de charger PCNA, la « pince », une protéine en forme d'anneau mais qui à la différence du facteur β de *E. coli*, est ici trimérique. PCNA recrute à son tour la polymérase δ , l'enzyme répliquative eucaryote. Il est à noter que chez l'humain, la polymérase ϵ ne semble jouer aucun rôle, alors qu'elle est essentielle à *S. cerevisiae* (Araki *et al.* 1992). Deux protéines sont ensuite impliquées dans la digestion complète des amorces ARN, la RnaseHI et FEN1, une 5'→3' exo/endonucléase (Murante *et al.* 1996). Les trous ainsi générés sont ensuite comblés par la polymérase δ et les fragments d'Okazaki ligués entre eux par l'ADN ligase I.

1.3 Remarques et conclusion

Ce paragraphe consacré à la réplication n'a bien entendu rien d'exhaustif, et de nombreuses données, en particulier sur le plan structural, n'ont pas été abordées. Le but est simplement d'offrir un regard sur une des étapes essentielles de la maintenance des génomes, dans laquelle les protéines de la famille At-SBP40 sont susceptibles d'être impliquées. Pour terminer, nous allons rapidement passer en revue les thématiques actuelles concernant la réplication. Les aires de recherches les plus actives sont sans doute constituées par la caractérisation des séquences d'initiation et des protéines impliquées dans le contrôle de celles-ci. Un autre domaine est constitué par la recherche des interactions et de nouvelles protéines expliquant la remarquable cohésion de la fourche de réplication. Enfin, de nombreuses questions restent posées

sur l'interaction de tous ces facteurs avec ceux responsables de la compaction de l'ADN. Bien que la réplication ait déjà été l'objet d'un nombre considérable de publications, beaucoup de travail semble encore nécessaire avant de pouvoir prétendre à la compréhension complète du mécanisme.

2 La recombinaison

Les réarrangements génomiques de grande ampleur, comme les duplications, les délétions, les translocations ou encore les insertions, sont pour la plupart engendrés par trois classes de processus moléculaires, qui diffèrent par la longueur des séquences qu'ils utilisent. La recombinaison homologe (HR) fait intervenir des séquences homologues et est générée par des protéines spécifiques. L'activité centrale nécessaire, appelée RecA, est présente dans tous les organismes vivants et catalyse l'échange des brins d'ADN (Egelman 2000). Un second type de recombinaison, appelée illégitime, se caractérise par l'emploi de séquences de longueur inférieure à celles utilisables par RecA. Chez les procaryotes, la recombinaison illégitime peut résulter de la simple ligation de deux séquences sans lien apparent (Hanada *et al.* 2000). Chez les eucaryotes, ce mécanisme est catalysé par des enzymes spécialisées et appelé NHEJ (« Non Homologous End Joining ») (Lewis et Resnick 2000). Enfin, les recombinaisons entre des répétitions en tandems mettent en jeu des voies spécifiques et forment la troisième classe de processus (Gordenin et Resnick 1998).

Le type de recombinaison auquel nous avons pour l'instant été confrontés au cours de l'étude des protéines At-SBP40 met en jeu des séquences homologues de plus de 200pb. Même si on ne peut exclure que ces protéines puissent jouer un rôle dans d'autres voies, nous nous contenterons ici de détailler les réactions responsables de la recombinaison homologe, qui est impliquée dans différentes tâches nécessaires au bon fonctionnement cellulaire. Elle maintient à la fois l'intégrité et la plasticité des génomes, garantit le bon fonctionnement de la réplication et participe à la réparation. Elle est également responsable de l'échange de matériel génétique entre, d'une part, les organites et le noyau et de l'autre entre les organismes, ces deux processus étant vitaux pour l'évolution des populations (Vulic *et al.* 1997).

2.1 L'échange de brins, cœur du mécanisme

Le mécanisme à la base de la RH, l'échange de brins (« DNA Strand Exchange ») ou « Strand invasion ») est catalysé par la famille de protéines RecA/Rad51. Chez *E. coli*, RecA est la seule protéine présente et occupe de ce fait un rôle central (Cox 2003). Chez *S. cerevisiae*, quatre homologues de *recA* ont été identifiés, *RAD51*, *RAD55*, *RAD57* et *DMC1* (Sung *et al.* 2003). Le gène *RAD51* est un orthologue de

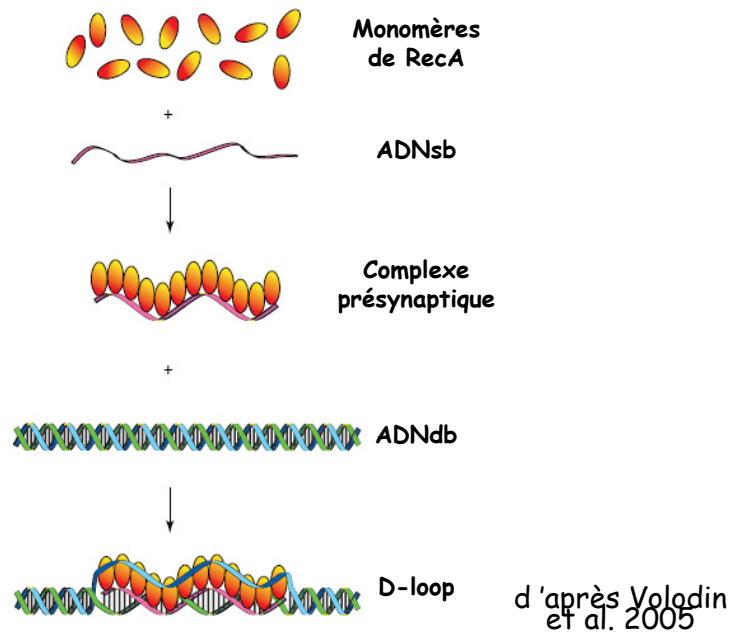


Figure i 1.3: Mode d'action des protéines de type RecA/Rad51

Des monomères de la protéine RecA fixent l'ADNsb de manière coopérative, conduisant à l'apparition d'un filament nucleo-protéique appelé complexe pré-synaptique. La reconnaissance d'une homologie entre l'ADNsb et la double hélice permet l'invasion de celle-ci par le nucléofilament, formant une structure appelée D-loop..

recA, l'hétérodimère Rad55-Rad57 facilite l'échange de brins catalysé par Rad51 en présence de la protéine RPA et DMC1 est un gène exprimé spécifiquement pendant la méiose. Chez *S. pombe*, on trouve cinq homologues de *recA* : *rhp51⁺*, *rhp55⁺*, *rhp57⁺*, *rlp1⁺* et *dmc1⁺* (Grishchuk et Kohli 2003). Enfin, chez les mammifères, 7 gènes de la famille sont présents, *RAD51*, *RAD51L1/B*, *RAD51L2/C*, *RAD51L3/D/TRAD*, *XRCC2*, *XRCC3* et *DMC1* (Liu *et al.* 1998). *RAD51L1/B*, *RAD51L2/C*, *RAD51L3/D/TRAD*, *XRCC2* et *XRCC3* sont des paralogues de *RAD51*, c'est à dire qu'ils dérivent vraisemblablement d'un ancêtre commun et qu'ils ont acquis des fonctions plus spécialisées que celles retrouvées chez les organismes moins élaborés. *DMC1*, comme c'est le cas chez la levure, est un gène spécifique de la méiose. Le gène *recA* de *E. coli* et son orthologue *RAD51* sont présents chez tous les organismes vivants, suggérant qu'ils sont essentiels à la recombinaison homologue (Kowalczykowski et Eggleston 1994). Enfin, le nombre de gènes paralogues semble varier proportionnellement à la taille du génome de l'organisme considéré.

La **Figure i1.3** présente le mécanisme central de la RH : la reconnaissance de l'homologie entre deux molécules d'ADN, catalysée par les protéines de la famille RecA/rad51. Un long filament nucléo-protéique, aussi appelé complexe pré-synaptique, est formé par la fixation, de manière hélicoïdale, autour de l'ADNsb des protéines RecA/rad51. Ce nucléo-filament incorpore de l'ADNdb de séquence homologue par un mécanisme qui n'est pas encore élucidé. Dans la structure caractéristique résultante, appelée D-loop (« Displacement-loop »), l'ADNsb contenu dans le nucléo-filament forme des paires de bases Watson-Crick avec le brin complémentaire du duplexe « cible » (Prevost et Takahashi 2003).

Un autre aspect des études sur la RH est constitué par des tests de son adaptabilité à tous les types de séquences, bien qu'elle ait été baptisée à l'origine par opposition à la recombinaison site-spécifique. En effet, les premières études *in vitro* pouvaient faire penser que la RH était un mécanisme séquence-indépendant. Toutefois, peu après sa découverte, différents groupes ont démontré, au moyen d'homopolymères (Cazenave *et al.* 1984) puis de séquences naturelles (Seitz et Kowalczykowski 2000), une préférence des protéines de la famille RecA/Rad51 pour des séquences riches en nucléotides G et T. D'autre part, plusieurs études ont confirmé cet état de faits en démontrant la distribution irrégulière des événements de recombinaison homologue sur les chromosomes eucaryotes. Cet effet est avéré chez plusieurs espèces comme la levure (Petes 2001) et l'humain (Kauppi *et al.* 2004). Bien que d'autres facteurs, comme l'état de la chromatine ou la présence de facteurs de transcription puissent être responsables d'une partie de ces observations, il semble évident à l'heure actuelle que la composition en bases du substrat a une influence sur ce mécanisme.

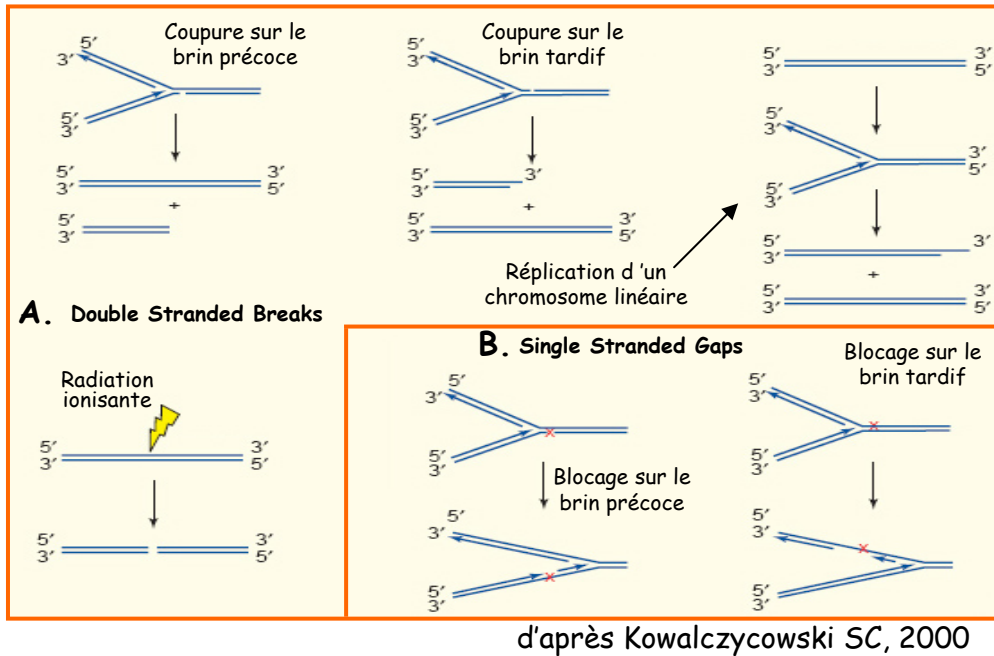


Figure i1.4: Génération de DSB (« Double Stranded breaks ») et de SSG (« Single Stranded Gaps »)

A. Les DSB peuvent être créés de différentes manières: par des radiations ionisantes, la rencontre d'une coupure sur l'un des deux brins par l'ADN polymérase ou encore simplement par la réplication d'un chromosome linéaire (voir Kass-Eisler K et Greider CW, 2000). **B.** Les SSG doivent leur origine à l'arrivée de la polymérase sur un obstacle (voir texte).

Figure i1.5 : voir page suivante ...

..

2.2 Les évènements produisant des cassures de l'ADN

La **Figure i 1.4** présente les différents types de lésions potentiellement réparables par le mécanisme RH. Des cassures double ou simple brin, DSB (« Double Stranded Breaks ») ou SSG (« Single Stranded Gaps ») peuvent être produites de deux manières distinctes : i) directement, c'est à dire par des radiations ionisantes comme les rayons UV et γ , ou encore par des agents mutagènes chimiques ii) indirectement, c'est à dire par le passage de la fourche de réplication sur un brin matrice endommagé. Une coupure du squelette de l'ADN sur le brin à synthèse continue a pour conséquence une DSB franche lors de la réplication (Kowalczykowski 2000). La même lésion sur le brin tardif produit quant à elle une DSB possédant une queue d'ADNsb 3' sortante. Enfin, la simple réplication sans encombre d'un chromosome produit une DSB portant une extrémité 3' en ADNsb qui doit en quelque sorte également être réparée (Kass-Eisler et Greider 2000). De la même manière, des coupures simple brin peuvent être causées par la réparation de dommages d'origine exogène (Evans *et al.* 2004) (Fang *et al.* 2002) (bases oxydées, alkylées, etc...) ou par le blocage de la fourche de réplication après la rencontre d'anomalies (par exemple des mésappariements) (Schofield et Hsieh 2003). Il peut paraître paradoxal que des lésions potentiellement létales soient nécessaires à un processus biologique aussi universel et indispensable que la recombinaison. Il semble donc qu'au cours de l'évolution, les organismes aient trouvé par ce biais le moyen de rendre profitable une situation pouvant à priori apparaître dommageable.

2.3 Le modèle DSBR chez *E.coli*

Nous allons à présent détailler le fonctionnement du mécanisme de la recombinaison homologue dans sa globalité, qui permet également de comprendre comment sont réparées les DSB, et appelé pour cette raison modèle DSBR (« Double Strand Break Repair model ») (Szostak *et al.* 1983). Ce processus peut être décomposé en quatre étapes : i) initiation, consistant en la prise en charge des extrémités franches en vue de la production d'ADNsb nécessaire à l'invasion de la double hélice par le nucléo-filament. ii) hybridation par homologie et échange des brins: la protéine RecA se fixe sur l'ADNsb, ce qui, comme nous l'avons vu dans le paragraphe précédent, permet la recherche de l'homologie sur une autre molécule d'ADNdb iii) extension de l'hétéroduplexe (« branch migration ») par la mise en œuvre du complexe protéique RuvAB et iv) résolution qui consiste en la cassure de la jonction par RuvC, qui clive certains des brins (West 1996).(voir **Figure i 1.5**).

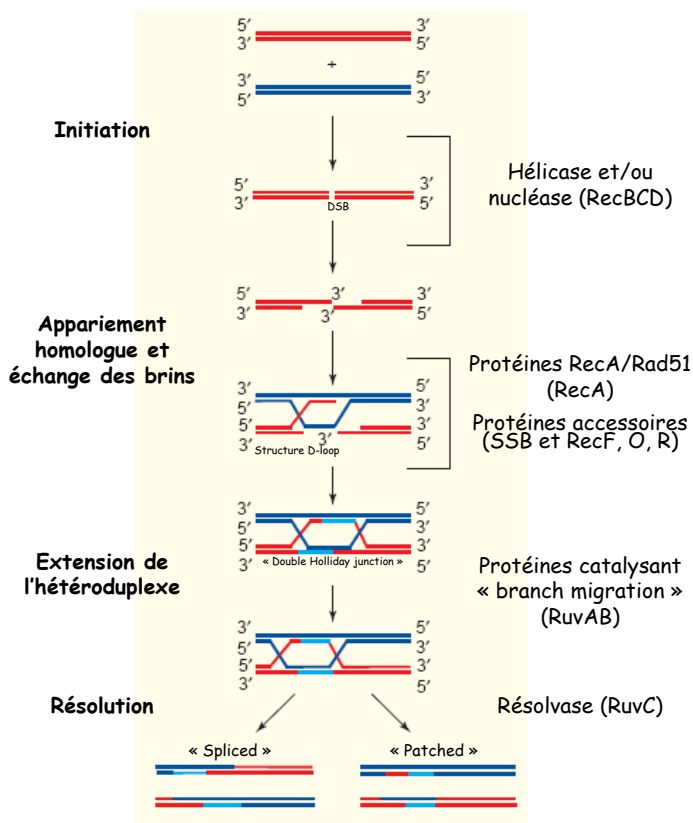


Figure i1.5: Modèle du DSBR (« Double Stranded Break Repair »)

Ce modèle est généralisable à la recombinaison homologue de la plupart des organismes. Les deux premières étapes sont également utilisées pour l'initiation de la réplication recombinaison-dépendante. Les enzymes mentionnées à côté du schéma sont celles employées par *E.coli*. L'ADN synthétisé lors de ce processus est représenté en bleu ciel.

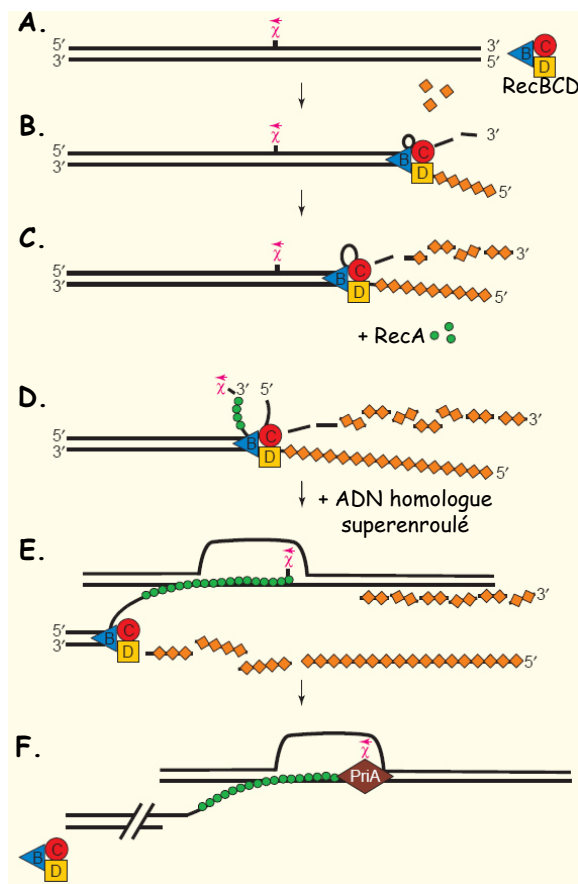


Figure i1.6: Les deux premières étapes de la recombinaison homologue

- A. Fixation de RecBCD à l'extrémité de l'ADNdb
- B. & C. Déroulement de l'ADNdb et dégradation préférentielle des extrémités 3'
- D. Reconnaissance de χ , atténuation de l'exo- 3'→5', activation de l'exo- 5'→3', chargement de RecA
- E. Appariement homologue et échange des brins
- F. Dissociation de RecBCD, fixation de PriA à la boucle D

Les deux figures sont tirées de Kowalczykowski SC, 2000

2.3.1 Modes de recrutement de RecA

2.3.1.a La voie RecBCD

La **Figure i 1.6** détaille la voie RecBCD de la première étape du modèle DSBR. Comme nous venons de le voir, l'enzyme RecBCD reconnaît une DSB franche. Simultanément à sa progression sur la double hélice, elle digère l'ADNsb au moyen d'une activité exonucléase 3'→5'. L'une de ses caractéristiques essentielles est sa capacité à moduler son activité lors de son passage sur un site χ (5'-GCTGGTGG-3') (Sourice *et al.* 1998). Cela prend la forme d'une abolition de l'activité exo- 3'→5', en faveur d'une faible activité exo- 5'→3', sans altération de l'activité hélicase (Anderson et Kowalczykowski 1997). La conséquence à ce changement est l'apparition d'ADNsb portant χ à son extrémité 3', ce qui représente la configuration idéale pour la fixation de RecA (voir paragraphe 2.1). Dès l'interaction avec χ , RecBCD dirige le chargement de RecA sur l'ADNsb en excluant la fixation des protéines SSB, afin d'empêcher leur activité inhibitrice (Arnold et Kowalczykowski 2000). En effet, des travaux antérieurs qui cherchaient à évaluer la capacité des SSB à stimuler les recombinaisons avaient conduit à l'observation de comportements opposés en fonction des conditions réactionnelles. Ces travaux montraient que la saturation de l'ADNsb par les SSB préalablement à l'incubation avec RecA avait un effet inhibiteur sur la recombinaison. D'un autre côté, si la recombinaison était chargée en priorité, l'addition de SSB au milieu réactionnel stimulait nettement le processus d'échange de brins par rapport à un témoin sans SSB (Sugiyama *et al.* 1997). Le chargement préférentiel de RecA est donc une fonction biologique essentielle de l'interaction χ -RecBCD. L'hypothèse actuelle pour expliquer ce phénomène semble être un changement conformationnel de la sous-unité D consécutif à la rencontre de χ (Dixon *et al.* 1994). Une fois RecA chargée, la recherche d'homologie peut avoir lieu et engendrer la formation de la boucle D. Là encore, la protéine SSB favorise l'échange, en se fixant cette fois au brin déplacé, empêchant ainsi les brins de la double hélice envahie de se ré-hybridiser (Mazin et Kowalczykowski 1998).

2.3.1.b La voie alternative RecF

Nous allons à présent passer en revue le second mode de recrutement de RecA, la voie RecF, en charge de la réparation des SSG (Tseng *et al.* 1994) qui peut également intervenir dans la réparation des DSB. Cette voie, moins caractérisée pour l'instant que la voie RecBCD, implique les protéines RecA, RecF, RecG, RecJ, RecN, RecO, RecQ, RecR, RuvA, RuvB, RuvC et les protéines SSB (Kowalczykowski *et al.* 1994). Plusieurs données prouvent que cette voie est essentielle pour la quasi-totalité des événements de recombinaison autres que ceux engendrés par les DSB : i) la

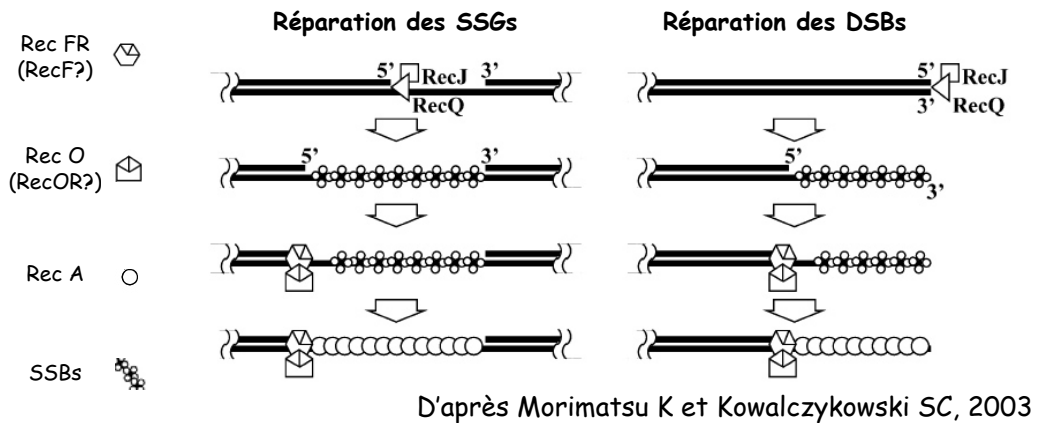


Figure i 1.7: Recrutement de RecA lors de la voie de RecF

La cassure est traitée par les enzymes Rec J et RecQ, qui permettent l'obtention d'une région d'ADNs_b suffisamment longue pour l'initiation des étapes suivantes du processus. Les protéines RecFOR jouent le rôle de médiateurs: elles sont capables de charger RecA sur une région d'ADNs_b recouverte par les SSB. Il faut rappeler le fait que la voie RecF est « normalement » en charge des SSG ; elles n'interviennent sur les DSB que lorsque la voie Rec BCD est inutilisable.

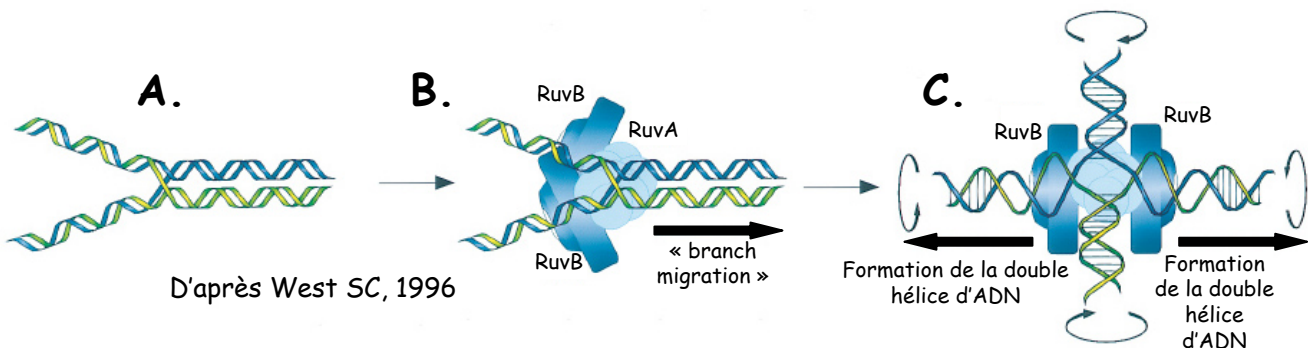


Figure i 1.8: Mode de fonctionnement de RuvAB

Les « crossing-over » formés pendant la recombinaison (A.) sont spécifiquement reconnus par RuvA, conduisant à l'assemblage du complexe RuvAB à la jonction de Holliday. L'hypothèse est que ce complexe, via l'activité ATP-ase ADN-dépendante de RuvB, soit en charge du déplacement de la jonction (B.). L'activité conjuguée des deux enzymes constitue une « pompe » moléculaire, agissant dans un même plan (C.). Pour des raisons de clarté, l'ADN est schématisé au premier plan, alors qu'il se situe bien entendu à l'intérieur de l'anneau (hexamère) constitué par l'assemblage des sous-unités RuvB.

réparation des dommages dus aux UV, produisant à la fois des DSB et des SSG, nécessite à la fois les voies RecBCD et RecF (Kowalczykowski *et al.* 1994) ii) la recombinaison plasmidique requiert la voie RecF (Kolodner *et al.* 1985) iii) la recombinaison pendant la conjugaison et à d'autres sites que les DSB utilise la voie RecF (Smith 1991). En parallèle, des expériences utilisant des souches de *E.coli* dont la voie RecBCD était inactivée par des mutations ont permis de constater que la voie RecF peut être activée pour prendre en charge les DSB (Lloyd et Buckman 1985). Dans ces cellules, l'efficacité de la recombinaison par la voie RecF est équivalente à celle de la voie RecBCD dans des cellules sauvages. Récemment, des expériences *in vitro* ont démontré que le recrutement de RecA pouvait être effectué de la même manière sur les SSG et les DSB (voir **Figure i 1.7**). L'hélicase RecQ (Umezumi *et al.* 1990) et l'exonucléase 5'→3' RecJ (Lovett et Kolodner 1989) prennent en charge la cassure, en se fixant sur une extrémité 5', conduisant à l'apparition d'une extrémité 3' sous forme d'ADNsb, aussitôt recouverte par la SSB. L'extrémité 5' rentrante résultante est fixée par RecF, puis RecR et enfin RecO. Les protéines RecFOR permettent alors d'initier la synthèse du nucléofilament de RecA malgré la présence des protéines SSB. Le nucléofilament peut alors s'allonger, dans la direction 5'→3' (Morimatsu et Kowalczykowski 2003).

2.3.2 « Branch migration » et résolution de la jonction

Le complexe protéique ATP-dépendant RuvAB est une hélicase responsable de la phase de migration des brins de la recombinaison, en faisant migrer la région de crossing-over, après la formation de la jonction de Holliday (West 1996) (voir **Figure i 1.8**). RuvA dirige le complexe vers la jonction, permettant à RuvB de s'assembler sous forme d'hexamère autour des deux bras opposés de la jonction. L'ADN est ensuite tiré hors du complexe RuvAB par translocation dans des directions opposées pour permettre la migration des brins. La protéine RecG possède une activité similaire, mais agit en direction inverse afin de dissocier les jonctions (Whitby *et al.* 1993). La perte de cette activité stimule la réplication recombinaison-dépendante (Hong *et al.* 1995).

L'endonucléase RuvC qui clive spécifiquement la jonction afin de séparer les brins (Connolly *et al.* 1991), agit conjointement avec RuvAB. L'enzyme RuvABC présente une forte affinité pour la structure formée par la jonction de Holliday et l'action des trois protéines est hautement coordonnée (Eggleston *et al.* 1997). Bien que la protéine Ruv C agisse tardivement dans la recombinaison, elle semble également avoir un rôle dans l'initiation de la réplication recombinaison-dépendante (Seigneur *et al.* 1998), en agissant sur les fourches de réplication bloquées ayant régressé en jonction de Holliday. Ce clivage est un autre moyen de créer des DSB pouvant servir

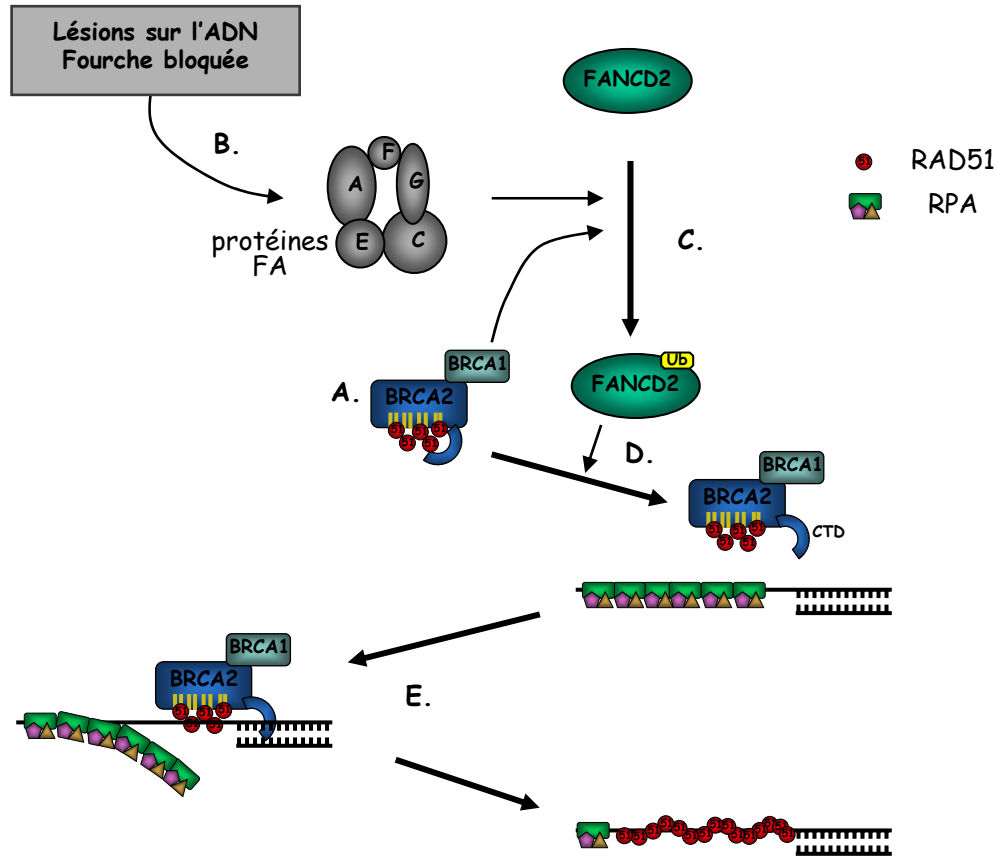


Figure i1.9: Mode d'action de BRCA2

A. Dans des noyaux non endommagés, BRCA2 « séquestre » RAD51 qui est alors inactive. Les protéines de l'anémie de Fanconi (FA), FANCA, FANCC, FANCE, FANCF et FANCG sont présentes sous forme d'un complexe. **B.** En réponse à une lésion sur l'ADN ou durant la réplication, le complexe FA peut être activé. **C.** Cela qui engendre la monoubiquitylation (Ub) de FANCD2. Ce processus de modification nécessite également la présence de BRCA1. **D.** Il est probable que ce soit ce signal qui engendre une modification post-transcriptionnelle du domaine C-terminal (CTD) de BRCA2, permettant à RAD51 de se fixer à l'ADNsb, vraisemblablement par le biais des « OB-folds » de BRCA2. **E.** Lorsque la protéine RPA a été déplacée, BRCA2 peut décharger RAD51, qui forme un nucléofilament hélicoïdal sur l'ADNsb.

de site d'entrée pour RecBCD. Avec cette capacité, RuvC, semble donc également avoir un rôle d'initiateur de la recombinaison au niveau de sites où les fourches de réplication sont bloquées.

2.4 Les données concernant les eucaryotes

Comme chez tous les organismes, l'étape clé de la recombinaison eucaryote est catalysée par les protéines de la famille RecA/RAD51. Des observations en microscopie électronique (Yu *et al.* 2001) et l'étude d'une structure de faible résolution obtenue par cristallographie (Pellegrini *et al.* 2002) ont montré que la protéine RAD51 humaine forme également un nucléofilament en association avec l'ADNsb. Le mode de formation de celui-ci présente par contre des différences par rapport au cas de *E.coli*. En effet, RAD51 possède une affinité plus faible pour l'ADNsb que pour l'ADNdb (Mazin *et al.* 2000). La formation du nucléofilament nécessite de ce fait la présence d'autres enzymes, et plusieurs protéines capables de remplir cette fonction ont déjà été identifiées. On peut citer RAD52 (Sung 1997), qui est à la fois capable de stimuler le processus d'échange catalysé par RAD51 en présence de RPA (New *et al.* 1998) et l'hybridation de deux molécules d'ADNsb indépendamment de RAD51 (Mortensen *et al.* 1996). Des études chez la levure et la souris ont montré que l'efficacité de l'invasion est également dépendante d'une autre protéine, RAD54 (Van Komen *et al.* 2000), qui favoriserait l'invasion par le biais de changements topologiques de la double hélice cible (Sigurdsson *et al.* 2002) et qui aurait aussi un rôle dans le recyclage de RAD51 après formation de la boucle D (Solinger *et al.* 2002).

Le cas le plus intéressant est sans doute celui des protéines BRCA1 et BRCA2, appelées suppresseurs de tumeurs, car des mutations dans les gènes *BRCA1* et *BRCA2* prédisposent les individus qui les portent à des cancers du sein et des ovaires (Venkitaraman 2002). Les mécanismes impliquant ces protéines ont pour cette raison été l'objet d'un grand nombre d'études au cours des dernières années, et un modèle pour leur mode d'action peut à présent être envisagé (voir **Figure i 1.9**). Sa compréhension nécessite de détailler quelque peu les propriétés de la structure de BRCA2. Elle porte dans sa partie centrale huit motifs d'une trentaine d'acides aminés, appelés BRC, dont la structure tridimensionnelle mime celle d'une région impliquée dans la capacité d'oligomérisation de RAD51 (Pellegrini *et al.* 2002). De plus, on trouve en C-terminal trois domaines « OB-fold » (voir Chapitre II), conférant la capacité de fixation aux acides nucléiques (Yang *et al.* 2002). Son mode de fonctionnement, encore hypothétique pour l'instant, pourrait être le suivant (Venkitaraman 2002): i) lorsqu'aucun processus de recombinaison n'est nécessaire, elle serait en charge de séquestrer RAD51 dans le noyau sous forme

inactive, par le biais des domaines BRC ; ii) les signaux induits par des lésions sur l'ADN, par l'intermédiaire de modifications post-traductionnelles, engendrent la translocation BRCA2-dépendante de RAD51 sur les lieux endommagés où BRCA2 régule la formation du nucléofilament ; iii) lorsque les réparations ont été effectuées, BRCA2 est capable d'empêcher la formation ou l'élongation des nouveaux nucléofilaments.

Les mécanismes moléculaires responsables des étapes suivantes de la recombinaison sont peu connus chez les eucaryotes. Par exemple, dans l'attente d'une confirmation expérimentale, plusieurs théories existent pour envisager le mode de capture du second brin après formation de la boucle D (Paques et Haber 1999). D'autre part, les équivalents fonctionnels des protéines RuvABC sont longtemps restés introuvables. Très récemment, des études menées sur des cellules HeLa ont démontré que RAD51L2/C était indispensable aussi bien à la migration des brins qu'à la résolution de la jonction de Holliday (Liu *et al.* 2004). Le niveau de compréhension du mécanisme dans son ensemble demeure toutefois beaucoup plus faible que chez *E. coli*.

On peut enfin formuler quelques remarques sur la façon dont est mis en route le mécanisme de réponse aux lésions sur l'ADN. Il n'existe pas chez les eucaryotes de système comparable à la réponse SOS. Les protéines impliquées dans la réparation des lésions sont présentes à taux constant et de manière diffuse dans le noyau durant tout le cycle cellulaire. Elles sont relocalisées selon les besoins, comme après une irradiation ionisante, sous forme de complexes subnucléaires appelés « foci » (Scully *et al.* 1997). Des foci contenant RAD51 peuvent se former après différents types de traitements et les co-facteurs RAD52, RAD54 (Essers *et al.* 2002), RPA (Raderschall *et al.* 1999) ou les suppresseurs de tumeurs BRCA1 (Scully *et al.* 1997) et BRCA2 (Chen *et al.* 1998) peuvent également y être observés. Enfin, ces foci sont dynamiques, c'est à dire que les cinétiques de localisation des différentes protéines sont variables, comme l'illustre l'incapacité de lignées cellulaires n'exprimant pas BRCA2 à accumuler des foci contenant RAD51 en réponse à des lésions sur l'ADN (Godthelp *et al.* 2002).

3 Conclusion

Comme nous venons de le voir, la recombinaison homologe est centrée autour du mécanisme d'échange des brins, catalysé par la famille de protéines RecA/RAD51, et retrouvé chez tous les organismes vivants. Les modes de régulation de celui-ci et les protéines modulatrices impliquées sont néanmoins très différents, et apparaissent de plus en plus élaborés au fur et à mesure de l'évolution. Les nombreuses informations

A. DSBR

B. Réplication recombinaison-dépendante

C. Recombinaison répliation-dépendante

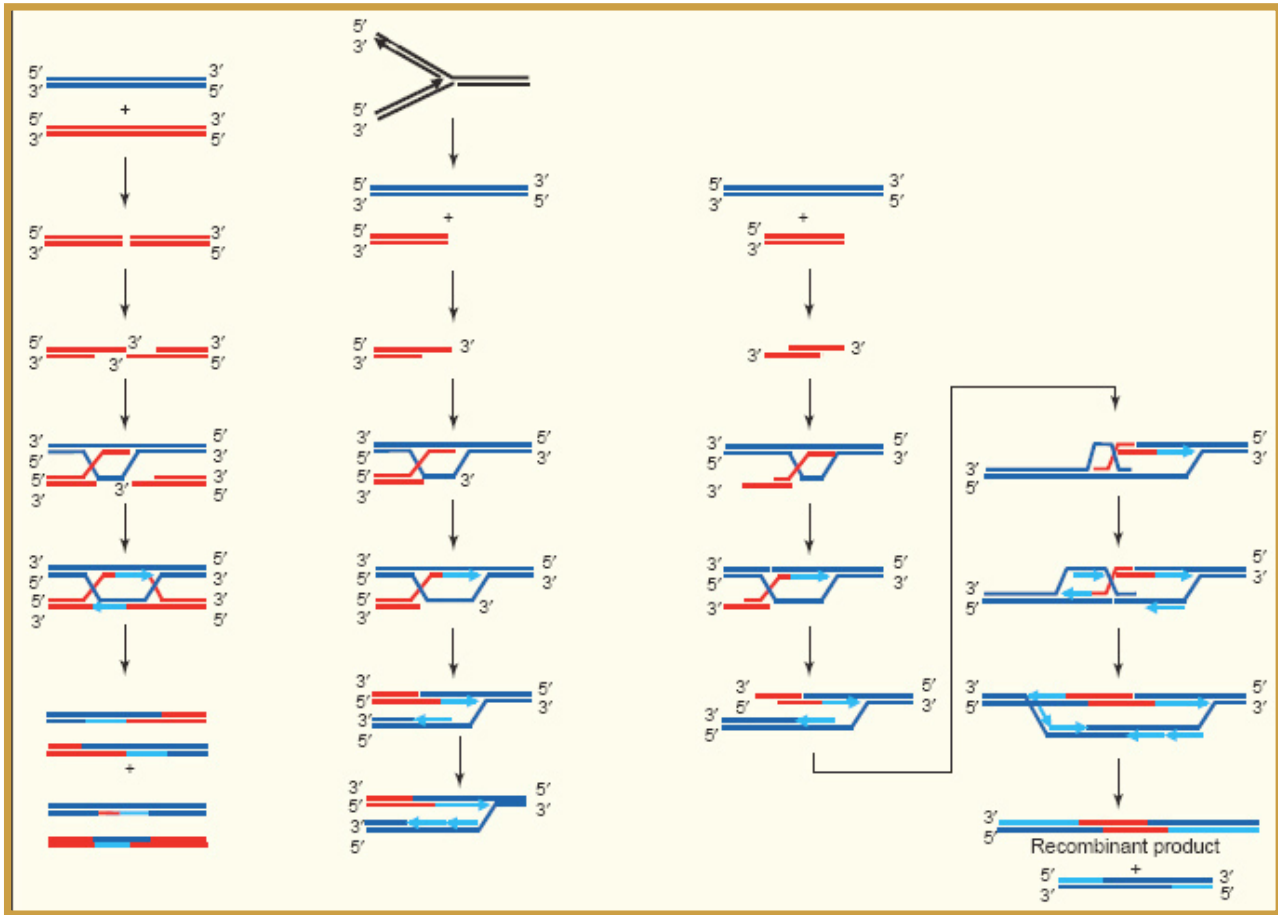


Figure i 1.10: Modèles de l'utilisation des DSB pour la maintenance des génomes

- A. Recombinaison résultant de la réparation d'un DSB. L'invasion de la double hélice homologue intacte par chacune des extrémités 3' en ADNsb ne sont pas forcément simultanées
- B. Le bon déroulement de la réplication est dépendant de la recombinaison. L'exemple présenté est celui d'une cassure sur la matrice du brin à synthèse continue, qui génère un DSB. Si la cassure se trouve sur le brin tardif (non représenté, voir **Figure i 1.4**), il est possible qu'un DSB portant une extrémité 3' en ADNsb (ne nécessitant donc pas de traitement avant invasion) soit générée.
- C. Exemple d'un modèle de recombinaison dépendant de la réplication montrant l'intégration par conjugaison d'un fragment d'ADNdb linéaire (rouge) dans un chromosome (bleu). Le traitement des deux extrémités du fragment est présenté en un seule étape, mais ces évènements ne sont pas obligatoirement simultanés.

d 'après Kowalczykowski SC, 2000

présentées ici avaient avant tout pour but de mettre en lumière le rôle primordial de l'équilibre RecA/SSB (ou RAD51/RPA), qui est responsable de la poursuite ou non du processus de recombinaison.

Au cours de cette partie dédiée à la recombinaison, nous avons essentiellement entrevu son rôle dans la réparation des lésions sur l'ADN. Cela s'explique par le fait qu'il s'agisse là de sa « raison d'être », et que les autres processus dans lesquels elle est impliquée nécessitent également, et peut-être paradoxalement, la présence de ces lésions. Comme nous l'avons vu, les grands axes de la maintenance des génomes, la réplication, la recombinaison et la réparation sont étroitement liés, étant interdépendants et utilisant les mêmes protéines. Pour finir, nous allons donc brièvement voir comment la recombinaison homologue peut, à partir d'un DSB, jouer son rôle dans la redistribution des gènes durant la méiose et faire redémarrer une fourche de réplication (voir **Figure i 1.10**). Il est à noter que le modèle présenté pour ce dernier cas est vraisemblablement celui employé pour l'initiation de la réplication, une fonction occupée par la recombinaison dans certains systèmes (comme par exemple chez le phage T4 (Mosig 1987)). Enfin, on peut encore ajouter que les processus conduisant à l'orientation vers l'une ou l'autre de ces voies restent pour l'heure inconnus, et impliquent sans doute de nombreuses protéines qui n'ont pas encore été identifiées.

CHAPITRE II: LES PROTEINES LIANT L'ADN SIMPLE BRIN (SSB)

Il est apparu au cours de la caractérisation fonctionnelle des protéines de la famille *At*-SBP40, que celles-ci étaient capables de fixer les acides nucléiques, avec une nette préférence pour l'ADNsb. Il semble donc nécessaire de présenter dans cette introduction les informations disponibles concernant les protéines de la classe de SSB, qui comme nous avons eu l'occasion de nous en rendre compte durant le chapitre précédent, est un composant essentiel des toutes les étapes du métabolisme de l'ADN. Elles sont en effet impliquées dans la réplication (Brill et Stillman 1991), la recombinaison (Alani *et al.* 1992) et la réparation (Alani *et al.* 1992), et de nombreuses études ont montré leur capacité à stimuler les activités de protéines catalytiques comme les polymérases (Robbins *et al.* 2004), les hélicases (Seo et Hurwitz 1993) ou les recombinases (Komori et Ishino 2001). L'interaction SSB/ADNsb se caractérise par une très faible (voire inexistante) spécificité de séquence, et a vraisemblablement pour buts principaux de protéger l'ADNsb de l'attaque des nucléases, d'empêcher la formation de structures secondaires et de former une structure particulière avec cet ADNsb, qui facilitera la reconnaissance par les enzymes catalytiques (Perales *et al.* 2003). Au vu de ces caractéristiques, il n'est pas surprenant de constater que les SSB soient ubiquitaires : des protéines fixant l'ADN simple brin ont été caractérisées dans les trois domaines du vivant: Archéobactéries (Robbins *et al.* 2004), Eubactéries (Lohman et Ferrari 1994) et Eucaryotes (Wold 1997) , mais également chez certains virus, comme l'adénovirus (Tucker *et al.* 1994) et les bactériophages (Alberts et Frey 1970). Nous allons donc détailler ici les propriétés de ces protéines, avec pour but d'offrir une vue globale de leurs caractéristiques en matière de fonction et de structure.

1 Propriétés générales

Parmi les protéines fixant l'ADNsb, trois ont fait l'objet d'un nombre d'études particulièrement important, ce sont celles codées par le bactériophage T4 (gp32), *E. coli* (*Eco*ssb) et *H. sapiens* (hRPA). Nous allons commencer par comparer ces trois protéines pour des propriétés caractéristiques de l'interaction SSB/ADNsb.

1.1 Affinité pour l'ADNsb et abondance dans les cellules

Certaines protéines dont la fonction biologique avérée est la fixation de l'ADNsb possèdent également la capacité de fixer, dans une moindre mesure, l'ADNdb et l'ARN. Un critère essentiel pour définir une protéine comme fixant spécifiquement l'ADNsb est donc défini par le rapport de ses affinités pour l'ADN sous forme simple et double brin. L'affinité relative (ADNsb/ADNdb) est d'au moins 10^8 pour la SSB de *E. coli* (Gasior *et al.* 2001), d'au moins 10^6 pour gp32 (Jensen *et al.* 1976) et d'environ 10^3 - 10^4 pour la protéine hRPA (Wold 1997). Les valeurs observées pour ce rapport présentent donc des variations considérables, et ont probablement des conséquences sur le mode de fonctionnement de ces protéines. Cela est illustré par le fait que la RPA humaine est capable, en plus de ses fonctions de SSB classique, d'interagir avec l'ADNdb flanquant l'origine de réplication de SV40 (Iftode et Borowiec 1997). Ce procédé, qui permet la poursuite de l'ouverture de la double hélice après l'action initiale de la protéine virale AgT, est une étape essentielle pour la réplication du virus (Iftode et Borowiec 1998).

Ces différences de spécificité pour l'ADNsb sont à mettre en parallèle avec une autre caractéristique permettant de définir le « taux d'activité » de liaison à l'ADNsb présent dans les différentes cellules, à savoir l'abondance de ces différentes protéines. Ce paramètre se définit comme le nombre de copies de la protéine présentes par cellule. L'abondance de hRPA a été évaluée à 10^4 - 10^5 (Wold 1997), celle de gp32 à 1 - 3.10^4 (Gasior *et al.* 2001) par cellule infectée et celle de *Eco*ssb à $0,3$ - 2.10^3 (Gasior *et al.* 2001). Deux remarques doivent être formulées à propos de ces observations. Tout d'abord, si l'on rapporte les valeurs constatées pour *Eco*ssb et gp32 au nombre de fourches de réplication attendues dans chacun des cas, ces deux protéines s'avèrent présentes à des taux équivalents : il y a suffisamment de *Eco*ssb et de gp32 pour couvrir environ 1400 nucléotides à chaque fourche de réplication (Chase et Williams 1986). De la même manière, l'abondance élevée de hRPA s'explique certainement par le nombre beaucoup plus élevé de zones d'initiation chez les eucaryotes par rapport aux procaryotes. En dépit de ces différences, toutes ces protéines ont une abondance que l'on peut qualifier de très élevée, ce qui est compréhensible au vu du type de fonction (protection de l'ADNsb) qu'elles ont à remplir dans les cellules.

1.2 Organisation multimérique et modes de fixation

La classe des SSB peut être divisée en deux sous-familles, si l'on s'intéresse à leur organisation multimérique lorsqu'elles fixent l'ADNsb : les protéines virales (gp32 et la protéine du gène 2.5 du bactériophage T7) agissent *in vivo* sous forme d'un monomère (Giedroc *et al.* 1987) (de 33 kDa pour gp32) alors que les SSB bactériennes

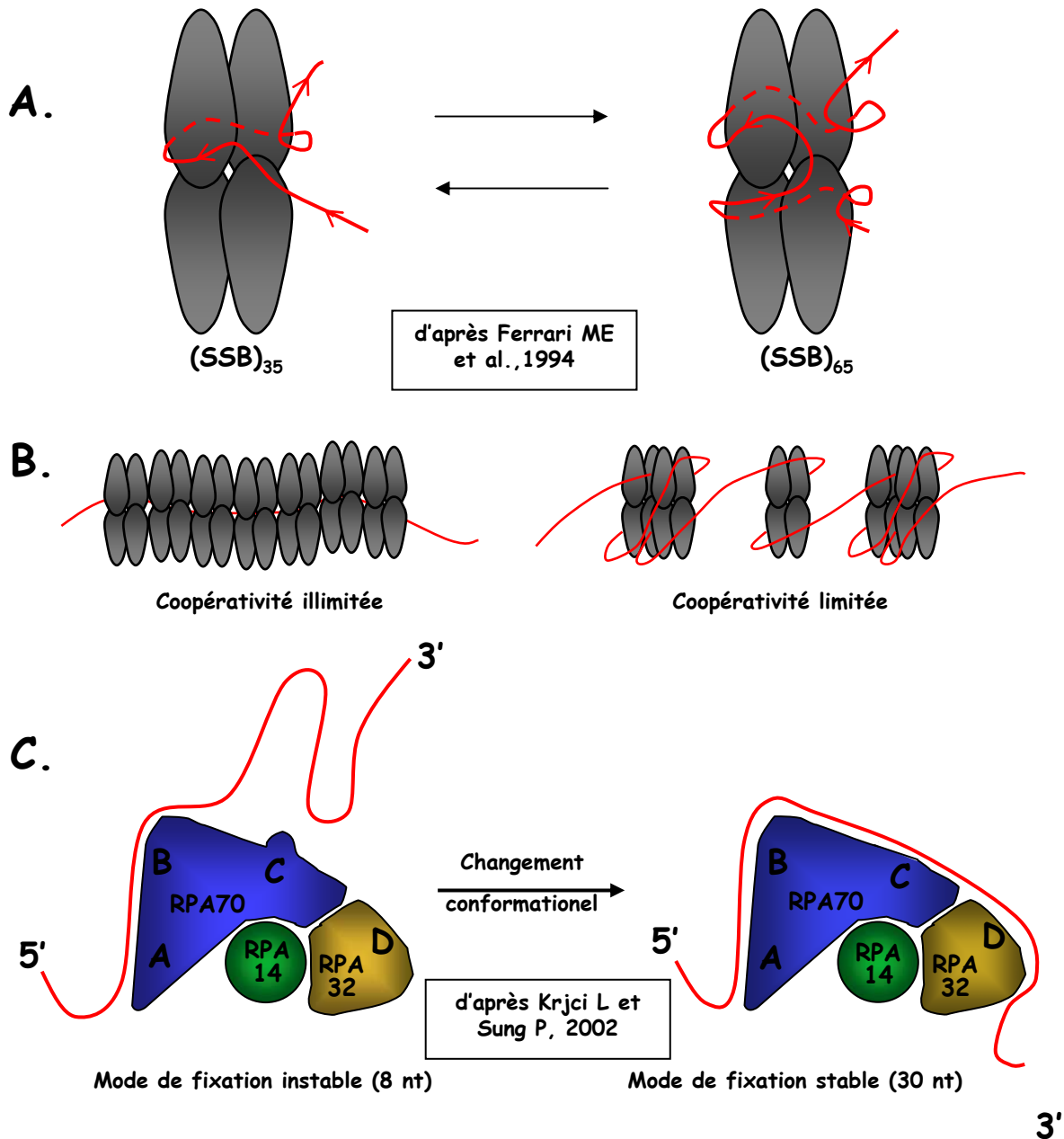


Figure i 2.1: Organisation multimérique et modes de fixation de *Ecosb* et RPA

A. Modes de fixation de *Ecosb*. Dans le mode de fixation $(SSB)_{35}$, l'ADNsb n'interagit qu'avec deux protomères du tétramère alors que dans le mode $(SSB)_{65}$, l'interaction implique les quatre protomères. **B. Modèles des deux types de coopérativité utilisés pour décrire la fixation d'*Ecosb* à l'ADNsb.** Le mode de fixation $(SSB)_{35}$ permet une coopérativité dite illimitée. La protéine peut former des interactions avec son plus proche voisin des deux côtés du tétramère. Le mode de fixation $(SSB)_{65}$, dans lequel les quatre protomères fixent l'ADNsb, limite la coopérativité à la formation de dimères de tétramères (c'est à dire d'octamères). Ce mode de fixation rend difficile la saturation complète de l'ADNsb. **C. Les deux modes de fixation de RPA.** Représentation schématique de l'hétérotrimère RPA, avec les quatre domaines de fixation désignés A, B, C et D. Le mode de fixation initial, impliquant 8 nucléotides, est instable et fait intervenir les domaines A et B. Un changement conformationnel du domaine C permet le passage au mode de fixation stable, faisant intervenir les quatre domaines, et interagissant avec 30 nucléotides de l'ADNsb.

et les RPAs eucaryotes présentent une structure quaternaire composée de quatre motifs de fixation aux acides nucléiques. *Eco*ssb forme un homotétramère de quatre sous-unités identiques de 19 kDa, ou protomères (Lohman et Ferrari 1994). RPA fixe quant à elle l'ADNsb sous la forme d'un hétérotrimère, composé chez la protéine humaine des protomères RPA70, RPA32 et RPA14, tous étant essentiels à la viabilité (Wold 1997). Comme chez *Eco*ssb, quatre motifs de fixation aux acides nucléiques sont présents, trois (A, B et C) sont portés par RPA70 et un (D) par RPA32, le rôle de RPA14 restant pour l'heure inconnu.

Ces modes d'organisation ont des conséquences directes sur l'interaction des différentes protéines avec les acides nucléiques. La protéine gp32 possède un seul mode de liaison à l'ADNsb, dans lequel le site de fixation occupe huit nucléotides (nt) de l'ADNsb cible (Prigodich *et al.* 1984). La situation est plus complexe pour *Eco*ssb et hRPA. La protéine bactérienne présente en effet au moins deux modes de fixation différents (voir **Figure i 2.1A**). Des études *in vitro* ont montré qu'à des concentrations faibles de sels (< 10 mM NaCl) (Ferrari *et al.* 1994), deux protomères fixent 35 nucléotides d'ADNsb (mode de fixation appelé (SSB)₃₅) alors qu'à des concentrations plus élevées ($\geq 0,2$ M NaCl) (Lohman et Overman 1985), quatre protomères se lient à 65 nt (mode (SSB)₆₅). Ce dernier cas induit la compaction de l'ADNsb, en raison de l'enroulement de celui-ci autour du tétramère (Bujalowski *et al.* 1988). L'hypothèse actuelle est que ces deux modes puissent être employés *in vivo*, en fonction de la situation rencontrée par *Eco*ssb. La situation chez les eucaryotes ressemble à celle trouvée chez les bactéries, dans la mesure où deux modes de fixation sont également présents (voir **Figure i 2.1C**). Un premier mode est constitué par la liaison de l'ADNsb par les domaines A et B de la sous-unité RPA70. Dans celui-ci, l'ADNsb est engagé de façon polaire, préférentiellement par son extrémité 5' (de Laat *et al.* 1998). Un changement conformationnel, de nature inconnue pour l'instant, réoriente le domaine C, conférant la capacité à l'ensemble des quatre domaines de l'hétérotrimère de se lier au substrat, dont 30 nt entrent alors en contact avec la protéine. Le mode de fixation impliquant 8 nt est instable, mais représente vraisemblablement la voie d'entrée de l'ADNsb dans les domaines de fixation. Il semble constituer un intermédiaire obligatoire pour le passage au mode stable, c'est à dire celui faisant intervenir 30nt.

1.3 La coopérativité

Un paramètre essentiel de l'interaction SSB/ADNsb est la coopérativité (ω). Elle représente la probabilité pour une molécule de SSB de se fixer sur une molécule d'ADNsb déjà fixée par une SSB, plutôt que sur une molécule d'ADNsb isolée. Une coopérativité positive peut résulter d'interactions protéines-protéines, comme cela

semble être le cas pour gp32 (Giedroc *et al.* 1990), ou simplement de la distorsion de l'ADNsb consécutive à la fixation d'une première SSB. De manière générale, les SSB sont coopératives, mais ce paramètre varie grandement en fonction des protéines. Il existe, de plus, deux modèles, appelés limités et illimités, pour décrire la coopérativité observée chez les SSB. Dans le modèle de coopérativité illimité, les interactions protéines-protéines peuvent avoir lieu des deux côtés d'un tétramère, permettant la formation d'amas de protéines saturant l'intégralité de l'ADNsb. Ce modèle s'applique aux interactions entre l'ADNsb avec gp32, mais également à *Eco*ssb lorsque celle-ci interagit selon le mode (SSB)₃₅ (voir **Figure i2.1B**). Les valeurs de coopérativité qui ont été déterminées pour ces deux protéines sont de $\omega_{\text{illim.}} = 2.10^3$ et $\omega_{\text{illim.}} = 10^5$, respectivement. Le modèle de coopérativité limité a quant à lui été développé pour décrire les aspects coopératifs uniques de *Eco*ssb observée en microscopie électronique (Chrysogelos et Griffith 1982) en mode (SSB)₆₅ (voir **Figure i2.1B**). L'hypothèse à la base de ce modèle est que, en l'enroulant complètement, l'ADNsb empêche le tétramère d'interagir avec ses voisins comme il le ferait en mode (SSB)₃₅. De ce fait, le niveau de complexité maximal observé pour le mode (SSB)₆₅ est l'octamère. Curieusement, la coopérativité dans ce cas a été trouvée indépendante du pH ou du type et de la concentration du sel présent dans le milieu, mais dépendante de la température. Les valeurs suivantes ont été obtenues : à 10°C, $\omega_{\text{illim.}} = 170 \pm 130$, à 25°C, $\omega_{\text{illim.}} = 420 \pm 80$ et à 37°C, $\omega_{\text{illim.}} = 1000 \pm 400$ (Lohman et Ferrari 1994). La coopérativité est également dépendante du taux de saturation de la molécule d'ADNsb cible, plus celle-ci est saturée, plus la proportion d'octamères est importante (Bujalowski et Lohman 1987). A 25°C, et pour un taux de saturation de 50 %, les tétramères sont également distribués entre les octamères et les tétramères. Dans le cas de la protéine RPA, les informations de coopérativité ont été moins approfondies et restent sujet à caution étant donné la variabilité constatée d'une espèce à l'autre. Par exemple, pour *H. sapiens*, $\omega_{\text{hRPA}} = 10-20$ et pour *S. cerevisiae*, $\omega_{\text{scRPA}} = 10^4-10^5$. En dépit de cette évidente incohérence, RPA aurait au minimum 10 à 20 fois plus de chance d'être observée à côté d'une autre molécule de RPA qu'isolée sur l'ADNsb. La coopérativité apparaît donc une caractéristique fondamentale des protéines fixant l'ADNsb.

2 L'« OB-fold » : un motif commun pour la fixation de l'ADNsb

Nous allons à présent nous intéresser aux propriétés structurales des motifs conférant aux SSB leur capacité à fixer l'ADNsb. Au cours de la dernière décennie, des progrès importants ont été réalisés dans ce domaine, tout d'abord par le biais d'expériences de biochimie et de mutagenèse dirigée (Curth *et al.* 1993), puis avec

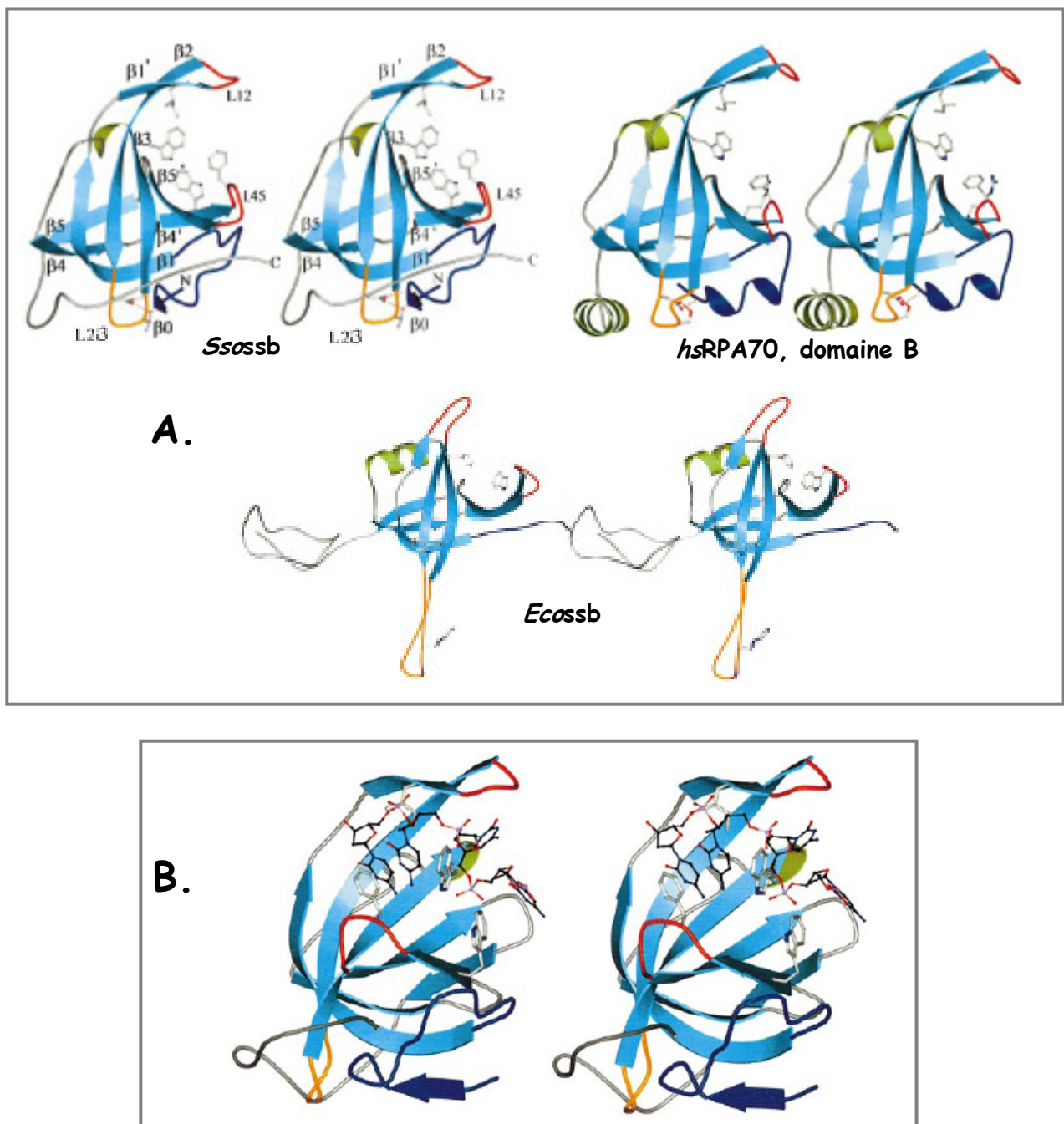


Figure i 2.2: Structure des « OB-folds » présents dans les trois domaines du vivant
Vues stéréographiques des domaines de fixation de *Ssossb*, *Ecosb*, du domaine B de *hsRPA70* (A.)
et d'une modélisation de *Ssossb* fixant une molécule d'ADNs (B.).

Les feuilletts β sont représentés en bleu, les hélices α en vert, les boucles L12 et L45 en rouge, la boucle L23 en orange et la région C-terminale en bleu foncé. Les résidus importants pour la fixation sont représentés selon le modèle « ball & stick ». Les feuilletts et les boucles mentionnés dans le texte sont annotés chez *Ssossb*.

d'après Kerr ID, 2003

l'obtention de structures cristallographiques, parmi lesquelles celles de gp32 (Shamoo *et al.* 1995), *Eco*ssb (Raghunathan *et al.* 2000), *hRPA* (Bochkarev *et al.* 1997) (Bochkareva *et al.* 2002) et *Sso*ssb (Kerr *et al.* 2003) (la SSB de *Sulfolobus sulfataricus*, une espèce appartenant au sous-embranchement des Crearchéobactéries). Leur analyse a conduit à l'observation que les SSB bactériennes et eucaryotes, dont les séquences primaires ne présentent aucune homologie détectable (Murzin 1993), ont par contre une structure tridimensionnelle semblable (Krejci et Sung 2002), appelée « OB-fold » (« Oligo-nucleotide/-saccharide/-peptide Binding fold ») (**Figure i2.2A**). Ce motif, renfermant en moyenne 100 acides aminés (Kerr *et al.* 2003), comporte cinq feuillets β torsadés pour former un tonneau β fermé et recouvert d'une hélice α , localisée entre les troisième et quatrième feuillets. Depuis sa découverte, les caractéristiques définissant l'« OB-fold » ont évolué : i) l'hélice α est en fait une structure variable (Bycroft *et al.* 1997) ; ii) les boucles L12 et L45 peuvent fermer ou ouvrir le tonneau β (Bochkareva *et al.* 2002), en présence ou en absence d'ADNsb, respectivement et enfin, iii) le motif peut servir de squelette à des structures plus complexes, dans lesquelles d'autres motifs (par exemple un « Zn-finger ») peuvent être insérés.

Si la structure globale du motif est commune à tous les règnes du vivant, certaines différences subsistent entre les modes de fixation de l'ADNsb observés chez les eucaryotes et les bactéries (voir **Figure i2.2A**). Chez les eucaryotes, l'ADNsb est enfermé à l'intérieur du tonneau β par l'intermédiaire des boucles L12 et L45, alors qu'il s'enroule autour du motif chez les bactéries, au moyen d'une interaction avec un résidu Trp (Trp40 chez *Eco*ssb) de la boucle L23 (Raghunathan *et al.* 2000), laquelle n'intervient pas dans la fixation chez les eucaryotes. Les boucles L45 et L23 sont également impliquées chez *E.coli* dans la formation du tétramère, ce qui explique la coopérativité limitée du mode de fixation(SSB)₆₅. Malgré ces différences, la façon dont est lié l'ADNsb est identique et implique deux types d'acides aminés (voir **Figure i2.2A et B**) : i) des résidus basiques (Lys 62 et 73 chez *Eco*ssb, ceux de la boucle L45 chez *hRPA*70) forment des liaisons hydrogène avec les groupements phosphates de l'ADNsb et ii) des résidus aromatiques (Ile30, Phe79, Trp75 et Trp56 chez *Sso*ssb) qui s'empilent sur les bases de l'ADNsb. Cependant, tous les « OB-folds » ne fixent pas l'ADNsb, comme l'a montré la structure globale de *hRPA* (Bochkareva *et al.* 2002). Cette protéine en comporte six (quatre pour la sous-unité RPA70 (70N, A, B et C) et un sur chacune des sous-unités RPA32 (appelé D) et RPA14). La structure cristallographique a démontré que seuls les « OB-folds » comportant au moins les acides aminés équivalents à Trp56 et Phe 79 chez *Sso*ssb (qui s'empilent sur les bases de l'ADNsb) étaient capables de se lier au substrat. Les « OB-

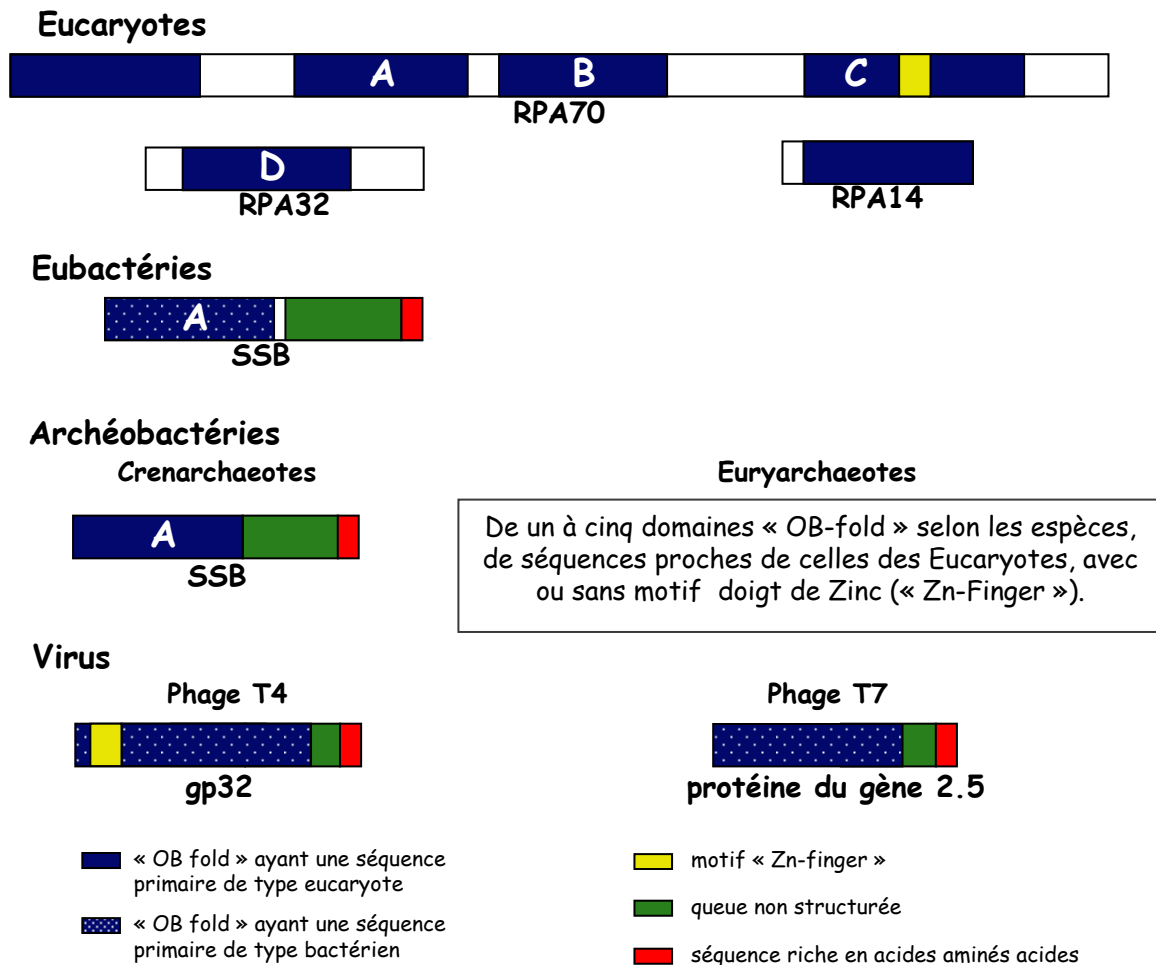


Figure i2.3: Structure des SSB dans les différents domaines du vivant et chez les virus

Les eucaryotes ont une SSB hétérotrimérique appelée RPA, comportant six motifs « OB-folds » (boîtes bleues), dont quatre (A, B, C et D) participent à la fixation de l'ADNsb. Le motif C de RPA70 contient une insertion, portant la signature du motif doigt de Zinc (boîte jaune), $C_{X2}C_{X(8-14)}C_{X2}(C/H)$.

Chez les bactéries, la SSB forme un homotétramère, avec un seul « OB fold » (dont la séquence primaire ne présente aucune homologie avec celle des OB fold eucaryotes) par peptide. Le tiers C-terminal de la SSB de *E.coli* (*Ecosb*) est une queue très flexible (boîte verte), terminée par une région riche en acides aminés (boîte rouge). Cette queue ne participe pas à la fixation de l'ADNsb mais joue un rôle dans les interactions protéines-protéines. Elle est impliquée dans la coopérativité et est essentielle au recrutement des partenaires.

La seule SSB identifiée chez les Crenarchaeotes est celle de *Sulfolobus sulfataricus*. Son motif « OB-fold » a une structure primaire proche de celle des Eucaryotes, mais l'organisation globale de la protéine est la même que chez *E.coli*. Chez les Euryarchaeotes, la situation est complexe. Certaines espèces présentent une organisation très proche de celle de eucaryotes alors d'autres sont des cas uniques comme *Methanococcus jannaschii* (quatre OB folds et un motif Zn-finger sur un seul peptide)

Les protéines virales liant l'ADNsb sont proches des SSB bactériennes, tant du point de vue de l'organisation générale que de leur structure primaire. La protéine gp32 possède en plus un motif « Zn-finger » à son extrémité N-terminale, précédée d'une courte séquence codant un feuillet β de l'« OB-fold ».

« folds » RPA70N et RPA14 n'ont pas de tels acides aminés, et de ce fait seuls les motifs A, B, C et D sont responsables de la fixation de l'ADNsb.

Comme nous l'avons vu précédemment, les « OB-folds » ne présentent pas systématiquement d'homologie au niveau de la séquence primaire. De ce fait, aucun programme n'est capable à l'heure actuelle de prédire la formation d'un « OB-fold » à partir d'une séquence primaire fortement éloignée de celles qui ont été caractérisées. Néanmoins l'ensemble des données décrites ci-dessus ont permis de détecter la présence de nouvelles protéines susceptibles de fixer l'ADNsb chez d'autres espèces, en particulier chez les Archéobactéries. Celles-ci sont composées, entre autres, des Crenarchéobactéries (chez lesquelles seule *SsoSsb* a été identifiée pour l'instant) et des Euryarchéobactéries. Dans ce sous-embranchement, de nombreux candidats ont été découverts, présentant l'originalité d'être composés d'un nombre différent d'« OB-folds » selon les espèces. La **Figure i2.3** montre l'organisation des différents types de SSB ayant été découverts à ce jour. On peut constater que des SSB d'origines différentes présentent des caractéristiques structurales et fonctionnelles communes, mais qui ne sont pas toujours reflétées par une similarité de séquence. Quoiqu'il en soit, ces protéines descendent vraisemblablement d'un ancêtre commun, et la variété observée dans l'arrangement des « OB-folds » est sans doute le reflet de trajectoires évolutives divergentes (Bochkarev et Bochkareva 2004).

3 Conclusion

Nous avons évoqué au cours de ce deuxième chapitre les caractéristiques principales des protéines de la classe des SSB, qui peuvent donc être définies de la manière suivante : elles sont abondantes, fixent l'ADNsb avec une forte affinité ($K_d \approx 10^{-9}$ nM en moyenne), de manière coopérative, sans spécificité de séquence et au moyen d'un « OB-fold ». L'existence de cas particuliers ne remplissant pas ces critères généraux a toutefois été rapportée. Un exemple intéressant est celui de la protéine Pot1 (« Protection of telomeres »), qui fixe l'ADNsb des régions télomériques de manière séquence-spécifique, avec une préférence pour les régions riches en guanines (Baumann et Cech 2001). La structure de cette protéine associée à l'ADNsb a récemment été communiquée (Lei *et al.* 2003). Elle montre que Pot1 utilise également un « OB-fold », indiquant qu'en fonction des acides aminés intervenant dans l'interaction, ce motif est capable d'une spécificité de séquence.

CHAPITRE III: LES GENOMES DES ORGANITES

Ce chapitre sera consacré à un bilan des connaissances concernant les mécanismes de maintenance des génomes mitochondriaux (chondriomes) et chloroplastiques (plastomes). Si certaines des protéines de la famille *At-SBP40* semblent adressées au chloroplaste, le seul phénotype visible auquel nous ayons pour l'instant été confronté est dû à un réarrangement de l'ADN mitochondrial (ADNmt). Etant donné ses propriétés très particulières, par rapport à celles des autres espèces ou même du chloroplaste, nous allons essentiellement axer cette partie sur le génome mitochondrial des plantes supérieures, en mentionnant les autres situations lorsque cela nous apparaîtra nécessaire.

1 Situation générale

Après avoir longtemps été un sujet de controverse, l'origine endosymbiotique des organites est aujourd'hui communément acceptée (Margulis et Bermudes 1985). Selon cette théorie, basée sur la comparaison des séquences de la sous-unité 16S de l'ARNr chez différentes espèces, deux événements indépendants sont à l'origine du phénomène. Il y a plus de 10^9 années, un eucaryote primitif aurait acquis, par endocytose d'une α -Protéobactérie, les précurseurs des mitochondries. Les chloroplastes seraient quant à eux apparus plus tard, via la capture par le même procédé d'une Cyanobactérie (Dyall *et al.* 2004).

La caractéristique principale de l'évolution des génomes des organites est la perte massive de leurs gènes, qui sont transférés vers le noyau. Deux hypothèses sont généralement invoquées pour expliquer cette tendance. Tout d'abord, en se spécialisant dans la fonction respiratoire et la photosynthèse, les organites sont devenus le siège de nombreuses réactions d'oxydo-réduction, produisant des radicaux libres mutagènes. D'autre part, cette situation permet certainement un meilleur contrôle, nécessaire à la coordination de l'activité des deux voire des trois compartiments cellulaires eucaryotes contenant du matériel génétique. Le génome bactérien originel aurait ainsi été réduit d'environ 95 %, passant par exemple de 4 mpb pour 3000 gènes à 150 kpb pour 120 gènes si l'on considère le cas des chloroplastes des plantes supérieures (Timmis *et al.* 2004). Il en va de même pour les mitochondries : le génome de *Rickettsia prowazekii* (Gray *et al.* 1999) (la bactérie actuelle la plus proche de l'endosymbiote de départ) comporte 1,1 mpb codant pour 834 protéines, alors que celui du protiste *Reclinomonas americana* (Lang *et al.* 1997)

Gènes	Nombre de gènes						
	A.th	M.po	P.wi	C.cr	R.am	S.ce	H.sa
Métabolisme mitochondrial							
Complexes respiratoires							
Complexe I	9	8	9	7	12	0	7
Complexe II	0	2	0	3	3	0	0
Complexe III	1	1	1	1	1	1	1
Complexe IV	3	3	3	3	3	3	3
Complexe V	3	4	4	3	5	3	2
Maturation des cytochromes <i>c</i>	5	5	0	0	4	0	0
Machinerie transcriptionnelle							
Protéines ribosomales	7	16	13	4	27	1	0
ARNr	3	3	3	2	3	2	2
ARNt	22	29	26	23	26	24	22

Figure i 3.1: Comparaison du contenu de différents génomes mitochondriaux

A.th *Arabidopsis thaliana*; **M.po** *Marchantia polymorpha*; **P.wi** *Prototheca wickerhamii*;
C.cr *Condrus crispus*; **R.am** *Reclinomonas americana*; **S.ce** *Saccharomyces cerevisiae*;
H.sa *Homo sapiens*

Figure i3.2 : voir
page suivante ...□□

..

(qui possède le génome mitochondrial le plus « ancien ») a une taille de 69 kpb et contient 97 gènes. La **Figure i 3.1** montre les gènes conservés dans différents génomes mitochondriaux (chondriomes) représentatifs de la situation générale. On s'aperçoit que seuls ceux codant pour les sous-unités III, IV et V du complexe respiratoire sont présents chez toutes les espèces. L'hypothèse couramment admise pour justifier cette persistance est que certaines protéines ne pourraient pas être correctement exprimées dans le cytosol ou importées dans les mitochondries. On peut considérer (Palmer 1997) le cas des hydrogénosomes, qui sont totalement dépourvus de génomes, comme la forme ultime du processus d'élimination du matériel génétique des organites engagé au cours de l'évolution.

Les chondriomes ne codent donc, en général, que pour une très faible proportion des enzymes (1000 à 3000) que l'on estime faire partie du protéome mitochondrial (Emanuelsson *et al.* 2000). Les autres peptides ainsi que certains ARNt sont en fait codés par le noyau et adressés à la mitochondrie. Les protéines, traduites dans le cytoplasme, sont agrémentées d'une séquence d'adressage, le plus souvent N-terminale, leur permettant d'entrer dans la mitochondrie via un complexe protéique composé des protéines TOM, TIM (« Translocase of the Outer/Inner Membrane ») et de chaperones (Bauer *et al.* 2000). La situation étant similaire dans les chloroplastes, les plantes ont eu à développer des systèmes complexes. En effet, une proportion significative des protéines à destination des organites ont un double adressage, à la fois chloroplastique et mitochondrial. Ce double adressage peut être effectué de plusieurs manières : i) par le biais d'un peptide d'adressage « ambigu », ii) de sites d'initiation de la traduction alternatifs, iii) de sites multiples de transcription, iv) du « splicing » alternatif ou encore v) de modification post-traductionnelles (Peeters et Small 2001) (Silva-Filho 2003). Ces mécanismes élaborés constituent un exemple de l'interconnexion entre les deux organites, qui malgré une origine endosymbiotique différente, semblent présenter une évolution convergente.

2 Composition en séquences des chondriomes embryophytes

L'ADNmt des végétaux possède des caractéristiques évolutives surprenantes, en regard des situations retrouvées chez la plupart des autres espèces. La plus évidente est sans doute l'augmentation de sa taille, coïncidant avec l'apparition des plantes terrestres (embryophytes) (Turmel *et al.* 2002). La **Figure i3.2** illustre bien ce phénomène : alors que le nombre de protéines codées ne change pas fondamentalement selon les espèces (hormis le cas particulier de *R. americana*, évoqué plus haut), la taille des génomes augmente de manière considérable chez les plantes. Le plus petit chondriome embryophyte, celui de *Marchantia polymorpha*, est

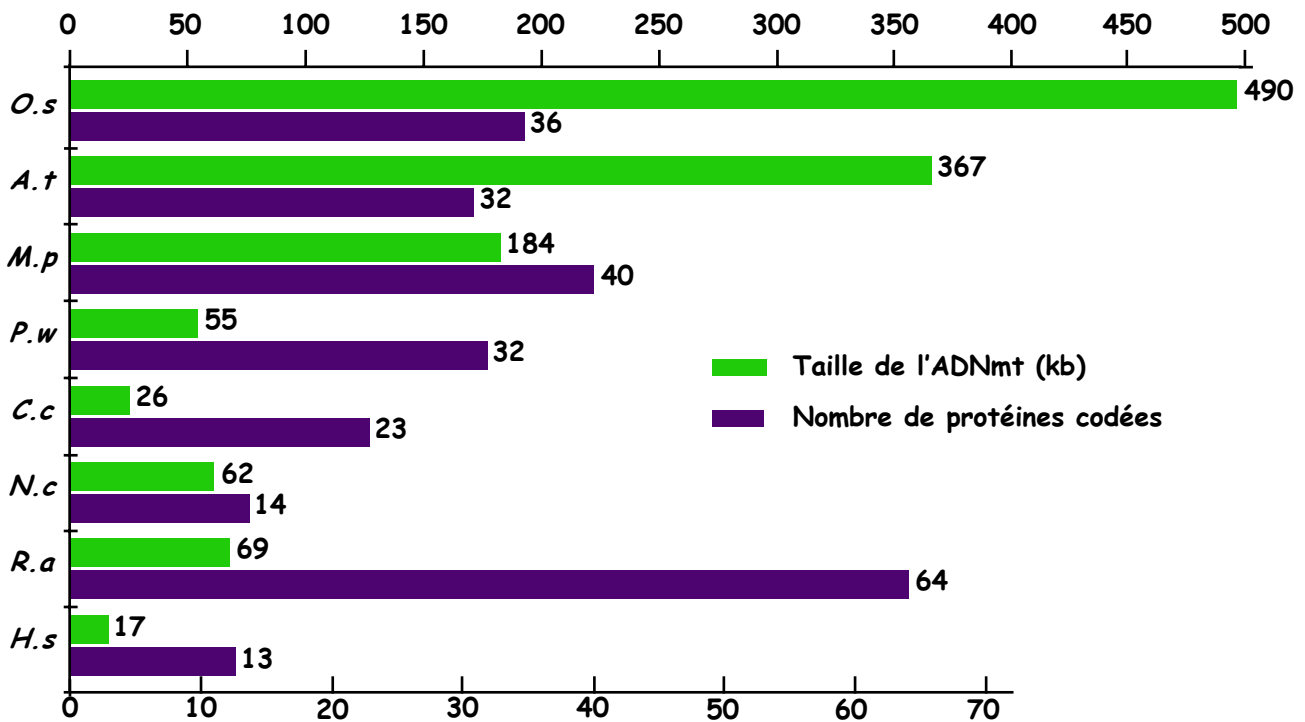


Figure i 3.2: Tailles et capacité codante relatives des génomes mitochondriaux

L'échelle du haut indique la taille des génomes en kb et celle du bas le nombre de protéines codées

O.s *Oryza sativa*; *A.t* *Arabidopsis thaliana*; *M.p* *Marchantia polymorpha*; *P.w* *Prototheca wickerhamii*; *C.c* *Chondrus crispus*; *N.c* *Neurospora crassa*; *R.a* *Reclinomonas americana*; *H.s* *Homo sapiens*

Caractéristiques	<i>Nta</i>	<i>Ath</i>	<i>Bna</i>	<i>Bvu</i>	<i>Osa</i>
CM (pb)	430,597	366,924	221,853	368,799	490,520
A + G (%)	55.0	55.2	54.8	56.1	56.2
Grandes séq. répétées (pb) ^a	34,532	11,372	2,427	32,489	127,600
Séq. uniques ^b					
séq codantes	39,206 (9.9%)	37,549 (10.6%)	38,065 (17.3%)	34,499 (10.3%)	40,065 (11.1%)
introns	25,617 (6.5%)	28,312 (8.0%)	28,332 (12.9%)	18,727 (5.6%)	26,238 (7.2%)
ORFs ^c	46,773 (11.8%)	37,071 (10.4%)	20,085 (9.2%)	54,288 (16.1%)	12,009 (3.3%)
séq. chloro	9,942 (2.5%)	3,958 (1.1%)	7,950 (3.6%)	2.1%	22,593 (6.2%)
autres	274,527 (69.3%)	248,662 (69.9%)	124,994 (57%)	65.9%	262,015 (72.2%)
Gènes	60	55	53	52	56

d'après Sugiyama et al. 2005

Figure i 3.3: Caractéristiques générales de l'ADNmt des Angiospermes

Nta *Nicotiana tabacum*; *Ath* *Arabidopsis thaliana*; *Bna* *Brassica napus*; *Bvu* *Beta vulgaris*; *Osa* *Oryza sativa*

^a Seules les grandes séquences répétées sont prises en compte. La taille indiquée est celle d'une copie de chaque répétition; ^b Une copie de chaque répétition est incluse; ^c A partir de 300 pb

déjà dix fois plus grand que celui des humains. Ce rapport augmente à mesure que l'on suit le cours de l'évolution, jusqu'à atteindre 120 (2000 kpb) pour les plantes à fleurs de la famille des Cucurbitaceae (Ward *et al.* 1981).

Il paraît donc approprié, afin de comprendre la nature et l'origine des séquences ayant engendré de telles différences, de se pencher sur le contenu des génomes mitochondriaux. A l'heure actuelle, sept chondriomes embryophytes ont été séquencés : ceux de *Marchantia polymorpha* (Oda *et al.* 1992), *Arabidopsis thaliana* (Unsold *et al.* 1997), *Beta vulgaris* (Kubo *et al.* 2000), *Oryza sativa* (Notsu *et al.* 2002), *Brassica napus* (Handa 2003), *Nicotiana tabacum* (Sugiyama *et al.* 2005), le dernier en date étant celui du blé, *Triticum aestivum* (453 kpb) (Ogihara *et al.* 2005). Quelques unes des caractéristiques représentatives du contenu génétique des génomes mitochondriaux angiospermes ont été rassemblées dans la **Figure i3.3**. La première observation concerne la quantité de séquences répétées, qui peut atteindre 25 % de l'ADNmt chez le riz. Ces séquences ont des conséquences directes sur la structure du génome et seront détaillées au cours du paragraphe suivant. D'autre part, environ 80% du génome sont occupés par des séquences non codantes. Celles-ci sont pour la plupart de provenance inconnue, le reste étant d'origine chloroplastique et nucléaire. Enfin, en plus des gènes, de nombreuses ORFs (« Open Reading Frame ») ont été repérées, occupant en moyenne 10% du génome. On trouve par exemple chez *Arabidopsis* 156 ORFs de plus de 100 codons (Marienfeld *et al.* 1999). Aucune fonction n'a pour l'instant pu leur être attribuée, même si l'expression de certaines d'entre elles a été démontrée chez le riz (Notsu *et al.* 2002). Elles ne sont pas non plus conservées, même entre des espèces très proches, comme *Arabidopsis* et *Brassica napus* (Handa 2003).

Cette dernière observation est symptomatique de l'organisation générale des chondriomes des plantes supérieures, qui connaissent un taux de réarrangement de leur matériel génétique particulièrement élevé d'une espèce à l'autre (Palmer et Herbon 1988). Cette plasticité contraste avec la situation rencontrée chez les chloroplastes. En effet, les génomes chloroplastiques (plastomes), avec leur forme circulaire composée d'une grande et d'une petite région présentes à un exemplaire et séparées par des répétitions inversées (IR, « Inverted Repeats »), ont une structure remarquablement conservée, et ce jusqu'aux algues primitives (Lemieux *et al.* 2000). La plupart des groupes de gènes chloroplastiques et l'arrangement de ces derniers apparaissent ainsi fixés depuis des temps reculés. L'aspect hétéroclite, même entre des espèces proches, des chondriomes végétaux est donc d'autant plus surprenant et pose la question du moment auquel cette situation est apparue au cours de l'évolution. Le groupe phylogénétique le plus proche de l'embryophyte primitif *Marchantia* est celui des Charales, dont fait partie *Chara vulgaris* (Turmel *et al.* 2003). Le séquençage du génome de cette algue, d'une taille de 67 kpb (par rapport aux 184 kpb que compte

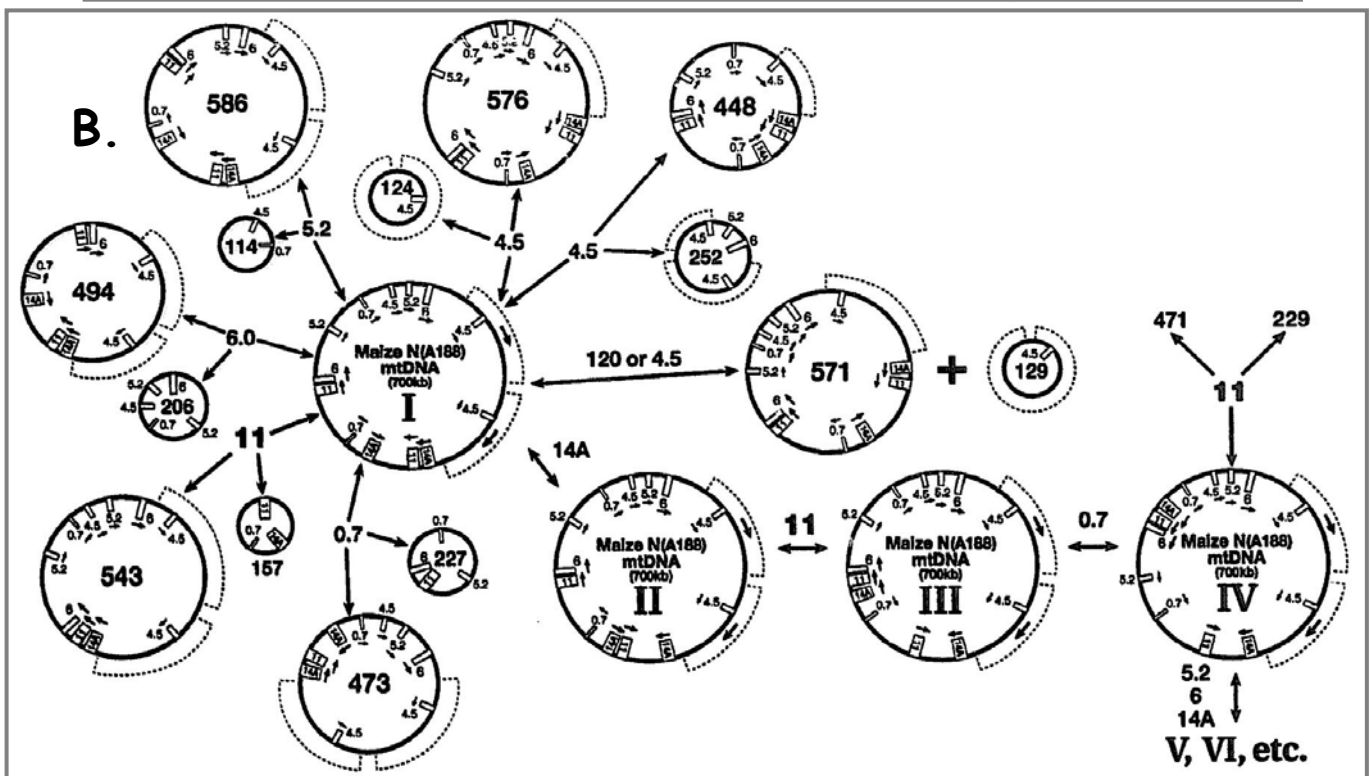
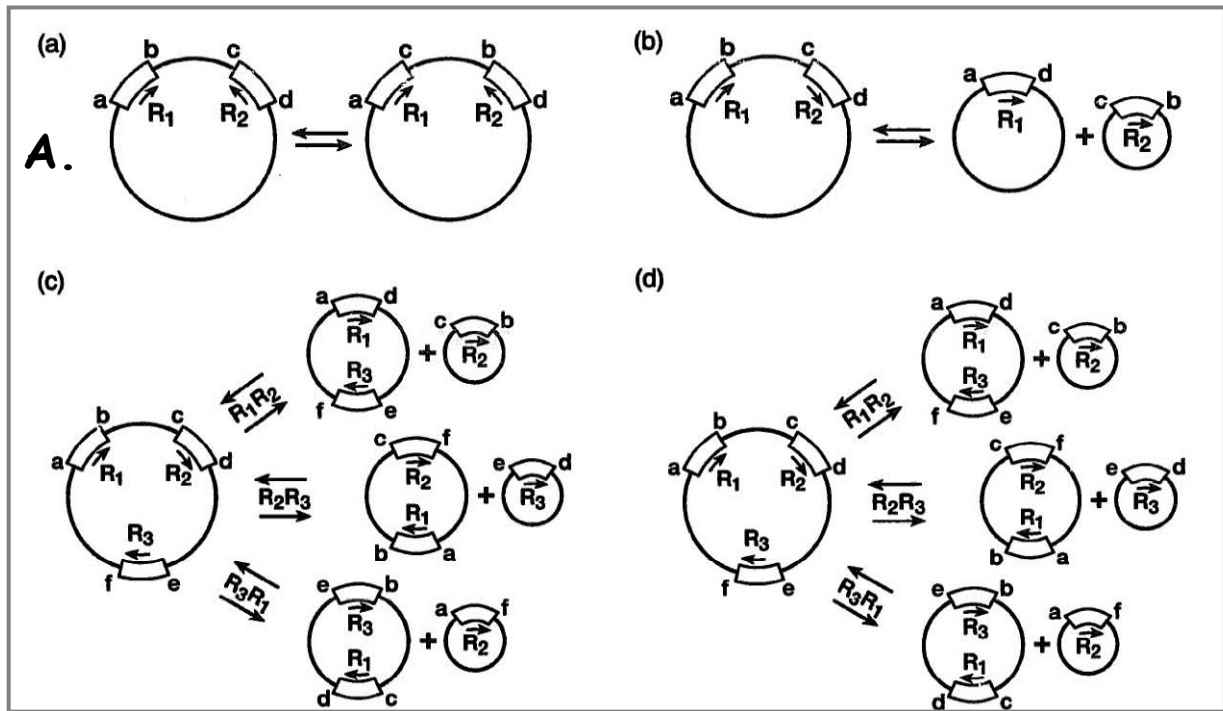


Figure i 3.4: Conséquences des répétitions sur les chondriomes

A. Exemples de l'effet de différentes combinaisons d'IRs et de DRs sur une molécule circulaire. (a.) Une IR (b.) Une DR (c.) Trois DRs (d.) Deux IRs et deux DRs.

B. Structure multipartite potentielle du chondriome du maïs NA. La forme isomérique I correspond à la molécule ayant été identifiée expérimentalement. Les chiffres situés à côté des flèches indiquent la taille de la répétition impliquée.

Voir le texte pour les explications.

d'après Fauron et al. 1995

celui de *Marchantia*), a permis de lier l'origine des caractéristiques des chondriomes des plantes supérieures à celle du groupe des embryophytes (Turmel *et al.* 2002). Le nœud phylogénétique suivant, liant *Marchantia* aux mousses et aux trachéophytes, se caractérise en effet déjà par un gain différentiel de nombreux introns selon les espèces (Knoop 2004). Cette première mise en évidence de particularités chez les chondriomes végétaux est suivie de la variation de nombreux autres paramètres: augmentation de la taille (Turmel *et al.* 2002), haute fréquence de recombinaison (Knoop 2004), apparition du *trans*-splicing (Malek et Knoop 1998), de l'édition (Hiesel *et al.* 1994), entrée d'ADN d'origine étrangère ou encore fréquences auxquelles les gènes sont transférés vers le noyau (Adams *et al.* 2000)), tout au long de l'évolution des embryophytes. Les mécanismes particuliers de la maintenance et de l'expression des génomes mitochondriaux des plantes pourraient donc devoir leur origine aux nécessités de l'adaptation à la vie terrestre.

3 Conséquences des séquences répétées sur les chondriomes végétaux

(Fauron *et al.* 1995)

Les données provenant des cartes de restriction des chondriomes végétaux conduisent à leur représentation sous forme d'une molécule circulaire, appelée cercle maître et renfermant l'intégralité du matériel génétique. Si ces cartes physiques ne reflètent pas la structure *in vivo* de l'ADNmt, elles permettent néanmoins de comprendre l'effet dynamique de la recombinaison sur ces génomes. Nous allons au cours de ce paragraphe, présenter des modèles basés sur ces cartes circulaires. Comme nous l'avons vu précédemment, l'ADNmt des plantes, en plus de sa grande taille, a pour particularité de contenir de nombreuses répétitions. Certaines d'entre elles participent à la dynamique du génome, alors que d'autres sont inactives, sans qu'il y ait à cela des raisons évidentes. Les répétitions peuvent être en orientation directe (DR, « Direct repeat ») ou indirecte (IR), ces deux situations ayant des conséquences radicalement différentes sur l'organisation du génome. La **Figure i3.4A** présente les conséquences sur le génome de recombinaisons entre différents types de répétitions. Une paire d'IRs génère par réarrangement une forme isomérique de la molécule qui les porte, caractérisée par un retournement (« flip-flop ») de la séquence située entre les répétitions (a). Une recombinaison entre une paire de DRs engendre deux molécules subgénomiques, contenant chacune une copie de la répétition (b). La présence de trois IRs peut conduire à l'apparition de six subgénomiques (c). Enfin, la présence de trois répétitions, dont l'une est inversée peut engendrer quatre isomères de la molécule de départ et deux jeux de molécules subgénomiques (d).

La **Figure 3.4B** montre la carte physique du chondriome de 700 kpb du maïs N(A188) (Fauron et Casper 1994), contenant sept jeux de répétitions de tailles différentes (0.7, 4.5, 5.2, 6, 11, 14 et 120 kb). Elle représente la forme isomérique I du cercle maître, qui contient six paires de DRs et une paire d'IRs (répétition 14A). Une recombinaison faisant intervenir l'IR 14A a pour effet de retourner la répétition 11, engendrant ainsi l'isomère II (chez lequel la répétition 11 est une IR). Cette IR 11 peut à son tour engendrer un isomère, la forme III, etc... Les molécules subgénomiques représentées sont dérivées des seules DRs de l'isomère I. Les répétitions 4.5, prises deux par deux, donnent lieu, en cas de recombinaison, aux molécules subgénomiques de tailles 124 et 576 kb, 252 et 448 kb ou encore 571 et 129 kb. De la même manière, des recombinaisons peuvent faire intervenir les DRs 5.2, 6, 11, 0.7 et 120 kb, engendrant les molécules subgénomiques indiquées. Certaines de ces molécules subgénomiques contiennent également des DRs, qui peuvent elles aussi recombiner pour donner lieu à des molécules plus petites. En fait, le nombre de molécules subgénomiques double à chaque fois qu'une nouvelle répétition recombinogène est découverte. De la même manière, d'autres molécules subgénomiques non représentées peuvent être créées à partir des autres formes isomériques du cercle maître, augmentant ainsi la diversité.

Ce type d'analyse illustre la façon dont le nombre, la position et l'orientation des répétitions recombinogènes déterminent la complexité de l'organisation des chondriomes végétaux, dictent leur fluidité et possèdent le potentiel d'influer sur leur fonction. Il y a deux conséquences principales à l'organisation multipartite de l'ADNmt. Premièrement, un nombre important de formes isomériques doit être présent pour chaque génotype mitochondrial. Chaque isomère contient l'intégralité de l'information génétique, mais selon une orientation et un ordre différents. De ce fait, les désignations IR et DR sont interchangeable et propres aux molécules auxquelles les répétitions appartiennent (voir l'exemple de la répétition 11 décrit plus haut). Deuxièmement, la taille des molécules engendrées par une répétition va varier en fonction de l'isomère du cercle maître auquel la répétition appartient. Par exemple, la répétition 11 génère des molécules subgénomiques de 157 et 543 kb dans l'isomère I et des molécules de 471 et 229 kb dans l'isomère IV. Les différences dans l'orientation des séquences peuvent avoir des conséquences importantes, par exemple en connectant ou en déconnectant un gène de son promoteur. Les variations de la structure multipartites sont susceptibles d'être employées comme mécanisme d'amplification pour produire la quantité nécessaire d'un gène requis à certains stades du développement.

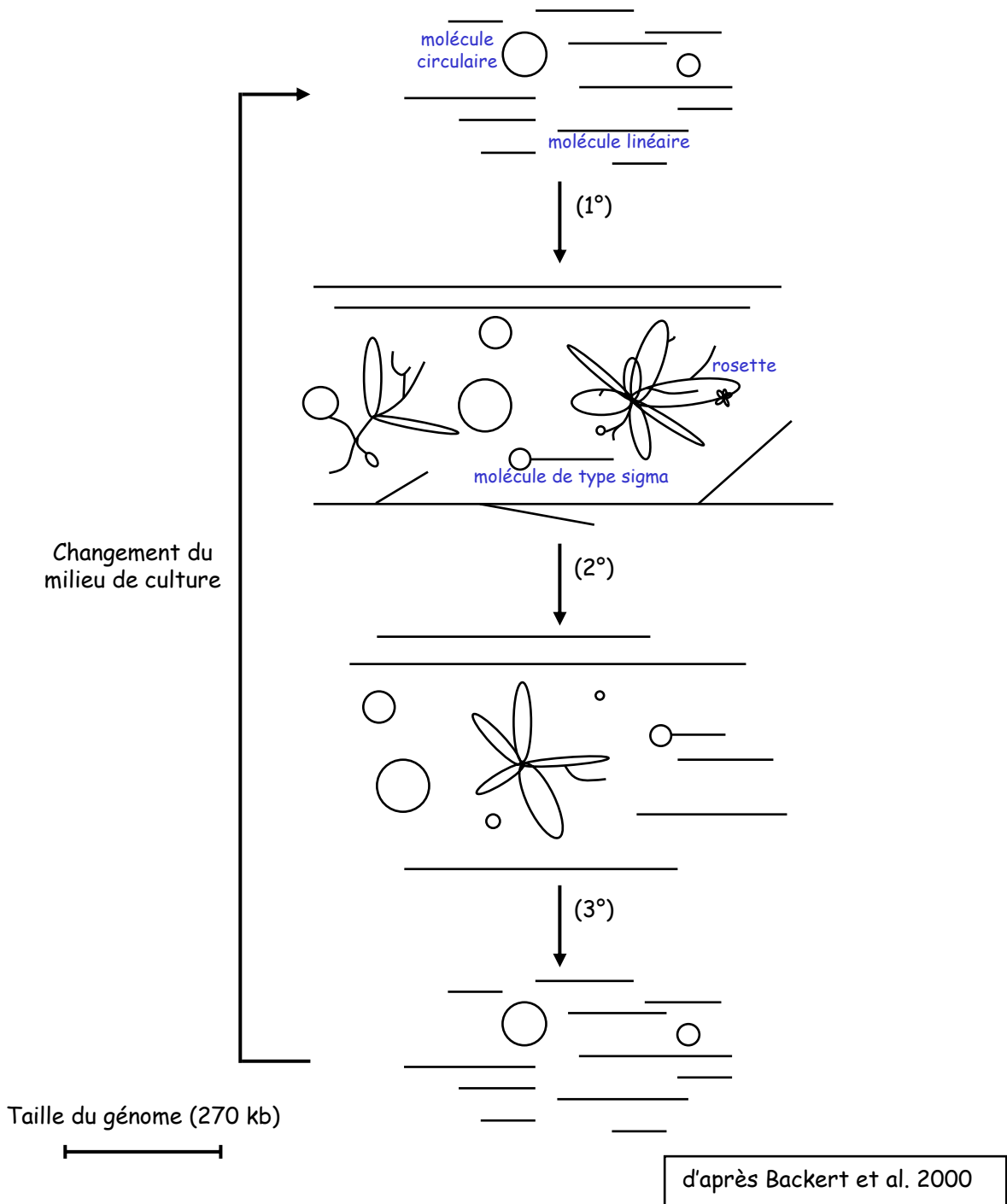


Figure i 3.5: Changements structuraux du génome mitochondrial de cultures cellulaires de *Chenopodium album* durant un cycle de croissance

1° Durant les premiers stades de la croissance, l'ADNmt semble être répliqué par des mécanismes faisant intervenir la recombinaison, conduisant à l'apparition de concatémères complexes, appelés rosettes, de taille plusieurs fois supérieure à celle du génome. 2° La réplication par rolling circle des molécules subgénomiques (faisant intervenir des structures de type sigma) et le début de la résolution des concatémères a lieu pendant la phase logarithmique de croissance. 3° La résolution finale des cocatémères en molécules de 40 à 200 kb, principalement linéaires a lieu durant le phase stationnaire.

4 L'organisation in vivo et les modes de réplication des organites

Jusqu'au milieu des années 1990, la question de l'existence *in vivo* des molécules circulaires, et en particulier celle du cercle maître, a été source de controverse (Bendich 1993). Les premières expériences de microscopie électronique (EM) qui avaient pour but de visualiser la présence de telles structures ont étudié les chondriomes des animaux. La confirmation de leur présence, ajoutée au fait que les génomes des ancêtres bactériens possédaient eux aussi un génome circulaire ont suffi à persuader une majorité qu'il s'agissait là d'une caractéristique générale. Ainsi, durant de nombreuses années, toutes les observations rapportant l'existence de molécules linéaires ont ainsi été déconsidérées. Ces résultats étaient imputés à la fragmentation mécanique ou enzymatique de supposées molécules circulaires pré-existantes au cours de l'expérience, en vertu de la théorie dite des « broken circles » (Bendich 2004). C'est avec l'arrivée de nouvelles techniques comme la PFGE (« Pulsed Field Gel Electrophoresis ») (Skelly et Maleszka 1989) et l'imagerie mobile (observation en microscopie de l'ADN durant l'électrophorèse) (Jacobs *et al.* 1996), permettant un traitement plus doux des échantillons, que l'idée de molécules linéaires fut enfin acceptée. Ces techniques ont révélé la nature essentiellement linéaire de plusieurs chondriomes de levures (Maleszka *et al.* 1991) de plantes (Backert *et al.* 1995) et des plastomes (Shaver *et al.* 2005).

Nous allons à présent évoquer les connaissances actuelles concernant les mécanismes impliqués dans la réplication. Celles-ci sont plus avancées chez les mammifères et les levures (pour une revue, voir (Lecrenier et Foury 2000)) que chez les plantes. Une synthèse des informations recueillies pour le génome mitochondrial de *Chenopodium album* (Backert et Borner 2000), qui a été l'objet du plus grand nombre d'études sur ce point, celui est présenté **Figure i3.5**. Cette expérience, ayant pour but de mieux comprendre les mécanismes impliqués dans la réplication du génome mitochondrial, est basée sur l'analyse de l'occurrence de différents types de molécules en fonction de l'état d'avancement du cycle cellulaire. Nous allons commencer par détailler les caractéristiques des molécules ayant été observées. Les molécules circulaires ne représentent qu'une assez faible proportion de l'ADNmt et ont une taille toujours inférieure à celle du génome. Les molécules linéaires sont quant à elles largement représentées, et peuvent atteindre 780 kb, c'est à dire trois fois la taille du génome. Enfin, les molécules impliquées dans la réplication semblent être de deux types : i) des molécules appelées rosettes, pouvant atteindre une taille six fois supérieure à celle du génome et constituées de nombreux points de branchements évoquant des fourches de réplication et des fonctions de Holliday (voir Chapitre I) ii) des molécules de type sigma, caractéristiques d'une réplication de type « rolling circle », ayant des queues formées d'ADNdb (retrouvées lors de la réplication du

phage λ (Bastia et Sueoka 1975)) ou d'ADNsb (réplication du phage ϕ X-174 (Koths et Dressler 1978)). La présence de ces différentes molécules, ainsi que la découverte antérieure de la possibilité de changement de brin matrice (« strand switching ») lors du rolling circle (Backert 2000), suggère la coexistence de nombreux mécanismes différents pour la réplication des chondriomes végétaux. Il semble d'autre part que ceux-ci ne soient pas utilisés de la même façon. En effet, bien que tous les types de molécules puissent être observés à n'importe quel moment, certains sont plus représentés en fonction de l'avancement du cycle cellulaire. Les rosettes apparaissent peu après le transfert des cellules dans du milieu frais, alors que le taux de molécules sigma est plus important durant la phase exponentielle. Cela semble indiquer que les tous modes de réplifications existants ne soient pas utilisés pour le même usage, et est peut être représentatif de leur emploi différentiel *in vivo*, selon les conditions ou les tissus. D'autre part, une diminution du nombre de molécules linéaires coïncide avec l'augmentation du nombre et de la taille moyenne des rosettes. Cela suggère que les rosettes sont des structures dans lesquelles la réplication aurait été amorcée par le biais d'évènements de recombinaison entre les molécules circulaires et linéaires, et qui seraient par la suite traitées pour redonner lieu aux molécules initiales. Enfin, une étude préalable avait démontré la présence dans le chondriome de *C. album*, d'une proportion importante d'ADNsb (Backert *et al.* 1997). Celui-ci était entre autres présent à la plupart des extrémités des molécules linéaires, ce qui constitue un pré-requis fondamental pour l'amorçage du processus d'invasion de brins catalysé par les protéines de la famille RecA/RAD51. L'ensemble de ces données va donc dans le sens d'une coexistence de phénomènes de réplication recombinaison-dépendants (comme celui utilisé par le phage T4) et de réplication par rolling circle de types λ et ϕ X-174 dans les mitochondries de *C. album*. Un autre groupe ayant conduit le même type d'expériences sur le chondriome de *M. polymorpha* (Oldenburg et Bendich 2001) et le plastome de *Zea mays* (Oldenburg et Bendich 2004) sont arrivés aux mêmes conclusions, suggérant que la présence de tels mécanismes est généralisable à l'ensemble des organites des plantes supérieures.

5 Mutations et « substoichiometric shifting » chez les mitochondries des plantes supérieures

5.1 Les recombinaisons ectopiques

Nous avons évoqué au cours de deux derniers paragraphes, diverses propriétés des chondriomes des plantes supérieures. Ceux-ci semblent utiliser plusieurs mécanismes différents pour leur réplication et sont composés non pas d'un type de

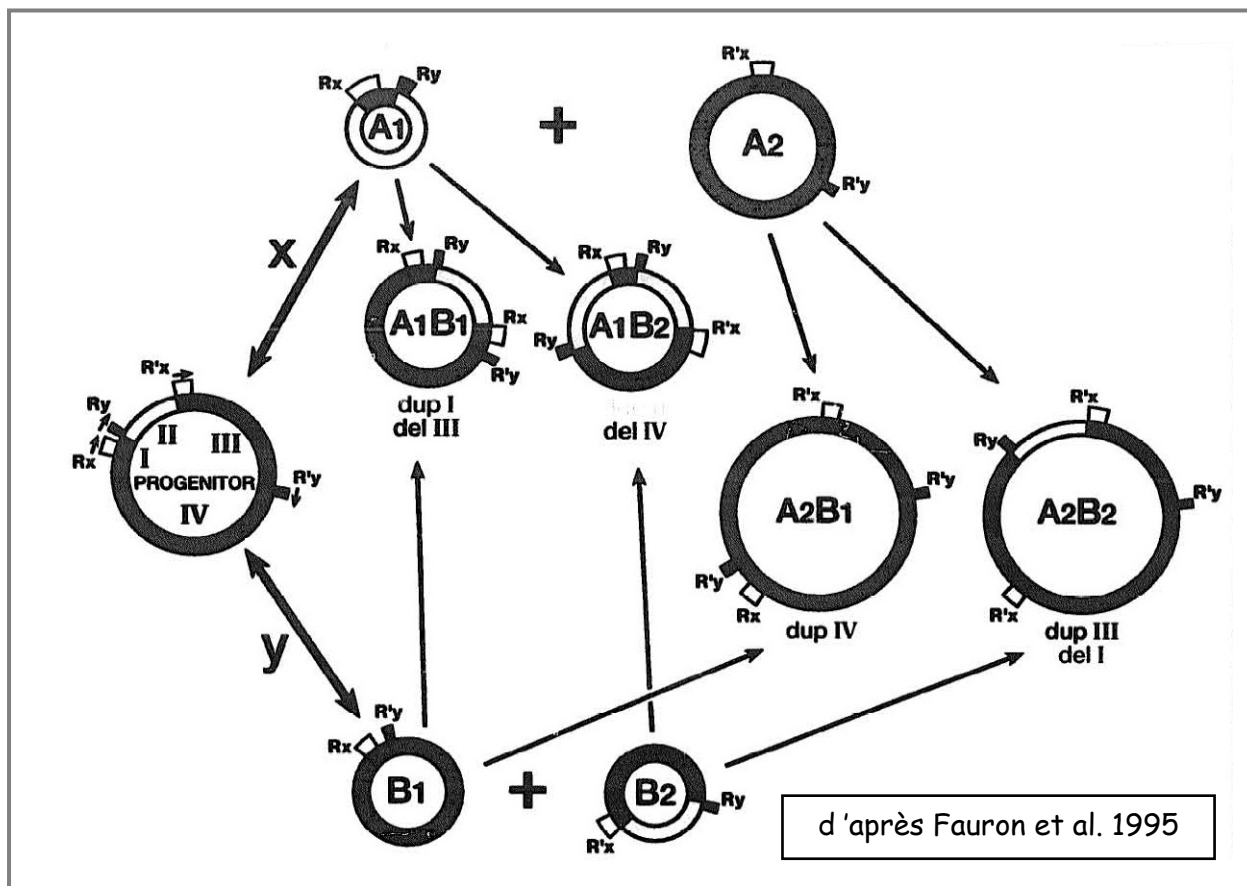


Figure i 3.6: Modèle expliquant la formation de nouveaux génomes stables à partir de recombinaisons entre deux paires de DR.

Un cercle maître progéniteur contient deux jeux de DR. La répétition X est impliquée dans la formation de la structure multipartite, tandis que la répétition Y n'est que très rarement active. Les événements de recombinaison entre deux copies de X (Rx et Rx') génèrent les subgénomés A1 et A2, et ceux entre deux copies de Y (Ry et Ry') les subgénomés B1 et B2. Une recombinaison intermoléculaire entre n'importe lesquels des subgénomés A1, A2, B1 ou B2 va créer de nouvelles molécules circulaires appelées A1B1, A1B2, A2B1 et A2B2. Chacun des nouveaux chromosomes contient une délétion et une duplication spécifique par rapport au progéniteur. Par exemple, le génome A1B1 a la région I dupliquée et la région III délétée. Ainsi, d'après ce modèle, n'importe laquelle des quatre séquences situées entre les répétitions peuvent être dupliquées ou supprimées. Si la région I contient un gène dont le produit est nécessaire en grande quantité au cours du développement, il serait avantageux de sélectionner le chromosome A1B1, dans lequel cette région est présente en deux exemplaires. Si au contraire, cette région devait être délétère pour l'organisme, on pourrait s'attendre que se soit le chromosome A2B2, dans lequel cette région est dupliquée, qui soit sélectionné. Ce dernier exemple illustre le cas du révertant pour le *cmsT*, ayant perdu le gène *urf 13*.

molécule bien défini, mais d'une multitude de molécules subgénomiques issues de phénomènes de recombinaison inter- et intramoléculaires entre les nombreuses répétitions directes indiquées précédemment. Ces répétitions peuvent être subdivisées en deux catégories : i) celles qui, actives en permanence, sont responsables de la structure multipartite du génome et ii) celles qui conduisent à des événements de recombinaison dits ectopiques, pouvant provoquer des réarrangements définitifs de l'ADNmt. Ces recombinaisons ectopiques peuvent avoir lieu entre des répétitions bien caractérisées, de grandes tailles (mais qui ne sont, pour une raison jusque là inconnue, pas fréquemment impliquées) mais aussi entre des séquences de tailles très inférieures (la plus petite séquence découverte ayant pu donner lieu à ce type recombinaison comportait seulement six nucléotides (Newton *et al.* 1990)).

Nous allons à présent illustrer l'intérêt des modèles basés sur les cartes physiques pour la compréhension de l'impact de ces recombinaisons sur les chondriomes, en prenant pour exemple un cas de stérilité mâle cytoplasmique, ou CMS (pour plus de détails, voir paragraphe 5.3.1). L'ADNmt *cmsT* du maïs porte une copie unique d'un gène chimérique, appelé *urf13* (Dewey *et al.* 1986). Ce gène code pour un polypeptide de 13 kDa, qui semble responsable du phénotype de stérilité mâle. L'analyse de plantes révertantes pour le phénotype après une étape de dédifférenciation a permis de démontrer que le retour à la fertilité provenait d'une double recombinaison mettant en jeu deux répétitions (Fauron *et al.* 1990). L'une de ces répétitions, longue de 4,6 kb, est impliquée dans la formation de la structure multipartite du génome. Une deuxième, d'une longueur de 58 pb, n'est pas impliquée dans la dynamique du génome et doit son activité aberrante au passage par l'étape de culture cellulaire. La **Figure i3.6** permet de comprendre comment ces deux répétitions engendrent l'abolition du phénotype en éliminant la région codant pour le peptide de 13 kDa. Ce modèle est général et applicable à d'autres exemples, comme celui de la CMS du haricot, (*Phaseolus vulgaris*) dans lequel le génotype de la lignée stérile comporte trois copies d'une séquence du génome progéniteur. Dans ce cas, un troisième événement recombinaison intervient, entre un des chromosomes A1B1, A1B2, A2B1 ou A2B2 avec n'importe lequel des subgénomés A1, A2, B1 ou B2.

5.2 « Substoichiometric shifting »

Le « substoichiometric shifting » est un autre phénomène propre aux chondriomes des plantes supérieures, permettant d'augmenter encore la diversité de leur contenu. Il se définit comme la capacité de ces génomes à moduler le taux de représentation des différentes molécules subgénomiques qui les composent. Ce mécanisme a été découvert en observant la présence d'un total de quatre environnements génétiques différents dans la région 5' du gène *atp9* chez de

nombreuses lignées de maïs (Small *et al.* 1987). Un voire deux environnements y sont majoritaires et définissent le type de chondriome (normal, N ou stérile, *cms-T*, *cms-C* et *cms-S*). Toutefois, des temps d'exposition plus longs permettent d'observer également la présence des autres types d'arrangements, alors que ceux-ci n'apparaissent pas sur les cartes physiques respectives. Le processus de « substoichiometric shifting » fut découvert en réalisant que ces molécules subgénomiques, présentes à un faible nombre de copies (sublimons) ne peuvent avoir été générées par des événements simples de recombinaison. Afin de comprendre l'origine du phénomène, les auteurs de cette découverte se sont ensuite intéressés à un chondriome de maïs « ancestral » (RU). Ils se sont aperçus de l'absence chez celui-ci d'une répétition de 12 kb que l'on trouve chez le maïs moderne. La cartographie du chondriome RU a permis de comprendre le mode de formation probable de cette répétition, qui nécessite au moins trois événements de recombinaison, faisant intervenir des sublimons (Small *et al.* 1989). Ces expériences ayant permis d'expliquer la raison d'être des sublimons, à savoir qu'ils permettent le maintien d'une variation génétique sous une forme silencieuse mais néanmoins utilisable, ont grandement contribué à la compréhension de l'évolution des délétions et des duplications chez les chondriomes végétaux. Il paraît important à ce stade de distinguer l'apparition d'une recombinaison du « substoichiometric shifting ». Bien entendu, les deux sont liés dans la mesure où les sublimons sont issus d'événements rares de recombinaison et qu'une recombinaison provoque une modification de la stoechiométrie. Néanmoins, le « substoichiometric shifting » ne se définit pas comme un événement de recombinaison, mais comme l'amplification, faisant probablement intervenir la réplication, de molécules pré-existantes.

Ce phénomène fut par la suite caractérisé chez plusieurs espèces : *Nicotiana tabacum* (Kanazawa *et al.* 1994), *Phaseolus vulgaris* (Mackenzie et Chase 1990) ou encore *Nicotiana glauca* (Vitart *et al.* 1992). Ces études ont permis d'affiner la connaissance du mécanisme. Elles ont conduit à l'observation que les changements engendrés peuvent bouleverser la stoechiométrie de façon très importante en une génération (Vitart *et al.* 1992) et que les sublimons pouvaient être présents à des taux aussi bas qu'une copie pour 100 ou 200 cellules (Vitart *et al.* 1992). Mais les progrès les plus intéressants en vue de la compréhension du mécanisme sont venus de la découverte de son contrôle par le génome nucléaire. Là encore, l'influence du noyau a pu être démontrée pour différents systèmes comme la lignée P2 du maïs (Kuzmin *et al.* 2005), le gène *MS4* chez *N. glauca* (Albert *et al.* 2003), le gène *Fr* chez *P. vulgaris* (Mackenzie et Chase 1990) et le gène *MSH1* chez *Arabidopsis* (Abdelnoor *et al.* 2003). Les modes d'action chez les deux systèmes les plus connus, la CMS de *P. vulgaris* et le phénotype MDL chez *Arabidopsis*, seront détaillés dans le paragraphe

suisant. Pour finir, on peut encore ajouter que si ces découvertes, dans leur ensemble, ont démontré que le « substoichiometric shifting » est utilisé par de nombreuses espèces et sans doute généralisable au moins aux angiospermes, plusieurs points restent à éclaircir, concernant notamment le mécanisme moléculaire mis en jeu et la façon dont il est contrôlé.

5.3 Exemples de phénotypes liés aux mutations sur l'ADNmt

À deux notables exceptions près, (chez *Pelargonium hortorum* (Cho *et al.* 2004) et les espèces du genre *Plantago* (Parkinson *et al.* 2005)), l'ensemble des séquences contenues dans les chondriomes des plantes supérieures présentent un taux de substitution de nucléotides particulièrement faible par rapport aux génomes nucléaires et chloroplastiques (Palmer et Herbon 1988). Ainsi, la totalité des phénotypes observés sont consécutifs aux réarrangements et aux délétions créés par des événements de recombinaison. Ils consistent pour la plupart, soit en des désordres morphologiques, soit en l'incapacité à produire du pollen fertile. Nous allons maintenant présenter deux gènes nucléaires impliqués dans le contrôle de la stoechiométrie des molécules subgénomiques et dont le dysfonctionnement conduit pour l'un à la CMS (le gène *Fr* du haricot) et pour l'autre à un phénotype morphologique (le gène *CHM* d'*Arabidopsis*).

5.3.1 La stérilité mâle cytoplasmique (CMS)

Nous n'allons faire ici qu'une brève incursion dans les mécanismes de la CMS, qui est pourtant une thématique majeure de la biologie végétale (Budar et Pelletier 2001) (Linke et Borner 2005). La stérilité mâle cytoplasmique, a en effet été décrite chez plus de 150 espèces. Elle se définit comme l'incapacité, héritée maternellement, à produire du pollen chez des individus normalement hermaphrodites. Cette incapacité engendre la coexistence, au sein d'une même espèce, d'individus hermaphrodites et femelles (atteints de stérilité mâle). Les plantes atteintes de CMS ont des organes mâles non fonctionnels, mais présentent des organes femelles et des parties végétatives normales. Selon les cas, différents stades du développement peuvent être touchés, du début de la formation de l'anthere à la deuxième mitose du pollen. Généralement, plus l'étape affectée est située tôt dans le développement, plus l'architecture de la fleur va être affectée (Linke et Borner 2005). Dans presque tous les cas de CMS, le phénotype coségrège avec une mutation de l'ADNmt qui se caractérise par l'apparition d'une nouvelle ORF, composée de parties d'autres gènes et/ou de séquences non codantes (Hanson et Bentolila 2004). Malgré l'identification de nombreuses séquences associées à la CMS, la manière dont elles conduisent aux dysfonctionnements

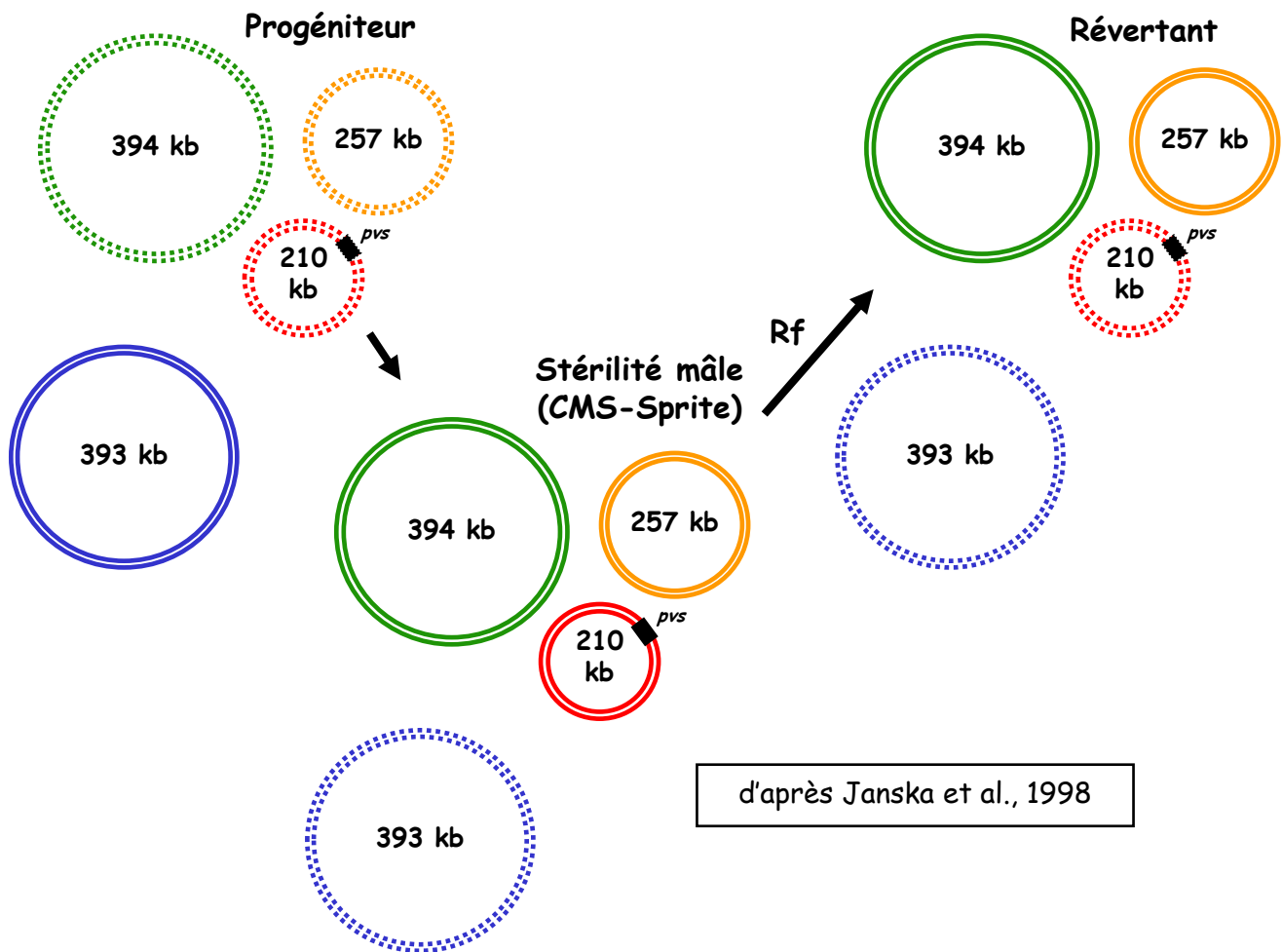


Figure i 3.7 : Modèle schématisant les variations stœchiométriques des chondriomes du haricot lors de l'apparition de la « CMS » et de la restauration de la fertilité

Les molécules présentes à un nombre de copies « normal » sont schématisées par les lignes continues, les sublimons par des pointillés. Voir le texte pour plus de détails.

observés reste inconnue. Une hypothèse est que ces réarrangements impliquent une baisse de l'efficacité des fonctions mitochondriales, laquelle serait signalée au noyau qui à son tour réprimerait les gènes nécessaires au développement des organes mâles. Un aspect fondamental de la CMS est le fait que le génotype mitochondrial responsable du phénotype a dans de nombreux cas pu être associé à un locus nucléaire modulant son effet. Certains des allèles ont la capacité de restaurer la fertilité pollinique, et peuvent pour cela agir à différents niveaux : sur les ARNm (ce qui semble constituer la majorité des cas (Touzet et Budar 2004)), sur le métabolisme (Liu et Schnable 2002) ou encore sur la stoechiométrie des molécules subgénomiques.

Le seul exemple de restauration de la CMS ayant lieu au niveau de l'ADN est celui du haricot. La stérilité y a été associée à l'expression d'une séquence mitochondriale codée par une région de 2,4 kb appelée *pvs* (« *Phaseolus vulgaris sterility sequence* »). Cette région contient au moins deux ORFs uniques, nommées *pvs-orf98* et *pvs-orf-239* (pouvant respectivement donner lieu à des peptides de 98 et 239 acides aminés). Seule l'expression de l'ORF *pvs-orf239* est détectée, produisant un peptide de 27 kDa, responsable du phénomène de stérilité. Celui-ci s'accumule dans les tissus transmetteurs (Abad *et al.* 1995) mais est rapidement dégradé dans les tissus végétatifs (Sarria *et al.* 1998). La cartographie du chondriome stérile du haricot montre une organisation sous forme de trois molécules recombinées et hautement redondantes de 394, 257 et 210 kpb, cette dernière portant la séquence *pvs* (voir **Figure i3.7**) (Janska et Mackenzie 1993). Ces trois molécules sont présentes sous forme substoechiométrique dans le chondriome du progéniteur, mais prédominant dans le chondriome conférant la stérilité. L'introduction d'un contexte nucléaire comportant le gène *Fr* conduit à un nouvel événement de « substoechiometric shifting », qui n'agit que sur la molécule de 210 kpb (Janska *et al.* 1998). Le faible nombre de copies de la molécule de 210 kpb chez le révertant (une pour 100-200 cellules) semble toléré en raison de la persistance du taux des deux autres molécules subgénomiques, qui à elles deux comportent tous les gènes essentiels. Pour finir, des expériences de PCR quantitative en temps réel sur des tissus provenant du révertant ont permis d'observer que la molécule de 210 kpb était présente à un nombre de copies plus élevé dans les tissus méristématiques des racines que dans un tissu végétatif utilisé comme témoin (Arrieta-Montiel *et al.* 2001). Cela suggère donc que les tissus transmetteurs sont le lieu privilégié de la rétention des sublimes, permettant de conserver toutes les séquences à transmettre à la descendance.

5.3.2 Le phénotype des feuilles panachées (« variegated »)

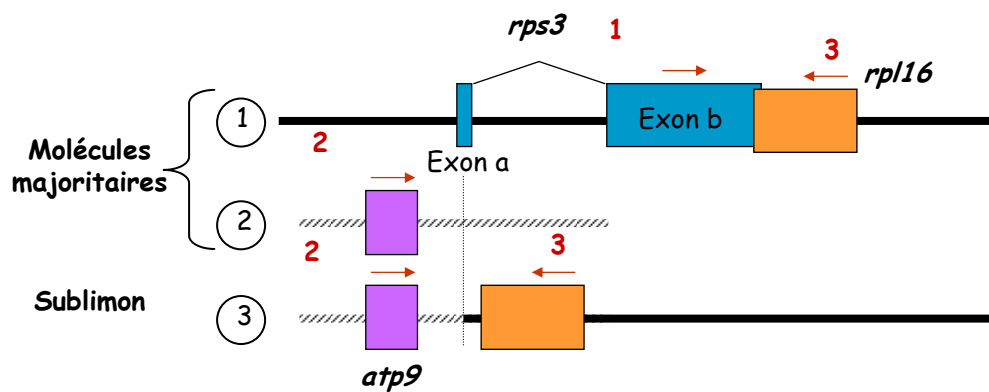


Figure i 3.8 : Schéma de la recombinaison conduisant à la molécule chimérique identifiée à un nombre anormalement élevé de copies chez les mutants *chm*.

La situation représentée est celle du chondriome de plantes sauvages. Chez les mutants, la situation est inversée, la molécule 3 est majoritaire.

L'appellation, « chloroplast mutator » est due à l'hypothèse initiale que la mutation affectait le plastome, en raison de tâches blanchâtres apparaissant sur les feuilles. Il s'agit toutefois bien d'une mutation mitochondriale, le phénotype étant probablement dû à des interactions défectueuses entre les organites, comme cela a été documenté pour un des mutants NCS du maïs (Newton et Coe 1986). Le phénotype « variegated » affecte *Arabidopsis*, chez des plantes ayant subi une perte de fonction du gène *CHM*. Il a été découvert au moyen de trois allèles récessifs : deux d'entre eux sont issus de mutations induites au méthanesulfonate (*chm-1* et *chm-2*) et le dernier (*chm-3*) d'une étape de différenciation par culture cellulaire. Les plantes qui en sont atteintes portent des feuilles panachées, c'est à dire perclues de tâches blanchâtres, qui après quelques générations présentent une hérédité maternelle. Ce phénotype semble toutefois présenter un taux de pénétrance relativement faible, ce qui signifie que toutes les plantes affectées au niveau moléculaire ne vont pas développer de phénotype visible. Son apparition est accompagnée d'un réarrangement spécifique (Martinez-Zapater *et al.* 1992), provoquant l'amplification d'une molécule subgénomique comportant une séquence chimérique impliquant le gène *atp9* (Sakamoto *et al.* 1996) (voir **Figure i 3.8**). Le produit du gène *CHM* est de ce fait probablement impliqué dans le contrôle du « substoichiometric shifting » de la molécule chimérique. Le clonage positionnel du gène *CHM* a été réalisé, et sa localisation confirmée d'abord par la présence de mutations chez chacune des trois lignées de plantes portant un allèle *chm* puis par l'obtention de deux mutants d'insertion présentant le même phénotype moléculaire.

Le gène *CHM* est un orthologue du gène *MSH1* de la levure (qui semble également jouer un rôle dans la recombinaison mitochondriale) et donc un paralogue du gène *Mut-S* de *E.coli* (ayant un rôle dans la voie MMR, « Mismatch repair »), avec une homologie particulièrement importante pour une région comportant un domaine de fixation de l'ATP (Abdelnoor *et al.* 2003). Cette dernière caractéristique est intéressante, au vu de l'identification récente d'un autre homologue de *Mut-S*, présent cette fois chez *Helicobacter pylori*, et appelée HpMutS2 (Pinto *et al.* 2005). En effet, cette protéine, dont l'activité biochimique a été testée, présente une activité ATPase stimulée spécifiquement par des structures mimant des intermédiaires de recombinaison, par rapport à de l'ADNdb comportant ou non des mésappariements. Cette protéine a de plus été capable *in vitro* d'inhiber la réaction d'invasion de brins catalysée par RecA. L'ensemble de ces données suggère donc fortement un rôle de la protéine MSH1 dans le contrôle de la stoechiométrie des molécules subgénomiques.

5.4 Les tissus transmetteurs

L'ADNmt est donc constitué d'un arrangement hétérogène de molécules, tant du point de vue de la distribution des séquences que de sa structure. Le fait qu'une partie des mitochondries puissent contenir une quantité de matériel génétique inférieure à la taille du génome a été confirmé par plusieurs études, qui avaient pour objet des cultures de protoplastes (Arimura *et al.* 2004) ou des tissus végétatifs (Bendich AJ, communication orale). Cela conduit à s'interroger sur la façon dont est transmis à la descendance l'intégralité du patrimoine génétique mitochondrial. L'hypothèse, mise en évidence par l'étude de la CMS de *P. vulgaris*, est que les sublimons soient conservés préférentiellement dans les tissus transmetteurs (cela a été démontré pour le méristème racinaire apical chez *P. vulgaris*). Cela signifie donc qu'un mécanisme spécifique à ces tissus permet d'assurer la conservation des séquences présentes en quantité substoechiométrique. Une autre approche a également conduit à l'observation qu'un processus de fusion, au cours duquel la recombinaison était particulièrement active, avait lieu lors de la génération d'hybrides somatiques (Moriguchi *et al.* 1997). Une publication récente, faisant état des conséquences sur la mitochondrie d'un processus de dédifférentiation a permis de visualiser la façon dont s'opèrent les échanges de matériel génétique nécessaires au retour à la totipotence (Sheahan *et al.* 2005). Cette étude, basée sur l'observation de la colocalisation de marqueurs différents en fonction de la cellule dont proviennent les mitochondries, démontre que la dédifférentiation implique une intense activité de fusion des mitochondries, ou MMF (« Massive Mitochondrial Fusion »). Cette MMF a lieu dès le début de la mise en culture des protoplastes et conduit à la redistribution complète du contenu mitochondrial en 24 h. De plus, par coloration au DAPI (un marqueur permettant de visualiser l'ADN), il est possible de visualiser la redistribution des nucléoides (organisation naturelle de l'ADNmt en complexes nucléoprotéiques) consécutive au processus de fusion. Les nucléoides sont en effet invisibles dans certaines mitochondries lors de la mise en culture et présents partout lorsque la phase exponentielle de croissance démarre. Enfin, la MMF s'avère être un processus spécifique de l'acquisition de la totipotence, car elle n'était pas observée dans le cas de protoplastes déjà dédifférentiés. L'ensemble de ces données suggère donc que les tissus transmetteurs se doivent de contenir des mitochondries possédant l'intégralité des séquences de leur génome. Le fait que la MMF ait lieu au début de la mise en culture devient particulièrement intéressant si on le met en parallèle avec les mécanismes de répllication dans des cultures de cellules de *C. album* (Backert et Borner 2000), décrites au cours du paragraphe 4. En effet ces mécanismes diffèrent selon l'étape du cycle cellulaire à laquelle l'ADNmt est observé. Les premières structures qui apparaissent après la mise en culture sont les rosettes, ces molécules

complexes contenant de nombreux branchements et jonctions de Holliday. Il semble donc que le mécanisme employé par les tissus transmetteurs fasse intervenir la recombinaison et/ou la réplication recombinaison-dépendante. Au vu du nombre d'événements de recombinaison différents nécessaires à la création de certains sublimons (voir paragraphes précédents), le processus employé est vraisemblablement la réplication, initiée par l'invasion de brins, qui permet la ré-amplification de n'importe quelle molécule, même présente en quantité substoechiométrique.

MATERIELS ET METHODES

CHAPITRE I : MATÉRIELS ET OUTILS INFORMATIQUES

1 Matériel végétal

1.1 Plantes entières

Nous utilisons des *Arabidopsis thaliana* écotype Columbia 0. Elles sont cultivées soit en serre en conditions 12 heures de jour à 22°C - 12 heures à l'obscurité à 20°, à un taux d'humidité de 50% et une luminosité de 10 klx minimum, soit en logette en conditions 12 heures à la lumière à 20°C - 12 heures à l'obscurité à 18°C, avec un taux d'humidité de 80% et une luminosité de 10 klx.

1.2 Cultures in vitro

Les cultures d'*A. thaliana* *in vitro* sont réalisées sur du **milieu de culture Arabidopsis** en boîtes de Pétri. Ces boîtes sont placées dans une logette avec 16 h de jour à 22°C et 8h de nuit à 20°C.

Milieu de culture Arabidopsis : MS255 pH 5,6 (titré au KOH) - saccharose 0,5% (p/v) - agarose 0,9% (p/v) (Difco)

1.3 Cultures cellulaires

Les cellules d'*A.thaliana* sont cultivées dans du milieu Gamborg's B5 (Duchefa G0210) contenant 20 g/L de saccharose et 1 mg/L de 2,4 D à PH 5,7. Les cultures sont placées sur un agitateur rotatif (125 rpm, 25°C). Elles sont repiquées stérilement dans du milieu frais toutes les semaines. Des cultures âgées de 5 jours sont utilisées pour les préparations de mitochondries.

2 Souches bactériennes

2.1 E. coli

Trois souches dérivant de *E. coli* K12 (Yanisch-Perron *et al.* 1985) ont été utilisées pour les différents clonages. Il s'agit des souches SURE, XL1-Blue et DH5 α . Ces souches présentent des avantages multiples : un taux de recombinaison réduit

(augmentation de la stabilité de l'ADN), une perte de l'activité de l'endonucléase I (*endA1*) et une délétion de la partie N-terminale du gène codant pour la β -galactosidase (*lacZ M15*) qui abolit l'activité de la protéine correspondante. La transfection de ces souches par un plasmide portant la partie tronquée de ce gène restaure l'activité β -galactosidase et permet la sélection des bactéries recombinantes suivant leur couleur.

Le génotype de ces souches est le suivant :

SURE : $e14^-(McrA^-) \Delta(mcrCB-hsdSMR-mrr)171$ *endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kan^r) uvrC [F' proAB lacI^qZ Δ M15 Tn10 (Tet^r)]*

XL1-Blue : *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI^qZ Δ M15 Tn10 (Tet^r)]*

DH5 α : $F^- deoR endA1 hsdR17 (r_k-m_k^+) phoA supE44 \lambda^- thi-1 recA1 gyrA96 relA1 \Delta(lacZYA-argF) U169 (\phi 80\Delta lacZ \square M15)$

Pour la surexpression de protéines, nous avons utilisé la souche BL21(DE3) (Stratagène). Les BL21(DE3) sont des bactéries qui possèdent le gène de l'ARN polymérase T7 sur leur ADN génomique. Ce gène est sous le contrôle du promoteur *lacUV5*, ce qui permet une expression inductible de la polymérase par l'IPTG. De plus, ces bactéries sont déficientes en protéases OmpT et Lon, ce qui limite la dégradation des protéines surexprimées.

Génotype des BL21 (DE3) : *E. coli B F^- dcm ompT hsdS(r_B^-m_B^-) gal \lambda (DE3)*

Pour l'expression en très grande quantité (par exemple lorsque l'on envisage de purifier des protéines à partir de corps d'inclusions), nous avons utilisé la souche BL161, un dérivé des BL21(DE3) contenant un plasmide conférant la résistance à la kanamycine et codant pour deux ARNt^{Arg} peu abondants chez les bactéries (AGA et AGG).

2.2 *Agrobacterium tumefaciens*

La souche d'*Agrobacterium* utilisée est la souche désarmée GV3101::pMP90. Elle possède sur son chromosome le gène de la réplicase RK2, qui permettra la réplication du plasmide binaire pBINPLUS, et un gène de résistance à la rifampicine. Cette souche contient également le plasmide « helper » pMP90, qui provient du plasmide pTiC58, délété de son T-DNA. Il porte les gènes *vir* permettant le transfert du T-DNA vers les cellules végétales hôtes ainsi qu'un gène de résistance à la gentamicine.

3 Plasmides

3.1 pSK (-/+)

Ce vecteur, de type pBluescript, dérive de pUC19 (Stratagène). Il possède une origine de réplication de phage filamenteux. Le signe - ou + indique l'orientation du brin d'ADN qui est produit si la culture de bactéries est infectée par un phage « helper ». La cassette de clonage est flanquée par les séquences promotrices des ARN polymérases des phages T3 et T7, permettant la transcription des inserts. Ce vecteur porte en outre le gène de la β -lactamase qui confère aux bactéries transfectées la résistance à l'ampicilline (sélection positive des transformants). La cassette de clonage est insérée en phase dans le gène codant pour la partie N-terminale de la β -galactosidase, placé sous le contrôle du promoteur *lac*, lequel est inductible par l'IPTG. En absence d'insert, la complémentation de l'activité de la β -galactosidase permet, après induction, l'hydrolyse de l'X-gal (analogue du substrat naturel de l'enzyme) présent dans le milieu de culture. Cette hydrolyse conduit à la coloration en bleu des bactéries transfectées par un vecteur sans insert. Par contre, la présence d'un insert dans la cassette de clonage supprime la complémentation (changement de phase, protéine de fusion aberrante) et conduit à l'obtention de colonies blanches.

3.2 pRSETc

Le plasmide pRSETc (Invitrogen) est un vecteur d'expression de 2,9 kb dérivant de pUC. Le vecteur porte une origine de réplication *f1* ainsi que le gène de la β -lactamase qui confère la résistance à l'ampicilline. L'expression de protéines est médiée par le promoteur T7. La cassette de clonage se situe en aval d'une séquence codant pour un peptide de fusion N-terminal. Cette séquence comprend, après le codon initiateur, une étiquette hexahistidines, une séquence de stabilisation des transcrits du gène 10 des phages T7, l'épitope X-pressTM et une séquence codant pour le site de clivage de l'entérokinase. En amont du codon initiateur, le plasmide porte également un site de fixation des ribosomes pour améliorer la traduction.

3.3 PGEX

Nous avons utilisé le plasmide pGEX-4T-1 (Amersham Biosciences), un vecteur d'expression de 4,9 kb permettant d'ajouter la protéine GST ("Gluthation S-Transferase") en N-terminal de la protéine d'intérêt. Il comporte l'origine de réplication de pBR322 et le gène de la β -lactamase qui confère la résistance à l'ampicilline. L'expression est contrôlée par le promoteur *tac*, inductible à l'IPTG. Le

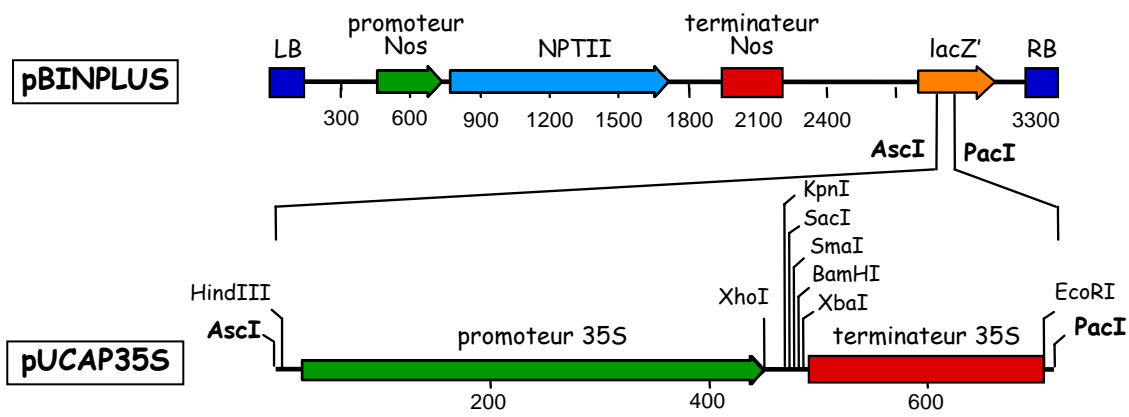


Figure m1: Carte des vecteurs pUCAP et pBINPLUS

plasmide porte également le gène *LacI*^q, codant pour un répresseur de la région *tac*, qui permet d'éviter des fuites de l'expression en absence d'inducteur. Enfin, la cassette de clonage est en phase (en 3') avec un gène codant pour la GST de *Schistosoma japonicum*.

3.4 pUCAP, pUCAP35S et pUCAP35S-TAP.

Le plasmide pUCAP est un pUC avec des sites rares *AscI* et *PacI* en bordure de la cassette de clonage. Ces sites sont présents dans le plasmide pBINPLUS. Contrairement à pBINPLUS, pUCAP est un plasmide à grand nombre de copies et de petite taille. On se sert donc de pUCAP pour réaliser les constructions facilement avant de les transférer dans pBINPLUS par digestions *AscI* et *PacI* en vue de transformer du matériel végétal (voir **Figure m1**). Dans cette optique de transformation, le vecteur, alors appelé pUCAP35S, a été pourvu du promoteur 35S du virus de la mosaïque du chou-fleur. Le promoteur est situé entre les sites *AscI* et *PacI*, avant la cassette de clonage. Ce promoteur permettra, une fois l'ADN-T transféré via *Agrobacterium*, une expression forte de l'insert dans la plante. Enfin, une dernière modification de ce vecteur a consisté en l'ajout entre les sites *KpnI* et *BamHI* d'une séquence comportant les peptides FLAG (DYKDDDDK) et CBP ("Calmodulin Binding Peptide" ; KRRWKKNFIAVSAANRFKKISSSGAL), séparés par un bras de trois Glycines, de façon à obtenir des fusions C-terminales de nos protéines d'intérêt.

3.5 pBINPLUS

Le plasmide binaire pBINPLUS (van Engelen *et al.* 1995) (basé sur le vecteur pBIN19 (Bevan 1984)), d'une taille de 11,7 kb, a été utilisé pour la transformation des plants *d'Arabidopsis* et l'infiltration de feuilles de *N.tabacum*. Il comporte deux origines de réplication. Le site de réplication originellement présent dans pBIN19, RK2, entraîne un faible taux de réplication. Celui-ci est augmenté d'un facteur 4 par l'ajout du site *ColE1*. Ce plasmide comporte également le gène *NPTIII*, qui code pour la néomycine phosphotransférase (conférant la résistance à la kanamycine), sous le contrôle d'un promoteur bactérien. Il contient de plus les séquences flanquantes de l'ADN-T (LB et RB) qui permettent, via *Agrobacterium*, le transfert de l'information contenue entre elles dans le génome nucléaire du matériel végétal. Entre ces séquences flanquantes, on trouve le gène *lacZ'*, permettant de réaliser la sélection blanc/bleu. Ce gène est interrompu par une cassette de clonage avec les sites *AscI* et *PacI* aux extrémités. On y trouve aussi le gène *NPTII* qui confère la résistance à la kanamycine sous le contrôle du promoteur et du terminateur NOS du gène de la nopaline synthase (voir **Figure m1**). Cette résistance permettra la sélection du matériel végétal transformé. Dans pBINPLUS, contrairement à pBIN19, l'extrémité 3' de ce gène se

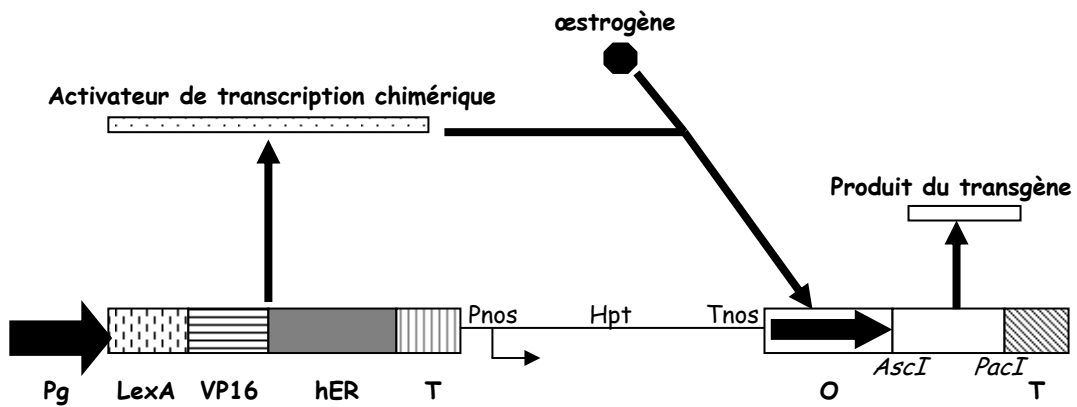


Figure m2 : Plasmide pER8

Pg: Promoteur synthétique (G 10-90) ; **LexA**: domaine de liaison à l'ADN ; **VP16**: domaine activateur de transcription ; **hER**: région régulatrice du récepteur d'œstrogène ; **T**: terminaison ; **O**: opérateur ; **Pnos et Tnos**: promoteur et terminateur de la nopaline synthase ; **Hpt**: gène de résistance à l'hygromycine ; **AscI** et **PacI**: sites d'insertion du transgène.

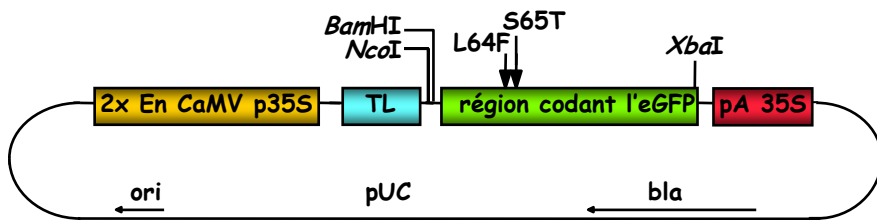


Figure m3 : Vecteur pCK-GFP3 (4,45 kb)

trouve juste à côté de la bordure gauche de L'ADN-T. Comme l'ADN-T est transféré de droite à gauche, les plantes résistantes à la kanamycine auront vraisemblablement intégré la totalité de celui-ci.

3.6 pER8

Le plasmide binaire pER8 (11,5 kb), permet de contrôler l'expression de l'ADNc cloné dans la plante par une induction à l'œstradiol (Zuo *et al.* 2000). Ce vecteur comporte le gène résistance à la spectinomycine qui permet la sélection lorsque l'hôte est bactérien et le gène de résistance à l'hygromycine pour la sélection des plantes transformées. La construction à exprimer est réalisée, comme pour pBINPLUS, dans le vecteur pUCAP puis transférée dans la cassette XVE de pER8 après digestion par *AscI* et *PacI*. L'activateur chimérique XVE (voir **Figure m2**) est constitué d'une fusion du domaine de fixation à l'ADN du répresseur bactérien LexA avec le domaine acide de transactivation VP 16 et la région régulatrice du récepteur d'œstrogènes humain (*hER*). L'activité transactivatrice de XVE, dont l'expression est sous le contrôle du promoteur constitutif fort G 10-90, est strictement régulée par les œstrogènes.

3.7 PBI 101

Le plasmide binaire pBI 101 (Clontech) permet de fusionner une région potentiellement promotrice en amont du gène de la beta-glucuronidase (GUS). L'activité promotrice pourra ainsi être rapportée par la synthèse *in vivo* de GUS, qui précipite dans le cytoplasme des cellules où elle est produite. La présence de GUS dans les tissus est révélée par une réaction colorimétrique suivie d'une observation microscopique *in toto* ou sur coupes.

3.8 pCK-GFP3

Le vecteur utilisé pour permettre l'expression transitoire des protéines de fusion avec l'eGFP ("Green Fluorescent Protein") est le plasmide pCK-GFP3 (Bevan 1984). Ce vecteur est dérivé de pUC (**Figure m3**). Il contient le gène de l'eGFP, qui est issu de modifications apportées à celui de la GFP en vue de l'adapter à l'usage des codons des eucaryotes et d'augmenter la fluorescence de la protéine (L64F et S65T) (Youvan et Michel-Beyerle 1996). Des sites *NcoI* et *BamHI*, situés en 5' du gène, permettent le clonage en phase avec celui-ci pour l'expression des protéines de fusion. L'expression dans un système eucaryotique végétal est assurée par deux éléments promoteurs de l'ARN 35S du CaMV, ainsi que par une séquence TL, activatrice de traduction, dérivée du génome du virus de la mosaïque de la luzerne. Les bactéries portant le vecteur sont sélectionnées grâce à la résistance à l'ampicilline (gène *bla*).

3.9 pCR[®]2,1-TOPO

Le vecteur pCR[®]2,1-TOPO (3,9 kb) possède des extrémités 3'dT sortantes ainsi que 2 sites de restriction *EcoRI* flanquant le site d'insertion. Ce plasmide est utilisé pour le clonage direct d'un fragment de PCR. Il contient le gène qui confère la résistance à l'ampicilline.

4 Oligodésoxyribonucléotides

Les oligodésoxyribonucléotides de synthèse sont fournis par diverses sociétés. La température de fusion est estimée par la formule suivante :

$$T_m (\text{°C}) = 2 \times (A+T) + 4 \times (C+G)$$

ou calculée plus précisément avec celle-ci :

$$T_m (\text{°C}) = 81,5 + 0,41\%(G+C) + 16,6 \cdot \log[\text{sels}] - 675 / (A+C+G+T) - 0,65\%(\text{formamide}).$$

5 Outils informatiques

5.1 Banques de données

Voici une liste non exhaustive des URLs les plus couramment utilisées :

NCBI : <http://www.ncbi.nlm.nih.gov/>

ExPASy : <http://www.expasy.ch/>

« In silico analysis techniques » : <http://fosi.immunesig.org/bioinformaticsprotids.html>

ABIM : <http://www.up.univ-mrs.fr/wabim/english/logligne.html>

The « *Arabidopsis* Information Ressource » (TAIR) : <http://www.arabidopsis.org/>

5.2 Prédictions informatiques d'adressage vers les organites

Il existe plusieurs sites Internet qui permettent de prédire dans quel(s) compartiment(s) subcellulaire(s) sera adressée une protéine. De manière générale, il est préférable de recouper les résultats des prédictions.

Predotar : <http://urgi.infobiogen.fr/predotar/french.html>

TargetP : <http://www.cbs.dtu.dk/services/TargetP/>

MitoP2 : <http://ihg.gsf.de/mitop2/start.jsp/>

iPSort : <http://hc.ims.u-tokyo.ac.jp/iPSORT/>

Wolf PSORT : <http://wolfsort.seq.cbrc.jp/>

CHAPITRE II : METHODES

1 Purification de mitochondries

1.1 Mitochondries de pommes de terre et de chou-fleur.

Ce protocole est une combinaison des techniques décrites par Neuburger et Struglics (Neuburger *et al.* 1982) (Struglics *et al.* 1993). Toutes les étapes se déroulent à 4°C. Les pommes de terre sont pelées et coupées en morceaux. Le jus, environ 2 L pour 7 kg de tubercules en est ensuite extrait au moyen d'une centrifugeuse à jus (Moulinex) dans 1 L de **tampon d'extraction 3x** concentré. Le pH de l'extrait est ajusté à pH 7.5 par addition de K₂HPO₄ 1 M. L'extrait décante ensuite une dizaine de minutes pour culotter l'amidon, avant d'être filtré sur une toile de Nylon de 96 µm de diamètre de maille. L'étape suivante consiste à enrichir la solution en mitochondries : par des centrifugations différentielles, à basse vitesse (10 minutes à 2000 g pour sédimenter les débris cellulaires et les noyaux) et à haute vitesse (10000 g pendant 20 minutes pour culotter les mitochondries). Les mitochondries sont enfin purifiées sur des gradients de Percoll (autoformés à partir d'une **solution de gradient**) lors d'une centrifugation à 40000 g pendant 1h30. La bande correspondant aux mitochondries est récupérée et lavée deux fois dans du **tampon de lavage**. Le rendement est d'environ 50 mg d'équivalent protéines de mitochondries pour 5 kg de tubercules de pommes de terre. Pour une utilisation ultérieure, les mitochondries sont congelées dans l'azote liquide (cryoprotectant : 7.5% d'éthylène glycol) et conservées à -80°C.

Pour la purification de mitochondries de chou-fleur, le protocole est identique à celui des mitochondries de pommes de terre. Avant broyage, les feuilles et la majorité des tiges du chou-fleur sont éliminées pour ne conserver que les ébauches florales. Un chou-fleur produit environ 300 mL de jus. L'étape de décantation n'est pas nécessaire car les choux-fleurs ne contiennent pas autant d'amidon que les pommes de terre. Le rendement chez le chou-fleur est d'environ 10 mg d'équivalent protéines de mitochondries pour un chou-fleur.

Tampon d'extraction 3x : mannitol 0.9 M - diphosphate de sodium 90 mM pH 7,5 - EDTA 6 mM - PVP 25k 2,4% (p/v) - cystéine 0,15% (p/v) - glycine 15 mM - BSA 0,9% (p/v) β-mercaptoéthanol 6 mM

Tampon de lavage : mannitol 0.3 M - phosphate de potassium 10 mM pH 7,5 - EDTA 1 mM - BSA 0,1% (p/v) - glycine 5 mM

Solution de gradient : saccharose 0,3 M - phosphate de potassium 10 mM pH 7,5 - EDTA 1 mM - BSA 0,1% (p/v) - Percoll 28% (v/v)

1.2 Fractionnement subcellulaire et purification de mitochondries d'*Arabidopsis*

Ce fractionnement suit un protocole décrit par Sakamoto (Sakamoto *et al.* 2000). Les cellules d'une culture cellulaire d'*Arabidopsis* sont récoltées par passage à travers un filtre de nylon de 100 µm de maille. Des protoplastes sont obtenus par digestion de la paroi dans du tampon enzyme contenant 1% (p/v) cellulase RS, 0,1% (p/v) pectolyase Y-23 pendant 3 h à 25°C sous agitation à 45 rpm. La suspension est centrifugée 10 min à 800 g et le culot de protoplastes est lavé deux fois dans le **tampon enzyme**. Les protoplastes sont remis en suspension dans du **tampon d'extraction** puis cassés par 3 passages à travers un filtre de nylon de 30 µm de maille. Une première centrifugation de 5 min à 100 g permet d'éliminer les noyaux et débris cellulaires. Deux cycles de centrifugations différentielles sont effectués ensuite : 5 min à 3000 g et 10 min à 16000 g. Après la centrifugation à 3000 g, le culot contient une fraction enrichie en chloroplastes, qui sont purifiés sur un gradient discontinu à 40% et 80% de Percoll. Le surnageant de la centrifugation à 16000 g est conservé comme témoin de la fraction cytosolique. Le culot de cette centrifugation, enrichi en mitochondries, est repris dans un petit volume de **tampon de lavage**, chargé sur un gradient de Percoll discontinu (18%-23%-40%) et centrifugé 15 min à 20000 g. Les mitochondries purifiées sont récupérées à l'interface 18%-23% et lavées dans du tampon de lavage.

Tampon enzyme : Mannitol 0,4 M - MES-KOH 3,5 mM pH 5,7

Tampon d'extraction : Saccharose 0,4 M - Phosphate de potassium 30 mM pH 7,5 – EDTA 2 mM - BSA 0,1% - DTT 2 mM

Tampon de lavage : Saccharose 0,3 M – Phosphate de potassium 30 mM pH 7,5 – EDTA 1 mM - BSA 0,1%

2 Préparation des bactéries compétentes

Cinq mL d'une culture d'une nuit d'*Escherichia coli* sont ajoutés à 500 mL de milieu LB. La suspension bactérienne est incubée à 37°C jusqu'à acquisition d'une DO₆₀₀ de 0,5 à 0,8. Ensuite, la culture est refroidie dans la glace pendant 30 min et passe par deux étapes successives de centrifugation de 5 min à 5 000 g et de resuspension dans de l'eau stérile froide, d'abord dans un volume de 500 mL, puis dans 250 mL. Après une troisième centrifugation de 5 min à 5 000 g, le culot de bactéries est repris dans 40 mL de glycérol à 10%, centrifugé et concentré dans 4 mL

de glycérol à 10%, puis aliquoté en fractions de 40 μ L. Les aliquotes sont congelées et gardées à -80°C.

Les *Agrobacterium* poussent quant à elles à 28°C pendant une nuit dans du milieu LB additionné des antibiotiques adéquats (Rifampicine, Kanamycine et Gentamycine ou Tetracycline). Après une centrifugation de 20 min à 3000 x g et 4 °C, le surnageant est éliminé et remplacé par un volume équivalent de glycérol 10 %. L'étape précédente est répétée, puis les bactéries sont resuspendues dans 1/500^è du volume initial et congelées en aliquotes de 35 μ L à - 80 °C.

3 Techniques communes aux différents acides nucléiques

3.1 Quantification des acides nucléiques

Pour doser une solution d'acides nucléiques tout en s'assurant de l'absence de contamination protéique, nous utilisons une méthode spectrophotométrique. L'absorbance de la solution est mesurée à 260 nm et à 280 nm (A_{260} et A_{280}) pour un trajet optique de 1 cm. Le rapport A_{260}/A_{280} nous renseigne sur la pureté de la solution (un bon rapport est compris entre 1.7 et 2.0). Pour l'évaluation de la concentration en acides nucléiques, on utilise les correspondances suivantes :

1 A_{260} équivaut à 50 μ g/mL d'ADNdb

1 A_{260} équivaut à 33 μ g/mL d'ADNsb

1 A_{260} équivaut à 40 μ g/mL d'ARN ou d'ADN simple brin.

3.2 Séparation par électrophorèse

3.2.1 Electrophorèse non dénaturante en gel d'agarose

Les molécules d'ADN peuvent être séparées en fonction de leurs tailles par migration électrophorétique. L'ADN est chargé avec 1/3 de volume de **tampon de charge** sur un gel d'agarose (entre 0.6% et 2.5% (p/v) dans du **tampon TEB**). La migration de l'ADN s'effectue dans du tampon TEB sous une tension de 100 V. Les fragments d'ADN sont ensuite visualisés sous UV grâce au bromure d'éthidium dissous dans le gel (50 μ g pour 100 mL).

Tampon TEB : Tris-HCl 90 mM pH 8,0 - EDTA 2 mM - acide borique 90 mM

Tampon de charge : glycérol 50% (p/v) - SDS 1% (p/v) - EDTA 1 mM - xylène cyanol 0,1% (p/v) - bleu de bromophénol 0.1% (p/v)

3.2.2 Electrophorèse dénaturante en gel d'agarose/formaldéhyde

Ce type de gel permet de séparer les ARN en fonction de leur taille et indépendamment de leur structure. Les ARN (1 volume) sont dénaturés 5 min à 65°C dans la **solution de dénaturation**, refroidis rapidement dans la glace et chargés sur un gel d'agarose 1.2% (p/v), formaldéhyde 2,2 M et **tampon MOPS** après addition d'un dixième de volume de **tampon de charge**. La migration s'effectue sous une intensité de 25 mA pendant 12-15 h dans du tampon MOPS.

Tampon de charge : glycérol 50% (v/v) - EDTA 1 mM - bleu de bromophénol 0,4% (p/v)

Tampon MOPS : MOPS 25 mM pH 7,0 - acétate de sodium 5 mM - EDTA 1 mM

Solution de dénaturation (finale) : tampon MOPS x 1 - formamide 40% (v/v) - formaldéhyde 7% (v/v)

3.2.3 Electrophorèse dénaturante en gel de polyacrylamide

La migration sur gel de polyacrylamide permet la séparation aussi bien de l'ADN (gel de séquence) que de l'ARN. Le gel est préparé à partir d'une solution d'acrylamide/N,N' méthylène bisacrylamide (19/1) à 7 M urée dans du tampon TEB. La polymérisation du gel est catalysée par l'addition d'APS 10% (0.1% p/v final) et de TEMED (0.035% v/v final). La migration s'effectue sous environ 25 mA (à ajuster en fonction de l'épaisseur du gel). Le bleu de bromophénol et le xylène cyanol du **tampon de charge** constituent des témoins de migration (l'équivalence en nt dépend de la concentration du gel). Les acides nucléiques sont ensuite visualisés par autoradiographie s'ils sont marqués ou sous UV après coloration au bromure d'éthidium (BET).

Tampon de charge : formamide désionisée 95% (v/v) - EDTA 1 mM - xylène cyanol 0,1% (p/v) - bleu de bromophénol 0,05% (p/v)

3.2.4 Purification d'acides nucléiques par électrophorèse

3.2.4.a A partir d'un gel d'agarose

Le fragment d'ADN est purifié à partir d'un gel d'agarose à 0,6 %. Après la localisation du fragment d'intérêt sous lumière UV (grâce au BET), la partie de gel qui le contient est excisée. Pour les fragments d'ADN de taille supérieure à 0,3 kb, l'extraction est réalisée au moyen du kit Concert (Gibco-Brl) : dissolution du gel au perchlorate de sodium, rétention de l'ADN sur membrane de silice, lavée ensuite avec une solution éthanolique et élution au **tampon TE**. Pour extraire les fragments de

taille inférieure à 0,3 kb, nous avons utilisé le kit Ultrafree-DA (Millipore) : basé sur la force centrifuge.

Tampon TE : Tris-HCl 10 mM pH 8 - EDTA 1 mM

3.2.4.b *A partir d'un gel de polyacrylamide*

La réaction de synthèse des oligonucléotides à une efficacité d'environ 98-99 % pour chaque incorporation, ce qui signifie que 60 % de la quantité totale d'un 40-mer auront la bonne taille (98,5⁴⁰) et que les 40 % restants mesureront entre 1 et 39 nt. Cela permet de prendre la mesure de l'importance de purifier les oligonucléotides employés comme sondes pour étudier les interactions ADN-protéines. Après séparation sur gel de polyacrylamide, les fragments contenant les bandes d'ADN d'intérêt sont excisés, broyés, puis l'ADN en est élué par une incubation (la nuit à 42 °C) dans du tampon d'éluion (0,5 M acétate d'ammonium / 0,1 mM EDTA). Les fragments de gel sont retenus sur un filtre de coton stérile et l'ADN est précipité à l'éthanol en présence d'un entraîneur. Le culot est ensuite lavé et repris dans le volume adéquat de tampon TE (10mM Tris-HCL pH(8,0) ; 1mM EDTA).

4 **Techniques relatives à l'ADN**

4.1 **Techniques de clonage**

4.1.1 Digestion par des enzymes de restriction

Les digestions d'ADN par des enzymes de restriction sont réalisées dans les conditions préconisées par le fournisseur.

4.1.2 Ligation

Lors de la préparation du vecteur et dans le cas où celui-ci ne serait digéré que par une seule enzyme de restriction, on prendra soin de déphosphoryler l'extrémité 5'-phosphate (traitement à la phosphatase alcaline de veau (CIP, pour "Calf Intestine Phosphatase") afin d'éviter la recircularisation du vecteur sans insert. Ce problème ne se pose pas si le vecteur est digéré par deux enzymes de restriction (extrémités non cohésives). L'insert est digéré par des enzymes de restriction qui présentent des extrémités cohésives avec celles du vecteur.

La réaction de ligation est incubée 2h à température ambiante ou toute la nuit à 14°C. Le milieu de réaction contient :

- 20 ng de vecteur

- insert d'ADN, dans un rapport stœchiométrique insert/vecteur compris entre 3 et 5.

- **tampon de ligation 5x (GibcoBRL)**

- 1 U de T4 ADN ligase (GibcoBRL)

Tampon de ligation 5x (GibcoBRL): Tris-HCl 250 mM pH 7,6 - MgCl₂ 50 mM - ATP 5 mM - DTT 5 mM - PEG 8000 25% (p/v)

4.1.3 Transformation par électroporation des *E. coli* et des *Agrobacterium*.

Après la ligation, 2 à 50 ng de plasmide sont mélangés avec 40 µL d'une suspension d'*E. coli* compétentes. Les bactéries sont placées dans une cuve à électroporation refroidie dans la glace et soumises à une décharge électrique (25 µF ; 2,5 V ; 200 Ω) dans un électroporateur (Bio-Rad Gene pulser) qui perméabilise transitoirement la paroi bactérienne et fait pénétrer le vecteur plasmidique. Immédiatement après l'électroporation, les bactéries sont reprises dans 1 mL de **milieu LB** froid et mises à incuber pendant 30 min à 37°C. L'incubation permet la régénération de la paroi bactérienne et l'expression des gènes de résistances codés par le vecteur. Puis 10 à 150 µL de bactéries sont étalés sur une boîte de Pétri contenant du milieu LB - agar 1,5% (p/v) et l'antibiotique permettant la sélection. Si l'on utilise un plasmide avec une sélection blanc/bleu, on ajoute sur la boîte avant l'étalement des bactéries 10 µL d'IPTG 100 mM et 40 µL de X-Gal 2% dans du diméthyl formamide (p/v). Les boîtes sont ensuite incubées une nuit à 37°C pour le développement des colonies.

Pour les *Agrobacterium*, le principe est le même que pour les *E. coli* à la différence que les conditions d'électroporation sont les suivantes 25 µF ; 2,5 V ; 400 Ω. Les agrobactéries sont ensuite étalées sur milieu LB agar contenant de la kanamycine (50 µg/mL), de la rifampicine (50 µg/mL) et de la gentamicine (10 µg/mL) et mis en culture 2 jours à 28°C.

Milieu LB (Lennox L broth, Sigma) : Tryptone 10 g/L - extrait de levure 5 g/L - NaCl 5 g/L

4.2 Préparation de l'ADN plasmidique

4.2.1 Préparation d'ADN plasmidique à petite échelle (miniprep)

Pour préparer une petite quantité d'ADN, nous avons employé la technique de la lyse alcaline (Birnboim et Doly 1979). Deux millilitres de suspension bactérienne sont centrifugés 5 min à 1000 g. Le culot bactérien est resuspendu dans 100 µL de **tampon A** (glucose 50 mM - Tris-HCl 25 mM pH 8 - EDTA 10 mM). On ajoute ensuite 200

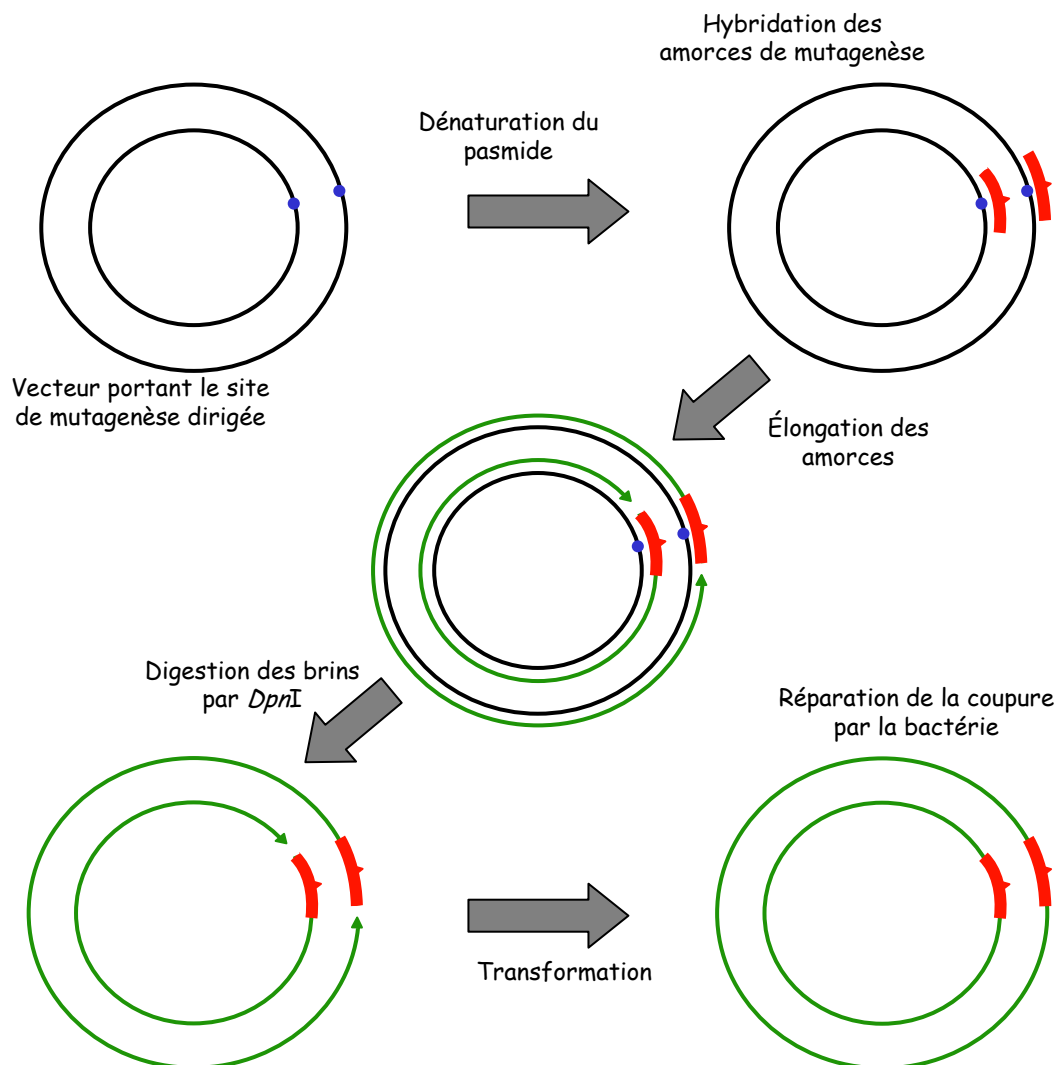


Figure m 4: Les diérentes étapes de la technique de mutagenèse par la « Pfu DNA polymerase »

Le point de mutation désiré apparaît en bleu, les amorces portant la mutation en rouge et les brins néo-synthétisés en vert

μL d'une **solution de lyse** (NaOH 0,2 M – SDS 1 %). Après 5 min, 150 μL d'acétate de potassium 3 M pH 5,2 sont ajoutés et les tubes sont mis dans la glace pendant 5 min. Après 10 min de centrifugation à 15000 g, les acides nucléiques contenus dans le surnageant sont extraits au phénol/chloroforme (1v/1v). On réalise ensuite une précipitation éthanolique par addition d'acétate d'ammonium pH 5,3 (300 mM final) puis de 2,5 volumes d'éthanol absolu. Après centrifugation 10 min à 15000g, le culot est lavé 2 fois à l'éthanol 70% et resuspendu dans 50 μL de **tampon TE** ou d'eau avec 100 $\mu\text{g}/\text{mL}$ RNase.

Tampon TE : Tris-HCl 10 mM pH 8 - EDTA 1 mM

4.2.2 Préparation d'ADN plasmidique à moyenne et grande échelle (midiprep et maxiprep)

Afin d'obtenir des quantités plus importantes d'ADN, nous sommes partis de 50 à 500 mL de culture bactérienne. La purification a été faite à l'aide de kits Jet star Midi 2.0 et Maxi 2.0 (Genomed) basés sur lyse alcaline, suivie de la fixation de l'ADN sur résine échangeuse d'anion, et de l'éluion aux sels à pH basique. Toutes les étapes de ces purifications sont réalisées dans les conditions préconisées par le fabricant.

4.3 Mutagenèse dirigée

Deux types de mutagenèse ont été réalisés : i) certains mutants de délétion ont été obtenus en utilisant des amorces portant des sites de restriction en 5' (cette méthode ne nécessitant que des étapes de PCR et de clonage, décrits par ailleurs, nous ne la détaillerons pas) et d'autres, ii) parmi lesquels les mutants ponctuels, ont été obtenus en utilisant la technique de mutagenèse schématisée en **Figure m4**, utilisant la « *Pfu* DNA polymerase » (Clontech). Elle peut servir à modifier un codon comme à en ajouter ou en déléter plusieurs. Chacun des brins du plasmide non coupé à modifier est amplifié par PCR unidirectionnelle. Au cours de chaque élongation (10 à 20 min), l'enzyme copie chaque brin entièrement et vient buter sur l'amorce (elle n'a pas d'activité exonucléase 5'→ 3'). Le produit de PCR est ensuite traité par *DpnI*, spécifique des brins méthylés (le plasmide utilisé, produit dans la souche DH5 α , est donc *dam* méthylé) et reconnaissant la séquence 5'-G^{m6}ATC-3'. Cela permet donc l'élimination spécifique du brin parental non muté. Après transformation dans une bactérie de souche DH5 α , le vecteur muté est réparé puis amplifié par celle-ci.

4.4 Séquençage de l'ADN

Les séquences d'ADN sont réalisées par le service de séquençage de l'Institut (M.Alioua) qui utilise l'appareil "3100 ABI Prism Genetic Analyser" (Applied Biosystems/Hitachi).

4.5 Amplification par PCR ("Polymerase Chain Reaction")

Il s'agit d'une amplification exponentielle d'une séquence d'ADN comprise entre deux amorces oligonucléotidiques, réalisée par une série de cycles de type dénaturation-hybridation-élongation. L'ADN polymérase utilisée, la *Taq* ADN polymérase, est une enzyme thermorésistante extraite de la bactérie *Thermophilus aquaticus* qui reste active aux températures élevées nécessaires à la dénaturation de l'ADN. La réaction est réalisée dans le milieu suivant :

- 5 pg d'ADN cible
- 50 pmol de chaque oligonucléotide amorce
- dNTP 200 μ M chacun
- 1.5 à 2.5 mM de $MgCl_2$
- 5 μ L de **Tampon *Taq* PCR 10x (GibcoBRL)**
- 2 U de *Taq* ADN polymérase
- H_2O qsp 50 μ L

Tampon *Taq* PCR 10x (GibcoBRL) : Tris-HCl 200 mM pH 8,4 - KCl 500 mM
- EDTA 1mM

4.6 Marquage d'oligonucléotides

Les oligonucléotides qui nous sont livrés sont déphosphorylés à leur extrémité 5'. Il est possible de les marquer par incubation 30 min à 37°C dans le milieu réactionnel suivant :

- 2 à 5 pmol d'oligonucléotide
- 20 μ Ci [γ - ^{32}P]ATP (3000Ci/mmol, NEN Life Science, soit 7 pmol de [γ - ^{32}P]ATP)
- 1 μ L Tampon PNK 10x (Biolabs)
- 5 mM DTT
- 5 U de T4 polynucléotide kinase (Biolabs)
- H_2O q.s.p. 10 μ L

Le volume est ajusté à 200 μ L, les oligonucléotides sont extraits au phénol/chloroforme puis la radioactivité non incorporée est éliminée par chromatographie d'exclusion sur une colonne de Séphadex G50 (1 mL) préalablement

équilibrée dans du tampon TE. La colonne est séchée par centrifugation 1 minute à 200 g. La solution est déposée et éluée dans les mêmes conditions.

Tampon PNK 10x (Biolabs) : Tris-HCl 700 mM pH 7,6 - MgCl₂ 100 mM

4.7 Marquage d'ADNdb

4.7.1 Par Amorçage aléatoire (« random priming »)

De 10 à 100 ng d'ADN sont dénaturés à 95°C pendant 5 min puis portés dans la glace. On ajoute alors 50 ng d'oligonucléotides hexamères, 200 µM de dATP, dTTP et dGTP, 50 µCi de [α -³²P]dCTP (NEN Life Science, 3000Ci/mmol, 10 µCi/µL), 0,25 % (p/v) de BSA, 10 µL de **tampon "random priming" 5x**, 2 U de Klenow et de l'eau pour arriver à un volume final de 50 µL. Le tout est incubé 1 heure à 37°C. La radioactivité non incorporée est éliminée par chromatographie d'exclusion comme décrit pour le marquage des oligonucléotides. L'activité de la sonde est déterminée par comptage Cerenkov : un µL de sonde est compté dans un volume total de 100 µL d'eau sur le canal tritium.

Tampon "random priming" 5x : Tris-HCl 250 mM pH 8,0 - MgCl₂ 25 mM - β -mercaptoéthanol 10 mM

4.7.2 Par PCR classique

Il s'agit dans ce cas d'obtenir une sonde ADNdb ayant une séquence particulière. Les conditions sont celles d'une PCR classique, hormis pour le rapport [α -³²P]dCTP / dCTP froid, qui est ajusté en fonction du nombre de nucléotides marqués que l'on souhaite incorporer par molécule de sonde.

5 Techniques relatives à l'ARN

5.1 Extraction des ARN totaux

Le protocole utilisé ici est très largement inspiré de (Goodall *et al.* 1990). C'est une méthode standard d'extraction au thiocyanate de guanidium. Environ 1 g de matériel végétal est broyé dans l'azote liquide. La poudre obtenue est reprise dans 3 mL de **tampon d'extraction d'ARN** puis la solution est vigoureusement agitée. On y ajoute alors :

- 0.2 mL d'acétate de sodium 3 M pH 4,8
- 3 mL de phénol saturé à l'eau

- 0.6 mL de chloroforme/alcool isoamylique (24/1)

La solution est mélangée par agitation puis laissée dans la glace pendant 15 min. La phase aqueuse est récupérée après centrifugation (30 min 25000 g) puis les acides nucléiques sont précipités par addition d'un volume d'isopropanol. Les acides nucléiques sont alors lavés à l'éthanol 70%, séchés puis resuspendus dans de l'eau stérile. Il est préférable de renouveler l'extraction au phénol/chloroforme afin d'améliorer la qualité des acides nucléiques. Après précipitation, lavage à l'éthanol 70 %, séchage et resuspension, la concentration des acides nucléiques est finalement déterminée au spectrophotomètre.

Tampon d'extraction d'ARN : thiocyanate de guanidium 4 M - citrate de sodium 25 mM pH 7,0 - sarcosyl 0,5% (p/v) - β -mercaptoéthanol 100 mM

5.2 Traitement ADNase « RNase free »

Après extraction, les acides nucléiques obtenus peuvent être traités à l'ADNase I d'*E. coli*. Le milieu réactionnel contenant les acides nucléiques, le **tampon ADNase 1x** et 2 U de DNase « RNase-free » (Promega) est incubé 30 min à 37°C. Une élimination complète de l'ADN peut nécessiter plusieurs traitements successifs. Après traitement, l'enzyme est extraite au phénol/chloroforme et les ARN sont précipités à l'éthanol.

Tampon ADNase 1x : Tris-HCl 40 mM pH 7,5 - MgCl₂ 6 mM - CaCl₂ 10 mM - NaCl 10 mM

5.3 Réaction de transcription inverse

Les transcriptases inverses ou ADN polymérase ARN dépendantes sont des enzymes capables de synthétiser un ADN complémentaire (ADNc) à partir d'une matrice d'ARN. Comme les autres ADN polymérase, les transcriptases inverses requièrent une amorce pour démarrer la synthèse. Ces amorces peuvent être des oligonucléotides spécifiques qui permettent la synthèse d'ADNc d'un ARN particulier. Avec des oligo-dT ce sont tous les ARNm qui seront transcrits en ADNc. Enfin il est possible d'utiliser des hexamères aléatoires qui s'hybrideront avec tous les ARN sans distinction. Après la transcription inverse, les ADNc obtenus sont amplifiés par PCR et la transcription inverse constitue ainsi la première étape de la RT-PCR (« Reverse Transcription-PCR »).

Pour la réaction, 2 μ g d'ARN sont utilisés. Ces ARN sont tous d'abord dénaturés 5 min à 65°C en présence de 20 à 100 pmol d'amorces. Après la dénaturation, le mélange est plongé immédiatement dans la glace afin d'éviter que les ARN ne se restructurent. Cette solution ensuite incubée 40 min. à 37°C, après avoir ajouté pour 20 μ L finaux de réaction :

- 10 mM de chaque dNTP
- 20 mM DTT
- 4 µL de **tampon RT 5x (Promega)**
- 10 U de MMLV RT (« Moloney Murine Leukemia Virus Reverse Transcriptase » ; Promega)

De par leur aspécificité, les hexamères aléatoires peuvent s'hybrider en plusieurs endroits et ainsi amorcer l'élongation en différents sites, créant divers fragments du même ARN. Lorsque les hexamères aléatoires sont utilisés comme amorces, il faut réaliser, après la réaction de transcription inverse, une ligation pour liguer les différents fragments obtenus à partir du même ARN entre eux de sorte à n'avoir qu'un ADNc par molécule d'ARN. Pour se faire, les extrémités 5' des différents fragments sont d'abord phosphorylées par un traitement d'un quart d'heure avec 10 unités de T4 polynucléotide kinase en présence d'ATP. Puis on ajoute 2 U de T4 ADN ligase pendant 45 min à 37°C.

Après la transcription inverse, un dixième de la réaction servira alors de matrice pour une réaction de PCR afin d'amplifier les ADNc synthétisés.

Tampon RT 5x (Promega) : Tris-HCl 250 mM pH 8,3 - KCl 150 mM - MgCl₂ 40 mM - DTT 50 mM

5.4 Transfert d'ARN sur membrane : « northern blot »

5.4.1 Transfert de l'ARN

Les ARN sont séparés sur un gel dénaturant (agarose/formaldéhyde), puis transférés par capillarité sur une membrane de Nylon Hybond N+ (Amersham) par un flux de **tampon SSC 10x** pendant la nuit. La fixation des acides nucléiques sur la membrane est effectuée par un pontage covalent aux UV (1200 Joules, "UV Stratalinker 2400" de Stratagène).

Tampon SSC 10x : citrate tri-sodique 150 mM pH 7,0 - NaCl 1,5 M

5.4.2 Hybridation de la sonde

La membrane est tout d'abord pré-hybridée dans le **tampon d'hybridation** pendant 4 heures. Après une dénaturation 5 min à 95°C suivie d'un refroidissement brutal dans la glace, la sonde (environ 5.10⁶ cpm) est ajoutée au tampon d'hybridation. L'hybridation s'effectue à 65°C pendant la nuit. Le filtre est ensuite lavé une première fois dans du tampon SSC 2x - SDS 0,1 % (p/v) puis trois fois 10 min à 65°C dans du tampon SSC 0,5x - SDS 0,1 % (p/v). Un dernier lavage est enfin effectué dans du tampon SSC 0,1x - SDS 0,1 % (p/v) et la membrane peut être mise à exposer.

Tampon d'hybridation : Phosphate de sodium 0,5 M pH 7,2 - SDS 7 % (p/v)

6 Techniques relatives aux protéines

6.1 Quantification

6.1.1 Dosage à 280 nm

L'absorption maximale des acides aminés aromatiques se situe au voisinage de 280 nm. C'est par une mesure de l'absorbance à cette longueur d'onde que l'on suit habituellement le profil protéique lors d'une étape de chromatographie. Une unité d'absorption à 280 nm pour un trajet optique de 1 cm correspond approximativement à une concentration en protéines de 1 mg/mL.

6.1.2 Méthode de Bradford

Cette méthode est basée sur la modification de la longueur d'onde d'absorption du bleu de Coomassie G-250 en milieu acide après fixation sur les acides aminés (de 460 à 595 nm) (Bradford 1976). Un mL de **réactif de Bradford** et 100 µL d'extrait protéique à quantifier sont mélangés et l'absorbance est mesurée à 595 nm. Cette mesure de DO est ensuite comparée à une gamme étalon de BSA. Il est à noter que les échantillons contenant des détergents comme le SDS ou le triton X-100 en quantités importantes ne peuvent être quantifiés par cette méthode car ces derniers interfèrent avec le réactif de Bradford.

Réactif de Bradford : Bleu de Coomassie G-250 0,01% (p/v) - éthanol 4,7% (v/v) - acide orthophosphorique 8,5% (v/v)

6.1.3 Kit RC DC protein Assay (Bio-Rad)

La méthode RC DC est principalement utilisée dans le cas de solutions de protéines contenant des détergents. Ce dosage est basé sur la réaction des protéines avec une solution alcaline de tartrate de cuivre et le réactif de Folin suivant la méthode de Lowry (Lowry *et al.* 1951). Les protéines réagissent avec le cuivre en milieu alcalin ce qui leur permet par la suite de réduire le réactif de Folin qui prend alors une coloration bleue absorbant au maximum à 750 nm. La quantification est réalisée suivant la procédure du fournisseur et par rapport à une gamme étalon de BSA effectuée en parallèle dans le même tampon que celui de l'échantillon à doser.

6.2 Techniques de concentration

6.2.1 Précipitation au TCA

Les protéines précipitent en présence d'acide trichloroacétique (TCA) 10% au froid (30 min à -20°C). Une centrifugation à 15000 g pendant 20 min permet de les sédimenter. Le culot est ensuite lavé 2 fois à l'acétone, séché puis resuspendu dans un volume adéquat. Pour bien resuspendre les protéines, il convient éventuellement de neutraliser le TCA restant. Ce type de précipitation présente l'inconvénient d'un traitement acide fort, et il est déconseillé d'utiliser les protéines ainsi concentrées pour autre chose que l'analyse sur gel de polyacrylamide.

6.2.2 Précipitation au méthanol/chloroforme

Ce type de précipitation est particulièrement indiqué dans le cas d'échantillons riches en lipides et/ou détergents. On ajoute à l'extrait protéique (un volume), 4 volumes de méthanol froid et un volume de chloroforme froid. La solution est agitée puis incubée 15 min dans la glace. Après addition de 3 volumes d'eau froide et une agitation forte au vortex de 30 secondes, les protéines sont à nouveau mises dans la glace pendant 5 min puis centrifugées 5 min à 15000 g. La solution présente deux phases : chloroforme et méthanol-eau. Les protéines se trouvent à l'interface. Après élimination de la phase supérieure (méthanol-eau), le chloroforme est éliminé par addition de 9 volumes de méthanol. Une nouvelle centrifugation de 1 min à 15000 g permet de sédimenter les protéines. Ce culot est séché à l'air puis les protéines sont resuspendues dans le volume adéquat.

6.2.3 Ultrafiltration (cellules de concentration Amicon)

Le principe de ces cellules est une ultrafiltration sélective. Ainsi les molécules de taille supérieure à la taille d'exclusion de la colonne sont retenues par la membrane hydrophile, tandis que l'eau et les petites molécules (sels par exemple) traversent cette membrane. Ces cellules sont très utilisées après certaines étapes de chromatographie qui peuvent avoir tendance à diluer l'échantillon. Dans la pratique, l'échantillon est simplement chargé dans la cellule et centrifugé à 1500 g jusqu'à la concentration désirée.

6.3 Fractionnement par électrophorèse en conditions dénaturantes

6.3.1 SDS-PAGE (SDS-Polyacrylamide gel electrophoresis)

Les protéines sont séparées par migration électrophorétique dans un gel de polyacrylamide après dénaturation par des détergents, des réducteurs et une élévation de la température. Les protéines chargées négativement par le SDS migrent en fonction de leur masse (Laemmli *et al.* 1970). Le gel est composé d'un **gel de séparation** sur lequel est coulé un **gel de concentration**. Les protéines sont reprises dans du **tampon de charge** et déposées sur le gel après dénaturation à 95°C pendant 5 min. La migration est conduite dans du **tampon Laemmli** sous un ampérage de 20 mA pendant 2 heures environ. Après la migration, le gel peut être coloré de diverses manières et séché. Le gel peut également être transféré sur une membrane de PVDF immédiatement après la migration (voir paragraphe 7.3.1). Les tailles apparentes des protéines sont estimées par rapport à un marqueur de poids moléculaire (Bio-Rad).

Tampon de charge : Tris-HCl 0,1 M pH 6,8 - glycérol 10% - SDS 2% - β -mercaptoéthanol 3% - Bleu de Bromophénol 0,1% (p/v)

Gel de concentration : polyacrylamide 5% (avec un rapport en acrylamide/bisacrylamide de 37,5 :1) - Tris-HCl 0,125 M pH 6,8 - SDS 0,1% (p/v) - APS 0,1% (p/v) - TEMED 0,05% (v/v)

Gel de séparation : polyacrylamide 5 à 15% (avec un rapport en acrylamide/bisacrylamide de 37,5:1) - Tris-HCl 0,275 M pH 8,8 - SDS 0,1% (p/v) - APS 0,1% (p/v) - TEMED 0,05% (v/v)

Tampon Laemmli : Tris-HCl 25 mM pH 8,3 - glycine 0,25 M - SDS 0,1% (p/v)

6.3.2 Gel à deux dimensions (2D)

Le gel 2D combine deux techniques de séparation des protéines : l'IEF et le SDS-PAGE. Il permet donc de séparer les protéines selon deux critères: leurs points isoélectriques tout d'abord et leurs masses ensuite. Il est ainsi possible d'obtenir un degré de séparation assez important pour un mélange complexe.

6.3.2.a Préparation des échantillons

Pour un gel, il convient d'utiliser entre 200 et 400 μ g de protéines. Ces protéines sont précipitées dans 10 volumes d'acétone à -80°C sur la nuit. Elles sont ensuite culottées par centrifugation pendant 5 minutes à 20000 g. Le culot est alors lavé avec 0,1 M d'acétate d'ammonium dans du méthanol froid (4°C). Après une nouvelle centrifugation de 5 minutes à 20000 g, le culot est lavé dans de l'acétone 80%. Suit une autre centrifugation, et il est mis à sécher à l'air pendant 30 minutes puis est repris

dans 500 µL de **tampon de réhydratation**. La dissolution du culot s'effectue en chauffant à 37°C. Pour finir, après une dernière centrifugation de 5 minutes à 20000 g, le surnageant est dosé par RC-DC (Méthodes 2.6.1.3).

Tampon de réhydratation : Urée 8 M - CHAPS 4% (p/v) - Ampholytes 0,2% (p/v) - DTT 20 mM

6.3.2.b *Première dimension*

Pour cette séparation, nous utilisons des gels à gradient de pH immobilisés (3-10 linéaire) pré-coulés (IPG strips, Bio-rad). Ces IPG-strips sont fabriqués à partir de dérivés d'acrylamide comprenant des groupements libres d'acides carboxyliques et d'amines tertiaires qui sont polymérisés avec de l'acrylamide et du Bis. Avec cette technique, le gradient ne peut plus être déplacé au cours de l'électrophorèse.

Les strips sont tout d'abord réhydratés avec l'échantillon préparé comme décrit précédemment. Pour cela, l'échantillon est déposé dans un plateau de focalisation, puis le strip est disposé par-dessus en évitant les bulles. Le tout est recouvert d'huile minérale (Bio-Rad) pour éviter l'évaporation. Le plateau de focalisation est placé dans l'appareil Bio-rad Protean IEF cell pour effectuer la réhydratation du strip à 50 V pendant 11 à 16 heures. Une fois la réhydratation achevée, des buvards humides sont placés aux extrémités du strip en vue de la focalisation. Celle-ci est conduite à 80000 Vhrs jusqu'à atteindre un voltage de 10000 V

6.3.2.c *Deuxième dimension*

Après l'IEF, les protéines sont réduites et alkylées avant la deuxième dimension. Pour cela, le strip est placé 15 minutes sous agitation dans le **tampon d'équilibration 1** puis 20 minutes dans le **tampon d'équilibration 2**. Les tampons d'équilibration sont préparés extemporanément. La deuxième dimension est ensuite réalisée dans un gel de séparation par SDS-PAGE (voir paragraphe 6.3.1). Pratiquement, après l'ionisation-réduction dans les tampons d'équilibrations, le strip est simplement placé au-dessus du gel de séparation en évitant soigneusement les bulles d'air, puis il est recouvert d'agarose 0,5% (p/v) dans du tampon Laemmli. La migration est ensuite conduite comme pour un gel SDS-PAGE.

Tampon d'équilibration 1 : Urée 6 M - SDS 2% (p/v) - Tris-Hcl 0,375 M pH 8,8 - Glycérol 20% (v/v) - DTT 130 mM

Tampon d'équilibration 2 : Urée 6 M - SDS 2% (p/v) - Tris-Hcl 0,375 M pH 8,8 - Glycérol 20% (v/v) - Iodoacétamide 135 mM - bleu de bromophénol 0,005% (p/v)

6.4 Détection

6.4.1 Coloration au bleu de Coomassie

Après migration, les gels sont colorés dans une **solution de coloration** sous agitation lente constante au moins 1 h, puis décolorés par un bain de **solution de décoloration**.

Solution de coloration : acide acétique 10% (v/v) - éthanol 25% (v/v) - bleu de Coomassie 0,125% (p/v)

Solution de décoloration : acide acétique 10% (v/v) - éthanol 20% (v/v)

6.4.2 Coloration au bleu colloïdal

La coloration au bleu colloïdal est plus sensible que la coloration au bleu de Coomassie classique : elle permet de détecter des quantités de protéines de l'ordre de 30 ng. De plus, les protéines colorées par cette technique peuvent facilement être utilisées par la suite en spectrométrie de masse.

Cette coloration est très simple à mettre en œuvre. Une fois la migration terminée, le gel est incubée dans une **solution de fixation** pendant au moins 20 minutes puis il est placé pendant une nuit dans la **solution de coloration**, avant d'être décoloré à l'eau.

Solution de fixation : Méthanol 45% (v/v) - acide acétique 1% (v/v)

Solution de coloration : Méthanol 34% (v/v) - $(\text{NH}_4)_2\text{SO}_4$ 17% (p/v) - acide acétique 0,5% (v/v) - bleu de Coomassie 0,1% (p/v)

6.4.3 Coloration au nitrate d'argent

La coloration au nitrate d'argent est plus sensible que la coloration au bleu de Coomassie et au bleu colloïdal puisqu'elle permet de détecter des protéines à partir de 0,2 ng. Elle présente cependant le désavantage d'être moins "universelle" car certaines protéines ne réagissent que très mal à ce type de coloration.

Après migration, le gel est placé une heure dans une **solution de fixation**. Il est ensuite lavé dans la **solution de lavage** pendant 30 minutes. Le gel est alors passé trois minutes dans une solution de KMnO_4 0,01% (p/v) préparé extemporanément. Il est de nouveau lavé 10 min dans la solution de lavage puis dans une solution à 10% d'éthanol et enfin juste dans de l'eau MilliQ. Le gel est ensuite baigné 15 à 20 minutes dans une solution d' AgNO_3 0,2% (p/v) réalisée avant usage. Les excès de nitrate d'argent sont enlevés par un rinçage rapide de 20 secondes dans de l'eau MilliQ puis le gel est placé en conditions alcalines dans une solution de K_2CO_3 10% (p/v) pendant une minute. La

coloration apparaît alors après 2 à 10 min passées dans une **solution de révélation**. Quand la coloration désirée est obtenue, la réaction est arrêtée dans une solution d'acide acétique à 1% (v/v). Le volume des différents bains est à adapter à la taille du gel.

Solution de fixation : CuCl_2 2% (p/v) - TCA 12% (p/v) - méthanol 50% (v/v)

Solution de lavage : éthanol 10% (v/v) - acide acétique 5% (v/v)

Solution de révélation : K_2CO_3 2% (p/v) - formaldéhyde 0.02% (v/v)

6.5 Surexpression

6.5.1 Dans *E.coli*

Les bactéries sont mises en préculture dans du milieu LB contenant 100 $\mu\text{g/mL}$ d'ampicilline et 2% (p/v) de glucose. Le lendemain, cette préculture est utilisée pour ensemercer au centième un milieu contenant de 2 à 500 mL de LB-amp (à 100 $\mu\text{g/mL}$ d'ampicilline) et 2% glucose. Cette culture est incubée à 37°C sous agitation jusqu'à obtenir une absorbance à 600 nm de l'ordre de 0,7. L'expression de la protéine recombinante est alors induite par addition d'IPTG (entre 0.1 mM et 1 mM final). Après 2 h d'induction, les bactéries sont centrifugées à 4000 g et le culot bactérien est repris dans un volume adéquat de **tampon de lyse** (environ 3 mL de tampon par gramme de poids sec), additionné de lysozyme. Pour ne pas dénaturer les protéines, le travail s'effectuera à 4°C à partir de cette étape.

Les bactéries sont ensuite lysées suivant le volume, soit par sonication pour les petits volumes, soit par deux passages successifs dans une presse de French à 1200 psi pour les grands volumes. L'extrait est alors clarifié par centrifugation pendant 20 min à 15000 g. Les protéines solubles se trouvent dans le surnageant et les protéines insolubles, les agrégats insolubles (corps d'inclusion) ainsi que les débris cellulaires culottent. Avant la purification, il convient de vérifier la solubilité de la protéine surexprimée par SDS-PAGE voire par western en utilisant les anticorps monoclonaux anti-pentahistidines (Qiagen).

Remarque : Si les protéines surexprimées sont insolubles, il est possible de les resolubiliser partiellement dans du tampon de lyse à 8 M urée.

Tampon de lyse : Tris-malate 50 mM pH 8,3 - NaCl 1 M - glycérol 10% (p/v) - Triton X-100 0,05% (v/v) - β -mercaptoéthanol 10 mM

6.5.2 Dans *Nicotiana benthamiana*

Des plantules de *N.benthamiana* âgées de trois semaines ont été utilisées comme système d'expression, dans un but purement technique, expliquant la présence d'un

végétal dans cette partie dédiée aux méthodes concernant les protéines. Le principe et les constructions utilisées sont présentées dans le chapitre II des résultats. Nous allons donc simplement détailler ici la façon de procéder pour l'infiltration. Une culture d'*Agrobacterium* ayant poussé jusqu'en fin de phase logarithmique est culottée et reprise dans un volume équivalent de $MgCl_2$; 100 μ M acétylsyringone, puis laissée immobile pendant 4 à 24 h, afin d'induire les gènes de virulence permettant le transfert de l'ADN-T. Les deux souches employées (l'une contient la construction d'intérêt, l'autre celle codant pour la protéine P19 du TBSV ("Tomato Bushy Stunt Virus"), supprimant le phénomène de « gene silencing » (Voinnet *et al.* 2003)) sont ensuite mélangées. Cette suspension bactérienne est forcée à travers les stomates, sur la face inférieure des feuilles, à l'aide d'une seringue sans aiguille. Au bout de 4 jours, les parties infiltrées sont récupérées et testées par SDS-PAGE pour l'expression de la protéine d'intérêt.

6.6 Purification des protéines surexprimées

6.6.1 Protéines exprimées dans *E.coli*

6.6.1.a Protéines en fusion avec l'étiquette de six histidines

La protéine surexprimée est purifiée par chromatographie d'affinité sur une colonne de résine Ni-NTA (Qiagen). En effet, les 6 histidines présentes à l'extrémité C-terminale de la protéine recombinante (Matériels 1.3) confèrent une affinité extrêmement forte pour les ions Ni^{2+} de la résine ($K_d=10^{-13}$). Cette purification se fait, soit dans des conditions qui maintiennent la protéine sous forme native (pour des tests d'activité par exemple), soit dans des conditions dénaturantes.

Toutes les étapes de la purification s'effectuent à 4°C. Le surnageant obtenu après clarification est incubé avec la résine Ni-NTA à 50% (v/v) préalablement lavée 3 fois avec du tampon de lyse. La quantité de résine utilisée pour la purification est ajustée en fonction du volume de bactéries cultivées et de sa teneur approximative en protéine surexprimée sachant que théoriquement 1 mL de résine peut fixer entre 5 et 10 mg de protéine. Le tout est laissé sous agitation lente 1 h à 4°C. La solution est ensuite déposée sur une colonne équipée d'un verre fritté. La matrice est alors lavée extensivement avec le tampon de lyse (15 mL pour 500 μ L de résine) dans un premier temps, puis avec le **tampon d'équilibration** (même volume). L'élution de la protéine recombinante se fait par compétition avec de l'imidazole qui est un analogue structural de l'histidine. Pour éliminer les protéines se fixant de manière aspécifique, une élution par étape est réalisée avec du **tampon d'élution** à 50 mM, 100 mM, 150 mM et 200

mM d'imidazole avec environ 2 volumes de colonne. Les fractions sont ensuite analysées par SDS-PAGE et le résultat peut-être confirmé par western en utilisant les anticorps monoclonaux anti-pentahistidines (Qiagen).

Tampon d'équilibration : Hepes 25 mM pH 7,5 - NaCl 150 mM - glycérol 5% (p/v) - β -mercaptoéthanol 10 mM

Tampon d'éluion : Hepes 25 mM pH 7,5 - NaCl 250 mM - glycérol 5% (p/v) - β -mercaptoéthanol 10 mM - imidazole de 50 à 200 mM

Les protéines resolubilisées par le tampon de lyse 8 M urée peuvent être chargées sur la colonne après une centrifugation de 20 min à 15000 g afin d'éliminer les agrégats. La purification se déroule de manière analogue à celle réalisée en conditions non dénaturantes mais avec des tampons à 8M urée.

6.6.1.b Protéines en fusion avec la GST

Le surnageant obtenu après clarification est incubé avec le volume approprié de résine « Glutathione Sepharose 4B » (Pharmacia Biotech) (laquelle comporte des molécules de glutathion fixées covalamment aux billes de sépharose), préalablement lavée avec du tampon **PBS 1X**, sous agitation lente à température ambiante pendant 30 min. La solution est appliquée sur une colonne, lavée avec 10 volumes de matrice de PBS 1X et éluée avec un volume de **tampon d'éluion**.

Tampon PBS 1X: 0,1 M NaCl; 3 mM KCl; 10 mM Na₂ PO₄, 1,8 mM KH₂PO₄ (pH 7,3)

Tampon d'éluion: 10 mM Glutathion réduit dans 50 mM Tris-HCl (pH 8,0)

6.6.2 Protéines exprimées dans *N.benthamiana*

La méthode de purification des protéines fusionnées à l'étiquette TAP (voir **Figure r2.3**) est identique pour celles exprimées dans *A.thaliana* et *N.benthamiana*. La technique a été mise au point dans la levure. Dans notre cas, la première étape consiste en un broyage à l'azote des tissus végétaux, repris ensuite dans du tampon de broyage (20 mM MOPS pH 7,5 ; 150 mM NaCl ; 1 % (v/v) Triton X-100 ; inhibiteurs de protéases Complete sans EDTA (Roche Applied Science)). Le surnageant d'une centrifugation à 20000 x g de 10 min est incubé avec des billes de Calmoduline (Stratagene) pendant 1 h à 4 °C en présence de 2 mM CaCl₂. Les billes sont lavées trois fois dans du tampon de lavage (20 mM MOPS pH 7,5 ; 150 mM NaCl ; 0,2 % Tween 20) contenant 2 mM CaCl₂, puis les protéines en sont éluées avec du tampon de lavage additionné de 5 mM d' EGTA. L'éluat est ensuite incubé avec des billes couplées aux anticorps FLAG M2 (Sigma) 30 min à 25 °C pour la deuxième étape de purification. Les billes sont lavées trois fois dans du tampon de lavage et les protéines

éluées dans du tampon d'élution (10 mM MOPS pH 7,5 ; 50 mM NaCl ; 100 g/mL peptide FLAG).

6.7 Purification des corps d'inclusion

Ce sont les protéines exprimées dans la souche *E.coli* BL161 en fusion avec l'étiquette de six histidines qui ont été utilisées dans le cadre de cette méthode. Le culot bactérien est repris dans du **tampon de lyse** à raison de 4 mL par g de cellules. Celles-ci sont resuspendues à l'aide d'un « Waring blender », lysées par deux passages à la presse de French et le mélange est ensuite homogénéisé au moyen d'un sonicateur. La suspension est clarifiée par une centrifugation d'1 h à 22000 x g à 4 °C, et le culot repris dans du **tampon de lavage** (5 mL par g de cellules initialement présentes). Une autre centrifugation a alors lieu, de 30 min à 22000 x g et 4°C. Le surnageant est éliminé et le culot repris dans du tampon de lavage. Cette étape est encore répétée deux fois, puis le culot final est repris dans le tampon de lavage sans urée ni Triton. Ce culot est ensuite conservé à – 80 °C ou traité directement en vue de la solubilisation des protéines qu'il contient.

6.8 Couplage de protéines à une matrice de Sépharose-CNBr activé (système « CNBr-activated Sepharose 4B » ; Amersham biosciences)

Ce système permet le couplage covalent d'un ligand à une matrice inerte de Sépharose. La matrice de Sépharose 4B établit des liaisons covalentes avec les amines primaires du ligand via l'activation du CnBr. Les groupements activés de la colonne qui n'ont pas réagi avec le ligand sont ensuite bloqués par un excès de tampon qui contient des amines primaires de petite taille (glycine, éthanolamine ou tampon Tris par exemple). Le ligand peut être une protéine (ce qui est par exemple le cas lorsque l'on purifie des anticorps) ou des acides nucléiques (cas de la caractérisation de protéines ayant une affinité pour les acides nucléiques).

Les protéines sont d'abord dialysées dans des boudins de dialyse contre du **tampon de couplage** pendant 2x2 heures. Pendant ce temps, la matrice est activée et lavée trois fois avec 1 mM d'HCl pendant 15 minutes sous agitation. Après la dialyse, les protéines sont couplées à la matrice pendant une heure à température ambiante ou sur la nuit à 4°. Suite au couplage, la matrice est culottée à 12000 g pendant 30 secondes puis est lavée trois fois avec du tampon de couplage. Les groupes actifs restés libres sont ensuite bloqués avec des amines primaires (Tris-HCl 0,1 M pH 8 - NaCl 0,5 M) pendant 2 heures à température ambiante. Enfin la colonne est lavée avec du **tampon S1** puis du **tampon S2** afin d'éliminer l'excédent de protéines non couplées

Tampon de couplage : NaHCO₃ 0,1 M pH 8,3 – NaCl 0,5 M

Tampon S1 : NaAc 0,1 M pH 4 – NaCl 0,5 M

Tampon S2 : Tris 0,1 M pH 8 – NaCl 0,5 M

6.9 Séquençage N-terminal des protéines

L'IBMP disposait d'un service de séquençage par dégradation d'Edman qui utilisait le système "473A Peptide sequencer" (Applied Biosystems). Ce système permet la détermination de la séquence N-terminale à partir d'un seuil de 50 pmoles de protéine purifiée (responsable: M. Le Ret, IBMP). Après la dégradation récurrente de l'extrémité N-terminale de la protéine selon la méthode d'Edman, la nature des acides aminés est ensuite déterminée par HPLC ("High Pressure Liquid Chromatography"). La procédure employée pour préparer un échantillon à séquencer est la suivante : les protéines sont séparées par SDS-PAGE puis transférées comme décrit pour le western blot (voir paragraphe 7.3.1). La membrane est colorée au rouge Ponceau et décolorée à l'eau. La bande qui correspond à la protéine à séquencer est découpée (maximum 6 mm²) puis lavée au méthanol et abondamment rincée à l'eau. Le fragment de membrane est donné tel quel au service de séquençage.

7 Techniques d'immunologie

7.1 Production d'anticorps polyclonaux

Les anticorps sont produits contre une protéine recombinante surexprimée dans *E. coli*, purifiée par SDS-PAGE et transférée sur membrane de nitrocellulose. Après dissolution de la membrane dans du DMSO et addition d'adjuvant complet de Freund (1 vol.) la solution qui renferme quelques dizaines de mg de protéine est injectée à un lapin pour la production d'anticorps. Les injections suivantes (3 à 5) sont réalisées toutes les deux semaines dans les mêmes conditions mais avec de l'adjuvant incomplet. Les prélèvements du sang contenant les anticorps polyclonaux sont effectués une semaine après chaque injection.

7.2 Purification d'anticorps

Les sérums de lapin obtenus à l'animalerie reconnaissent souvent plusieurs protéines. Pour obtenir des anticorps spécifiques, ils nécessitent d'être purifiés contre la protéine d'intérêt couplée à du Sépharose. Le principe de la purification est basé sur la diminution de l'affinité des anticorps pour la protéine à des pH extrêmes.

Pour ce faire, 2 mL de sérum sont d'abord centrifugés 10 minutes à 10000g. Le surnageant est ensuite passé lentement sur une colonne contenant la protéine préalablement couplée à du Sépharose (voir paragraphe 6.7). La colonne est lavée avec 20 mL de **tampon L1**. Les anticorps retenus sont ensuite élués en conditions acides avec 3 mL de **tampon E1**. L'éluat est récupéré dans 1M Tris pH8 à raison de 200 µL par mL de tampon E1. Suite à l'éluion, le pH de la colonne est équilibré avec 10 mL de **tampon L2**. Après cela, les anticorps sont élués en conditions basiques avec 3 mL de **tampon E2** et récupérés dans 200 µL de Tris HCl 1M pH 7,5. La colonne est ensuite lavée avec 10 mL de tampon L1 pour rétablir un pH de 7,5 dans la colonne.

Il arrive qu'au cours des éluions à pH extrêmes, des protéines soient relarguées de la colonne. Pour éviter cela, avant de réaliser la purification, un premier cycle de purification est réalisée sur la colonne sans anticorps. Les protéines mal fixées et susceptibles d'être relarguées vont ainsi être éliminées et ne se retrouveront pas éluées avec les anticorps.

Tampon L1 : Tris 10 mM pH 7,5 - NaCl 0,3 M

Tampon E1 : Glycine 0,1 M pH 2,5 - NaCl 0,3 M

Tampon L2 : Tris 10 mM pH 8,8 - NaCl 0,3 M

Tampon E2 : Triéthylamine 0,1 M pH 11,5 - NaCl 0,3 M

7.3 Western blot

7.3.1 Transfert des protéines sur membrane

Après séparation sur un gel de polyacrylamide, les protéines sont électrotransférées en milieu semi-sec (**tampon de transfert**) sur membrane de PVDF (Immobilon-P; 0,45 µm; Millipore). Celle-ci est précédemment activée au méthanol, puis rincée à l'eau et enfin au tampon de transfert. Le transfert est réalisé pendant 45 minutes sous 1mA par cm² dans un appareil de transfert Scie-plas. La membrane est ensuite colorée dans une **solution de coloration** puis décolorée au méthanol 50 %, acide acétique 7 % pour révéler le marqueur de poids moléculaire et les protéines. La membrane est finalement complètement décolorée au méthanol. La coloration peut également se faire au rouge Ponceau S (décoloration à l'eau).

Tampon de transfert : Tris 48 mM - glycine 39 mM

Solution de coloration : bleu de Coomassie R250 0,1% - méthanol 40% - acide acétique 10%

7.3.2 Fixation des anticorps

Suite à la décoloration, la membrane est équilibrée et les sites aspécifiques de fixation saturés dans un **tampon de blocage**, pendant 1 heure à température ambiante et sous agitation. La membrane est alors incubée en présence de l'anticorps primaire spécifique de la protéine d'intérêt (dilué de 500 à 50000 fois dans du **tampon anticorps**), pendant une heure à température ambiante (ou pendant la nuit à 4°C). Puis la membrane est rincée 3 fois au **TBS-T** pendant 15 min pour éliminer l'excédent d'anticorps primaire. Elle est ensuite mise en présence de l'anticorps secondaire (anti-lapin ou anti-souris suivant la provenance des anticorps primaires) couplé à la peroxydase (diluée 10000 fois dans du tampon anticorps) pendant une heure à température ambiante. Enfin la membrane est lavée 3 fois 15 min au TBS-T. Les anticorps sont alors prêts à être révélés.

TBS-T : Tris-HCl pH 7.4 20 mM - NaCl 150 mM - Tween 20 0,2% (p/v)

Tampon de blocage : TBS-T - lait écrémé en poudre 5% (p/v)

Tampon anticorps : TBS-T - lait écrémé en poudre 2,5% (p/v)

7.3.3 Révélation

La révélation des anticorps est faite par ECL (« Enhanced Chemical Luminescence »). Elle consiste en une réaction catalysée par la peroxydase de Raifort couplée aux anticorps secondaires. Cette réaction entraîne la production de lumière qui permet de visualiser le lieu de fixation des anticorps secondaires. La révélation se fait selon les indications du fournisseur (Amersham). L'émission de lumière est détectée par une exposition de la membrane sur un film photographique. Les anticorps peuvent être décrochés par incubation 30 min à 50°C dans une solution contenant 62,5 mM de Tris-HCl pH 6,7, 2% de SDS (p/v) et 100 mM de β -mercaptoéthanol.

7.4 Immunoprécipitation

L'extrait de protéines mitochondriales est préparé à partir d'1 à 2 mg de mitochondries. Celles-ci sont diluées dans 500 μ L de TBS-T puis lysées par 3 cycles de congélation-décongélation rapide dans l'azote liquide. Le lysat est ensuite centrifugé à 20000 g et le surnageant récupéré. Cette opération est répétée deux fois, pour éliminer les débris membranaires et les protéines insolubles. L'extrait protéique obtenu est ensuite mis en présence des anticorps primaires sous agitation (les conditions de température, de temps et la quantité sont à définir pour chaque anticorps). Cent microlitres de TBS-T - protéine A couplée à la sépharose 5% (p/v) (Protéine A – Sépharose CL4B ; Sigma) sont alors ajoutés. Le tout est incubé pendant 15 minutes à température ambiante sous agitation ou 4 heures à 4°C. Les billes de

Sépharose sont ensuite précipitées par centrifugation (5 secondes à 4000 g), le surnageant est éliminé puis les billes sont lavées trois fois avec 1 mL de TBS-T. Les anticorps sont finalement décrochés de la protéine A avec 50 µL de **tampon d'élution**.

Tampon d'élution : glycine 100 mM pH 2,5 - NaCl 50 mM

8 Techniques relatives au matériel végétal

8.1 Transformation d'*Arabidopsis* par *Agrobacterium*

La transformation des *Arabidopsis* s'effectue sur les jeunes fleurs et pour améliorer le rendement de transformation, les premières hampes sont coupées en dessous du premier nœud ce qui va multiplier le nombre de hampes florales par plante.

Lorsque les plantes sont au bon stade de développement pour la transformation, une préculture de 3 mL d'agrobactéries contenant le plasmide avec la séquence d'intérêt dans du milieu **LB - antibiotiques** est mise à incuber sur la nuit à 28°C. Le lendemain, 0,7 mL de cette préculture servent à ensemercer 500 mL du même milieu. Les agrobactéries sont cultivées dans les mêmes conditions et lorsque la DO de la culture atteint 0,8, les cellules sont culottées par centrifugation pendant 20 min à 5000 g. Le culot est resuspendu dans 10 mL de MgCl₂ 10mM. Les agrobactéries sont ensuite diluées dans 500 mL de saccharose 5% (p/v) et Silwet 0,02% (p/v) pour retrouver une DO de 0,8. Les hampes florales d'*Arabidopsis* sont ensuite plongées rapidement dans ce milieu et les plantes sont maintenues en conditions très humides pendant une journée sous un dôme. Le taux d'humidité est réduit progressivement avant qu'elles ne retournent aux serres. Les *Arabidopsis* sont alors cultivées jusqu'à leur complète dessiccation et les graines sont récoltées. Les graines transformées sont ensuite sélectionnées *in vitro* sur un milieu contenant l'antibiotique de sélection (kanamycine, hygromycine). Pour cela, les graines sont placées tout d'abord à 4°C et à l'obscurité pendant plusieurs jours afin de lever la dormance. Les graines sont ensuite stérilisées en milieu gazeux chloré pendant 12 heures. Ce milieu est réalisé dans une cloche fermée dans laquelle, à un récipient renfermant 100 mL d'eau de Javel car(5% de chlore actif), ont été rapidement ajoutés 3 mL d'acide chlorhydrique concentré. Après élimination de l'atmosphère chlorée à la trompe à eau, les graines sont laissées 30 min sous une hotte à flux laminaire pour éliminer le chlore résiduel. Les graines sont reprises dans 8 mL de **top agar** et sont ainsi étalées sur boîtes contenant du **milieu de culture *Arabidopsis***.

Milieu LB – antibiotiques : LB - kanamycine 50 µg/mL – gentamicine 10 µg/mL – rifampicine 50 µg/mL

Top agar : agarose 0,15% (p/v)

Milieu de culture *Arabidopsis* : MS255 pH 5,6 (titré au KOH) – saccharose 0,5% (p/v) agarose 0,9% (p/v) (Difco) – kanamycine 100 µg/mL (de la vancomycine (500 µg/mL) ou de la carbenicilline (20 µg/mL) sont également ajoutées pour éliminer les agrobactéries encore présentes dans les graines).

8.2 Croisement entre deux lignées d'*Arabidopsis*

Les *Arabidopsis* sont des plantes autogames. La fécondation a lieu dans le bouton floral avant son ouverture. Pour réaliser un croisement entre deux lignées d'*Arabidopsis*, il faut donc polliniser la lignée femelle avec le pollen de la lignée mâle avant l'ouverture des boutons floraux. En pratique, lorsque les *Arabidopsis* de deux lignées à croiser arrivent en fleurs, les fleurs de la lignée femelle sont éliminées et seuls quelques boutons floraux encore fermés sont conservés. Ceci permet, d'une part, de ne pas mélanger par la suite les graines provenant du croisement des graines sauvages, mais aussi de favoriser les boutons floraux conservés. Ces boutons floraux vont être disséqués de sorte à éliminer les étamines et ne laisser que les pistils. Les pistils mis à nu sont induits par contact avec des étamines de la lignée mâle. Les siliques des pistils fécondés sont prélevés à leur maturité pour récupérer les graines.

8.3 Transfection par bombardement

Le bombardement consiste à projeter sur les cellules des particules de tungstène sur lesquelles l'ADN est précipité.

8.3.1 Préparation des particules

Pour 5 bombardements, 6 µg d'ADN plasmidique sont additionnés de 40 µL d'une suspension de billes de tungstène (diamètre moyen de 1 µm) à 30 mg/mL dans du glycérol à 50 % (v/v). L'ensemble est agité énergiquement (vortex) pour une bonne homogénéisation. Puis 40 µL de CaCl₂ 2,5 M et 16 µL de spermidine 0,1 M sont ajoutés à la suspension qui est mise ensuite sous forte agitation pendant 30 minutes à 4°C. Les particules sont enfin lavées 3 fois avec 200 µL d'éthanol absolu froid et reprises dans 30 µL d'éthanol.

8.3.2 Bombardement

La boîte contenant la feuille de *Nicotiana benthamiana* est placée à 11 cm du canon dans une chambre sous vide (0,9 bar). Six microlitres de particules préparées sont déposés dans la seringue et projetés par une libération d'hélium à 7 bars pendant

25 ms. Les boîtes sont ensuite remises dans leurs conditions de cultures originales. Cinq heures après le bombardement, les feuilles sont transférés sur milieu solide isotonique. Le lendemain, les zones susceptibles de contenir des cellules transformées sont recherchées sous la loupe. Ces zones sont découpées, posées sur lame, face supérieure de la feuille côté lame. Une goutte d'eau est déposée sur la face inférieure et la feuille est grattée délicatement avec la lamelle pour percer la cuticule et faciliter la pénétration de l'eau dans les tissus.

8.4 Observation au microscope

Pour pouvoir comparer la localisation des mitochondries à celle des protéines d'intérêt, les cellules ou les feuilles ont été co-transfectées avec le plasmide pCK-mRFP (« mitochondrial Red Fluorescent Protein »), qui a été réalisé au laboratoire à partir du plasmide pCK-GFP3 (Menand *et al.* 1998). Il permet l'expression de la protéine RFP (« Red Fluorescent Protein »), fusionnée à la préséquence d'adressage mitochondriale de la protéine COXIV (sous unité IV de la cytochrome c oxydase) de levure. La RFP est codée par le gène *DsRED1* de *Discosoma sp* (ClonTech). Elle présente les mêmes avantages que la GFP, mais ses spectres d'excitation et d'émission sont décalés et son émission maximale est dans le rouge. Pour la co-transfection, le plasmide pCK-mRFP est mélangé au plasmide codant la protéine de fusion à l'eGFP avant la préparation des billes de tungstène en vue du bombardement.

Le microscope confocal utilisé est un microscope inversé à balayage laser Zeiss LSM 510. La lumière d'excitation va exciter les fluorochromes de l'échantillon qui vont renvoyer la lumière à une longueur d'onde différente (longueur d'onde d'émission). Le microscope possède différents jeux de laser, filtres d'excitation, et miroirs dichroïques qui permettent d'observer des fluorochromes de diverses natures. Les différentes longueurs d'ondes d'excitation et d'émission des fluorochromes utilisés ou observés sont données ci-dessous.

<i>Fluorochromes</i>	<i>Excitation maximale</i>	<i>Emission maximale</i>
eGFP	489 nm	508 nm
RFP	558 nm	583 nm
Chlorophylle	450 nm	660 nm

8.5 Analyse des fusions-GUS

8.5.1 Préparation des échantillons

Les pièces végétales à examiner sont isolées et immergées 10 à 15 min dans de l'acétone (90%). Après élimination de l'acétone, les pièces sont rincées dans du tampon 50 mM phosphate pH 7,2 et plongées dans la **solution de coloration GUS**. Un vide (environ 600 mm Hg pendant 2 x 15 min ou plus si les tissus flottent dans le liquide) est appliqué pour infiltrer la solution dans les tissus. La coloration se développe 12 heures à 37°C et à l'obscurité. La chlorophylle est ensuite enlevée par une série de rincages à l'éthanol: 10 min 50% éthanol, 10 min 70%, et 95% jusqu'à décoloration complète. Les tissus sont ensuite conservés dans de l'éthanol 70% jusqu'à observation.

Solution de coloration GUS: 50 mM phosphate pH 7,2; EDTA 10 mM; Triton X100 0,1%; K₃Fe(CN)₆ 1 mM; K₄Fe(CN)₆ 1 mM; X-gluc 0,4 mg/mL.

8.5.2 Réalisation de coupes pour l'observation au microscope de l'expression GUS

Les tissus sont fixés la nuit par immersion dans du **FAA** avec une infiltration sous vide de 2 x 15 min. Ils sont ensuite lavés 30 min dans de l'éthanol 50%, 30 min dans de l'éthanol 70% puis colorés la nuit à 4°C dans de l'éthanol 95% renfermant de l'éosine (0,1%). Les échantillons sont rincés 2 x 30 min dans de l'éthanol 95% puis placés 2 x 1 heure dans de l'éthanol 100%, 15 min dans un mélange Histoclear/éthanol 100% (1 vol/1 vol), 3 x 20 min dans de l'Histoclear, 2 heures dans un mélange Histoclear/Paraplast (1 vol/1 vol) à 60°C et enfin 4 jours dans du Paraplast à 60°C avec changement 2 fois par jour. Des coupes de 10 microns sont réalisées dans les tissus inclus dans le Paraplast à l'aide d'un microtome Leica et déposées sur les lames d'observations puis incubées une nuit à 40°C, ce qui permet la fixation sur le verre traité. Les échantillons sont déparafinés par 2 x 25 min dans du xylène, puis recouverts d'Eukitt qui permet de sceller une lamelle sur la préparation. Après séchage, les lames sont observées à l'aide d'un microscope Nikon E 800.

FAA: Ethanol 50%, formaldéhyde 10% (v/v), acide acétique glacial 5% (v/v).

8.6 Etablissement de cultures de cellules transgéniques

Le protocole que nous avons utilisé pour tenter d'établir des cultures cellulaires transgéniques d'*Arabidopsis* est adapté de ((Mathur J et Koncz C 1998)).

8.6.1 A partir de graines

Des aliquotes d'environ 1000 graines sont stérilisés selon la méthode décrite dans le paragraphe 8.1. Les graines sont ensuite mises à germer dans des boîtes contenant du milieu MS (Sigma, M-5524) 0,5X ; 0,8 % agar ; 50 µg/mL kanamycine à 25°C, avec 16 h de jour et 8 h de nuit. Au bout de 15 à 20 jours, les plantules sont placées dans 10 mL de milieu liquide MS avec 3 % de sucrose, et agitées à 120 rpm, 25°C, et la même photopériode. Au bout de deux semaines, les racines sont récupérées en éliminant tous les tissus verts, pour obtenir des morceaux d'environ 2 mm. Ceux-ci sont ensuite placés dans 15 mL de **milieu de suspension**, et agitées à 140 rpm à l'obscurité. Au bout de trois semaines, les cellules obtenues sont filtrées à travers un tamis (850 ou 350 µm), décantées pendant 30 min pour éliminer le maximum de milieu et remises en suspension dans du milieu frais. Les cellules sont par la suite repiquées toutes les deux semaines, dans l'attente d'obtenir une densité utilisable. On peut alternativement placer les graines directement dans du **milieu de suspension**, ce qui a pour effet de générer des cellules isolées sans passer par l'étape précédemment décrite. On rafraîchit alors le milieu une première fois lorsque les cellules ont atteint une densité suffisante, ce qui prend environ un mois. Cette deuxième méthode s'est dans notre cas avérée tout aussi efficace que la première. Nous n'avons jamais observé la synthèse de nos protéines d'intérêt, même si la construction était visible tout au long de la culture.

Milieu de suspension (pH5,8, pour un litre): 4,33g de milieu MS ; 2 X Vitamines B5 (réf Gamborg) ; 30 g sucrose ; 0,5 mg acide 2,4-dichlorophénoxyacétique (2,4-D) ; 2,0 mg acide indole-3-acétique (IAA) ; 0,5 mg 6(~,~)dimethylallylamino)-purine riboside (IPAR)

8.6.2 A partir de cals

Des fragments d'environ 2mm provenant de racines, de tiges et de feuilles ont été placés sur un milieu d'induction de cals (milieu de Gamborg B5, glucose 2 %, agar 0,8 %, MES 0,5 g/L, 2,4-D 0,5 mg/L, kinétine 0,05 mg/L). Quel que soit le tissu d'origine des cals, là encore, les protéines n'ont jamais pu être détectées, en dépit de la présence du transgène.

9 Autres méthodes

9.1 Extraction simultanée d'ARN, d'ADN et de protéines au trizol

Le matériel végétal est tout d'abord broyé dans l'azote liquide. Pour un volume de broyat, trois volumes de trizol sont ensuite ajoutés. La solution est homogénéisée et laissée cinq minutes à température ambiante pour bien permettre la dissociation des complexes nucléoprotéiques. 0,8 volume de chloroforme est alors additionné dans la solution. Après mélange, le tout est laissé 5 minutes à température ambiante puis centrifugé 15 minutes à 12000 g. La centrifugation permet la séparation de la phase aqueuse contenant les ARN de la phase organique avec à l'interface une couche de débris végétaux. L'interface et la phase organique contiennent les protéines et l'ADN.

Trizol: Phénol saturé (H₂O) 38% (v/v); guanidine thiocyanate 0,8 M; ammonium thiocyanate 0,4 M; sodium acétate pH 5 0,1M; glycerol 5%

9.1.1 Isolation des ARN

La phase aqueuse est récupérée et les ARN sont précipités par addition de 2 volumes d'isopropanol (par rapport au volume de broyat de départ). La précipitation se fait durant 10 minutes à température ambiante puis les ARN sont culottés par centrifugation pendant 8 minutes à 12000 g. Le culot est ensuite lavé dans de l'éthanol à 75%. Après le lavage, le culot est repris dans 400 µL d'eau Milli-Q stérile pour être encore purifié par séparation au phénol/chloroforme suivie d'une précipitation éthanolique.

9.1.2 Isolation de l'ADN

Après prélèvement de la phase aqueuse, 1,2 volume d'éthanol 100% (par rapport au volume de broyat de départ) sont ajoutés dans l'interface et la phase organique, afin de précipiter l'ADN. La précipitation est réalisée pendant 2 minutes à température ambiante et l'ADN est sédimenté pendant 5 minutes à 2000 g. Après la centrifugation, le surnageant est récupéré pour en extraire les protéines. Le culot d'ADN est lavé deux fois dans 4 volumes de citrate de sodium 0,1 M, éthanol 10% pendant 30 minutes à température ambiante puis centrifugé à 2000 g pendant 5 minutes. L'ADN est alors lavé dans 6 volumes d'éthanol 75% pendant 20 minutes à température ambiante. Après centrifugation 5 minutes à 2000 g le culot est séché à l'air 5 minutes puis est repris dans 8 mM de NaOH.

9.1.3 Isolation des protéines

Les protéines contenues dans le surnageant récupéré au cours de l'isolation de l'ADN sont précipitées avec 6 volumes d'isopropanol (par rapport au volume de broyat de départ). La précipitation s'effectue pendant 10 minutes puis les protéines sont culottées par centrifugation pendant 10 minutes à 12000 g. Le culot est ensuite lavé 3 fois dans 6 volumes d'hydrochloride de guanidine 0,3 M - 95 % d'éthanol. Chaque lavage s'effectue pendant 20 minutes et est suivi par une centrifugation de 5 minutes à 7500 g. Finalement le culot est lavé une dernière fois dans 6 volumes d'éthanol 100 %. Suite au lavage, nous effectuons une centrifugation 5 minutes à 7500 g et le culot protéique est resuspendu dans 10 M urée, 50 mM DTT. Si les protéines se resuspendent mal, il convient de bouillir l'échantillon 3 minutes.

9.2 Test d'importation *in vitro* de protéines dans les mitochondries

9.2.1 Transcription et traduction *in vitro*

Le précurseur de la protéine dont nous souhaitons étudier l'importation dans les mitochondries est synthétisé par une transcription et une traduction couplées dans un lysat de réticulocytes de lapin en présence de méthionine marquée au ³⁵S. Nous utilisons le système "TNT Coupled Reticulocytes Lysate System" de Promega en respectant les recommandations du fabricant. En général, la réaction est faite avec 1 µg d'un plasmide contenant le promoteur d'une ARN polymérase (généralement T3 ou T7) et l'ADNc complet de la protéine d'intérêt.

9.2.2 Test d'importation *in vitro* dans les mitochondries

Le test d'importation *in vitro* de protéines dans les mitochondries est réalisé selon le protocole décrit pour RPS10 (Wischmann et Schuster 1995). Après leur purification, les mitochondries de pomme de terre (200 µg par expérience) sont diluées dans du **tampon d'importation** en présence de 4 mM d'ADP. Les protéines radiomarquées sont alors ajoutées et la solution est incubée 30 min à 25°C sous agitation. Pendant ce laps de temps, les protéines vont être importées dans les mitochondries. Le milieu réactionnel est ensuite déposé sur un **coussin de saccharose à 27%** (1 mL). Le coussin permet de séparer le milieu réactionnel des mitochondries par centrifugation (10 min à 12000 g). Le surnageant est éliminé et l'on ne conserve que le culot de mitochondries. Les mitochondries sont reprises dans du tampon de charge pour protéines, bouillies 5 minutes et résolues par SDS-PAGE. Le gel est ensuite séché et mis à l'exposé au contact d'un film photosensible.

A côté de cette expérience d'importation, différents contrôles sont réalisés en parallèle, entre autres pour vérifier que ce transport est actif. Le premier contrôle est un traitement à la protéinase K (10 µg/mL inhibée ensuite au PMSF 1 mM). Elle est ajoutée après l'incubation des protéines marquées avec les mitochondries. Dans ce contrôle, les protéines importées sont protégées par les membranes mitochondriales et doivent donc être résistantes au traitement. La vérification du caractère actif ou passif de l'import se fait en additionnant de la valinomycine (0,25 µM) aux mitochondries avant incubation avec les protéines marquées. Ce traitement a pour effet de neutraliser le potentiel de membrane. Dans le troisième contrôle, les mitochondries sont mises en présence de triton X-100 (0,5%) après l'importation, puis un traitement à la protéinase K est effectué. Dans ces conditions, les membranes mitochondriales doivent être dissoutes et toutes les protéines y compris les mitochondriales doivent être dégradées. Cette expérience permet de s'assurer que la protéine n'est pas résistante à la protéinase K et ainsi qu'elle est bien importée dans les mitochondries.

Tampon d'importation : mannitol 250 mM - Hepes 20 mM pH 7,5 - KCl 80 mM - K₂HPO₄ 1 mM - ATP 1 mM - malate 1 mM - DTT 1 mM - NADH 2 mM

Coussin de saccharose 27% : Tris HCl 20 mM pH 7,5 - saccharose 27% - EDTA 1 mM pH 8 - K₂HPO₄ 100 mM - BSA 1 mg/mL

9.3 Southwestern

Le but de cette technique est de révéler les protéines pouvant lier l'ADN. Après séparation par SDS-PAGE, l'échantillon est transféré sur une membrane de PVDF (voir paragraphe 7.3.1). Etant de ce fait sous forme dénaturée, il est nécessaire de permettre aux protéines de retrouver une conformation dans laquelle le domaine de liaison aux acides nucléiques est fonctionnel. Ceci est réalisé par dilution de l'agent dénaturant dans du **tampon de renaturation** (3 bains de 30 min puis une incubation d'une nuit, le tout à 4°C). Ensuite, après un bain de 15 min dans du **tampon de blocage**, les protéines sont mises en présence d'une sonde ADN radiomarquée (10000 cpm) dans du **tampon d'incubation**, durant 2 h à température ambiante. On réalise enfin 4 lavages de 10 min dans du **tampon de lavage**. L'émission des rayons β est détectée par autoradiographie ou quantifiée au Phosphorimager, à l'aide du logiciel MacBas).

Tampon de renaturation : Tris-HCl (pH 7,5) 100 mM ; Nondiet P40 0,1 %

Tampon de blocage : BSA 5 % ; Tris-HCl (pH 7,5) 10 mM ; Acétate de Mg 5 mM ; DTT 2 mM ; Triton X-100 0,01 %

Tampon d'incubation : Tris-HCl (pH 7,5) 10 mM ; Acétate de Mg 5 mM ; DTT 2 mM ; Triton X-100 0,01 %

Tampon de lavage : Tris-HCl (pH 7,5) 10 mM ; Acétate de Mg 5 mM ; DTT 2 mM

9.4 Le gel retard ou EMSA (« Electromobility Shift Assay »)

La première étape consiste à soigneusement nettoyer (avec un détergent) et sécher (à l'éthanol 95 %) les plaques. Les concentrations en acrylamide utilisées se sont situées entre 5,5 (sonde de 70 nt) et 9 % (sonde de 25 nt). Par exemple, pour un volume de 25 mL, la composition d'un gel à 7 % est la suivante : 2,5 mL de Tris-glycine 10X ; 3,75 mL d'acrylamide (39 : 1) à 40 % ; 300 µL d'APS 10 % ; 30 µL de TEMED ; H₂O qsp. 25 mL. Une pré migration à 4 °C de 30 min (8-10 V/cm) permet d'assurer au gel une température constante lorsqu'il contiendra l'échantillon. Les incubations sonde-protéine ont été réalisées dans des volumes de 16 µL, dont la composition finale typique (après avoir testé différents paramètres) était la suivante : 30 mM Tris-HCl pH (7,5) ; 10 % glycérol ; 1 mM β-mercaptoéthanol ; 80 mM NaCl ; 2mM MgCl₂ ; inhibiteurs de protéases (Complete sans EDTA; Roche Applied Science) ; 500 à 30000 cpm de sonde et 1-5 ng de protéine. Les temps d'incubations testés sont allés de 2 min à température ambiante à 2 h à 4 °C, sans que l'on constate de différence notable. Les gels sont chargés en prenant soin d'ajouter un témoin de migration contenant du bleu de bromophénol. La migration, à 120 V, dure environ 3 h pour une sonde de 70 nt. Enfin, le gel est démoulé, déposé sur une feuille de papier Wattmann humide et séché à 80 °C pendant 40 min. L'émission des rayons β est détectée par autoradiographie ou quantifiée au Phosphorimager, à l'aide du logiciel MacBas).

9.5 Test de recombinaison *in vitro*

Le but de cette expérience est de tester l'influence d'une protéine d'intérêt (dans notre cas d'*At*-SBP40b) sur la capacité de la protéine RecA de *E.coli* à initier la formation de la boucle D, première étape du processus de recombinaison homologue.

9.5.1 Synthèse d'ADNsb marqué

Dans un premier temps, l'oligonucléotide R1 (voir ANNEXE) a été phosphorylé à l'aide d'ATP non radioactif et de la T4 polynucléotide kinase (voir paragraphe 4.6 de la partie Matériels et Méthodes). Une PCR a ensuite été réalisée à l'aide des amorces R2 (non phosphorylée), R1 (phosphorylée) et du plasmide pUCAP35S en guise de matrice, donnant lieu à un fragment d'ADNdb de 500 pb. Ce fragment (200 ng) a enfin été digéré durant 2 heures à 37°C à l'aide de 10 U d'exonucléase λ (qui digère le brin phosphorylé dans le sens

5'→3'), dans du **tampon de digestion λ**, afin d'obtenir une séquence d'ADNsb de 500 pb complémentaire à la partie amplifiée du plasmide pUCAP35S. Ce dernier fragment est par la suite marqué à l'aide de la T4 polynucléotide kinase et d'ATP radioactif pour obtenir la sonde utilisée lors du test de recombinaison.

Tampon de digestion λ : 67 mM glycine NaOH pH 9,4 ; 2,5 mM MgCl₂ ; 50 µg/mL BSA

9.5.2 Hybridation

Des quantités constantes d'ADNsb marqué (2 ng) et d'ADNdb froid (pUCAP35S ; 100 ng) ainsi que des quantités variables de RecA de *E.coli* (Sigma ; de 0 à 3 µg) et de la protéine à tester (AtSBP40b ; de 0 à 3 ng) ont été incubées en présence ou en absence d'ATP (2,5 mM) pendant 2 heures à 37 °C dans du **tampon de recombinaison**. Les protéines sont par la suite éliminées par l'ajout de SDS (1 % P/V final) et d'EDTA (10 mM) et les acides nucléiques sont séparés sur gel d'agarose 1 %. A l'issue de la migration, le gel est séché et les résultats analysés par autoradiographie.

Tampon de recombinaison : 10 mM acétate de magnésium ; 25 mM Tris acétate pH 7,5 ; 1 mM DTT ; 5 % glycérol.

RESULTATS

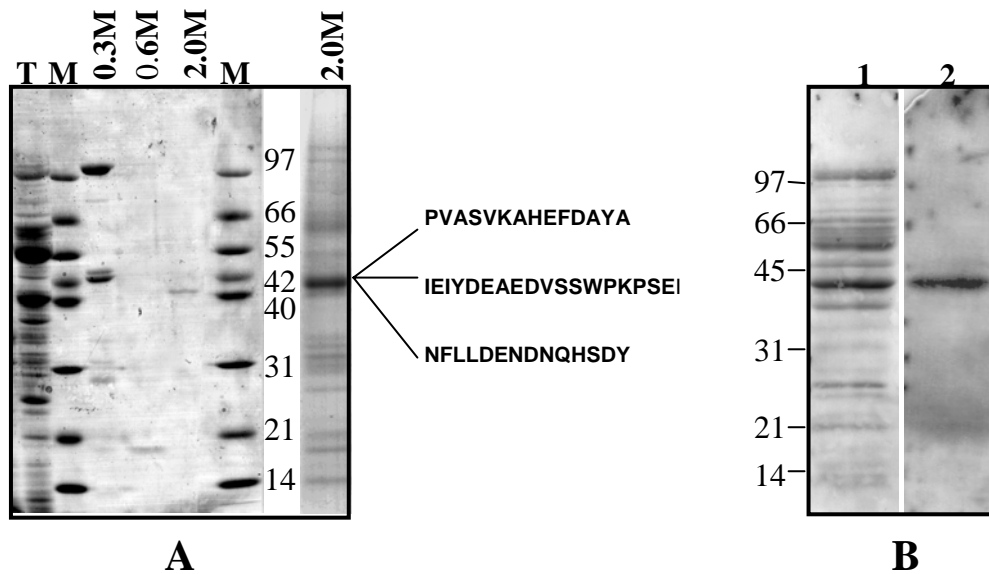


Figure r1.1: Analyse des protéines mitochondriales de pomme de terre par chromatographie d'affinité sur ADBsb et par southwestern

A- Après précipitation au sulfate d'ammonium (entre 30 et 80%), les protéines solubles d'un lysat de mitochondries de pommes de terre retenues sur une colonne d'ADNsb en présence de 0,1 M NaCl ont été éluées successivement par 0,3, 0,6 et 2 M NaCl. Des aliquotes des fractions éluées ont été analysées par SDS-PAGE et colorées ainsi qu'une aliquote de l'extrait total (T). A droite (2 M) la totalité de la fraction 2 M a été déposée sur gel en vue d'un séquençage des peptides détectés. Les séquences obtenues à partir de la bande colorée voisine de 40 kDa sont indiquées (Vermel et col. 2002).

B- Les protéines solubles de mitochondries de pommes de terre ont été fractionnées sur SDS-PAGE, transférées sur membrane et hybridées avec de l'ADNsb radioactif (oligonucléotide **S1** marqué en 5' à l'aide de γ ATP). 1: coloration de la membrane; 2: autoradiogramme.

CHAPITRE I : IDENTIFICATION DE PROTEINES MITOCHONDRIALES LIANT L'ADN SIMPLE BRIN CHEZ LES PLANTES SUPERIEURES

Au cours de l'introduction, les connaissances actuelles concernant la gestion des patrimoines génétiques mitochondriaux ont été détaillées. Mais ces données sont essentiellement descriptives et de nombreuses questions restent posées, notamment au niveau moléculaire. Afin de combler ce vide, deux approches distinctes sont envisageables. La plus courante consiste à rechercher, par des moyens informatiques, des candidats comportant des domaines de liaison aux acides nucléiques ainsi que des séquences d'adressage vers les organites. La principale lacune de cette méthode est qu'elle ne permet pas d'identifier les protéines possédant des motifs nouveaux ou trop divergents de leurs homologues connus. Nous avons donc privilégié une approche de type biochimique, avec comme objectif d'identifier des candidats ayant une fonction au sein de mécanismes spécifiques des génomes mitochondriaux des plantes supérieures.

1 Identification de SSB chez *Solanum tuberosum* par chromatographie d'affinité sur matrice d'ADNsb

Cette approche a été entreprise au laboratoire en 1999 par Matthieu Vermel (Vermel *et al.* 2002), pour détecter des protéines qui possèdent une affinité pour l'ARN. En effet, il n'est pas possible d'utiliser une colonne avec une matrice de type ARN dans le but de purifier un extrait de protéines totales, principalement pour des raisons d'instabilité. Des travaux antérieurs (Li et Sugiura 1990) avaient montré que l'utilisation d'une matrice d'ADNsb, mimant certaines des propriétés de l'ARN, permettaient la purification d'une partie des protéines capables de fixer celui-ci.

Le support utilisé était une matrice de cellulose, couplée à de l'ADNsb sans spécificité de séquence (Sigma, réf. D8273). Les protéines isolées après élution à des concentrations croissantes de sels ont été analysées par électrophorèse en conditions dénaturantes (voir **Figure r1.1 A**) et leur identification réalisée par le microséquençage de leur extrémité N-terminale. Cette méthode a permis l'identification de nombreuses protéines préalablement décrites comme liant l'ARN ou l'ADN, mais a également conduit à l'obtention de peptides non caractérisés. Ceux-ci ont servi à définir un ensemble d'oligonucléotides dégénérés susceptibles d'initier des

réactions de RT-PCR. Parmi les ADNc obtenus, l'un d'eux s'est révélé particulièrement intéressant, car il correspond à une protéine de 40 kDa, abondante dans la fraction éluée à 2.0 M de NaCl lors de la chromatographie d'affinité (voir **Figure r1.1 A**). Cette protéine ne comporte d'autre part pas de motif connu de fixation aux acides nucléiques, mais porte à son extrémité C-terminale une séquence de 50 acides aminés présente à deux exemplaires. Enfin, les prédictions informatiques d'adressage indiquent la présence d'une extrémité N-terminale susceptible de la diriger vers les organelles.

2 Confirmation de la capacité de liaison à l'ADN :

En complément à la chromatographie d'affinité, un extrait de protéines totales de mitochondries de pommes de terre a été fractionné par SDS-PAGE et transféré sur membrane, laquelle a été incubée avec des sondes ADN double et simple brin radiomarquées. Un signal unique a été détecté à 40 kDa (voir **Figure r1.1 B**), augmentant l'intérêt porté au candidat de même taille identifié par chromatographie d'affinité. Cette protéine sera appelée dans la suite *St*-SBP40 (*S. tuberosum* Single Strand-DNA Binding Protein 40 kDa).

Pour confirmer que le signal observé en southwestern était effectivement dû à *St*-SBP40, la partie de son ADNc codant pour la protéine mature (déterminée par des prédictions informatiques d'adressage) a par la suite été clonée dans le vecteur pRSETc. Cette construction conduit dans *Escherichia coli* à la synthèse d'une protéine de 41 kDa pourvue à son extrémité N-terminale d'une étiquette poly-histidine. Après purification par chromatographie d'affinité sur une colonne Ni-NTA, cette protéine est elle aussi capable de fixer l'ADN en southwestern.

3 Les orthologues de *St*-SBP40 :

3.1 Le cas d'*Arabidopsis thaliana*

Dans un premier temps, des recherches dans les banques de données ont été entreprises afin de trouver des homologues chez l'espèce modèle *Arabidopsis thaliana*. Elles ont permis d'identifier quatre protéines de cet organisme appartenant à la même famille que *St*-SBP40 (voir **Figure r1.2**). Ces protéines sont codées par les gènes nucléaires *At1g47720* (*At*-SBP40a), (*At*-SBP40b), *At5g44785* (*At*-SBP40c) et *At1g31010* (*At*-SBP40d). On peut remarquer qu'aucun d'eux n'est situé dans le groupe

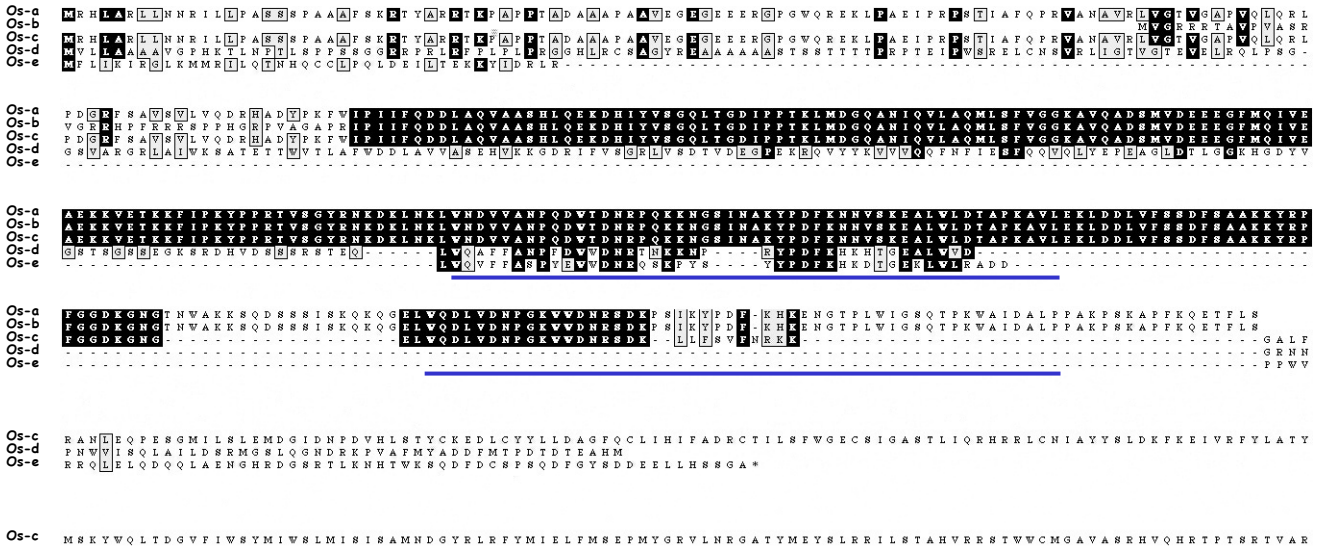


Figure r 1.3: Alignement des séquences des protéines *Os-SBP40*

Le motif PDF est souligné en bleu. La valeur définie comme seuil pour l'identité et la similarité est de 65 %. Les acides aminés identiques apparaissent sur fond noir et les acides aminés similaires sur fond gris. *Os-a Os-SBP40a* ; *Os-b Os-SBP40b* ; *Os-c Os-SBP40c* ; *Os-d Os-SBP40d* ; *Os-e Os-SBP40e*

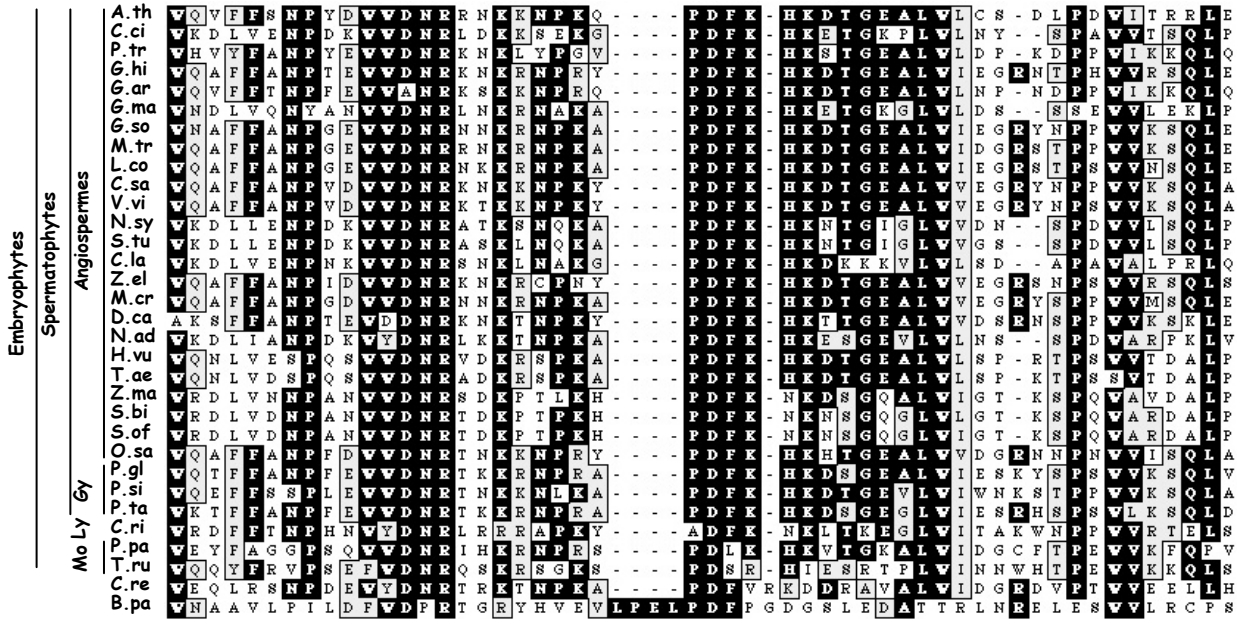


Figure r 1.4: La famille de protéines SBP40 chez les autres espèces

Un domaine PDF représentatif de chaque espèce est inclus dans l'alignement. La plupart des espèces pour lesquelles une EST présente une homologie significative sont représentées. Les résidus identiques et similaires sont respectivement sur un fond noir et gris.

A.th At-SBP40a ; *C.si Citrus sinensi* ; *P.tr Poncirus trifoliata* ; *G.hi Gossypium hirsutum* ; *G.ar Gossypium arboreum* ; *G.ma Glycine max* ; *G.so Glycine soja* ; *M.tr Medicago truncatula* ; *L.co Lotus corniculatus* ; *C.sa Cucumis sativus* ; *V.vi Vitis vinifera* ; *N.sy Nicotiana sylvestris* ; *S.tu Solanum tuberosum* ; *C.la Codonopsis lanceolata* ; *Z.el Zinnia elegans* ; *M.cr Mesembryanthemum crystallinum* ; *D.ca Dianthus caryophyllus* ; *N.ad Nuphar advena* ; *H.vu Hordeum vulgare* ; *T.ae Triticum aestivum* ; *Z.ma Zea mays* ; *S.bi Sorghum bicolor* ; *S.of Saccharum officinarum* ; *O.sa Oryza sativa* ; *P.al Picea glauca* ; *P.si Picea sitchensis* ; *P.ta Pinus taeda* ; *C.ri Ceratopteris richardii* ; *P.pa Physcomitrella patens* ; *T.ru Tortula ruralis* ; *C.re Chlamydomonas reinhardtii* ; *B.pa Bordetella parapertussis*

Gy Gymnospermes ; **Ly** Lycophytes ; **Br** Bryophytes

de gènes du chromosome 3 codant pour de nombreuses protéines potentiellement impliquées dans la maintenance du génome mitochondrial (Elo *et al.* 2003).

La **Figure r1.2** montre que les protéines de la famille SBP40 peuvent être divisés en trois régions distinctes : (a) un peptide N-terminal responsable de l'adressage et ne présentant pas d'homologie significative d'une protéine à l'autre; trois des protéines sont prédites à l'aide du programme TargetP (Emanuelsson *et al.* 2000) comme dirigées vers les organites : le chloroplaste pour *At*-SBP40a et d, la mitochondrie pour *At*-SBP40c, aucune information n'étant communiquée pour *At*-SBP40b. (b) une région centrale présentant des zones conservées et qui, dans la base de données TAIR (« The Arabidopsis Information resource ») possède des similarités avec la classe des SSB (logiciels Prosite et Pfam). (c) une région C-terminale composée d'un motif bien conservé présent à un ou plusieurs exemplaires. Ce motif, présent deux fois chez *St*-SBP40 est retrouvé une fois chez *At*-SBP40a, deux fois chez *At*-SBP40b et *At*-SBP40d et enfin trois fois pour *At*-SBP40c. C'est cette région de 50 acides aminés, que l'on nommera dans la suite motif PDF, qui présente le taux d'homologie le plus élevé entre les différentes protéines de la famille, avec de 40 à 80 % d'identité. Elle est en outre absente de toute autre protéine connue.

3.2 Les autres espèces

Nous nous sommes ensuite intéressés à l'autre plante supérieure dont le génome a été entièrement séquencé, le riz (*Oryza sativa*). La famille SBP40 y est effectivement présente, mais son organisation est sensiblement différente de ce que l'on trouve chez *Arabidopsis* (voir **Figure 1.3**). Quatre protéines homologues sont présentes dans la base données du NCBI (National Center for Biotechnology Information), dont trois (*Os*-SBP40a à c) provenant du même gène, *Os1g72040*, vraisemblablement par le biais d'un système d'épissage alternatif. Ces données sont toutefois à prendre avec réserve en raison du caractère récent des informations disponibles pour ce génome. La quatrième protéine, *Os*-SBP40d, est quant à elle codée par *Os3g43420*. Un autre homologue enfin n'est pas présent dans cette base de données, c'est une petite protéine de 141 acides aminés, *Os*-SBP40e, provenant d'un ARNm codant pour une autre protéine plus grande, expliquant qu'elle ne soit pas annotée.

La **Figure r1.4** montre un alignement des motifs PDF trouvés en réalisant une recherche parmi l'ensemble des EST disponibles. On remarque que ce motif se trouve très conservé chez toutes les plantes supérieures, des angiospermes aux fougères. Il a également été retrouvé chez les Bryophytes (*Tortula ruralis* et *Physcomitrella patens*). Un homologue, RB38, présent chez *Chlamydomonas reinhardtii*, possède une affinité pour l'extrémité 5' du transcrit chloroplastique *psbA* (Barnes *et al.* 2004). Enfin, l'homologue le plus proche parmi les bactéries se trouve dans le genre *Bordetella*, de

la classe des β -Protéobactéries, chez une famille de protéines à domaine Lip-A-acyltransférase. Pour celui-ci, le taux de conservation est beaucoup plus faible et ne garantit pas une réelle appartenance de ces protéines à la famille des SBP40. Aucune séquence proche n'a pu être trouvée chez d'autres bactéries, les protistes ou les animaux. Il apparaît donc que les SBP40 sont spécifiques des plantes.

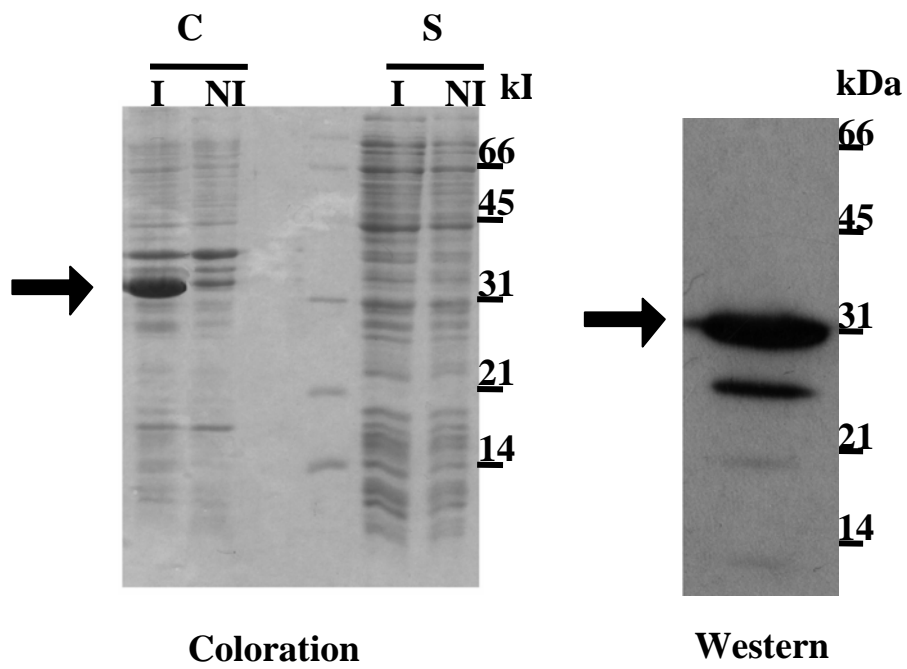


Figure n2.1: Expression de *At-SBP40a* dans *E. coli*

Les protéines d'*E. coli* obtenues après induction par IPTG (I) ou non (NI) sont analysées sur SDS-PAGE après centrifugation du lysat. C: culot; S: surnageant. Après transfert sur membrane, les protéines sont analysées par western avec des anticorps anti-His tag

CHAPITRE II: DEVELOPPEMENT D'OUTILS EN VUE DE L'ETUDE FONCTIONNELLE

Les informations collectées dans les bases de données ont donc permis de conforter l'intérêt porté à la famille de protéines SBP40. L'ensemble des données préliminaires suggère un rôle de ces protéines dans un mécanisme de maintenance des génomes des organites propres aux plantes supérieures. La mise en œuvre d'expériences en vue de confirmer ces hypothèses nécessite de disposer d'un certain nombre d'outils. Ce chapitre décrit les techniques employées et les résultats obtenus en vue de produire des quantités suffisantes de protéine surexprimée, des anticorps et des plantes génétiquement modifiées pour le niveau d'expression des différentes protéines de la famille.

1 Expression de protéines de fusion dans différents systèmes :

1.1 Dans *Escherichia coli*

Dans un premier temps, les parties des ADNc codant pour les protéines matures *At*-SBP40a à c, préalablement définies au moyen de programmes de prédiction d'adressage, ont été clonées dans le vecteur pRSETc. Ces constructions avaient pour but d'obtenir la synthèse dans *E. coli* de chacune des protéines pourvue d'une étiquette N-terminale poly-histidine en vue de sa purification par affinité. Des bactéries *E. coli* de souche BL21 ont été utilisées pour l'expression. Avec ce système, seule *At*-SBP40a a pu être exprimée, alors que les tentatives en vue d'obtenir *At*-SBP40b et c n'ont jamais conduit à l'obtention de polypeptides aux tailles attendues. Ce type de difficultés est fréquemment rencontré chez *E. coli* et peut être dû à de nombreux facteurs, comme la structure du gène, la stabilité et l'efficacité de traduction de l'ARN messager, la dégradation par les protéases, l'absence d'ARNt nécessaires à certains acides aminés ou encore la toxicité potentielle de la protéine exprimée pour la bactérie. Quelles que soient les souches bactériennes, les conditions de culture ou d'induction expérimentées, nous n'avons pas pu obtenir les protéines *At*-SBP40b et c en quantités détectables. Il semble que les bactéries ne survivent pas à la présence de ces protéines. Nous avons en conséquence concentré notre activité sur *At*-SBP40a.

L'étape suivante a consisté à tester la solubilité de cette protéine. Après fractionnement par centrifugation à 20000 x g d'un lysat de bactéries induites avec 500 µM d'IPTG, *At*-SBP40a est retrouvée dans la fraction insoluble (**Figure r2.1**). Il s'agit

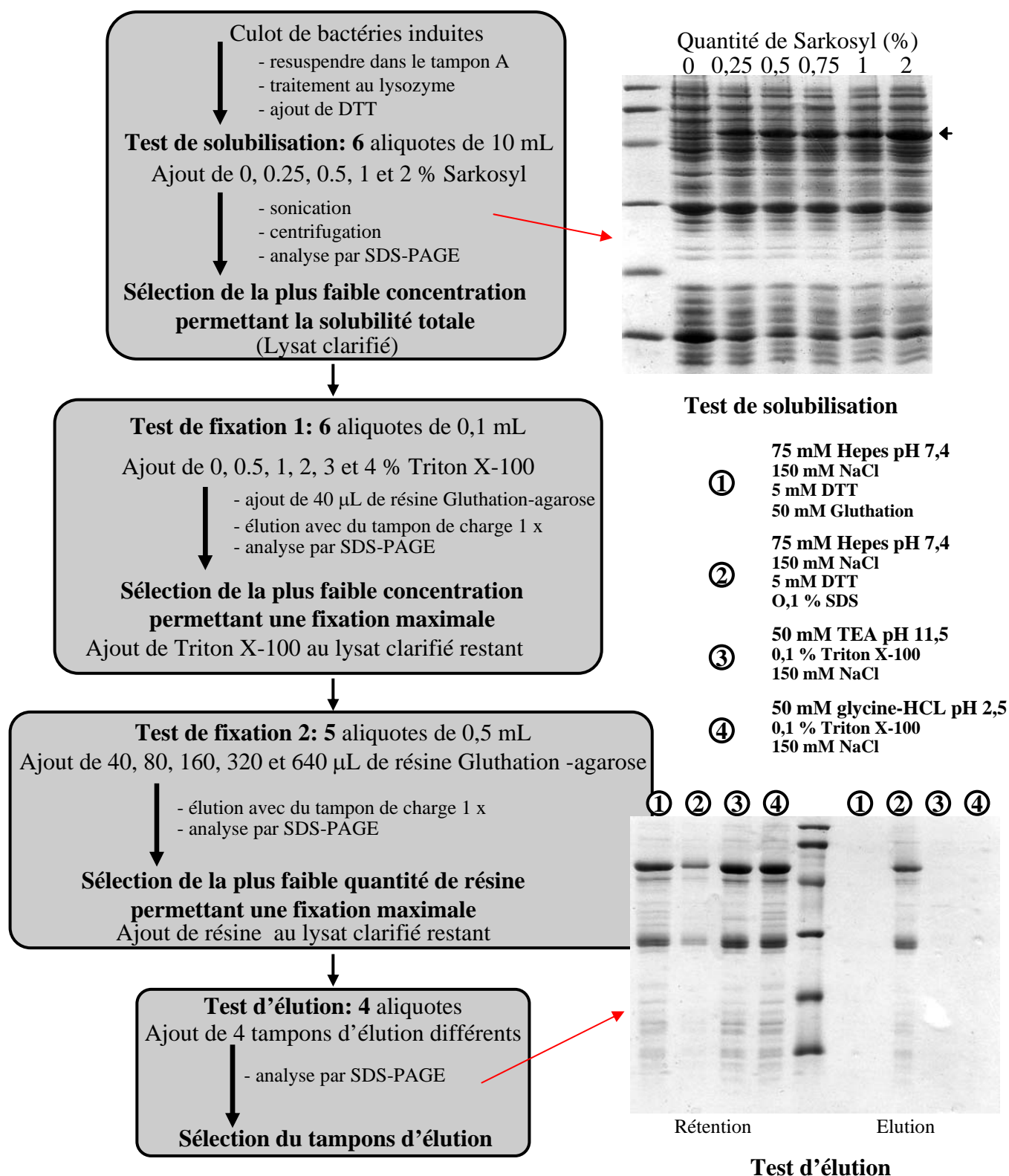


Figure r2.2: Test de solubilisation d'*A7-SBP40a-GST*

là encore d'une conséquence courante de la surexpression d'une protéine chez *E. coli*. Les protéines synthétisées en grande quantité ne sont pas efficacement prises en charge par les chaperonnes, et se retrouvent sous forme d'agrégats de protéines inactives appelés corps d'inclusion. Ce problème est fréquemment rencontré dans le cas où la protéine sur-exprimée contient des ponts dissulfure. Ceux-ci s'avèrent souvent cruciaux pour la formation de la structure tridimensionnelle, c'est-à-dire pour la stabilité des intermédiaires structuraux au cours de la synthèse protéique ainsi que pour garantir la stabilité nécessaire au produit final. Or, *At*-SBP40a ne contient que deux cystéines, non conservées chez les autres protéines de la famille. Il est donc peu probable qu'une telle liaison soit la cause du mauvais repliement de la protéine lors de sa synthèse par les bactéries. Pour contourner le problème de l'insolubilité, plusieurs types de méthodes sont envisageables. Absente de la protéine native, l'étiquette ajoutée peut être à l'origine de l'insolubilité, on peut donc essayer de la changer. On peut aussi agir sur les conditions de croissance et d'induction des bactéries, ou bien encore sur la composition des tampons au cours des différentes étapes de purification. Une dernière possibilité consiste à tenter de purifier sélectivement les corps d'inclusion, puis à renaturer les protéines non repliées qui les constituent.

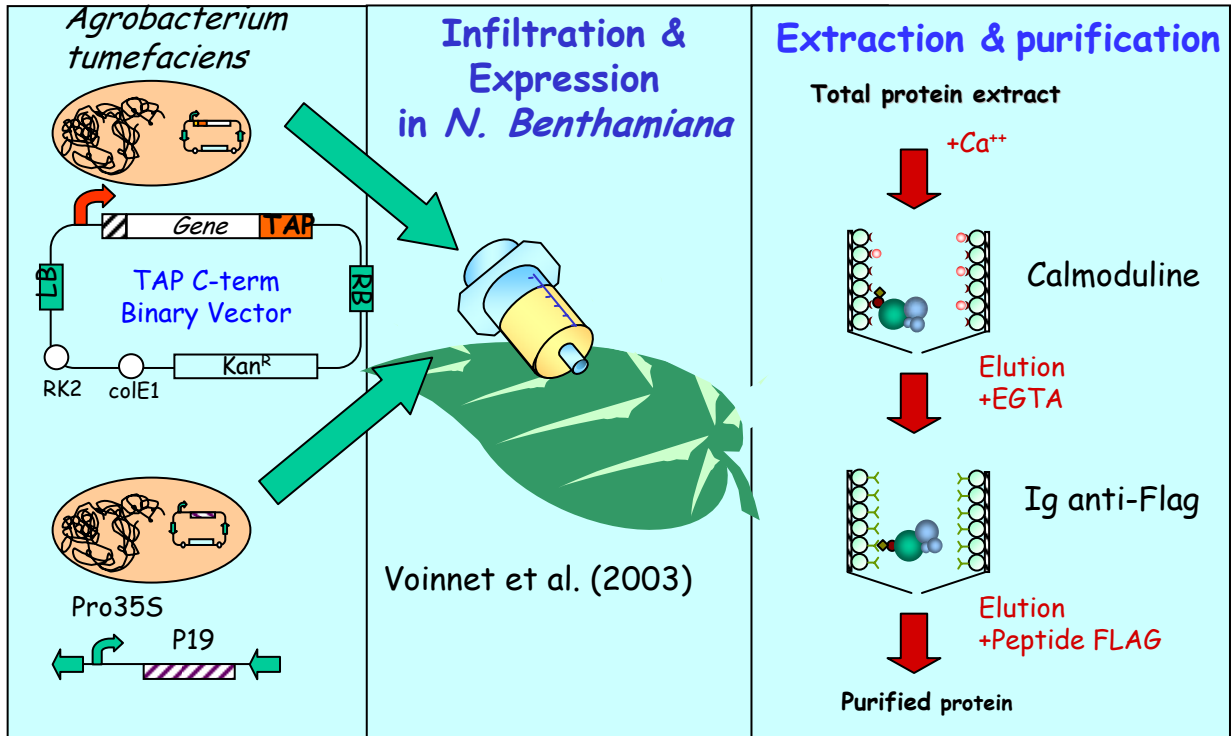
Nous avons commencé par changer les conditions de culture des bactéries, en vue de ralentir la traduction et donc de permettre aux chaperonnes un meilleur accès à la protéine exprimée. Différentes concentrations finales d'IPTG (de 5 à 250 μ M) ainsi que différentes températures de croissance des bactéries (de 20 à 30°C) ont été testées, sans conduire à l'apparition d'*At*-SBP40a dans la fraction soluble. Nous avons par la suite ajouté de l'éthanol (de 2 à 6%) au milieu de culture au moment de l'induction ce qui permet d'induire la réponse « heat shock », provoquant l'expression de protéases et de chaperonnes. Là encore, aucune augmentation du taux de protéine soluble n'a été constatée. La dernière tentative pour obtenir *At*-SBP40a sous forme soluble dans *E. coli* a consisté à modifier son étiquette en couplant la protéine avec un partenaire hydrophile ce qui tend à augmenter la solubilité de la protéine de fusion résultante. Nous avons donc cloné la séquence codant pour *At*-SBP40a dans le vecteur pGEX, afin d'ajouter la GST (Gluthation-S-transférase) à l'extrémité N-terminale de la protéine d'intérêt. Différentes conditions de culture et d'induction ont été testées, sans succès là encore.

Un protocole type (utilisant ici l'exemple de la fusion à la GST) qui a été créé pour déterminer les meilleures conditions en vue de la solubilisation d'*At*-SBP40a est présenté **Figure r2.2**. Différents paramètres ont été testés, en vue d'établir les conditions les plus efficaces pour chaque étape. Dans cet exemple, la protéine précipite sur la colonne, sauf dans le cas où le tampon d'élution contient du SDS (**Figure r2.2**, test d'élution 2). Comme nous l'avons évoqué plus haut, une alternative

consiste à tenter de renaturer les protéines inactives contenues dans les corps d'inclusion. Ceux-ci sont isolés au moyen de centrifugations différentielles et de lavages à des concentrations croissantes en agents chaotropiques et en détergents. Les protéines contenues dans ces agrégats sont ensuite solubilisées à l'aide d'urée ou de guanidine-HCl en concentration élevée. Enfin, la renaturation des protéines, étape la plus délicate, s'effectue par élimination des agents utilisés pour la solubilisation. Dans ce contexte, l'intérêt est d'obtenir un taux d'expression élevé. Nous avons donc utilisé des bactéries de souche BL161 transformées avec le vecteur pRSETc. L'induction de deux litres de milieu de culture a été réalisée à l'aide de 500 μ M (concentration finale) d'IPTG à une DO_{600nm} de 0,7. Trois heures après, les bactéries sont récoltées par centrifugation, aliquotées et congelées dans l'azote liquide puis conservés à $-80^{\circ}C$. Les différentes étapes de purification et de solubilisation des corps d'inclusion ont été réalisées en suivant le protocole décrit dans la partie Matériels et Méthodes. Au terme du dernier lavage, *At*-SBP40a représente environ 80% des protéines totales. Cette étape confirme la localisation de la protéine surexprimée dans ces agrégats et représente une première étape de purification concluante. Les protéines ont ensuite été resolubilisées l'aide de guanidine-HCl 6 M, plus efficace dans notre cas que l'urée. Nous disposons donc à ce stade d'une quantité de protéine (25 mg d'*At*-SBP40a solubilisée et purifiée obtenus par litre de culture) suffisante pour procéder à des essais de renaturation. Nous avons dans ce but testé différents protocoles:: dialyse en une ou plusieurs étapes, dilution rapide, élimination de l'agent chaotropique sur colonne Ni-NTA, reprise de l'échantillon dans 60 % de glycérol et élimination progressive de celui-ci, etc.... Quelles que soient la concentration d'*At*-SBP40a, la méthode et la composition du tampon de renaturation employées, nous n'avons jamais obtenu cette protéine sous forme soluble dans un environnement permettant de procéder à des tests d'activité. *At*-SBP40a précipite dès qu'elle se trouve dans un milieu contenant moins de 4 M de guanidine-HCl, 40% de glycérol ou 1 % de SDS.

1.2 Dans *Pichia pastoris*

Comme nous venons de le voir, toutes les tentatives d'obtenir les protéines *At*-SBP40 sous forme soluble dans *E. coli* ont échoué. Cet organisme procaryote n'est pas capable de générer des modifications post-traductionnelles comme la glycosylation, la phosphorylation, l'isomérisation *cis/trans* des prolines, etc... L'exemple de l'érythropoïétine (EPO), une protéine normalement glycosylée et qui exprimée dans *E. coli* ne parvient pas à acquérir sa conformation tridimensionnelle native (Narhi *et al.* 1991), permet d'envisager que l'absence de modifications post-traductionnelles soit la cause des problèmes rencontrés lors de nos tentatives d'expression dans cette bactérie. Les levures apparaissent à priori comme un bon moyen de contourner ce type de



TAP (4 kDa)

N Peptide Flag GGG CBD C

Figure r2.3: Expression dans *N. benthamiana*

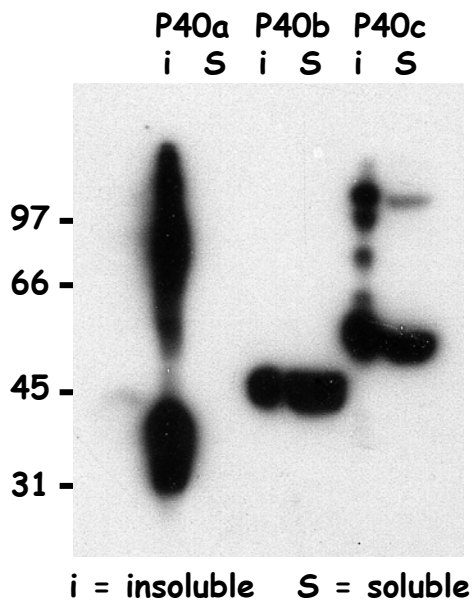


Figure r2.4: Analyse de la solubilité de *At-SBP40a, b* et *c* exprimées dans *N. benthamiana* par western à l'aide d'ig anti-flag.

difficultés et présentent également l'avantage de posséder une membrane nucléaire empêchant les protéines exprimées d'interagir avec l'ADN. Nous avons porté notre choix sur *Pichia pastoris*, chez qui l'utilisation de promoteurs forts permet en théorie l'obtention de rendements satisfaisants. Cela est illustré par une étude comparative (Lueking *et al.* 2000), dans laquelle des vecteurs d'expression pour *E. coli* et *P. pastoris* ont été testés en y clonant différents ADNc provenant d'une banque d'EST humains. Alors que toutes les protéines étaient solubles dans *P. pastoris*, la moitié étaient insolubles et 15 % n'étaient pas exprimées dans *E. coli*. Nous avons donc utilisé ce système pour tenter d'exprimer les protéines *At*-SBP40a et b, mais aucune d'elles n'a jamais pu être produite, pas même sous forme insoluble.

1.3 Dans *Nicotiana benthamiana*

Les deux systèmes d'expression décrits ci-dessus avaient été retenus sur la base de leur capacité théorique à produire des quantités importantes de protéines. Comme ils ne se sont pas révélés adaptés à notre cas, nous avons fait le choix d'expérimenter une expression *in planta*, offrant plus de garanties d'obtenir un minimum de protéine soluble. Cette méthode, décrite en **Figure r2.3**, est basée sur la co-infiltration de feuilles de tabac (*Nicotiana benthamiana*) par deux souches transformées d'une bactérie pathogène, *Agrobacterium tumefaciens*. L'une des souches renferme un plasmide portant une construction qui code pour une protéine inhibitrice du phénomène de « gene silencing », et l'autre souche un plasmide codant pour la protéine d'intérêt. En raison du faible taux d'expression généralement obtenu avec cette technique, nous avons pourvu les extrémités N-terminales de nos protéines de deux étiquettes, le « calmodulin binding peptide » (CBP) et le peptide Flag. Cela permet une double purification par affinité, sensée conduire à l'obtention de la protéine de fusion dépourvue de contaminants. Les résultats obtenus sont présentés en **Figure r2.4**. Les trois protéines testées sont exprimées avec succès, sous forme soluble pour *At*-SBP40b et c, alors qu'*At*-SBP40a est toujours insoluble. Par contre, les quantités de protéines obtenues sont très faibles. Après une purification par affinité sur une matrice de cellulose couplée à la calmoduline et au cours de laquelle moins de 20 % de la protéine exprimée est perdue, une plantule de tabac infiltrée permet d'obtenir en moyenne 400 ng de protéine. Cet ordre de grandeur nous a paru néanmoins suffisant pour permettre de procéder à des tests d'activité.

1.4 Bilan

L'ensemble des essais réalisés nous a donc permis d'obtenir trois des protéines de la famille, mais dans des conditions très différentes. *At*-SBP40a est disponible à un taux de pureté satisfaisant et en grande quantité, mais sous forme insoluble. Cet extrait

protéique a été utilisé pour des expériences de southwestern ainsi que pour la production d'anticorps. Les protéines *At*-SBP40b et c ont quant à elles été obtenues sous forme soluble, mais en quantité limitée, de l'ordre de quelques µg au maximum. Ces échantillons ont été testés lors d'expériences de gel retard.

2 Production d'anticorps polyclonaux et test des plantes transgéniques:

2.1 Production et purification des anticorps

La production d'anticorps nécessitant des quantités importantes de protéines (de l'ordre de 100 µg), pour l'injection aux lapins et la purification des sérums, nous avons utilisé *At*-SBP40a exprimée sous forme insoluble dans *E. coli*. Elle a été fractionnée par SDS-PAGE, transférée sur membrane, puis éluée. Après injection (voir Matériels et Méthodes), les sérums issus de différentes saignées ont été testés par western sur un extrait de protéines mitochondriales d'*Arabidopsis*. Tous ont engendré un profil complexe, composé de nombreuses bandes sans que certaines soient majoritaires. Il a donc paru nécessaire d'introduire une étape de purification par affinité de ces sérums.

Pour cela, nous avons utilisé une matrice de Sépharose-CNBr à laquelle est couplé le ligand. Cependant, comme la matrice de Sépharose établit des liaisons covalentes avec les amines primaires de l'épitope, via l'activation du CNBr, l'utilisation d'un échantillon solubilisé dans le guanidinium est impossible. Nous avons donc dialysé une aliquote d'*At*-SBP40a contre du SDS 1 % avant de pouvoir la coupler au Sépharose. Les anticorps ont ensuite été purifiés selon le protocole décrit dans le Chapitre Matériels et Méthodes, et deux éluations successives, à pH acide puis basique ont été réalisées. La spécificité des anticorps a dans un premier temps été testée sur la protéine recombinante exprimée dans *E. coli*. L'anticorps reconnaît très fortement cette protéine, alors qu'aucun signal n'est détecté pour un témoin négatif sans homologie de séquence avec *At*-SBP40a, hormis l'étiquette N-terminale de six histidines. Il s'agit là de la confirmation que les anticorps produits sont dirigés contre la séquence d'intérêt. Avant toute utilisation ultérieure, il convenait encore de déterminer s'ils étaient capables de reconnaître leur ligand dans un extrait végétal et de vérifier leur capacité à se fixer à d'autres protéines de la famille. Cela a pu être effectué en utilisant des plantes transgéniques, un autre outil en cours de développement en vue de l'étude fonctionnelle.

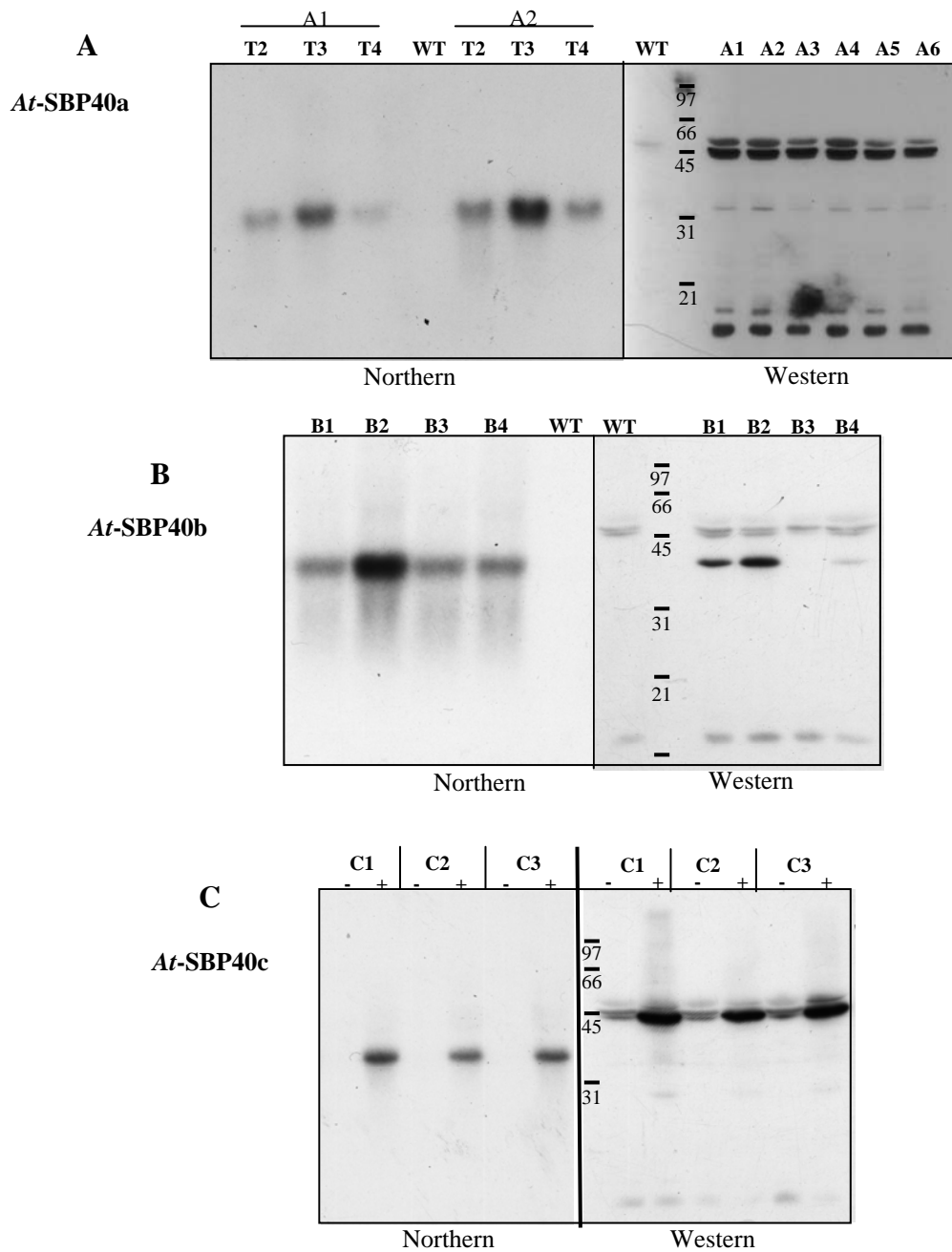


Figure r2.5: Expression stable de *At*-SBP40a, b et c dans des plantes transgéniques

Pour les Northern blots, les ARN (10 µg) sont hybridés avec une partie de la région codante du gène correspondant. Les anticorps purifiés sur *At*-SBP40a sont utilisés pour les westerns. WT: témoin sauvage.

A- *At*-SBP40a Northern: ARN issus de plantes des générations T2, T3 et T4 des lignées A1 et A2 (expression sous contrôle du promoteur 35S). Western sur les plantes T3 des plantes A1 à A6.

B- *At*-SBP40b Northern et western sur des plantes des lignées B1 à B4 (expression sous contrôle du promoteur 35S).

C- *At*-SBP40c Northern et western sur des plantes des lignées C1 à C3 induites (+) et non induites (-) par 10 µM d'œstradiol (expression sous contrôle du transactivateur inducible par l'œstradiol)

2.2 Plantes surexprimant les protéines At-SBP40

Deux types de vecteurs binaires ont été mis au point pour chaque protéine de la famille. L'un ne comporte que l'ADNc de la protéine d'intérêt alors que dans l'autre, cette séquence est agrémentée en 3' du système de purification en tandem utilisé lors des expériences d'expression dans le tabac (voir paragraphe 1.3). Ces deux approches sont complémentaires. La première présente l'avantage de ne pas risquer de produire une protéine inactive avec l'ajout d'un peptide supplémentaire à l'extrémité C-terminale. La seconde permet de se passer des anticorps pour purifier la protéine surexprimée, si ceux-ci ne s'avèrent pas efficaces, et garantit une pureté optimale. Après transformation des fleurs de plants d'*Arabidopsis* par *Agrobacterium*, les graines obtenues ont été mises à germer sur un milieu de sélection, et la présence des transgènes a été vérifiée par PCR sur l'ADN des plantules résistantes qui ont servi à produire des graines. Les descendants ont ensuite été soumis à des expériences de northern et de western blot en vue de vérifier leur aptitude à surexprimer les différentes protéines.

La **Figure r2.5** présente les résultats obtenus pour les plantes exprimant les protéines sans peptide de fusion avec pour *At-SBP40a* et *b* une expression constitutive sous le contrôle du promoteur 35S. Pour *At-SBP40c* l'expression est inducible par l'œstradiol sous le contrôle du transactivateur chimérique XVE. On voit que les ARNm correspondants aux trois protéines *At-SBP40a*, *b* et *c* ne sont pas détectés par northern dans les plantes sauvages ou non induites. Par contre les lignées de plantes transgéniques présentent synthétisent toutes ces ARNm. Dans le cas de *At-SBP40a* nous présentons une analyse des messagers produits dans différentes générations de plantes transformées propagées par autofécondation. La **Figure r2.5 A** semble indiquer que la quantité de messagers n'est pas constante au fil des générations et qu'elle semble être optimale en T3. En ce qui concerne les protéines, on constate que les anticorps anti- *At-SBP40a* purifiés s'avèrent capables de reconnaître l'ensemble des protéines de la famille parmi des protéines totales d'*Arabidopsis* (les deux éluions d'anticorps ont été testées et les anticorps élués à pH acide semblent plus efficaces car ils génèrent moins de bruit de fond pour un signal spécifique équivalent). Dans le cas d'*At-SBP40b* (**Figure r2.5 B**), deux des plantes surproduisant les messagers ne produisent pas ou très peu de la protéine correspondante. Enfin, après induction à l'œstradiol des lignées surexprimant *At-SBP40c* (**Figure r2.5 C**), on observe que le signal amplifié colocalise avec la bande inférieure d'un doublet déjà présent chez toutes les plantes non induites. *At-SBP40c* semble donc être la seule des trois SBP40 qui soit détectée chez les plantes sauvages et l'on peut ainsi émettre l'hypothèse qu'elle soit la plus fortement exprimée de la famille.

2.3 Mutants d'insertion

Le nouvel outil constitué par la mise à disposition de banques de mutants d'insertion (Gabi-kat, Salk) nous a dispensé de l'utilisation de mutants RNAi que nous avons produits et avec lesquels l'abolition de la synthèse des protéines *At*-SBP40 n'avait pas pu être mise en évidence, probablement en raison de leur faible taux d'expression naturelle. Du fait du caractère récent de cet outil, les mutants ont été disponibles les uns après les autres au cours de ces dernières années. A présent, on compte plusieurs insertions indépendantes disponibles pour chacun de nos gènes. Les mutants homozygotes correspondants ont été sélectionnés et propagés par autofécondation. Les phénotypes éventuels, la position des insertions ainsi que les informations qu'ils nous ont permis d'obtenir sont détaillés au cours du Chapitre III.

2.4 Conclusion

Nous nous sommes donc donné les outils nécessaires pour entreprendre l'étude fonctionnelle des *At* SBP40 : d'une part des anticorps purifiés reconnaissant toutes les protéines *At*-SBP40, caractéristique qui peut permettre des expériences d'immunodétection et d'immunoprécipitation avec chacune de ces protéines. Cependant, reconnaissant quatre protéines, ces anticorps sont susceptibles de donner des résultats difficilement interprétables. Les plantes transgéniques, d'autre part, qui peuvent être utilisées pour rechercher des partenaires nucléiques ou protéiques, et, à la condition de l'apparition d'un phénotype, pour déterminer dans quel type de mécanisme moléculaire ces protéines sont impliquées.

3 Cultures cellulaires transgéniques :

La décision de focaliser notre étude sur *Arabidopsis* a été prise pour disposer des avantages de ce modèle où la quantité d'outils et de données au niveau moléculaire est sans commune mesure avec les autres espèces végétales. Ce choix présente cependant pour nous un inconvénient majeur, puisqu'il s'avère peu adapté aux études biochimiques. Concrètement, l'obtention de mitochondries purifiées d'*Arabidopsis* à partir de plantes a un rendement extrêmement faible, même si on utilise uniquement les inflorescences où la proportion de mitochondries par rapport aux chloroplastes est plus élevée. Il est alors bien préférable de travailler à partir de cellules en culture qui permettent d'obtenir des mitochondries mieux purifiées avec un rendement environ 25 fois supérieur. Compte tenu de ces différents facteurs, nous avons décidé d'établir des cultures cellulaires à partir de nos plantes transgéniques. Il y a pour cela différentes manières de procéder. On peut initier la mise en culture à partir de cals, de graines

(voir Matériels et Méthodes) ou encore transformer des cultures de cellules déjà établies. Cette dernière technique ne permettant pas d'obtenir des lignées dans lesquelles 100% des cellules sont transformées, notre choix s'est porté les deux premières solutions. Les essais ont été effectués en utilisant les lignées déjà établies surexprimant les protéines *At*-SBP40b et c respectivement sous le contrôle du transactivateur chimérique XVE (inductible à l'œstradiol) et du promoteur 35S et avec les étiquettes CBP et Flag en fusion C-terminale. Les cals obtenus ont été testés avec succès par PCR pour la présence du transgène. En revanche, aucune expression n'a pu être détectée à l'aide d'anticorps anti-Flag. Les cultures cellulaires initiées directement à partir de graines ont été repiquées plusieurs semaines, sans toutefois obtenir des cellules à une densité satisfaisante et dénuées d'agrégats. Là encore, l'expression des protéines d'intérêt n'a pu être observée.

Si ces tentatives ont échoué, l'obtention de cultures cellulaires transgéniques reste un enjeu important pour l'étude de phénotypes moléculaires dans la mitochondrie d'*Arabidopsis*. Dans notre cas, cela aurait par exemple pu permettre de vérifier l'impact de l'abolition ou de l'augmentation de l'expression des protéines *At*-SBP40 sur le génome mitochondrial, de faciliter la localisation subcellulaire ou la recherche de partenaires nucléiques et protéiques.

A.

		<i>At</i> -SBP40a	<i>At</i> -SBP40b	<i>At</i> -SBP40c	<i>At</i> -SBP40d
Masse théorique (kDa)		30,5	41,7	49,9	43,9
pI théorique		8,9	6,3	8,7	9,3
Target P	Mt. score	0,336	0,177	0,808	0,555
	Cp. score	0,359	0,254	0,132	0,790
	Pred.	C/M	-	M	C/M
Predotar	Mt. score	0,41	0,35	0,35	0,24
	Cp. score	0,03	0,20	0,08	0,18
	Pred.	poss. M	poss. M	poss. M	poss. M
Wolf PSORT	Mt. score	5,5	0	6,5	3
	Cp. score	4,5	2	6	5
	Nu. score	4	12	1	8
	Pred.	C/M	Cyto.	C/M	-
MitoProt II		M (0,9520)	-	M (0,9996)	M (0,9745)
iPSORT		M	C	M	M
Consensus		M	C (?)	M	M

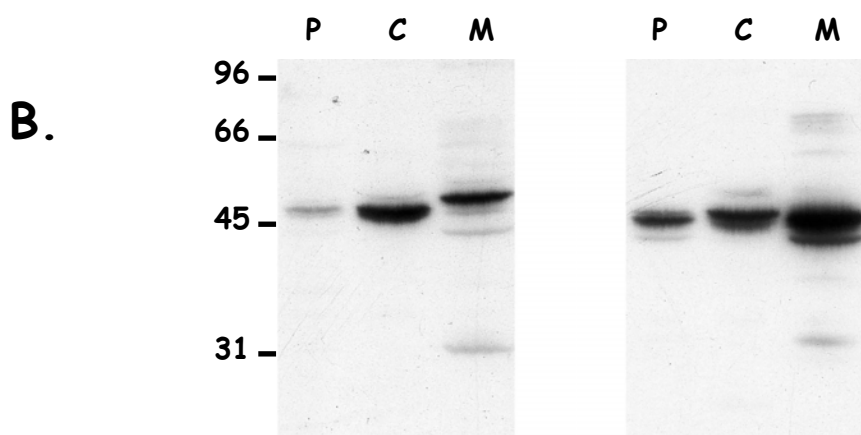


Figure r3.1: Adressage des protéines *At*-SBP40 vers les organites

A-Prédictions de l'adressage par analyse informatique. B- Western blot réalisé sur des fractions de mitochondries (M), de chloroplastes (C) ou de protoplastes (P) provenant de cellules de culture d'*A.thaliana*. Les anticorps utilisés ont été purifiés avec *At*-SBP40a et élués en conditions basiques (à gauche) ou acides (à droite).

CHAPITRE III: ETUDE FONCTIONNELLE

1 Confirmation de l'adressage aux organites :

Dans un premier temps, il a fallu s'assurer que les protéines *At*-SBP40, identifiées par leur homologie avec la protéine mitochondriale *St*-SBP40 étaient effectivement adressées aux organites. Nous avons commencé par utiliser différents programmes de prédiction d'adressage et effectué une synthèse des résultats obtenus (voir **Figure r3.1 A**). La quasi-totalité de ces programmes a confirmé un adressage aux organites pour l'ensemble des séquences testées, mais de nombreuses divergences subsistent quant à la destination mitochondriale ou chloroplastique. *At*-SBP40c est la seule pour laquelle tous ces programmes sont unanimes, lui accordant un adressage mitochondrial. *At*-SBP40a a elle aussi une forte probabilité d'être adressée à la mitochondrie. Pour les deux dernières, *At*-SBP40b et d, les résultats semblent trop divergents pour tirer une conclusion définitive. Si les programmes de prédiction d'adressage sont un moyen rapide d'obtenir une information, ils ne constituent bien entendu pas une preuve formelle de la localisation d'une protéine. De même, chacune des approches comme l'immunodétection, l'expression transitoire d'une protéine fusionnée à un rapporteur, ou encore l'importation *in vitro*, prises séparément, peuvent être source d'artéfacts. Il est donc nécessaire d'en entreprendre plusieurs et de comparer les données obtenues avant de pouvoir se prononcer de manière définitive.

1.1 Immunodétection

Des protoplastes issus de cellules de culture d'*Arabidopsis* ont servi à obtenir des fractions enrichies en mitochondries et en chloroplastes, lesquelles ont été utilisées comme substrat pour réaliser un western blot (voir **Figure r3.1 B**). Pour cela, les anticorps polyclonaux purifiés et dirigés contre l'ensemble des protéines de la famille SBP40 (voir CHAPITRE II) ont été employés. Les nombreuses bandes ainsi que les différences constatées en utilisant les anticorps élués à pH acide ou basique ne permettent pas de conclure formellement au cas par cas. On observe une bande présente sur les deux membranes à environ 31 kDa, qui pourrait correspondre à *At*-SBP40a et localisée comme attendu dans la fraction mitochondriale. La situation complexe aux alentours de 45 kDa avec au moins 4 bandes sur l'ensemble des deux fractions chloroplastiques et mitochondriales peut s'expliquer par la présence de modifications post-transcriptionnelles, voire de précurseurs non encore maturés,

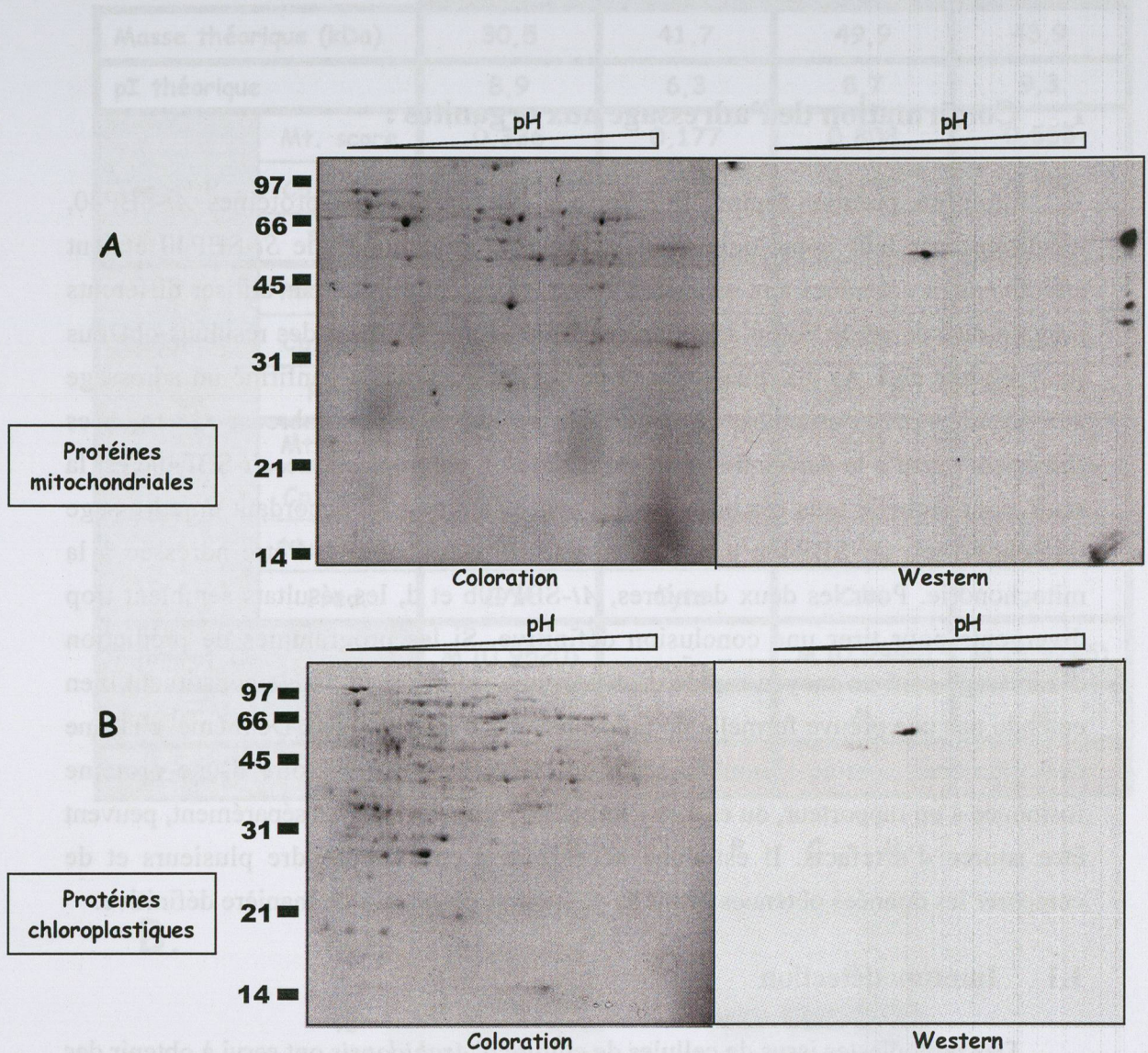


Figure n°3.2: Recherche des *At*-SBP40 parmi les protéines mitochondriales et chloroplastiques fractionnées sur gels bidimensionnels

Des extraits protéiques (1 mg) mitochondriaux (A) et chloroplastiques (B) issus de cultures de cellules d'*Arabidopsis* sont fractionnés dans un gradient de pH puis selon la taille des protéines. Après transfert et coloration les protéines sont analysées par western à l'aide des anticorps dirigés contre l'*At*-SBP40a

comme cela a été suggéré au paragraphe 2.3 du chapitre II. On retrouve toutefois une organisation générale semblable à ce qui a été décrit plus haut. La principale différence réside dans le fait que les protéines détectées ici sont toutes plus petites, d'environ 5 kDa. Ce décalage est sans doute dû au fait que les extraits protéiques ont été préparés dans des conditions différentes. En effet, les mitochondries et les chloroplastes issus de cultures de cellules ont simplement été incubés 3 minutes à 95°C en présence de tampon de charge, alors que les protéines totales des plantes transgéniques avaient été séparées par fractionnement au Trizol. La composition des tampons contenant les échantillons protéiques peuvent donc expliquer leur mobilité électrophorétique différente (Kruft *et al.* 2001).

Pour tenter d'attribuer les signaux de façon plus précise, nous avons réalisé des gels à deux dimensions. Les protéines contenues dans des fractions mitochondriales ou chloroplastiques issues de cultures de cellules d'*Arabidopsis* ont été séparées d'abord en fonction de leur point isoélectrique (pI), puis selon un SDS-PAGE. Le but de ces expériences était d'attribuer les signaux obtenus en se fiant aux pI déduits de la séquence primaire. En effet, ceux-ci diffèrent suffisamment d'une protéine à l'autre pour permettre, en théorie, l'assignation des signaux de western (voir **Figure r3.1**). La **Figure r3.2** présente les résultats obtenus à partir de fractions mitochondriales (**A**) et chloroplastiques (**B**). On observe uniquement dans la fraction mitochondriale, une série de signaux à des pI très basiques qui devraient correspondre aux protéines *At*-SBP40a, c et d. Ce résultat correspond parfaitement à ce qui était attendu en fonction des prédictions d'adressage. Le signal présent au centre de l'image, à un pI d'environ 7, et qui doit provenir comme prévu d'*At*-SBP40b, est retrouvé à la fois dans les échantillons mitochondriaux et chloroplastiques probablement du fait d'une contamination de la fraction mitochondriale par des chloroplastes. Pour clarifier la situation, nous avons fait analyser par spectrométrie de masse MALDI-TOF, les parties des gels correspondant aux signaux observés mais les quantités de protéines se sont avérées trop faibles pour conduire à l'identification de nos protéines par cette technique. Si une confirmation reste requise, ces expériences permettent tout de même de confirmer l'adressage aux organites des membres de la famille *At*-SBP40, à la mitochondrie pour trois d'entre elles et au chloroplaste pour la dernière, vraisemblablement *At*-SBP40b.

1.2 Importation in vitro

Le principe de cette technique est la mise en présence de mitochondries ou de chloroplastes purifiés avec la protéine d'intérêt synthétisée in vitro en présence d'un acide aminé détectable (radioactif par exemple). La technique employée au laboratoire utilise pour la synthèse de la protéine d'intérêt une transcription et une traduction

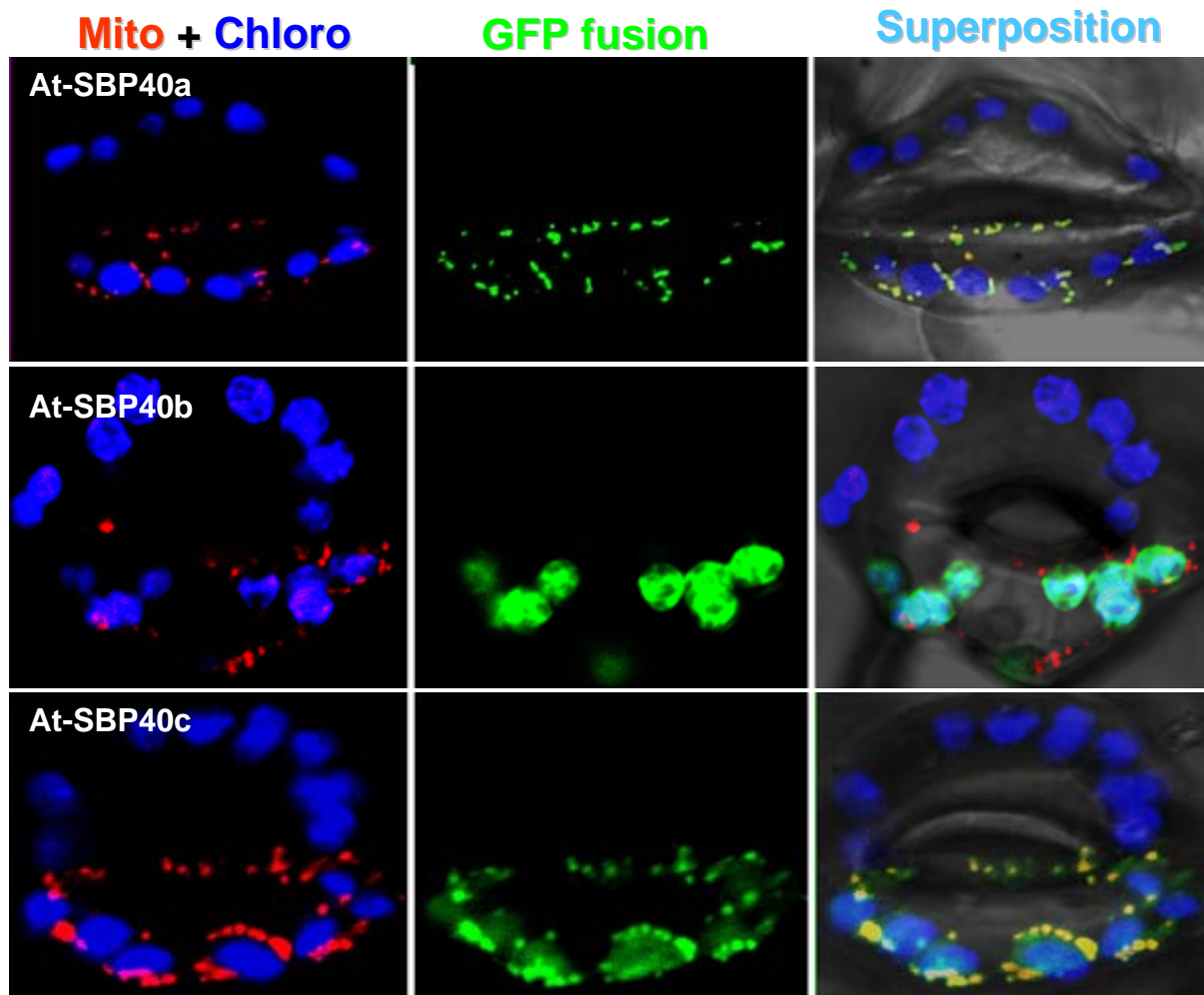
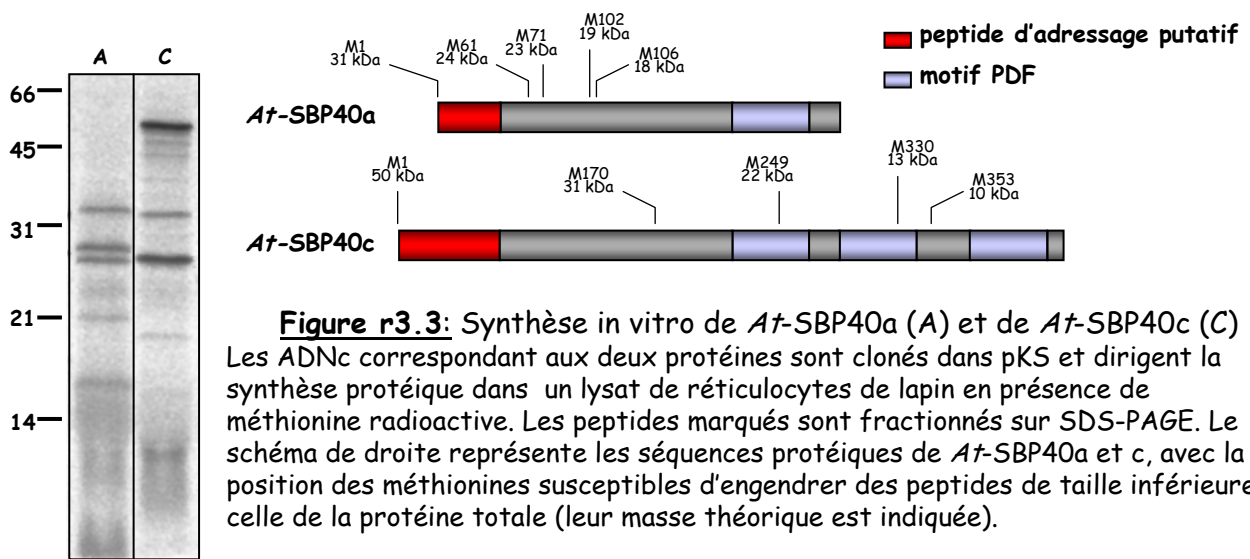


Figure r3.4: Expression transitoire dans des cellules de garde des fusions *At*-SBP40-eGFP. Les mitochondries apparaissent colorées en rouge, les chloroplastes en bleu et la protéine fusionnée à l'eGFP en vert.

couplées, dans un lysat de réticulocytes de lapin, en présence de méthionine marquée au ^{35}S . Les ADNc codant pour les précurseurs des protéines *At*-SBP40a et c ont été clonés dans le plasmide pKS. La **Figure r3.3** montre les produits de traduction radioactifs obtenus pour ces deux constructions ainsi que les sites potentiels d'initiation de la traduction. On constate que de nombreux polypeptides sont obtenus pour chacune des constructions. Certains au moins peuvent correspondre à des peptides dont la traduction serait initiée au niveau des méthionines présentes dans la phase de lecture. Les produits de traduction obtenus ont été incubés en présence mitochondries purifiées d'*Arabidopsis*, de choux fleur et de pomme de terre, mais aucun résultat clair n'en a résulté, en raison de la qualité insuffisante des polypeptides générés.

1.3 Expression transitoire dans des feuilles de tabac

Afin de confirmer les résultats obtenus par immunodétection, nous avons entrepris des expériences d'adressage *in vivo*. Celles-ci ont consisté en l'expression transitoire des protéines *At*-SBP40 fusionnées à une protéine rapportrice, l'eGFP (enhanced Green Fluorescent Protein), par bombardement de feuilles de tabac (*Nicotiana benthamiana*). La **Figure r3.4** montre les résultats obtenus sur des cellules de garde où seule la cellule du bas sur l'image a été transformée. La séquence N-terminale d'*At*-SBP40a dirige l'eGFP vers la mitochondrie et celle d'*At*-SBP40b vers le chloroplaste. Pour *At*-SBP40c, on observe une localisation mitochondriale, mais une fluorescence plus faible semble également présente dans les chloroplastes ce qui suggère la possibilité d'un double adressage vers les mitochondries et les chloroplastes.

1.4 Conclusion

Pris dans leur ensemble, les résultats obtenus ont permis de déterminer que les protéines de la famille *At*-SBP40 sont dirigées vers les organites. De plus nous avons pu attribuer une localisation mitochondriale à *At*-SBP40a et une localisation chloroplastique à *At*-SBP40b. La localisation mitochondriale d'*At*-SBP40c semble claire, mais son double adressage vers les chloroplastes demande encore confirmation. Des expériences d'importation *in vitro*, par exemple, seraient cependant encore souhaitables pour établir définitivement ces résultats.

2 Recherche de partenaires nucléiques :

Comme cela a été évoqué au CHAPITRE I, différentes expériences (chromatographie d'affinité et southwestern) ont démontré la capacité de *St*-SBP40 à

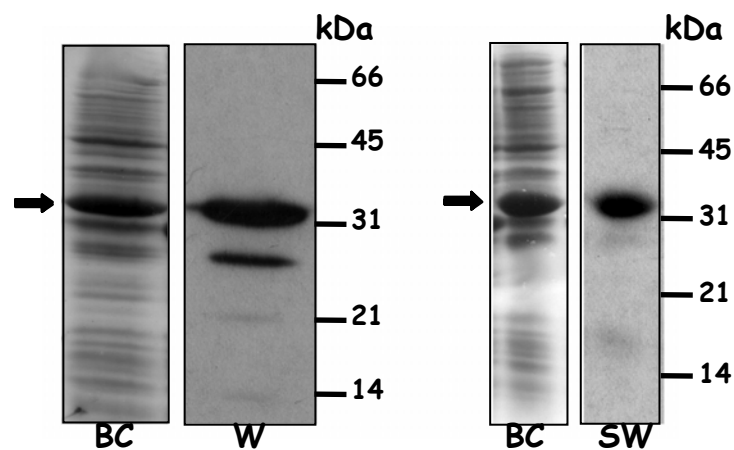


Figure r3.5: *At*-SBP40a est capable de fixer l'ADNs_b en southwestern

Les protéines *d'E. coli* surproduisant *At*-SBP40a sont fractionnées sur SDS-PAGE et transférées sur membranes. Les membranes sont révélées coloration au bleu de coumassie (BC), western à l'aide des anticorps anti-histag (W) et southwestern avec ADNs_b radioactif (SW). Les flèches pointent sur *At*-SBP40a.

fixer l'ADN. Ces résultats préliminaires, à l'origine de l'intérêt porté à la famille de protéines *At*-SBP40, contrastent avec des données publiées récemment (Barnes *et al.* 2004) et faisant état de l'affinité de la protéine RB38 de *Chlamydomonas* (la seule protéine hors les SBP40 renfermant le motif PDF), pour l'extrémité 5' non codante de l'ARNm chloroplastique *psbA*. Cette contradiction apparente soulève une série de questions à propos des déterminants de l'interaction entre les protéines d'*Arabidopsis* et les acides nucléiques. Dans cette partie, seront détaillées les approches employées ainsi que les résultats obtenus concernant la caractérisation des acides nucléiques susceptibles de remplir le rôle de ligand.

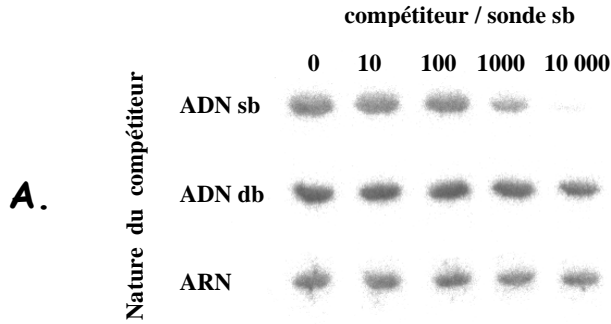
2.1 Détermination in vitro du type d'acide nucléique fixé

Le premier objectif a donc consisté à déterminer, parmi l'ARN, l'ADN double brin (ADNdb) ou l'ADN simple brin (ADNsb), lequel est fixé avec la plus grande affinité. Nous avons vu au cours du Chapitre II, qu'*At*-SBP40a, b et c, ont pu être exprimées sous forme de protéines de fusion. En raison des propriétés différentes de ces échantillons protéiques en matière de quantité et de solubilité, deux approches ont été menées de front, le southwestern pour *At*-SBP40a et le gel retard pour *At*-SBP40b et c.

2.1.1 Southwestern : le cas d'*At*-SBP40a

En premier lieu, nous avons cherché à vérifier la capacité d'*At*-SBP40a à fixer l'ADN dans ces conditions particulières. Le principe est de fractionner l'échantillon protéique à tester par SDS-PAGE, de transférer les protéines sur une membrane, puis de les renaturer afin de leur rendre une conformation dans laquelle elles sont capables de se lier à leur substrat. L'expérience de la **Figure r3.5** montre qu'au sein d'un extrait de protéines totales d'*E. coli* surexprimant *At*-SBP40a, une bande majoritaire est présente à une taille d'environ 33 kDa. Celle-ci colocalise avec le signal observé par western avec des anticorps monoclonaux dirigés contre l'étiquette penta-histidine, ce qui montre que cette protéine est bien *At*-SBP40a-His dont nous avons programmé la synthèse dans *E. coli*. Par southwestern, *At*-SBP40a se montre capable de fixer des sondes ADNsb (**Figure r3.5**) et ADNdb. En outre, les différentes sondes nucléiques utilisées (oligodésoxynucléotides divers ou parties de plasmides amplifiés par PCR) n'ont fait apparaître aucune variation notable de l'efficacité fixation à la protéine. *At*-SBP40a ne semble donc pas présenter de spécificité de séquence détectable dans les conditions du southwestern.

Cette mise en évidence de la capacité d'*At*-SBP40a à se lier aux acides nucléiques nous a donc permis de vérifier le bon fonctionnement de notre méthode.



Nature du compétiteur	Compétiteur / sonde			
	10	100	1000	10 000
ADN sb	82	90	26	3
ADN db	98	101	83	49
ARN	98	107	86	73

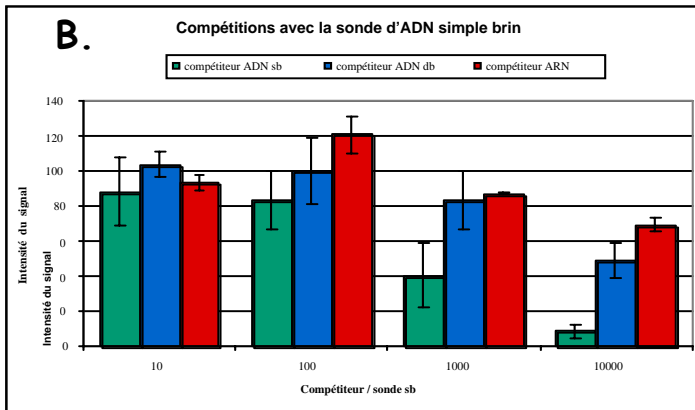


Figure r3.6: Affinité comparative d'*A7*-SBP40a pour les différents types d'acides nucléiques

Des quantités constantes d'*A7*-SBP40a sont transposées sur membrane puis hybridées avec une quantité constante de sonde ADNsb radioactive en présence de quantité croissante de compétiteur (ADNsb, ADNdb ou ARN). **A.** Une expérience type et quantification des résultats. **B.** Histogramme synthétisant les résultats de plusieurs expériences.

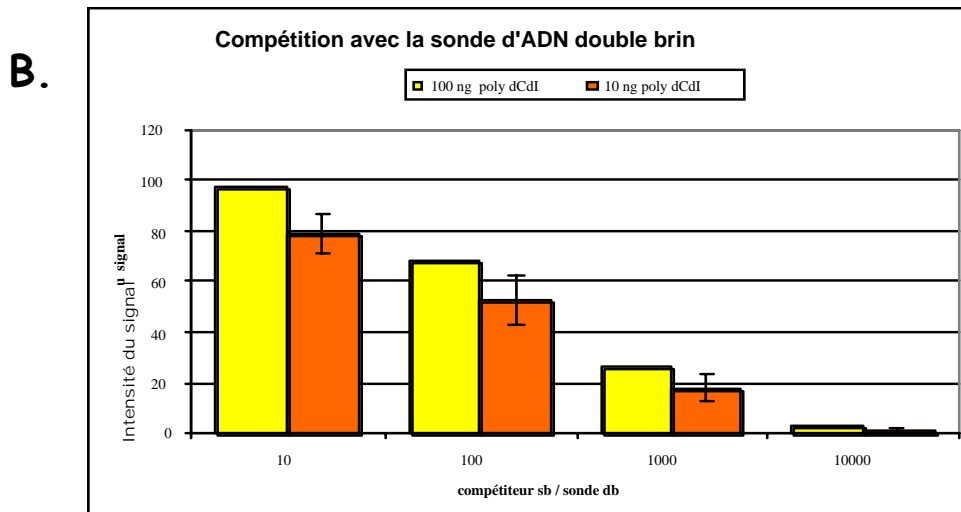


Figure r3.7: Influence du poly dCdI sur la mesure de l'affinité d'*A7*-SBP40 vis à vis de l'ADNdb.

Des quantités constantes d'*A7*-SBP40a sont transposées sur membrane puis hybridées avec une quantité constante de sonde ADNdb radioactive en présence de quantité croissante de compétiteur (ADNdb) et en présence de 10 ou 100 μ g de poly dCdI. **A.** Une expérience type et quantification des résultats. **B.** Histogramme synthétisant les résultats de plusieurs expériences.

Nous avons ainsi pu mettre en place une série d'expériences de compétition visant à comparer son affinité pour des acides nucléiques de natures différentes. Pour cela, des quantités identiques de protéine transférée sur membrane ont été mises en présence d'une quantité constante d'une sonde ADNsb et de quantités variables de compétiteurs. La **Figure r3.6 A** présente une expérience type et les résultats obtenus qui montrent une plus grande affinité d'*At*-SBP40a pour l'ADNsb. La **Figure r3.6 B** rassemble les résultats de plusieurs de ces expériences et montre les moyennes et les écarts-types des intensités des signaux enregistrés pour chaque quantité de compétiteur. Ces derniers se révèlent satisfaisants et permettent de dégager le classement décroissant suivant de l'affinité d'*At*-SBP40a pour les acides nucléiques : ADNsb>ADNdb>ARN.

Nous avons été étonnés par les quantités importantes de compétiteur nécessaires à l'atténuation du signal. Une explication est à rechercher dans l'évaluation de la quantité de sonde introduite dans la réaction. Cette quantité est déduite de la radioactivité incorporée dans la sonde après purification sur gel de polyacrylamide. Les molécules non marquées migrant au même endroit, elles sont co-purifiées avec la sonde radioactive. Le nombre de molécules d'acide nucléique contenues dans la solution de sonde radioactive est alors sous-estimé. Cela peut expliquer en partie l'importance des valeurs pour les rapports compétiteur / sonde. De plus, du poly dCdI (de l'ADN de synthèse constitué d'une succession des bases cytosine et inosine) est ajouté au bain d'hybridation, afin de diminuer les interactions aspécifiques et le bruit de fond. Il a également servi d'entraîneur pour précipiter la sonde purifiée sur gel. Ce composé n'influence pas les expériences de fixation lorsque les protéines étudiées possèdent une certaine spécificité vis-à-vis de la séquence reconnue (promoteur, activateur, etc...). Comme *At*-SBP40a ne semble pas avoir de telle spécificité, du moins dans le cadre du southwestern, il est possible qu'elle reconnaisse également le poly dCdI, qui se comporterait alors comme un compétiteur. La quantité d'ADN non marqué présente dans le milieu réactionnel sensé ne pas contenir de compétiteur pourrait donc être plus grande que prévue, et le compétiteur étudié, pour avoir un effet visible, devrait de ce fait être introduit en quantités beaucoup plus importantes.

Pour vérifier ces hypothèses, de nouvelles expériences ont été réalisées, dans lesquelles deux paramètres ont été modifiés. Tout d'abord, une sonde d'ADNdb a été synthétisée à l'aide du fragment de Klenow. Dans ce cas, toutes les molécules migrant à l'endroit d'excision de la bande ont statistiquement incorporé des nucléotides marqués et sont séparées de la matrice froide. De plus, les réactions de fixation sont réalisées dans des milieux comportant deux quantités de poly dCdI différentes (**Figure r3.7**). On observe une décroissance plus marquée du signal dans le cas où la quantité de compétiteur aspécifique est la plus faible. Cela confirme qu'en southwestern, *At*-SBP40a fixe également le poly dCdI ce qui implique l'utilisation de rapports

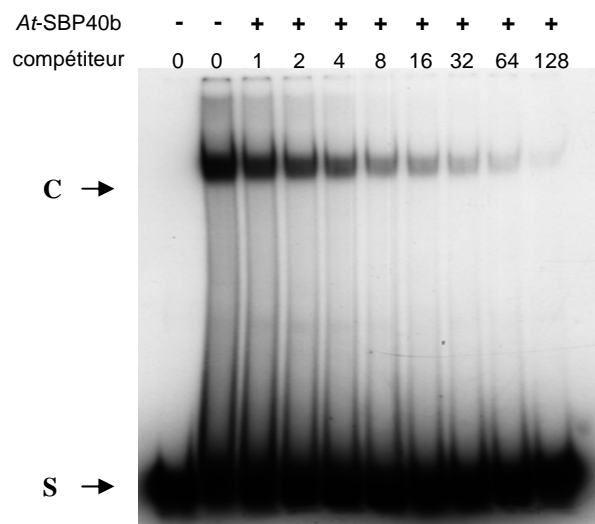


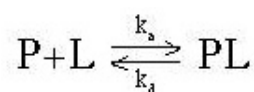
Figure n3.8: Test de l'activité d'*At*-SBP40b
 Des quantités identiques d'*At*-SBP40b et de sonde ADNsb sont incubées en présence de quantités croissantes de compétiteur (ADNsb non marqué). C: complexe; S: excès de sonde

compétiteur / sonde très élevés pour obtenir des résultats appréciables. Il faut remarquer que souvent les protéines qui interagissent avec l'ADN (SSB, replication protein A) le font sous forme multimérique ce qui est impossible lorsque la protéine est fixée sur une membrane comme dans le cas du southwestern. Cette limitation pourrait par exemple expliquer qu'*At*-SBP40a interagisse avec les trois types d'acides nucléiques testés ce qui semble peu probable dans les conditions réelles. Trois données importantes peuvent cependant être considérées comme acquises compte tenu de cette limitation: *At*-SBP40a possède effectivement une activité de fixation des acides nucléiques; elle présente une préférence pour l'ADN simple brin ; elle ne semble pas avoir de spécificité de séquence.

2.1.2 Gel retard : le cas d'*At*-SBP40b et c.

Le gel retard, ou EMSA (« Electrophoretic Mobility Shift Assay »), dont la première utilisation remonte à plus une vingtaine d'années (Garner, M.M. *et al.*, 1981), constitue la méthode la plus efficace pour l'étude des interactions ADN-protéines. Sa haute sensibilité (jusqu'à 10^{-18} mol d'ADN) en ont fait un outil largement utilisé dans de nombreuses études des mécanismes réplication, de réparation ou de recombinaison de l'ADN (Jing *et al.* 2004). Il possède en outre l'avantage d'être utilisable pour des analyses quantitatives et de cinétiques. Son principe de fonctionnement est basé sur le fait qu'une molécule présente une mobilité électrophorétique dans un gel non dénaturant différente selon qu'elle migre seule ou associée à une autre molécule. Par exemple une courte molécule d'ADN migrera plus rapidement libre qu'associée à une protéine.

Nous avons vu précédemment qu'*At*-SBP40b et c ont été obtenues sous forme de protéines de fusion solubles et purifiées, mais que les quantités disponibles sont relativement faibles. Le fait qu'une protéine soit soluble ne suffit pourtant pas à garantir qu'elle soit présente dans sa configuration tridimensionnelle active. La présence de l'étiquette TAP, permettant sa purification par affinité, peut par exemple engendrer une perte de l'activité. La première étape a donc consisté à vérifier qu'*At*-SBP40b et c purifiées sont capables de se lier à l'ADN en solution. La **Figure r3.8** montre un gel retard où des quantités identiques d'*At*-SBP40b et de sonde ADNsb ont été incubées en présence de quantités croissantes d'ADNsb non marqué. De la mise en présence de la protéine (P) et de son ligand (L) résulte l'apparition d'un complexe (PL), dont la quantité est dépendante de l'équilibre suivant:



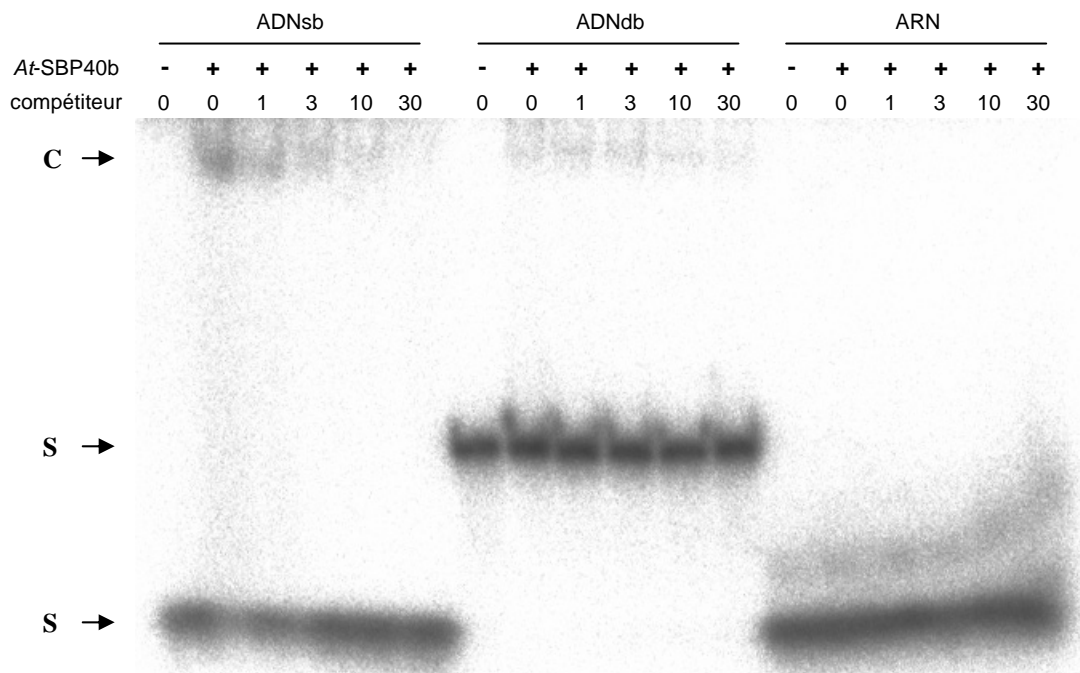


Figure r3.9: Expériences de compétition avec *At*-SBP40b

Des quantités identiques d'*At*-SBP40b et de sonde sont incubées en présence de quantités croissantes de compétiteur. Dans chaque expérience, la sonde est de même nature que le compétiteur. C: complexe; S: excès de sonde

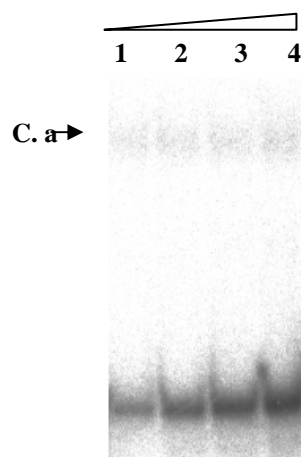


Figure r3.10: Influence de l'augmentation des quantités d'ADNdb

Incubation d'*At*-SBP40b et de sonde ADNdb dans les conditions de la Figure r3.9 sans compétiteur (1), et avec des quantités de sonde augmentées de 2 fois (2), 4 fois (3), 10 fois (4). C.a: complexe aspécifique.

où k_a est la constante d'association et k_d la constante de dissociation. Dans notre cas, deux bandes apparaissent, correspondant au complexe PL et à l'ADN libre (L), ce qui est la caractéristique d'une interaction forte entre la protéine et l'ADN, avec $k_a \gg k_d$. La traînée visible en contrebas du signal correspondant à PL est due à la dissociation qui s'opère de manière inévitable au cours de l'électrophorèse, l'ADN relargué ne pouvant rattraper les molécules de sonde non fixées. Enfin, la bande de très faible intensité visible au milieu du gel peut vraisemblablement être attribuée à une interaction faible, où $k_a < k_d$, mais qui serait stabilisée par la matrice de polyacrylamide via un effet de cage (Cann 1998). Ce phénomène caractérise une dissociation du complexe plus lente dans le gel qu'en solution, la matrice empêchant la diffusion libre des composants P et L, conduisant de ce fait à la stabilisation du complexe PL. On observe enfin une diminution de l'intensité de la bande correspondant à PL, proportionnellement à l'ajout d'ADNsb de même séquence que la sonde mais non marqué, confirmant que l'apparition de ce signal est due à la fixation d'*At*-SBP40a à l'ADN. La même expérience a été conduite pour *At*-SBP40c, mais celle-ci s'est révélée incapable de fixer la sonde, quelles que soient les conditions utilisées. Nous avons donc concentré notre attention sur *At*-SBP40b.

Comme cela avait été fait pour *At*-SBP40a, nous avons ensuite cherché à comparer l'affinité d'*At*-SBP40b pour l'ADNsb, l'ADNdb et l'ARN. Dans cette optique, trois sondes différentes ont été obtenues, une pour chaque type d'acide nucléique à tester. Celles-ci sont de même séquence, de même longueur et ne présentent pas de structure secondaire prévisible afin que l'efficacité de la fixation ne dépende que de la seule nature de l'acide nucléique. Des quantités identiques de protéine et de sonde ont été incubées en présence de quantités croissantes d'ADNsb froid (voir **Figure r3.9**). Lorsque le milieu réactionnel ne contient pas de compétiteur, *At*-SBP40b fixe environ 40 % de la sonde ADNsb, moins de 1 % de l'ADNdb et ne génère aucun signal détectable à la taille attendue pour le complexe en présence d'ARN. Cette analyse indique qu'*At*-SBP40b possède une activité de fixation des acides nucléiques spécifique à l'ADNsb. Si la présence de compétiteur froid ADNsb entraîne la diminution attendue dans le cas de la sonde ADNsb, le compétiteur ADNdb s'avère incapable d'abolir le signal dans le cas de la sonde ADNdb. Ce résultat suggère une interaction aspécifique d'*At*SBP40b avec l'ADNdb dans nos conditions.

La **Figure r3.10** présente une expérience réalisée en vue de tester cette hypothèse: en faisant varier la quantité de sonde d'un facteur 10 en présence d'une quantité de protéine identique à celle employée lors du test précédent, l'intensité du signal généré par le complexe reste effectivement à peu près constante. Cette observation correspond à une situation où les protéines présentes dans la réaction sont toutes déjà fixées à une molécule de ligand. Or, au vu des résultats décrits en **Figure**

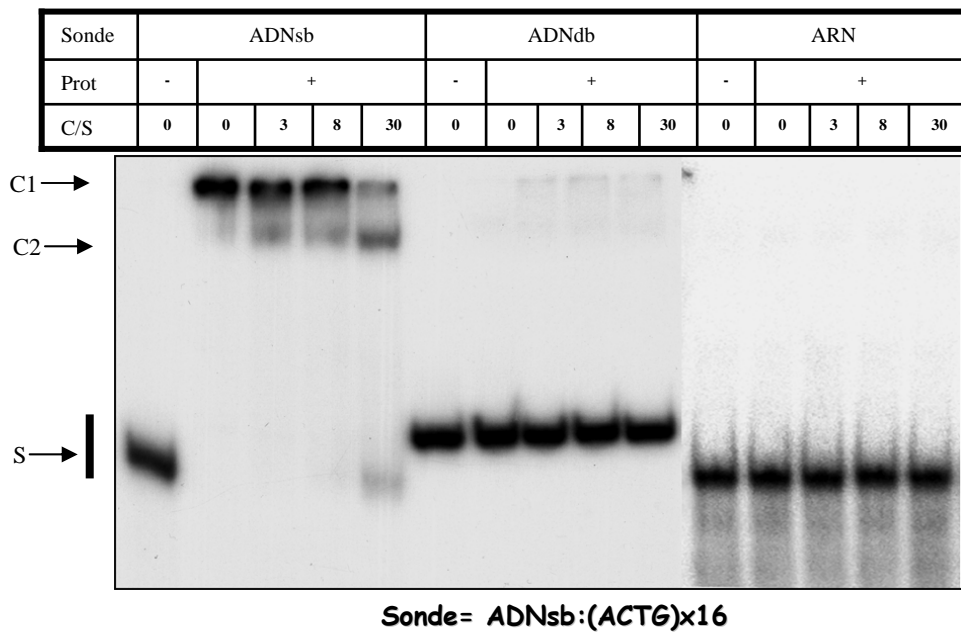


Figure r3.11: Expériences de compétition avec *At*-SBP40b en présence de poly dCdI. Des quantités identiques de sonde sont incubées avec de quantités croissantes de compétiteur (C/S: rapport compétiteur/sonde) en présence de 2,5 µg de poly dCdI et d'une quantité constante d'*At*-SBP40b. C1 et C2: complexes avec l'ADNsb; S: excès de sonde.

r3.9, nous ne sommes pas dans cette situation. La bande supplémentaire observée en présence d'ADNdb est donc vraisemblablement due à une interaction aspécifique d'*At*-SBP40b avec la sonde ou à une protéine qui n'aurait pas été éliminée au cours des deux étapes de purification par affinité.

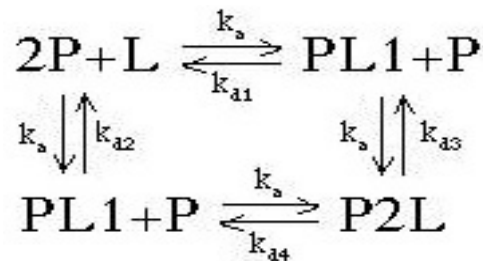
Le moyen le plus adéquat pour vérifier cette théorie est d'ajouter un compétiteur aspécifique au milieu réactionnel. Ce type de composé est fréquemment utilisé pour diminuer le bruit de fond lorsque la protéine d'intérêt possède une spécificité de séquence, ce qui n'est pas le cas ici. Des compétiteurs de différentes provenances, synthétiques ou non, peuvent être employés : ADN de sperme de saumon, ADN de thymus de veau, poly dCdl ou poly dAdT. Etant donné le manque d'informations dont nous disposons à ce stade concernant une éventuelle spécificité de séquence d'*At*-SBP40b, le compétiteur le plus adapté pour cette expérience préliminaire semble être le poly dCdl que nous avons par ailleurs utilisé dans les expériences de southwestern (voir 2.1.1.). Celui-ci possède en effet les caractéristiques les plus éloignées de celle de l'oligonucléotide utilisé comme sonde. La **Figure r3.11** présente les résultats obtenus lors d'un gel retard réalisé en présence de 2,5 µg de poly dCdl par réaction. On observe cette fois en absence de compétiteur que seul l'ADNsb est fixé, confirmant de manière définitive qu'*At*-SBP40b est une protéine de type SSB. On remarque clairement que l'ARN n'est pas du tout reconnu par *At*SBP40b comme aurait pu le suggérer la comparaison avec la protéine RB38, la seule protéine connue présentant des motifs PDF hormis les SBP40. Dans le cas de la sonde ADNdb, on détecte un faible signal dont l'intensité semble augmenter proportionnellement à la quantité d'ADNsb froid ajoutée. Ce signal est d'autant plus important que la sonde est courte ce qui suggère qu'il est dû à l'instabilité de la molécule d'ADNdb. En effet, en absence de protéine, l'ajout en quantité importante d'ADNsb froid de même séquence que le brin marqué conduit au relarguage de ce dernier, ce qui engendre l'apparition d'ADNsb marqué. Dans le cas de l'expérience de la **Figure r3.11**, l'ADNsb marqué ainsi généré est fixé par *At*-SBP40b, expliquant la présence du signal apparaissant avec la sonde ADNdb.

Dans le cas de la sonde ADNsb et en présence de compétiteur, on remarque l'apparition d'un second signal (C2) correspondant à un complexe migrant plus rapidement que le complexe observé en absence de compétiteur (C1). L'intensité de ce second signal est proportionnelle à la quantité de compétiteur présent dans le mélange. Nous avons testé différentes sondes de tailles croissantes (24, 32, 40, 48 56 et 64 bases), ce qui nous a permis de montrer que ce phénomène est lié à la taille de l'oligonucléotide : seule la sonde de 64 nucléotides (utilisée dans la **Figure r3.11**) engendre l'apparition du second signal. Nous interprétons ces résultats en proposant que 64 nucléotides représentent une molécule de taille suffisante pour permettre la

fixation de deux molécules d'*At*-SBP40b et que lorsque les quantités d'ADNsb froid augmentent, la quantité de protéine présente devenant limitante (ce qui est confirmé par l'apparition de sonde libre pour le rapport C/S=30), il apparaît un autre complexe (C2) formé cette fois d'une seule molécule de protéine par oligonucléotide. On peut ainsi déduire qu'*At*-SBP40b engendre un encombrement d'environ 32 bases sur l'ADNsb.

Le profil des signaux générés avec l'oligonucléotide de taille 64, à savoir la présence d'une bande intermédiaire bien distincte, est caractéristique d'une fixation de type non coopérative. Cela signifie que la présence d'une molécule de protéine sur l'ADN ne favorise pas la fixation des molécules de protéines libres sur celui-ci plutôt que sur une molécule d'ADN libre.

Ces observations permettent donc de définir l'équilibre observé de la façon suivante :



2.1.3 Bilan

Comme nous venons de le voir, deux types d'expériences menées à l'aide de deux protéines différentes, *At*-SBP40a et b, ont conduit à un résultat identique. Des différences ont néanmoins été décelées, comme la capacité d'*At*-SBP40a à fixer également, mais avec moins d'affinité, l'ADNdb, l'ARN et le poly dCdI. Ces observations peuvent toutefois être dues à la méthode employée, le southwestern étant moins fiable que le gel retard. Les membres de la famille *At*-SBP40 apparaissent donc clairement comme des protéines de type SSB.

Le fait que RB38, homologue présent chez *Chlamydomonas*, fixe l'ARN et de surcroît de manière séquence-spécifique devient d'autant plus intéressant. Cela est sans doute dû à la présence d'un domaine de fixation aux acides nucléiques différent et appelle la caractérisation ce celui-ci chez les protéines d'*Arabidopsis*. Cette approche sera développée au cours du paragraphe 3.

2.2 Essai de caractérisation de l'activité des *At*-SBP40

La mise en évidence de la capacité d'*At*-SBP40a et b à fixer l'ADNsb constitue un premier indice dans le cadre de la caractérisation fonctionnelle de ces protéines. Cela permet d'envisager leur implication dans des mécanismes faisant intervenir ce

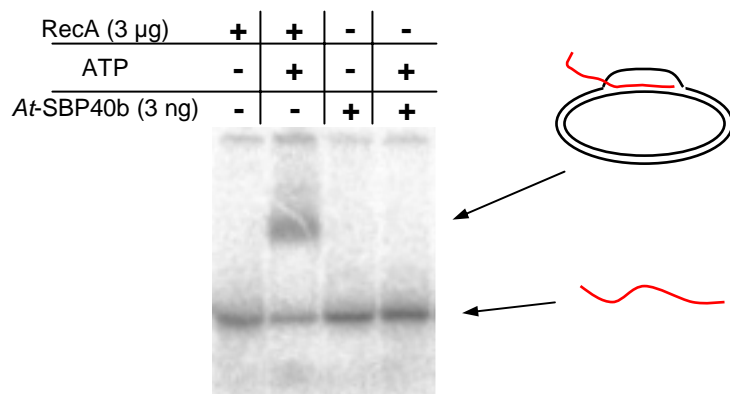


Figure r3.12: Test d'une activité de type RecA

A.t.-SBP40b et RecA d'*E. coli* sont incubées en présence d'ADNdb et d'ADNsb de séquences correspondantes avec ou sans ATP

type d'acide nucléique, comme la réplication, la recombinaison ou la réparation. Elles peuvent pour cela être en charge d'un processus enzymatique ou d'un simple rôle de protection de l'ADNsb généré lors des processus que nous venons d'évoquer.

Hormis le motif situé dans la partie centrale des séquences des différentes protéines de la famille et qui présente certaines caractéristiques du domaine SSB, les analyses informatiques n'ont pas conduit à l'identification d'un domaine connu. Nous avons en conséquence testé différentes hypothèses, comme la présence d'activités enzymatiques (RecA, ATPase) ou d'une spécificité de séquence.

2.2.1 Tests de recombinaison

Nous avons vu dans l'introduction que les génomes mitochondriaux sont soumis à une intense activité de recombinaison. La méthode employée pour tester une éventuelle activité de recombinaison consiste à vérifier si la protéine d'intérêt est capable de provoquer l'invasion d'une molécule d'ADNdb circulaire par de l'ADNsb (voir Matériels et Méthodes). La **Figure r3.12** montre qu'en présence d'ATP, la protéine RecA d'*E. coli* conduit comme attendu à l'apparition de la molécule complexe schématisée à côté du gel (pistes 1 et 2). Dans des conditions identiques, *At*-SBP40b ne conduit pas à la formation de ce complexe (pistes 3 et 4), quelle que soit la durée d'exposition du film. Nous ignorons le degré de pureté et d'activité de la RecA d'*E. coli* commerciale utilisée, mais les quantités de protéine introduites dans la réaction sont importantes (3 mg). Bien qu'*At*-SBP40b soit très probablement beaucoup plus purifiée que la protéine commerciale, les quantités disponibles sont beaucoup plus faibles (environ 1000 fois inférieures) et il est possible qu'une activité de recombinaison réduite ait pu nous échapper.

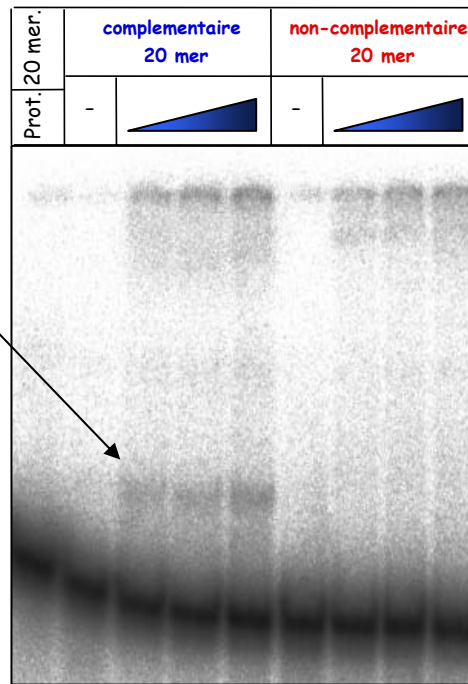
Pour pallier le facteur limitant que constitue la quantité utilisable d'*At*-SBP40b pour cette dernière expérience, nous avons développé une nouvelle méthode basée sur l'instabilité de la molécule d'ADNdb à faible concentration de sels et détaillée dans le paragraphe 2.1. En effet, une enzyme ne peut permettre une réaction thermodynamiquement impossible, mais l'accélère seulement, en diminuant l'énergie d'activation. Dans le cas présent, l'une des fonctions de RecA est de permettre l'ouverture de l'hélice afin de rendre le brin complémentaire à l'ADNsb invasif accessible. Nous avons incubé de l'ADNdb long de 64 bases dont un seul brin est radioactif avec un petit oligonucléotide (20 bases) complémentaire à l'extrémité 5' du brin marqué. Lorsque celui-ci est présent en concentration très importante (10^3 à 10^4 fois supérieure à celle de la sonde), on observe la formation d'une molécule dans laquelle le petit oligonucléotide est apparié au brin complémentaire de l'hélice :



Figure r3.13: Test de la capacité d'*At*-SBP40b à échanger des brins d'ADN

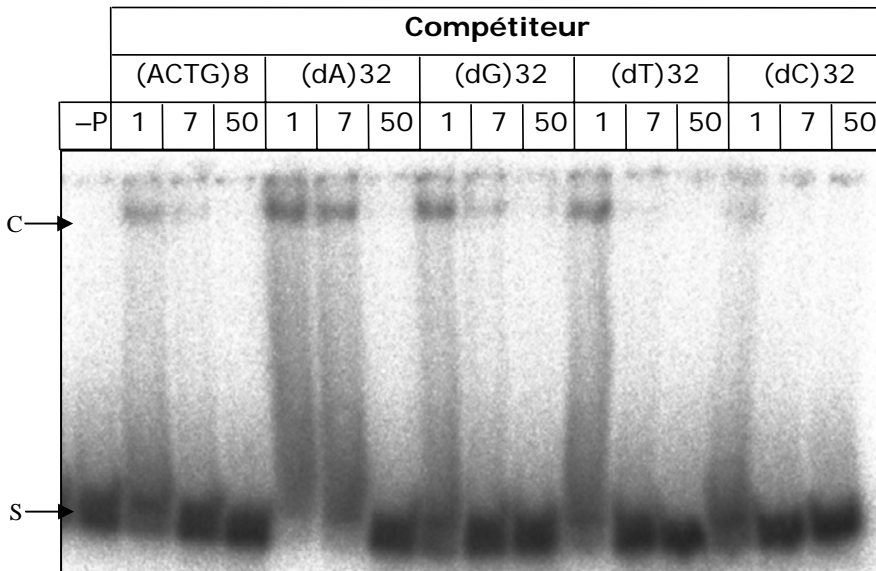
Des quantités croissantes d'*At*-SBP40b sont incubées avec une quantité constante de sonde ADNdb marquée sur un de ses brins et en présence d'un oligonucléotide complémentaire ou non.

sonde *
ADNdb



sonde							
fixation	+++	-	++	++	+++(+)	-	-

Figure r3.14: Sondes utilisées pour affiner la caractérisation de l'interaction *At*-SBP40b/ADN



Sonde: (ACTG)x8

Figure r3.15: Test de l'affinité d'*At*-SBP40b pour différents homodésoxyribonucléotides

Des quantités constantes de sonde et d'*At*-SBP40b et de sonde ADNsb ont été incubées en présence de quantités croissantes de compétiteur. -P: témoin sans protéine; C: complexe; S: sonde en excès.

En réduisant la concentration du petit oligonucléotide à un niveau juste inférieur à celui conduisant à l'apparition du complexe on dispose d'un système permettant de tester la capacité d'*At*-SBP40b à déplacer l'équilibre vers ce complexe et ainsi à mimer l'activité de recombinaison. La **Figure r3.13** montre que des quantités croissantes d'*At*-SBP40b conduisent à l'apparition d'un signal d'intensité croissante et correspondant au complexe comportant les trois oligonucléotides. Par contre, la même expérience réalisée à l'aide d'un petit oligonucléotide ne pouvant s'apparier à la sonde ne permet pas de détecter le signal provenant du complexe d'intérêt. Ce résultat montre qu'*At*-SBP40b favorise l'échange des brins entre ADNdb et ADNsb de séquences identiques et sans hydrolyse d'ATP. Il semble donc qu'*At*-SBP40b pourrait se conduire comme les protéines de type SSB qui, en se liant à l'ADNsb et en évitant la ré-hybridation des brins du duplex., facilitent l'action d'autres enzymes impliquées dans le métabolisme de l'ADN.

2.2.2 Spécificité pour une structure nucléique particulière

Avec toujours pour objectif d'affiner la caractérisation de l'interaction entre *At*-SBP40b et son ligand, nous avons synthétisé la série de sondes représentées **Figure r3.14**. Celles-ci sont inspirées d'une méthode développée pour l'étude d'une activité « Holliday junction resolvase » chez *S. pombe* (Oram *et al.* 1998). Dans notre cas, ces outils devaient permettre de tester une éventuelle préférence d'*At*-SBP40b pour l'une ou l'autre des extrémités 3' ou 5' de l'ADNsb, ou encore de trouver une affinité particulière pour une structure mimant une fourche de réplication ou un mésappariement. Les efficacités de fixation mentionnées pour chaque sonde, prises dans leur ensemble, conduisent à la conclusion que la protéine fixe toutes les structures contenant une région simple brin suffisamment longue. L'affinité très légèrement supérieure constatée pour la structure de type fourche par rapport à la molécule d'ADNsb linéaire ne permet pas d'affirmer avec certitude qu'*At*-SBP40b soit spécifique de ce type de structure.

2.2.3 Spécificité de séquence

Les expériences précédentes avaient été conduites de façon à limiter l'éventuelle influence de la séquence nucléotidique sur la fixation par la protéine pour pouvoir apprécier les autres paramètres. Pour caractériser la spécificité d'une protéine vis-à-vis d'un type particulier de séquence, l'expérience la plus aisée consiste à comparer son efficacité de fixation aux différents homopolymères, dans notre cas, le poly (dA), le poly (dT), le poly (dG) et le poly (dC). Dans l'expérience présentée **Figure r3.15**, la sonde commune à toutes les réactions est constituée de huit

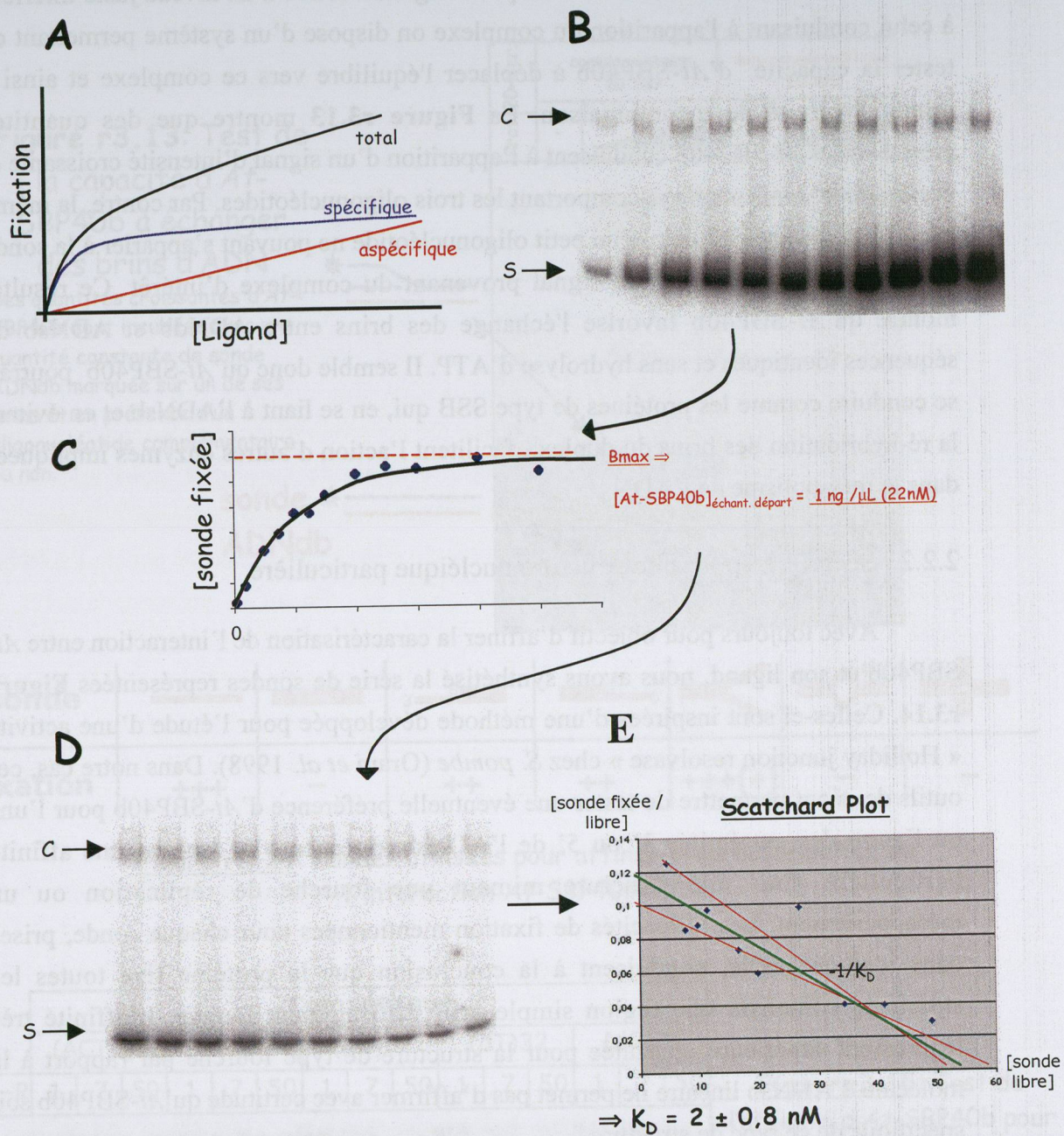


Figure r3.16: Détermination de la constante d'affinité d'*At*-SBP40b pour le ligand (ACTG) \times 8

- A- Schéma théorique illustrant les fixations spécifique et aspécifique d'un ligand.
- B- Incubation en présence de quantités croissantes de sonde. C: complexe= sonde fixée; S: sonde libre
- C- Courbe de saturation déduite de B.
- D- Equilibres dans la zone [sonde fixée] / [sonde libre] < 0, 15, permettant de déduire le K_D .
- E- Représentation de Scatchard déduite de D

répétitions du tétramère ACTG. L'oligonucléotide de 32 bases ainsi constitué ne peut former aucune structure, ce qui permet à l'interaction de ne dépendre que de la séquence primaire de l'acide nucléique. On observe que le poly (dG) et le poly (dT) sont fixés avec une affinité comparable à celle de l'oligonucléotide compétiteur de même séquence que la sonde. Le poly (dA) semble lui moins efficace pour abolir le signal dû au complexe alors que le poly (dC) présente une affinité supérieure. Ces résultats préliminaires suggèrent donc qu'*At*-SBP40b ait une préférence pour les séquences riches en cytosines. La signification biologique éventuelle de cette information sera évoquée dans la discussion. Un biais est toutefois constitué par l'absence dans ce type de test de séquences auxquelles la protéine est confrontée *in vivo*. Nous avons donc tenté d'entreprendre l'identification de celles-ci en les copurifiant avec la protéine surexprimée *in planta*. Ces expériences sont détaillées dans le paragraphe 2.3.

2.2.4 Calcul de la constante d'affinité

Dans notre cas, le premier obstacle à la détermination d'une constante d'affinité est la faible quantité disponible d'*At*-SBP40b. En effet, environ 1 μg de protéine est attendu après passage sur une matrice de Sépharose couplée à la calmoduline. Les méthodes classiques de dosage des protéines ne sont donc pas suffisamment résolutive pour être employées ici. Ce problème peut être évité en réalisant des expériences de saturation. Elles consistent à mesurer la fixation spécifique à l'équilibre et à différentes concentrations en ligand, afin de déterminer le nombre de sites de fixation présents. En plus de ces derniers, le ligand est susceptible de se fixer de façon aspécifique, ce qui conduit au biais schématisé **Figure r3.16 A**. Il est donc nécessaire de déterminer la proportion du signal dû à la fixation aspécifique en comparant les résultats obtenus avec et sans compétiteur aspécifique. Ce biais s'est avéré négligeable, et les résultats obtenus à l'aide de l'oligonucléotide constitué de huit répétitions de la séquence ACTG sont présentés **Figure r 3.16 B et C**. Le plateau est atteint lorsque tous les domaines de fixation à l'ADNsb sont occupés. Le signal provenant du complexe protéine-ADN devient dans ce cas représentatif du nombre de domaines présents dans la réaction. En comparant cette intensité à celle d'une quantité connue de sonde libre, on peut déduire la concentration de protéine présente dans la réaction. La concentration d'*At*-SBP40b dans l'éluat de la chromatographie sur la colonne de calmoduline déduite de ces expériences est de $1,0 \pm 0,2 \text{ ng}/\mu\text{L}$ soit environ 20 pM. La détermination de la valeur du K_D a été conduite selon la méthode de Scatchard, qui permet de linéariser les données en travaillant à des concentrations de protéines faibles par rapport à celle de la sonde. Une des expériences réalisées dans ce but et le graphique qui en découle sont présentés **Figure r3.16 D et E**. La pente de la

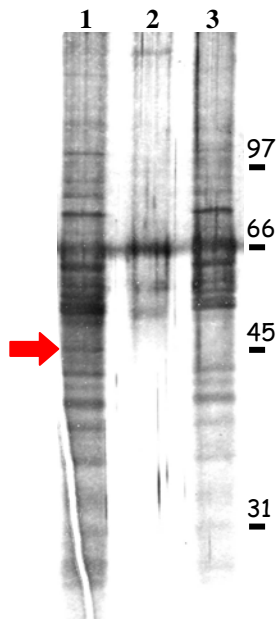


Figure r3.17: Essai d'immunoprécipitation d'*At-SBP40b*
 Analyse par SDS-PAGE des produits immunoprécipités à partir d'un extrait chloroplastique (500 µg) et à l'aide d'anticorps anti-*At-SBP40a* et de protéine A-Sépharose. (1)- Essai avec anticorps et protéine A-Sépharose. (2) Témoin sans extrait chloroplastique. (3) Témoin sans anticorps. La flèche pointe sur une protéine non présente dans les témoins et de la taille de l'*At-SBP40b*.

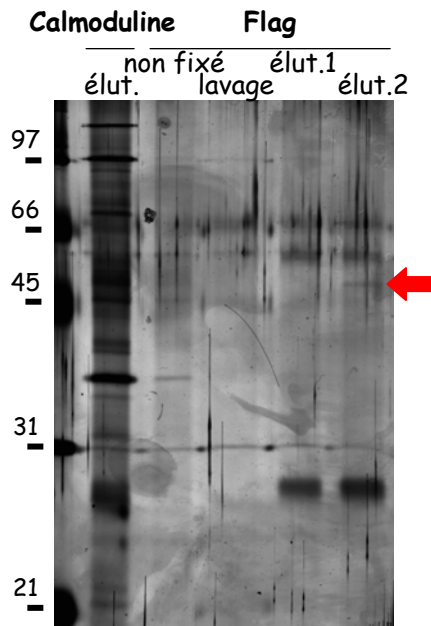


Figure r3.18: Purification à l'aide du système TAP d'*At-SBP40c* à partir de transformants stables d'*Arabidopsis*

Les protéines de l'extrait protéique brut (environ 50 µg) retenues sur la colonne de calmoduline sont éluées par EGTA puis déposées sur une colonne anti-flag. Après lavage, deux éluations avec le même tampon (80 puis 200 µl) sont réalisées successivement, afin d'éliminer le volume mort de la colonne (qui se trouve donc dans l'élué 1). Des échantillons issus des différentes étapes sont analysés par SDS-PAGE suivie d'une coloration au nitrate d'argent. La flèche pointe sur la bande qui correspond potentiellement à *At-SBP40c*.

droite correspond à la valeur de $-1/K_D$. La constante d'affinité caractérisant l'équilibre entre *At*-SBP40b et l'oligonucléotide constitué de huit répétitions du motif ACTG a été estimée par cette méthode à $1,9 \pm 0,8$ nM, ce qui se trouve dans l'ordre de grandeur (de 10 pM à 100nM) des constantes généralement observées pour les interactions ADN-protéines.

2.3 Recherche de partenaires *in vivo*

Deux types d'approches sont envisageables pour tenter de co-purifier les protéines *At*-SBP40 et les acides nucléiques auxquels elles sont fixées *in vivo*. On peut d'une part immunoprécipiter nos protéines à partir d'extraits mitochondriaux ou chloroplastiques purifiés mais, comme nos anticorps reconnaissent toutes les protéines de la famille (voir chapitre II paragraphe 2.2, **Figure r2.5**), il est possible que les résultats soient difficiles à interpréter. On peut d'autre part purifier les différentes protéines en fusion avec l'étiquette TAP à partir des plantes transgéniques qui les surproduisent. L'inconvénient ici, réside dans le fait que nous n'avons aucune certitude que le comportement de protéines de fusion surexprimées soit identique à celui des protéines endogènes.

2.3.1 Par immunoprécipitation

Nos anticorps, testés par western, ont été obtenus par injection à un lapin d'*At*-SBP40a dénaturée, purifiée sur SDS-PAGE, transférée sur membrane puis éluée. Il n'est donc pas certains qu'ils soient capables de reconnaître les protéines *At*-SBP40 non dénaturées en solution. Nous avons testé cette capacité par immunoprécipitation à l'aide de la protéine A-Sepharose sur un extrait chloroplastique (**Figure r3.17**). On constate qu'en présence d'anticorps, une bande supplémentaire faiblement visible apparaît à la taille attendue (45kDa) ainsi que d'autres bandes à des tailles supérieures (>100kDa). Les quantités trop faibles de substrat n'ont pas permis d'identifier formellement ces protéines, même par spectrométrie de masse MALDI-TOF. Les tentatives réalisées à partir d'extraits mitochondriaux n'ont quant à elles jamais révélé la présence d'une bande supplémentaire.

2.3.2 Par chromatographie d'affinité

La **Figure r3.18** montre un gel coloré au nitrate d'argent des différentes étapes de la purification d'*At*-SBP40c-TAP à partir de transformants stables chez *Arabidopsis*. Comme attendu, cette technique possède un rendement beaucoup plus

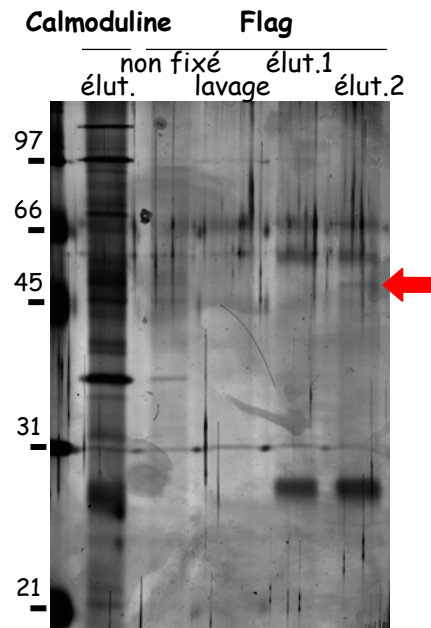


Figure r3.18: Purification à l'aide du système TAP d'*At*-SBP40c à partir de transformants stables d'*Arabidopsis*

Les protéines de l'extrait protéique brut retenues sur la colonne de calmoduline sont éluées par EGTA puis déposées sur une colonne anti-flag. Après lavage, deux éluations avec le même tampon (80 puis 200 μ l) sont réalisées successivement, afin d'éliminer le volume mort de la colonne (qui se trouve donc dans l'élué 1). Des échantillons issus des différentes étapes sont analysés par SDS-PAGE suivie d'une coloration au nitrate d'argent. La flèche pointe sur la bande qui correspond potentiellement à *At*-SBP40c.

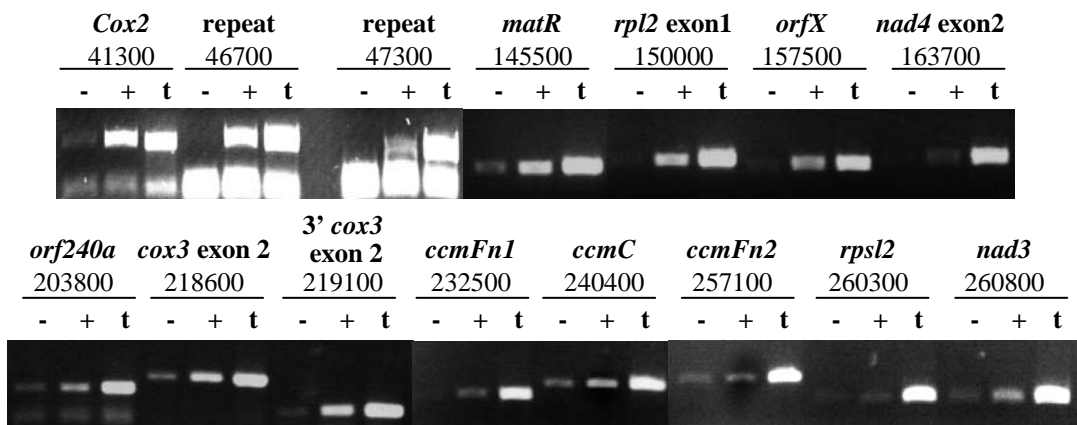


Figure r3.19: Recherche par PCR de la présence d'ADN mitochondrial accompagnant *At*-SBP40c purifiée par affinité.

L'extrait élué de la colonne anti-flag (Figure r3.20, élué 2) sert de matrice dans des réactions PCR visant les gènes mitochondriaux indiqués avec leurs positions dans le génome mitochondrial d'*Arabidopsis*. (+) correspond à l'extrait issu de transformants stables surexprimant *At*-SBP40c. (-) correspond à un extrait issu de plantes sauvages. (t) est un témoin positif réalisé à partir d'ADN total de plantes sauvages.

faible que la co-infiltration chez le tabac de la même construction avec un inhibiteur de « gene silencing ». Toutefois, l'objectif ici n'était pas d'obtenir une quantité maximale de protéine, mais un échantillon final suffisamment pur pour nous permettre d'identifier les séquences éventuellement fixées à la protéine d'intérêt. La purification recherchée a été obtenue puisqu'une seule protéine est détectable après deux chromatographies d'affinité et qu'elle a la taille attendue pour *At*-SBP40c (voir **Figure r3.18**). L'approche entreprise en vue de vérifier si, dans ces conditions, la protéine est toujours fixée à son substrat a consisté à tester par PCR différentes régions du génome mitochondrial. Après un traitement au phénol-chloroforme de l'éluat, l'ADN a été précipité et a ensuite servi de matrice à une série de PCR utilisant des amorces dirigées contre des régions définies arbitrairement du génome mitochondrial d'*Arabidopsis* (voir **Figure r3.19**). Le témoin négatif utilise comme substrat un éluat obtenu à l'aide du même type de purification, mais réalisé à l'aide de plantes ne surexprimant pas *At*-SBP40c. La moitié des réactions PCR sur l'extrait témoin négatif amplifient, bien que faiblement, un fragment ce qui signifie que de l'ADN se retrouve naturellement dans l'extrait final même en absence d'*At*-SBP40c surexprimée. Cependant dans l'extrait où elle est surexprimée, l'amplification est plus efficace dans 10 des 15 fragments mitochondriaux testés ce qui indique une quantité d'ADN entraîné plus grande. Le fait que les résultats varient selon les régions de l'ADN mitochondrial considérées est encourageant car cela suggère une certaine spécificité de fonction. Cependant il n'est pas exclu que l'ADN co-purifié avec *At*-SBP40c ait été fixé par la protéine après le broyage des plantes et non dans la mitochondrie. Il faudra donc encore vérifier que l'ADN retenu soit préférentiellement constitué d'ADN mitochondrial. Si tel était le cas, cette approche pourrait permettre un accès rapide à des informations essentielles concernant la fonction des protéines de la famille.

2.4 Conclusion

De nombreuses informations ont été apportées au cours du paragraphe 2 concernant le type d'acides nucléiques fixés par les protéines *At*-SBP40. Parmi celles-ci, la plus importante est sans doute le fait qu'elles semblent spécifiques de l'ADNsb. D'autre part, une approche a été initiée en vue de découvrir les régions des génomes mitochondriaux et chloroplastiques nécessitant leur présence. Si une telle spécificité pour certaines régions génomiques était avérée, leur recherche représenterait sans doute l'un des champs d'investigations les plus intéressants pour l'avenir.

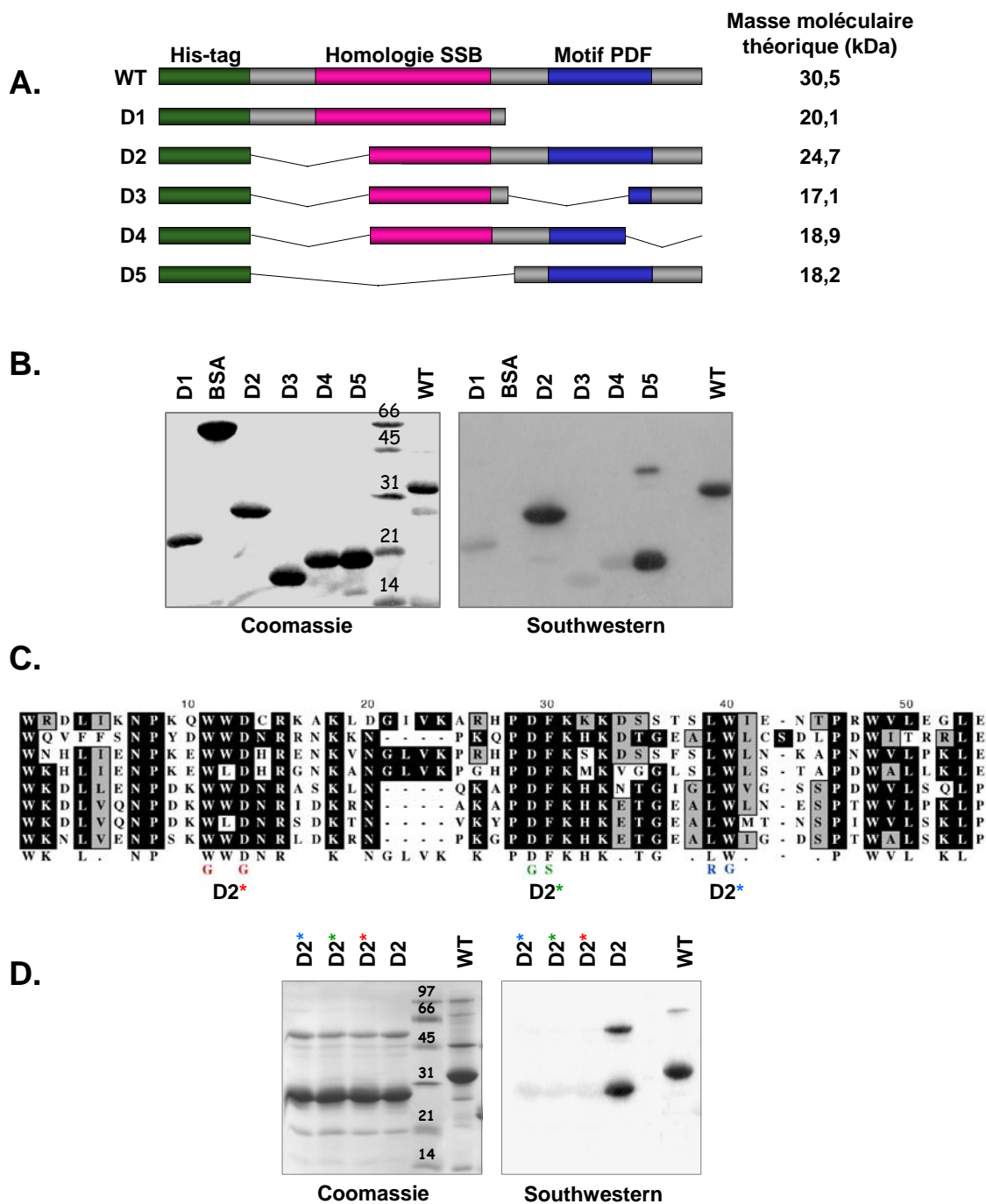


Figure r3.20: Caractérisation du domaine de fixation aux acides nucléiques d'*At-SBP40a*

A: Représentation schématique des différentes formes modifiées de *At-SBP40a*. La séquence codant pour la protéine mature (comportant selon les programmes de prédiction d'adressage les acides aminés 69 à 372) a été clonée dans le vecteur pRSETc à l'aide des sites *Hind*III et *Bam*HI. Les mutants de délétion ont été obtenus par PCR sur ce clone, à l'aide d'oligonucléotides flanquant les régions délétées et contenant des sites de restriction. Les régions délétées sont les suivantes: acides aminés 167 à 372 pour le mutant D1, 38 à 89 pour D2, 38 à 89 et 171 à 372 pour D3, 38 à 89 et 205 à 372 pour D4 et enfin 38 à 166 pour D5. **B:** SDS-PAGE sur *At-SBP40a* modifiée: coloration du gel et southwestern avec une sonde ADNs_b (amorce S1 marquée en 5'P) **C:** Alignement des motifs PDF des *At-SBP40a*, b et c. Les trois séries de mutations ponctuelles affectant D2 sont indiquées. **D:** SDS-PAGE sur la forme D2: coloration et southwestern.

3 Etude du domaine de fixation aux acides nucléiques :

Les programmes informatiques identifient dans la partie centrale des protéines *At*-SBP40 une séquence peptidique caractéristique des SSB. Cette classe de protéines est l'une de celles présentant la plus grande variabilité en termes de séquence primaire. Il nous a donc paru nécessaire de confirmer ces données par des moyens expérimentaux, avec à terme l'objectif de déterminer les acides aminés essentiels à la fonction. Cette étude a été menée par southwestern sur *At*-SBP40a qui présente l'avantage de posséder la structure primaire la plus simple avec un seul motif PDF et dont nous disposons en quantité bien que sous forme insoluble.

L'approche choisie consiste à vérifier par southwestern l'effet de différentes délétions dans la séquence d'*At*SBP40a sur son activité de fixation de l'ADN. La **Figure r3.20 A et B** présente les schémas des séquences protéiques et les données de southwestern. Seuls les polypeptides D2 et D5 ont une affinité pour la sonde équivalente à celle de la protéine entière. Ils ont en commun le motif PDF, alors que la séquence homologue aux SSBs est totalement absente chez D5. D'autre part, le mutant D1 comporte cette dernière séquence dans son intégralité mais ne possède qu'une activité de fixation résiduelle. Il semble donc que se soit le motif PDF qui confère aux *At*-SBP40 leur capacité de fixation aux acides nucléiques. En vue de tester cette hypothèse, nous avons procédé à des mutations plus ponctuelles. Cela a été réalisé en mutant des acides aminés parmi les plus conservés du domaine, d'après les informations apportées par l'alignement des séquences de tous les motifs PDF de la famille *At*-SBP40 (voir **Figure r3.20 C**). Le vecteur pRSET portant la construction codant pour D2 a servi de matrice aux trois clones mutants synthétisés (D2a, D2b et D2c). Pour chacun, deux acides aminés identiques chez toutes les protéines de la famille ont été transformés en résidus différant du point de vue de leur encombrement stérique et de leur polarité. L'analyse par southwestern de ces mutants ponctuels révèle qu'aucun d'eux n'est plus capable de fixer la sonde (**Figure r3.20 D**). Cela confirme donc que le motif PDF est responsable de l'activité de fixation de l'ADNsb.

Les données concernant *At*-SBP40a ont été obtenues au moyen du southwestern, méthode moins fiable que le gel retard. Toutefois cinq modifications différentes du domaine PDF dont trois ne touchent que deux acides aminés entraînent toutes une perte de la capacité à fixer les acides nucléiques. Ces résultats semblent suffisants pour affirmer que le motif PDF constitue la région responsable de la fixation des acides nucléiques. Cependant, les résultats obtenus avec les mutants ponctuels ne constituent pas la démonstration que les acides aminés mutés interagissent directement avec le ligand. En effet, il est probable que la nature des modifications soit suffisante pour modifier la structure spatiale du motif PDF. La vérification de son intégrité passe par

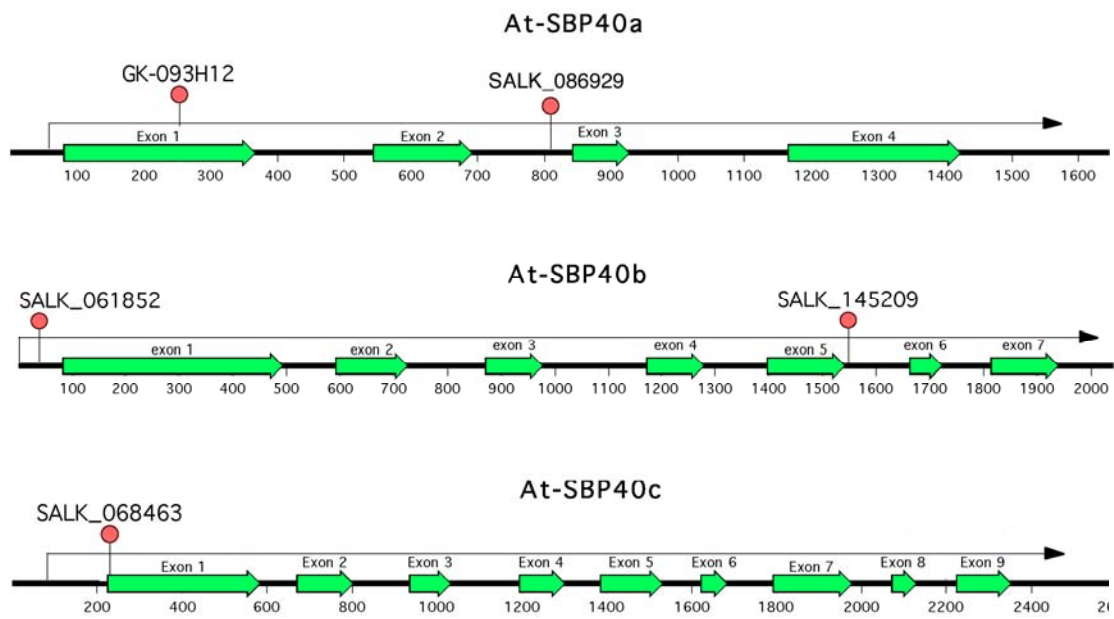


Figure r3.21: Position des insertions d'ADN-T dans les gènes des *At-SBP40*

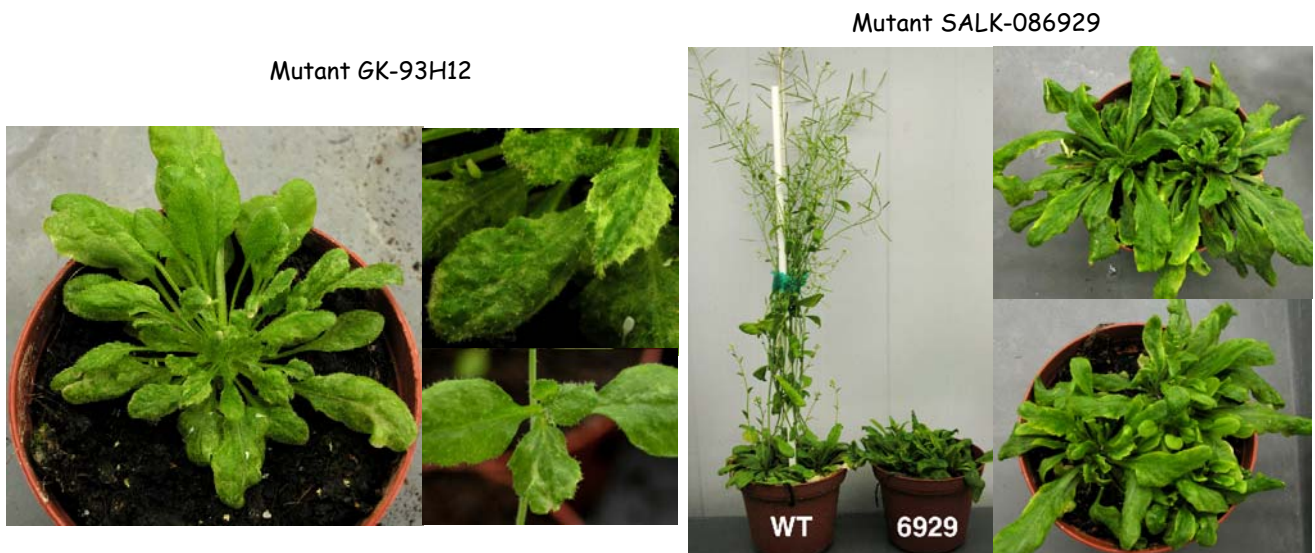


Figure r3.22: Phénotypes des plantes portant une insertion dans le gène codant pour *At-SBP40a*
 Les deux mutants sont caractérisés par un même retard de croissance et des feuilles déformées et bigarrées.

des expériences de type dichroïsme circulaire, qui ne nous sont pas accessibles en raison de l'insolubilité de cette protéine.

4 Mise en évidence de la participation d'At-SBP40a à la maintenance du génome mitochondrial :

Comme nous l'avons vu au Chapitre II, des mutants d'insertion ont été obtenus pour les différentes protéines de la famille (deux pour *At-SBP40a* et *b*, un pour *At-SBP40c*). La position des insertions est représentée **Figure r3.21**. Bien que les ARNm complets ne soient pas détectables par RT-PCR, confirmant l'abolition de l'expression des transcrits respectifs, les premières générations d'homozygotes ne présentaient pas de phénotype visible. A partir des générations T4-T5, les deux mutants affectés dans l'expression d'*At-SBP40a* (M1=Gaby-kat 093H12 et M2=Salk-086929, tous deux dans un écotype Col-0) ont commencé à présenter des perturbations : retard de croissance, déformation des fleurs voire stérilité dans certains cas et production de graines non viables. Une distorsion des feuilles ainsi que des taches jaunâtres, touchant en particulier les feuilles colinéaires, sont également apparues (voir **Figure r3.22**). Ces derniers symptômes étaient pour nous particulièrement intéressants, car ils ressemblent à ceux décrits pour les mutants *mshI* (Abdelnoor *et al.* 2003). En effet, ceux-ci présentent dans leur génome mitochondrial des quantités anormalement grandes de certaines molécules d'ADN recombinées qui n'existent chez la souche sauvage qu'à l'état sub-stœchiométrique. Nous avons donc testé les mutants M1 et M2 pour une anomalie caractérisée dans les mutants *mshI* et résultant d'une recombinaison dans la région 3' du gène *atp9*. Un southern blot réalisé à l'aide d'ADN total de fleurs des plantes les plus affectées et d'une sonde visant la région codante d'*atp9* (voir **Figure r3.23A**) met en évidence des changements de stœchiométrie et l'apparition de nouveaux fragments chez les plantes mutantes, par rapport au sauvage. Un test par PCR a également confirmé que ces changements résultent d'une recombinaison dans la même région que celle affectant les mutants *mshI* (Sakamoto *et al.* 1996). Ces derniers, issus d'un écotype Landsberg, doivent leur phénotype à l'amplification d'une molécule subgénomique. Dans le contexte Landsberg, cette molécule doit son origine à une recombinaison provenant de la présence d'une répétition de 7 paires de bases. Les mutants M1 et M2 proviennent par contre d'un écotype Col-0, chez lequel le génome mitochondrial n'a pas été séquencé. En se basant sur la séquence mitochondriale déterminée dans l'écotype C24 (Unsold *et al.* 1997) il a été possible à l'aide d'expériences de PCR de proposer un modèle des recombinaisons caractéristiques des mutants M1 et M2. On constate que la recombinaison met en jeu une séquence de 248 pb (RA), présente en 3' d'*atp9* et répétée dans la région en amont de *cox3* (voir **Figure**

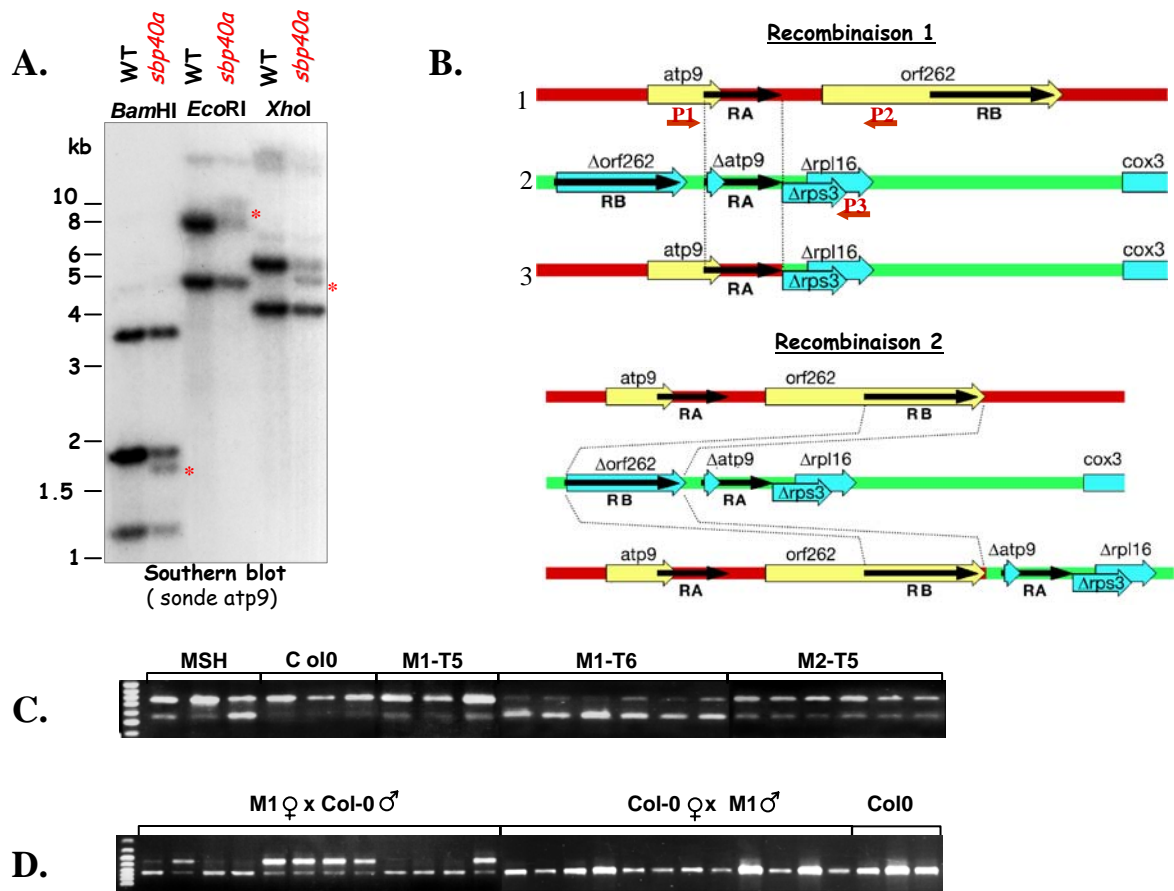


Figure r3.23: Mise en évidence changements de stœchiométrie dans l'ADN mitochondrial chez les mutants d'insertion de *At-SBP40a*

A- L'ADN total de plante sauvage (wt) et du mutant d'insertion *93H12* (*sbp40a*) digérés par *Bam*HI, *Eco*RI et *Xho*I a été hybridé avec une sonde *atp9*. Les fragments nouveaux apparaissant chez le mutant sont indiqués.

B- Recombinaison 1. Dans l'écotype *Col-0* la même séquence RA se trouve en 3' d'*ATP9* (1) et en amont de *cox3* (2). La recombinaison (3) autour de RA est favorisée chez les mutants de *SBP40a*. Les amorces P1 et P2 permettent d'amplifier par PCR la forme non recombinée, alors que P1 et P3 détectent la forme recombinée. Recombinaison 2: elle met en jeu la séquence RB dans l'environnement voisin de RA et a été mise en évidence dans une seule plante mutante.

C- Test de la recombinaison 1 par PCR: le fragment de 748 nt obtenu avec les amorces P1/P2 correspond à la forme non recombinée et le fragment de 546 nt (amorces P1/P3) à la forme recombinée. Illustration des résultats obtenus avec le mutant *93H12* (M1) aux générations T5 et T6, le mutant *585929* (M2) à la génération T5, le mutant *546763* (*MSH*) et le sauvage *Col-0*.

D- Analyse de la recombinaison 1 dans des plantes issues des retro-croisements entre *93H12* (M1) et *Col-0*

r3.23 B). Cependant dans une seule plante nous avons amplifié un fragment montrant qu'il existe également dans de rares cas une amplification mettant en jeu la région répétée RB (voir **Figure r3.23 B**).

Pour avoir accès à la proportion de molécules recombinées par rapport aux molécules sauvages, nous avons employé un test par PCR (voir **Figure r3.23 C**) au moyen de trois amorces (schématisées en **Figure r3.23 B**). Chez des plantes sauvages *Col-0*, le produit de PCR issu de la molécule recombinée (P1-P3, bande du bas) est presque invisible. A titre de comparaison nous avons également testé un mutant d'insertion (Salk 546763) dans le gène *MshI* et qui présente clairement l'ADN recombiné. A la génération T5 du mutant M1 le fragment recombiné est présent et devient majoritaire à la génération suivante, même si seulement 5 plantes sur 21 présentent un phénotype anatomique ou développemental clairement visible à ce stade. Les résultats sont identiques pour la génération T5 du mutant M2.

Nous avons également réalisé des croisements entre des mutants de la génération T5 et des plantes sauvages (voir **Figure r3.23 D**). En utilisant du pollen provenant d'une plante mutante pour féconder une plante sauvage, nous pouvons constater que les 39 plantes filles testées possèdent un ADNmt identique à celui des plantes sauvages. Cette observation n'est pas surprenante, en raison de l'hérédité maternelle avérée de la mitochondrie chez *Arabidopsis* (Martinez-Zapater *et al.* 1992) (Sakamoto *et al.* 1996). Le croisement réciproque revient à replacer des mitochondries renfermant de l'ADNmt mutant dans un contexte nucléaire renfermant de nouveau un gène *SBP40a* fonctionnel. Dans ce cas, on observe pour 15 des 19 plantes issues de la descendance une présence toujours anormalement haute de la molécule recombinée ce qui montre que la réactivation de la fonction *SBP40a* ne semble pas suffisante pour restaurer totalement la stœchiométrie de l'ADN mitochondrial tout au moins dans la première génération après la réinsertion du gène. De nouveaux croisements des plantes analysées dans la **Figure r3.23 D** avec du pollen de *col 0* ainsi que le suivi de ces mêmes plantes après autofécondation sont en cours pour déterminer si la restauration de l'ADN mitochondrial se fait dans les générations suivantes où si l'effet de la mutation perdure.

5 Etude de l'expression dans les organes :

L'étude de l'expression des gènes codant pour les protéines de la famille *At-SBP40* est rendue délicate par le fait que leurs ARNm ne sont pas suffisamment abondants pour être détectés par northern blot. Ceci est illustré par leur faible taux de représentation dans les banques d'ESTs (3 séquences pour *At-SBP40a*, 7 pour *At-SBP40b*, 9 pour *At-SBP40c* et 1 pour *At-SBP40d*). L'analyse de données publiées de

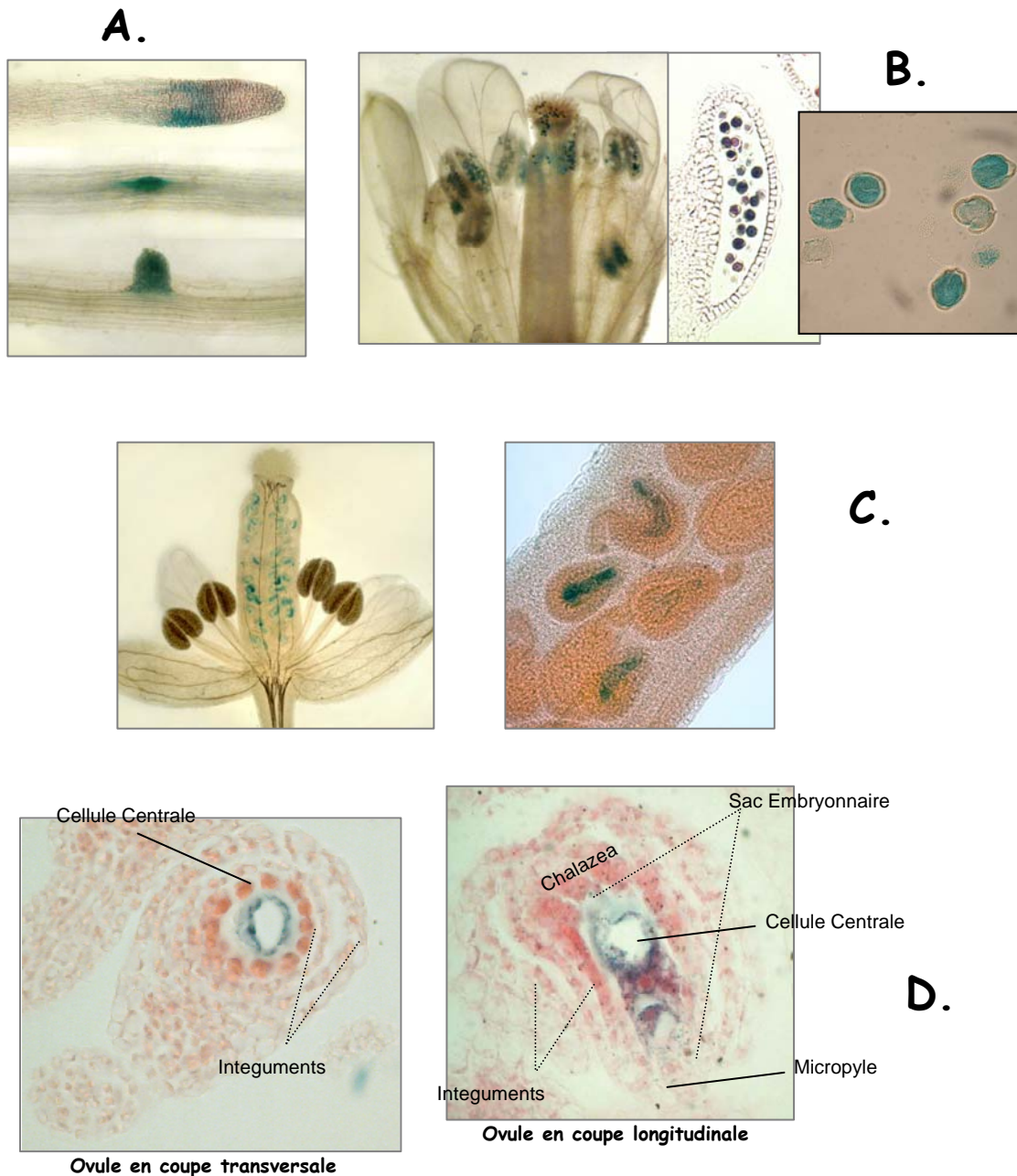


Figure r3.24: Zones d'expression d'*At-SBP40a* visualisées à l'aide du gène rapporteur *GUS*

Illustration des résultats obtenus avec plusieurs lignées de plantes transformées avec une fusion *GUS*-intergène précédant *At-SBP40a*. A- Plantules de 18 jours: expression dans les racines primaires (observation rare) et au démarrage des racines secondaires. B- Fleurs matures: expression dans les grains de pollen. C- Fleurs jeunes: expression dans les ovules. D- Coupes longitudinale et transversale d'ovules: expression dans le sac embryonnaire et la cellule centrale.

« microarrays » n'a pas révélé la présence d'évènements de régulation développementaux particuliers probablement parce que le niveau basal d'expression est trop faible pour que l'on puisse détecter d'éventuelles variations. Seule la RT-PCR permet de détecter les messagers des *At-SBP40* et nous avons cherché à évaluer leurs quantités relatives selon différents organes de la plante : feuilles, fleurs, tiges, racines par RT-PCR quantitative. Cette technique est délicate à mettre en œuvre et nous n'avons pas toujours obtenu des réactions d'amplification présentant des efficacités suffisantes pour donner des résultats comparatifs. Dans l'ensemble nous n'avons pas perçu de différences remarquable dans l'expression d'un organe à l'autre.

Les analyses menées à ce point n'étaient pas susceptibles de mettre en évidence d'éventuelles variations dans l'expression des *At-SBP40* si ces protéines ont une action délimitée dans l'espace (nombre restreint de cellules) et le temps (moment précis de développement). Pour vérifier cette éventualité, nous avons entrepris l'étude de l'expression des gènes au moyen de plantes transgéniques transformées de façon stable à l'aide de fusion promoteurs-gène rapporteur *GUS*.

5.1 Le cas d'*At-SBP40a*

La séquence située en 5' d'*At-SBP40a* (1074 nucléotides entre le codon stop du gène précédent, *Atlg47719* et le codon d'initiation de *Atlg47720*) a été fusionnée avec le gène codant pour la β -glucuronidase (*GUS*). Chez des plantes transformées âgées de quelques jours, l'expression du gène *GUS* est tout d'abord visible dans les zones méristématiques donnant lieu aux racines secondaires (voir **Figure r3.24 A**) et disparaît lorsque celles-ci s'allongent. Dans certaines plantes, la région située dans la zone d'élongation cellulaire, derrière l'apex de la racine principale apparaît également colorée. Aucune expression du marqueur n'est visible dans les tissus végétatifs.

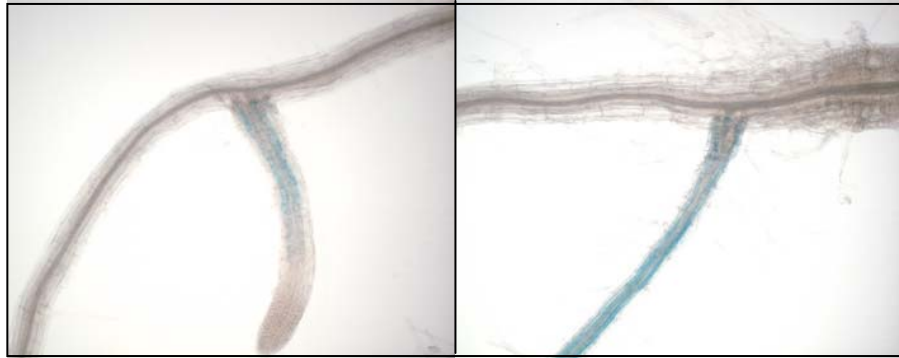
Dans les fleurs matures, l'expression du marqueur *GUS* est visible dans les grains pollen (voir **Figure r3.24 B**). Enfin, dans des fleurs plus jeunes, une coloration intense est présente dans les ovules non encore fécondés (voir **Figure r3.24 C**). Cette coloration semble plus précisément localisée au niveau du sac embryonnaire, présentant une cellule centrale nettement colorée (voir **Figure r3.24 D**).

L'expression du gène codant pour *At-SBP40a* apparaît donc limitée à la fois dans le temps et dans l'espace, et spécifique des tissus dits de transmission, les zones méristématiques et les gamètes.

5.2 Le cas d'*At-SBP40b*

Comme dans le cas d'*At-SBP40a*, la séquence correspondant à l'intergène en amont d'*At-SBP40b* a été fusionnée avec le gène *GUS* et, bien que les résultats soient encore préliminaires, l'expression de ce gène a été essentiellement trouvée dans les

A



B



C

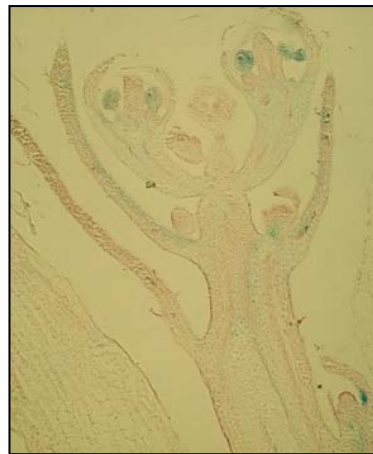


Figure r3.25: Zones d'expression d'*At-SBP40b* visualisées à l'aide du gène rapporteur *GUS*

Illustration des résultats obtenus avec plusieurs lignées de plantes transformées avec une fusion *GUS*-intergène précédant *At-SBP40b*. A- Racines secondaires: expression au long de la racine lorsque celle ci dépasse 1 cm. B- Ovaire: expression dans les funicules. C- Bouton floral: expression rare dans les anthères, la face intérieure des feuilles primordiales.

Figurer3.26 : voir page suivante ...□□

..

racines et les fleurs. En ce qui concerne les racines (**Figure r3.25 A**), l'expression est généralement trouvée dans les premières racines secondaires, mais seulement lorsqu'elles ont une certaine taille (environ 1 cm) et la coloration s'étend sur presque toute la longueur (l'extrémité n'est pas colorée). Dans une même plantule toutes les racines secondaires ne sont pas colorées, ce qui montre que l'expression est limitée dans le temps. Dans les fleurs la coloration est détectable au niveau des funicules qui supportent les ovules (**Figure r3.25 B**). Par ailleurs, dans un bouton floral situé à l'aisselle d'une feuille nous avons trouvé de nombreux tissus colorés: anthères, face intérieure de feuilles primordiales (**Figure r3.25 C**).

L'expression du gène codant pour *At-SBP40b* semble concerner des régions de la plantes plus étendues qu'*At-SBP40a*, mais toujours des tissus en croissance. Nous n'avons pas encore examiné un nombre suffisant de plantes différentes mais l'expression apparaît également limitée dans le temps et l'espace.

5.3 Le cas d'*At-SBP40c*

Dans les plantes que nous avons analysées, la séquence intergénique située en amont du gène *At-SBP40c* dirige l'expression du gène rapporteur *GUS* dans les fleurs et, rarement, à la base des feuilles, probablement au niveau du départ d'une nouvelle hampe (**Figure r3.26 A**). Aucun signal n'est visible au niveau des racines quel que soit leur degré de développement. Dans les fleurs la coloration est visible dans les ovules et probablement au niveau de la cellule centrale comme le montre la **Figure r3.26 B** et **C**. Là encore l'expression apparaît ponctuelle et limitée dans le temps ce qui constitue une caractéristique commune aux *At-SBP40*.

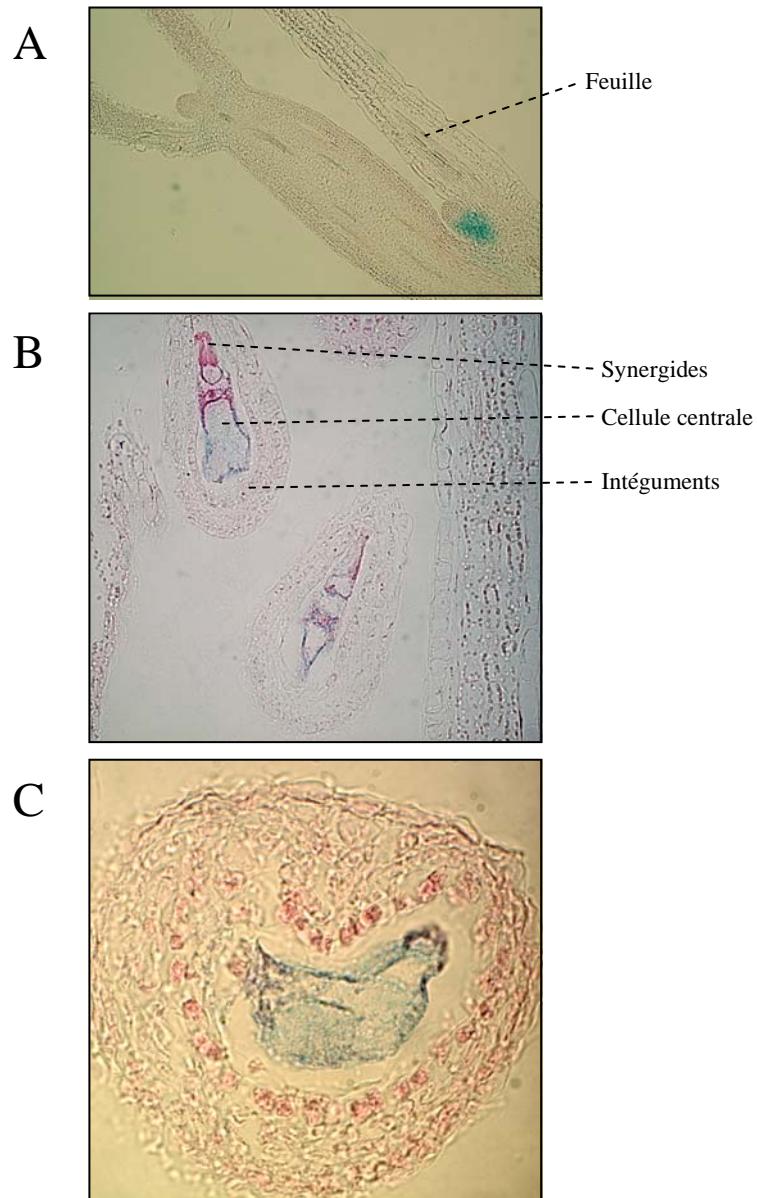


Figure r3.26: Zones d'expression d'*At-SBP40c* visualisées à l'aide du gène rapporteur *GUS*

Illustration des résultats obtenus avec plusieurs lignées de plantes transformées avec une fusion *GUS*-intergène *At-SBP40c*. A- Base de feuille: expression au niveau d'un bourgeon. B- Ovules coupe longitudinale: expression dans le sac embryonnaire et de la cellule centrale. C- Ovule coupe transversale.

DISCUSSION ET PERSPECTIVES

Cette dernière partie sera consacrée à la mise en commun de toutes les données collectées au cours ma thèse. L'objectif est de livrer une analyse globale et objective du travail réalisé, en montrant quelles avancées importantes ont été réalisées et en cherchant à définir les parties du projet nécessitant un approfondissement en vue d'une meilleure compréhension du rôle des protéines *At-SBP40* dans les organites des plantes supérieures.

1 Le motif PDF

Le premier critère ayant permis de considérer les protéines *At*-SBP40 comme appartenant à une même famille était l'homogénéité retrouvée dans l'agencement de leur séquence primaire. Elles sont toutes composées d'un domaine homologue à l'« OB-fold » des SSB bactériennes dans leur partie centrale et d'un ou plusieurs motifs PDF en C-terminal. Elles présentent de plus une similarité de leurs séquences globales allant de 30 à 50 %, pour les protéines prises deux à deux. Ce taux monte jusqu'à 50 % d'identité si l'on ne considère plus que les parties les plus conservées de ces séquences, c'est à dire les domaines PDF. À cela est venu s'ajouter la capacité de fixation de l'ADNsb, qui a été vérifiée expérimentalement par deux techniques différentes, pour *At*-SBP40a et b. Même si les résultats obtenus avec *At*-SBP40a sont moins clairs que ceux provenant du gel retard, il paraît vraisemblable au vu des ces données, que toutes les protéines de la famille sont capables de fixer l'ADNsb. Si tel est le cas, la position de leur domaine de fixation à l'ADNsb doit être conservée sur la séquence primaire. Il s'agit donc soit du domaine « SSB-like » situé dans la partie centrale, soit du motif PDF. Les expériences de gel retard utilisées pour tester des mutants de délétion de *At*-SBP40b n'ayant pas fonctionné, la seule expérience dont nous disposons pour conclure sur ce point est l'analyse par southwestern du cas de *At*-SBP40a. Cette méthode est malheureusement réputée moins fiable que le gel retard et ce pour des raisons évidentes. La protéine à tester est immobilisée, ce qui lui interdit entre autres d'atteindre un statut d'oligomérisation qu'elle pourrait nécessiter pour sa fonction (alors que les SSB bactériennes fixent justement l'ADNsb sous forme de tétramère). Toutefois, au vu des résultats obtenus en southwestern, il ne fait aucun doute que le motif PDF possède à lui seul cette capacité. La composition de ce domaine, riche en acides aminés aromatiques et basiques, paraît en adéquation avec ces observations. Ces deux types de résidus ont en effet été mis en évidence par plusieurs structures cristallographiques (voir le Chapitre II de l'introduction) comme interagissant directement avec l'ADNsb. De façon surprenante, la région d'*At*-SBP40a homologue à l'« OB-fold » bactérien ne s'est quant à elle pas montrée capable de fixer la sonde en southwestern. Une explication simple à ce paradoxe apparent peut être apporté par l'examen de sa séquence primaire. On peut effectivement s'apercevoir qu'il manque à cette région les résidus équivalents aux Trp 54 et 88 et Phe 60 (Raghunathan *et al.* 2000) de *Eco*ssb aux positions attendues pour qu'ils puissent intervenir dans l'interaction (il s'agit pourtant là des acides aminés sensés être parmi les plus conservés de la séquence primaire des SSB). On peut donc en déduire que les acides aminés impliqués interagissant directement avec l'ADNsb sont portés par le

domaine PDF. Bien entendu, cela ne signifie pas qu'en solution, le domaine central ne soit pas lui aussi impliqué dans la fixation. On peut par exemple imaginer que la structure tridimensionnelle soit organisée autour d'un « OB-fold » central et que le ou les domaines PDF viennent en complément à celui-ci, afin d'augmenter l'affinité en apportant leurs résidus aromatiques et basiques. Une autre hypothèse, selon laquelle le domaine PDF serait capable d'engendrer à lui seul une structure tridimensionnelle capable de fixation est également envisageable. Si tel était le cas, celle-ci serait sans doute très différente de l'aspect de l'« OB-fold ». En effet, la séquence primaire moyenne de ce motif comporte plus de 100 acides aminés, c'est à dire le double du contenu du motif PDF, qui n'a donc pas la longueur nécessaire pour se structurer en tonneau β .

On peut déjà, à ce stade, tirer plusieurs conclusions. La première est que des expériences complémentaires, comme par exemple l'étude de mutants plus ponctuels d'*At*-SBP40b s'avèreront sans doute nécessaires pour déterminer la participation de la région centrale à la fixation, mais qu'elles ne remettront sans doute pas en cause l'implication des motifs PDF. Bien entendu, une réponse définitive à la question de la structure tridimensionnelle de ces protéines ne pourra être apportée que par le biais d'expériences beaucoup plus « lourdes » : RMN ou cristallographie. En assumant le fait que le domaine PDF interagisse directement avec l'ADNsb, on peut se permettre de revenir sur la fixation de l'ADNsb par *At*-SBP40a. En effet, le fait que le domaine PDF soit impliqué modifie quelque peu la façon de voir les choses. *At*-SBP40b, qui possède une affinité relative (ADNsb/ADNdb) très élevée et déterminée de manière fiable (par gel retard), possède deux domaines PDF. Cette observation rend difficilement envisageable le fait qu'une protéine possédant un domaine de moins puisse de son côté être capable de fixer l'ADNdb.

Prises dans leur ensemble, ces données nous semblent suffisantes pour permettre de considérer que les protéines de la famille *At*-SBP40 ont une fonction dans les mécanismes de la maintenance de l'ADN, par le biais de leur domaine PDF qui est responsable de leur capacité à fixer l'ADNsb. Un dernier commentaire s'impose tout de même à propos de ce motif, concernant un homologue des *At*-SBP40, la protéine RB38 de *Chlamydomonas* qui a été identifiée comme liant spécifiquement l'extrémité 5' du transcrit chloroplastique *psbA* (Barnes *et al.* 2004). Le haut pourcentage d'acides aminés identiques entre les domaines PDF de cette protéine et ceux des *At*-SBP40 excluant une évolution convergente, on peut imaginer que les *At*-SBP40 aient adapté leurs domaines aux particularités des chondriomes embryophytes, qui nécessitent sans doute plus d'enzymes pour la maintenance de l'ADN que ceux des algues. Ainsi, RB38 doit probablement sa spécificité à une composition différente en acides aminés

impliqués dans l'interaction, à la manière du cas de Pot1 (Baumann et Cech 2001), qui fixe l'ADNsb de manière séquence spécifique au moyen d'un « OB-fold ».

2 Points communs et divergences entre les At-SBP40 et les SSB

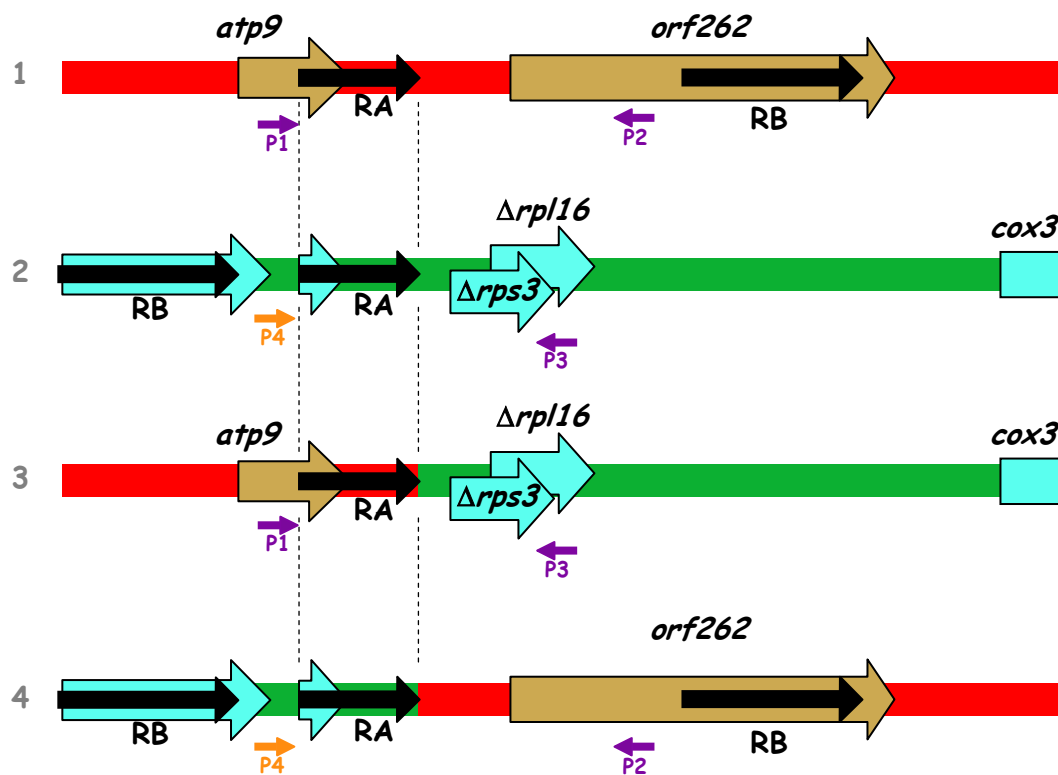
L'analyse des banques d'EST présentée au cours du Chapitre I a permis de constater que les protéines SBP40 sont uniquement présentes et exprimées chez les plantes. Cela permet de penser que ces protéines sont impliquées dans des étapes spécifiques de la maintenance des génomes des organites végétaux, l'ensemble des résultats obtenus leur attribuant en effet une localisation dans ces compartiments. Nous ne reviendrons pas ici sur la localisation au cas par cas qui a déjà été discutée lors de la présentation des résultats. En tenant compte des données fonctionnelles, seule *At-SBP40a* apparaît définitivement localisée: il s'agit d'une protéine mitochondriale. Des expériences complémentaires semblent donc nécessaires dans les autres cas. La recherche de la localisation a d'autre part mis en évidence les limites de nos anticorps, qui reconnaissent l'ensemble des protéines. L'obtention d'anticorps plus spécifiques peut donc également être classée dans la catégories des expériences à prévoir dans l'optique de la poursuite de cette étude.

Une caractéristique marquante de l'organisation de la famille *At-SBP40* est la présence à un nombre de copies variables du domaine PDF selon les protéines. Cette situation rappelle celle des RPA eucaryotes ou des SSB archéobactériennes, qui sont elles aussi composées de plusieurs domaines de fixation de l'ADNsb. Différents critères (par exemple les localisations différentes observées avec le gène *GUS* ou encore le phénotype observé seulement pour *At-SBP40a*) rendent toutefois peu probable le fait que les *At-SBP40* puissent être organisées sous forme d'hétérodimère. Les quatre protéines agiraient donc séparément dans le chloroplaste ou la mitochondrie. Elles peuvent de plus avoir des fonctions différentes ou relativement proches (en ayant par exemple de faibles différences d'affinité ou de spécificité de séquence). Cette dernière hypothèse expliquerait l'absence de phénotype constatée pour la plupart des mutants, qui serait ainsi due à la complémentarité des *At-SBP40*.

Cette question concernant leur fonction amène à essayer de comprendre ce qui distingue cette classe de protéines de celle des SSB. Elles possèdent en effet certaines propriétés (une séquence primaire qui semble dépourvue de domaine catalytique ; une capacité de fixation à n'importe quelle séquence, testée *in vitro* et *in vivo* ; un K_D de 2×10^{-9} M et une faible coopérativité apparente pour *At-SBP40b*) qui auraient tendance à les affilier à ce type de protéines. La principale différence que l'on ait pu observer réside dans leur faible abondance et leur expression spécifique de certains tissus. En effet, les SSB sont présentes à un grand nombre de copies dans les cellules, et même

dans des cas extrêmes (comme la réponse SOS chez *E. coli*), ne voient pas leur niveau d'expression augmenter de façon conséquente (Ferrari *et al.* 1994). Les expériences de fusions au gène *GUS* conduites avec les SBP40 paraissent démontrer exactement l'inverse. Cela devient compréhensible lorsque l'on prend en compte le fait qu'une SSB homologue à la protéine de *E. coli* a déjà été identifiée dans la mitochondrie d'*Arabidopsis*. Celle-ci semble en tous points identique à la protéine bactérienne, jusqu'à sa capacité de stimulation de *EcRecA* démontrée lors d'expériences *in vitro* (Edmondson *et al.* 2005).

L'ensemble de ces données permet donc d'envisager que le rôle des protéines *At*-SBP40 ne soit pas de protéger et de structurer l'ADNsb, mais qu'elles seraient en charge d'une mission probablement plus spécifique, dans les tissus dans lesquels leur expression est induite. Hormis le cas d'*At*-SBP40a, la seule piste que nous ayons pour l'instant à ce propos est l'expérience d'invasion de brins réalisée en gel retard avec *At*-SBP40b. Nous avons vu qu'elle favorise l'échange de brins à de faibles concentrations de sels, de manière ATP indépendante. Cette expérience ne permet pas de conclure qu'*At*-SBP40 ait une fonction de type RecA, mais on peut envisager qu'elle ait un rôle de médiateur, comme celui de Rad52 (Sung 1997) ou RecFOR (Morimatsu et Kowalczykowski 2003), en interagissant directement avec la protéine RecA mitochondriale (qui a été identifiée chez *Arabidopsis* (Khazi *et al.* 2003)), ou en retenant le brin non apparié de la boucle D. Nous avons également réalisé des expériences de compétition utilisant des homopolymères, qui ont montré une préférence de cette protéine pour les séquences riches en cytosines. Ce type de résultats peut s'avérer trompeur, le meilleur exemple d'artéfact étant sans doute celui de la SSB bactérienne. En effet, dans les mêmes conditions, *Ecosb* a une affinité 10^{18} fois supérieure pour le poly (dT) par rapport au poly (dC) (Ferrari *et al.* 1994), tout en ayant, *in vivo*, pour fonction de se fixer sans spécificité de séquence. Si on émet l'hypothèse que *Ecosb* se fixe dans un premier temps aux séquences riches en thymines, c'est sans doute sa coopérativité élevée qui a pour effet de permettre le recouvrement complet de larges portions d'ADNsb. Dans notre cas, en utilisant un oligonucléotide de 64 pb, la présence bien distincte d'une bande intermédiaire dans l'expérience de la **Figure r3.11** suggère que *At*-SBP40b soit quant à elle peu coopérative, voire pas du tout. Ce résultat demande à être confirmé par l'utilisation d'homopolymères, de façon à éliminer un biais éventuel dû à la séquence de la sonde employée dans ce cas. Il permet toutefois de suggérer que la spécificité de séquence de *At*-SBP40b soit transposable *in vivo*, ce qui ferait d'elle une protéine en charge du contrôle de régions particulières du génome. L'expérience la plus prometteuse en vue de déterminer si les *At*-SBP40 ont une telle spécificité est la co-purification par affinité, utilisant des plantes surexprimant les différentes protéines en fusion avec le système



Tests à réaliser pour déterminer le mécanisme moléculaire dans lequel intervient At-SBP40a

Les amorces employées pour l'instant sont représentées en violet (P1 à P3), celle permettant de tester le nombre de copies des molécules 2 et 4 apparaît en orange (P4). Le même type de PCR à trois oligonucléotides que celui utilisé jusqu'à présent, mais en utilisant cette fois P2, P3 et P4, permet d'observer la stoechiométrie des molécules 2 et 4. En la comparant à celle des molécules 1 et 3, on peut déterminer si le phénomène observé inverse les brins ou s'il initie la replication. Cette initiation impliquerait l'invasion de la répétition RA de la molécule 1 au niveau de la RA des molécules 2 ou 3.

TAP. Ce type de technique permettra peut être de mieux comprendre le rôle de la famille *At*-SBP40 dans les organites des plantes supérieures.

3 Quelle fonction pour *At*-SBP40a ?

Nous avons pu, au moyen de mutants d'insertion, attribuer une fonction à l'une des protéines de la famille, *At*-SBP40a, dans le contrôle de la stoechiométrie des molécules subgénomiques. Ce rôle paraît parfaitement adapté au profil de cette protéine, ayant une préférence pour l'ADNsb. On peut en effet supposer que son mode d'action consiste à se fixer sur les régions susceptibles d'être instables, et d'empêcher ainsi l'intervention de RecA. L'absence de *At*-SBP40a provoque en effet l'apparition en grande quantité d'une molécules chimériques, qui sont normalement présentes à un faible nombre de copies dans les cellules sauvages d'écotype Col-0. La présence de cette molécule chimérique, contenant le gène *atp9*, avait été préalablement caractérisée dans un autre système, à savoir les mutants *chm*. Ces mutants proviennent toutefois d'un écotype différent (Landsberg), dans lequel la recombinaison implique une répétition de sept paires de bases seulement. L'analyse par PCR des molécules chimériques intervenant dans notre cas a révélé que dans l'écotype Col-0, la répétition engendrant le réarrangement est bien plus grande (avec 248 pb). Cette différence au sein d'une même espèce reflète bien la diversité des chondriomes végétaux. On pourrait l'expliquer par un événement de recombinaison non conventionnelle impliquant la répétition de 248 pb qui aurait engendré la molécule subgénomique retrouvée dans l'écotype Landsberg. Chacune de deux molécules, conservée sous forme substoechiométrique, étant redondantes pour la majorité de leurs séquences, aurait été éliminée différemment dans les écotypes Landsberg et Columbia.

Le chondriome de Col-0 contient donc deux copies du gène *atp9*, l'une sur une molécule normale, présente chez le sauvage en quantité élevée, et l'autre sur une molécule chimérique, présente chez le sauvage en quantité substoechiométrique. Etant donné la « simplicité » de cette situation, dans laquelle une seule répétition intervient, on ne peut pas pour l'instant se prononcer sur le mécanisme faisant intervenir *At*-SBP40a. Il peut en effet s'agir d'un mécanisme de recombinaison ou de réplication recombinaison-dépendante. Cette situation relativement simple a par contre des avantages : elle devrait pouvoir nous permettre de par différents moyens, d'approfondir la fonction de cette protéine. Pour l'instant, nous n'avons utilisé pour déterminer les quantités respectives des molécules normales et chimériques, que les régions situées en 3' de la répétition (voir **Figure**). Si l'on examine également le taux de persistance de la région 5' de la molécule 2, on devrait pouvoir différencier un processus de réplication recombinaison-dépendant d'un processus de recombinaison.

En effet, la recombinaison classique inverse les brins, conduisant à l'apparition de la molécule 4. Si celle-ci apparaît, il s'agirait donc d'un processus de recombinaison et si elle n'apparaît pas il s'agirait plutôt d'un mécanisme faisant intervenir la réplication recombinaison-dépendante.

Un autre aspect vérifiable est l'adaptabilité de *At*-SBP40a à tous les types de séquences. Nous n'avons détecté que dans un seul cas, parmi les nombreuses plantes testées, une recombinaison faisant intervenir la répétition RB. Nous avons déjà mentionné le fait qu'*At*-SBP40b possède en gel retard une préférence pour les séquences riches en cytosines et une telle préférence d'*At*-SBP40a pourrait expliquer son mode d'action. Il paraît donc important de vérifier l'aspect général du chondriome des plantes mutantes, afin de déterminer si d'autres répétitions sont impliquées dans des recombinaisons apparaissant à un taux important. Cela sera fait prochainement par l'analyse comparative des transcrits présents chez les mutants et dans la souche sauvage ce qui devrait permettre de détecter les réarrangements qui affecteraient des zones transcrites. L'importance du contrôle de la recombinaison homologue prend tout son sens lorsque l'on s'intéresse à la séquence publiée du chondriome d'*Arabidopsis*. Celui-ci contient en effet 22 paires de répétitions 100 % identiques de plus de 100 pb (Unsel *et al.* 1997). Cela permet donc de comprendre qu'un contrôle relâché du processus d'invasion de brins peut avoir des conséquences dramatiques sur les génomes mitochondriaux. Différentes autres techniques sont encore envisageables afin d'affiner la compréhension du mécanisme. On peut par exemple utiliser les plantes transgéniques surexprimant *At*-SBP40a, qui n'ont pour l'instant été employées qu'à des fins « techniques », c'est à dire pour des expériences d'immunoprécipitation ou d'immunodétection. Dans celles-ci, l'expression est contrôlée par le promoteur constitutif 35S, ce qui permettra, si la surexpression a un effet, d'obtenir plus rapidement un phénotype moléculaire. On peut enfin regretter l'échec des cultures cellulaires transgéniques. La culture de cellules a en effet été d'une grande utilité pour l'observation des mitochondries et de leur contenu au cours des dernières années, ayant notamment permis de décrire les modes de réplication du chondriome de *C. album* (Backert et Borner 2000).

Le second résultat important concernant *At*-SBP40a est le fait que, très peu exprimée dans la plante entière (indétectable par northern et par western), elle le soit fortement dans les tissus transmetteurs (les méristèmes racinaires secondaires et les gamètes). Comme nous l'avons largement évoqué dans l'introduction, ces tissus sont responsables du contrôle de la transmission de l'intégralité du patrimoine génétique mitochondrial, vraisemblablement via un processus spécifique de réplication. Il s'agit à notre connaissance de la première fois qu'une protéine impliquée dans le contrôle de la stoechiométrie des molécules subgénomiques est identifiée comme exprimée

spécifiquement dans les tissus transmetteurs. Un dernier point à éclaircir concerne l'expression dans le pollen d'*At-SBP40a*. Cela paraît surprenant au vu de l'hérédité maternelle de la mitochondrie chez *Arabidopsis*. On peut expliquer ce phénomène en considérant que la caractéristique des tissus transmetteurs est d'être passés par une phase de dédifférenciation, et qu'*At-SBP40a* soit liée à cette phase. Les mécanismes intervenant ensuite, comme la diminution du taux de mitochondries par cellule et de la quantité d'ADN par mitochondries sont des processus particuliers faisant intervenir d'autres protéines. Le fait d'avoir observé *At-SBP40a* dans le pollen pourrait donc ne pas avoir une finalité particulière, sauf peut-être de parer à une éventuelle transmission paternelle accidentelle.

4 Conclusion générale

Nous pensons avoir démontré au cours de cette dernière partie que les travaux entrepris durant de cette thèse ont conduit à l'identification d'une famille de protéines spécifique des plantes, dont les membres ont en commun plusieurs caractéristiques : en plus de leur séquence primaire, on peut citer la présence d'un domaine particulier de fixation aux acides nucléiques leur conférant la capacité de fixer préférentiellement l'ADNsb ou encore leur implication vraisemblable dans des étapes de la maintenance des génomes cytoplasmiques. D'autre part, le membre de la famille pour lequel le plus de données ont été collectées, *At-SBP40a*, a été identifié comme impliqué dans le contrôle de la stoechiométrie des molécules subgénomiques composant le chondriome d'*Arabidopsis*. L'ensemble des informations disponibles concernant ce mécanisme nous paraît suffisant pour émettre l'hypothèse qu'il remplit cette fonction dans des tissus fondamentaux pour la transmission du matériel génétique (certains méristèmes et les gamètes), probablement en empêchant l'initiation aberrante d'un mécanisme de recombinaison ou de réplication recombinaison-dépendante.

BIBLIOGRAPHIE

- Abad A. R., Mehrtens B. J. et Mackenzie S. A.** (1995). Specific expression in reproductive tissues and fate of a mitochondrial sterility-associated protein in cytoplasmic male-sterile bean. *Plant Cell* **7**: 271-285.
- Abdelnoor R. V., Yule R., Elo A., Christensen A. C., Meyer-Gauen G. et Mackenzie S. A.** (2003). Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. *Proc Natl Acad Sci U S A* **100**: 5968-5973.
- Adams K. L., Daley D. O., Qiu Y. L., Whelan J. et Palmer J. D.** (2000). Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature* **408**: 354-357.
- Alani E., Thresher R., Griffith J. D. et Kolodner R. D.** (1992). Characterization of DNA-binding and strand-exchange stimulation properties of γ -RPA, a yeast single-strand-DNA-binding protein. *J Mol Biol* **227**: 54-71.
- Albert B., Lelandais C., Pla M., Leuret C., Vitart V., Mathieu C., Sihachakr D., Godelle B. et De Paepe R.** (2003). Amplification of *Nicotiana sylvestris* mitochondrial subgenomes is under nuclear control and is associated with phenotypic changes. *Genetica* **117**: 17-25.
- Alberts B. M. et Frey L.** (1970). T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature* **227**: 1313-1318.
- Anderson D. G. et Kowalczykowski S. C.** (1997). The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes Dev* **11**: 571-581.
- Antequera F.** (2004). Genomic specification and epigenetic regulation of eukaryotic DNA replication origins. *Embo J* **23**: 4365-4370.
- Araki H., Ropp P. A., Johnson A. L., Johnston L. H., Morrison A. et Sugino A.** (1992). DNA polymerase II, the probable homolog of mammalian DNA polymerase epsilon, replicates chromosomal DNA in the yeast *Saccharomyces cerevisiae*. *Embo J* **11**: 733-740.
- Arimura S., Yamamoto J., Aida G. P., Nakazono M. et Tsutsumi N.** (2004). Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proc Natl Acad Sci U S A* **101**: 7805-7808.
- Arnold D. A. et Kowalczykowski S. C.** (2000). Facilitated loading of RecA protein is essential to recombination by RecBCD enzyme. *J Biol Chem* **275**: 12261-12265.
- Arrieta-Montiel M., Lyznik A., Woloszynska M., Janska H., Tohme J. et Mackenzie S.** (2001). Tracing evolutionary and developmental implications of mitochondrial stoichiometric shifting in the common bean. *Genetics* **158**: 851-864.
- Backert S.** (2000). Strand switching during rolling circle replication of plasmid-like DNA circles in the mitochondria of the higher plant *Chenopodium album* (L.). *Plasmid* **43**: 166-170.
- Backert S. et Borner T.** (2000). Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant *Chenopodium album* (L.). *Curr Genet* **37**: 304-314.
- Backert S., Dorfel P. et Borner T.** (1995). Investigation of plant organellar DNAs by pulsed-field gel electrophoresis. *Curr Genet* **28**: 390-399.
- Backert S., Lurz R., Oyarzabal O. A. et Borner T.** (1997). High content, size and distribution of single-stranded DNA in the mitochondria of *Chenopodium album* (L.). *Plant Mol Biol* **33**: 1037-1050.

- Barnes D., Cohen A., Bruick R. K., Kantardjieff K., Fowler S., Efuet E. et Mayfield S. P.** (2004). Identification and characterization of a novel RNA binding protein that associates with the 5'-untranslated region of the chloroplast psbA mRNA. *Biochemistry* **43**: 8541-8550.
- Bastia D. et Sueoka N.** (1975). Studies on the late replication of phage lambda: rolling-circle replication of the wild type and a partially suppressed strain, Oam29 Pam80. *J Mol Biol* **98**: 305-320.
- Bauer M. F., Hofmann S., Neupert W. et Brunner M.** (2000). Protein translocation into mitochondria: the role of TIM complexes. *Trends Cell Biol* **10**: 25-31.
- Baumann P. et Cech T. R.** (2001). Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**: 1171-1175.
- Bell S. P. et Dutta A.** (2002). DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-374.
- Bendich A. J.** (1993). Reaching for the ring: the study of mitochondrial genome structure. *Curr Genet* **24**: 279-290.
- Bendich A. J.** (2004). Circular chloroplast chromosomes: the grand illusion. *Plant Cell* **16**: 1661-1666.
- Benkovic S. J., Valentine A. M. et Salinas F.** (2001). Replisome-mediated DNA replication. *Annu Rev Biochem* **70**: 181-208.
- Bevan M.** (1984). Binary Agrobacterium vectors for plant transformation. *Nucleic Acids Res* **12**: 8711-8721.
- Bhattacharyya S. et Griep M. A.** (2000). DnaB helicase affects the initiation specificity of Escherichia coli primase on single-stranded DNA templates. *Biochemistry* **39**: 745-752.
- Birnboim H. C. et Doly J.** (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513-1523.
- Bochkarev A. et Bochkareva E.** (2004). From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* **14**: 36-42.
- Bochkarev A., Pfuetzner R. A., Edwards A. M. et Frappier L.** (1997). Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**: 176-181.
- Bochkareva E., Korolev S., Lees-Miller S. P. et Bochkarev A.** (2002). Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *Embo J* **21**: 1855-1863.
- Bowers J. L., Randell J. C., Chen S. et Bell S. P.** (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol Cell* **16**: 967-978.
- Bradford M. M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.
- Brill S. J. et Stillman B.** (1991). Replication factor-A from Saccharomyces cerevisiae is encoded by three essential genes coordinately expressed at S phase. *Genes Dev* **5**: 1589-1600.
- Budar F. et Pelletier G.** (2001). Male sterility in plants: occurrence, determinism, significance and use. *C R Acad Sci III* **324**: 543-550.
- Bujalowski W. et Lohman T. M.** (1987). Limited co-operativity in protein-nucleic acid interactions. A thermodynamic model for the interactions of Escherichia coli single strand binding protein with single-stranded nucleic acids in the "beaded", (SSB)₆₅ mode. *J Mol Biol* **195**: 897-907.
- Bujalowski W., Overman L. B. et Lohman T. M.** (1988). Binding mode transitions of Escherichia coli single strand binding protein-single-stranded DNA complexes. Cation, anion, pH, and binding density effects. *J Biol Chem* **263**: 4629-4640.
- Bycroft M., Hubbard T. J., Proctor M., Freund S. M. et Murzin A. G.** (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**: 235-242.
- Cann J. R.** (1998). Theoretical studies on the mobility-shift assay of protein-DNA complexes. *Electrophoresis* **19**: 127-141.

- Cazenave C., Chabbert M., Toulme J. J. et Helene C.** (1984). Absorption and fluorescence studies of the binding of the recA gene product from *E. coli* to single-stranded and double-stranded DNA. Ionic strength dependence. *Biochim Biophys Acta* **781**: 7-13.
- Chase J. W. et Williams K. R.** (1986). Single-stranded DNA binding proteins required for DNA replication. *Annu Rev Biochem* **55**: 103-136.
- Chen J., Silver D. P., Walpita D., Cantor S. B., Gazdar A. F., Tomlinson G., Couch F. J., Weber B. L., Ashley T., Livingston D. M. et Scully R.** (1998). Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells. *Mol Cell* **2**: 317-328.
- Cho Y., Mower J. P., Qiu Y. L. et Palmer J. D.** (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc Natl Acad Sci U S A* **101**: 17741-17746.
- Chrysogelos S. et Griffith J.** (1982). *Escherichia coli* single-strand binding protein organizes single-stranded DNA in nucleosome-like units. *Proc Natl Acad Sci U S A* **79**: 5803-5807.
- Connolly B., Parsons C. A., Benson F. E., Dunderdale H. J., Sharples G. J., Lloyd R. G. et West S. C.** (1991). Resolution of Holliday junctions in vitro requires the *Escherichia coli* ruvC gene product. *Proc Natl Acad Sci U S A* **88**: 6063-6067.
- Cox M. M.** (2003). The bacterial RecA protein as a motor protein. *Annu Rev Microbiol* **57**: 551-577.
- Curth U., Greipel J., Urbanke C. et Maass G.** (1993). Multiple binding modes of the single-stranded DNA binding protein from *Escherichia coli* as detected by tryptophan fluorescence and site-directed mutagenesis. *Biochemistry* **32**: 2585-2591.
- de Laat W. L., Appeldoorn E., Sugasawa K., Weterings E., Jaspers N. G. et Hoeijmakers J. H.** (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* **12**: 2598-2609.
- Dewey R. E., Levings C. S., 3rd et Timothy D. H.** (1986). Novel recombinations in the maize mitochondrial genome produce a unique transcriptional unit in the Texas male-sterile cytoplasm. *Cell* **44**: 439-449.
- Din S., Brill S. J., Fairman M. P. et Stillman B.** (1990). Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells. *Genes Dev* **4**: 968-977.
- Dixon D. A., Churchill J. J. et Kowalczykowski S. C.** (1994). Reversible inactivation of the *Escherichia coli* RecBCD enzyme by the recombination hotspot chi in vitro: evidence for functional inactivation or loss of the RecD subunit. *Proc Natl Acad Sci U S A* **91**: 2980-2984.
- Dyall S. D., Brown M. T. et Johnson P. J.** (2004). Ancient invasions: from endosymbionts to organelles. *Science* **304**: 253-257.
- Edmondson A. C., Song D., Alvarez L. A., Wall M. K., Almond D., McClellan D. A., Maxwell A. et Nielsen B. L.** (2005). Characterization of a mitochondrially targeted single-stranded DNA-binding protein in *Arabidopsis thaliana*. *Mol Genet Genomics* **273**: 115-122.
- Egelman E.** (2000). A ubiquitous structural core. *Trends Biochem Sci* **25**: 183-184.
- Eggleston A. K., Mitchell A. H. et West S. C.** (1997). In vitro reconstitution of the late steps of genetic recombination in *E. coli*. *Cell* **89**: 607-617.
- Elo A., Lyznik A., Gonzalez D. O., Kachman S. D. et Mackenzie S. A.** (2003). Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the *Arabidopsis* genome. *Plant Cell* **15**: 1619-1631.
- Emanuelsson O., Nielsen H., Brunak S. et von Heijne G.** (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**: 1005-1016.
- Essers J., Houtsmuller A. B., van Veelen L., Paulusma C., Nigg A. L., Pastink A., Vermeulen W., Hoeijmakers J. H. et Kanaar R.** (2002). Nuclear dynamics

- of RAD52 group homologous recombination proteins in response to DNA damage. *Embo J* **21**: 2030-2037.
- Evans M. D., Dizdaroglu M. et Cooke M. S.** (2004). Oxidative DNA damage and disease: induction, repair and significance. *Mutat Res* **567**: 1-61.
- Fang L., Davey M. J. et O'Donnell M.** (1999). Replisome assembly at oriC, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. *Mol Cell* **4**: 541-553.
- Fang Y. Z., Yang S. et Wu G.** (2002). Free radicals, antioxidants, and nutrition. *Nutrition* **18**: 872-879.
- Fauron C., Casper M., Gao Y. et Moore B.** (1995). The maize mitochondrial genome: dynamic, yet functional. *Trends Genet* **11**: 228-235.
- Fauron C. M. et Casper M.** (1994). A second type of normal maize mitochondrial genome: an evolutionary link. *Genetics* **137**: 875-882.
- Fauron C. M., Havlik M. et Brettell R. I.** (1990). The mitochondrial genome organization of a maize fertile cmsT revertant line is generated through recombination between two sets of repeats. *Genetics* **124**: 423-428.
- Fay P. J., Johanson K. O., McHenry C. S. et Bambara R. A.** (1981). Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli*. *J Biol Chem* **256**: 976-983.
- Ferrari M. E., Bujalowski W. et Lohman T. M.** (1994). Co-operative binding of *Escherichia coli* SSB tetramers to single-stranded DNA in the (SSB)₃₅ binding mode. *J Mol Biol* **236**: 106-123.
- Fuller R. S., Funnell B. E. et Kornberg A.** (1984). The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* **38**: 889-900.
- Gasior S. L., Olivares H., Ear U., Hari D. M., Weichselbaum R. et Bishop D. K.** (2001). Assembly of RecA-like recombinases: distinct roles for mediator proteins in mitosis and meiosis. *Proc Natl Acad Sci U S A* **98**: 8411-8418.
- Geraghty D. S., Ding M., Heintz N. H. et Pederson D. S.** (2000). Premature structural changes at replication origins in a yeast minichromosome maintenance (MCM) mutant. *J Biol Chem* **275**: 18011-18021.
- Giedroc D. P., Keating K. M., Williams K. R. et Coleman J. E.** (1987). The function of zinc in gene 32 protein from T4. *Biochemistry* **26**: 5251-5259.
- Giedroc D. P., Khan R. et Barnhart K.** (1990). Overexpression, purification, and characterization of recombinant T4 gene 32 protein₂₂₋₃₀₁ (g32P-B). *J Biol Chem* **265**: 11444-11455.
- Godthelp B. C., Artwert F., Joenje H. et Zdzienicka M. Z.** (2002). Impaired DNA damage-induced nuclear Rad51 foci formation uniquely characterizes Fanconi anemia group D1. *Oncogene* **21**: 5002-5005.
- Goodall G. J., Wiebauer K. et Filipowicz W.** (1990). Analysis of pre-mRNA processing in transfected plant protoplasts. *Methods Enzymol* **181**: 148-161.
- Gordenin D. A. et Resnick M. A.** (1998). Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability. *Mutat Res* **400**: 45-58.
- Gray M. W., Burger G. et Lang B. F.** (1999). Mitochondrial evolution. *Science* **283**: 1476-1481.
- Grishchuk A. L. et Kohli J.** (2003). Five RecA-like proteins of *Schizosaccharomyces pombe* are involved in meiotic recombination. *Genetics* **165**: 1031-1043.
- Hanada K., Iwasaki M., Ihashi S. et Ikeda H.** (2000). UvrA and UvrB suppress illegitimate recombination: synergistic action with RecQ helicase. *Proc Natl Acad Sci U S A* **97**: 5989-5994.
- Handa H.** (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res* **31**: 5907-5916.
- Hanson M. R. et Bentolila S.** (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* **16 Suppl**: S154-169.

- Henricksen L. A., Carter T., Dutta A. et Wold M. S.** (1996). Phosphorylation of human replication protein A by the DNA-dependent protein kinase is involved in the modulation of DNA replication. *Nucleic Acids Res* **24**: 3107-3112.
- Herendeen D. R. et Kelly T. J.** (1996). DNA polymerase III: running rings around the fork. *Cell* **84**: 5-8.
- Hiesel R., Combettes B. et Brennicke A.** (1994). Evidence for RNA editing in mitochondria of all major groups of land plants except the Bryophyta. *Proc Natl Acad Sci U S A* **91**: 629-633.
- Hong X., Cadwell G. W. et Kogoma T.** (1995). Escherichia coli RecG and RecA proteins in R-loop formation. *Embo J* **14**: 2385-2392.
- Iftode C. et Borowiec J. A.** (1997). Denaturation of the simian virus 40 origin of replication mediated by human replication protein A. *Mol Cell Biol* **17**: 3876-3883.
- Iftode C. et Borowiec J. A.** (1998). Unwinding of origin-specific structures by human replication protein A occurs in a two-step process. *Nucleic Acids Res* **26**: 5636-5643.
- Jacobs M. A., Payne S. R. et Bendich A. J.** (1996). Moving pictures and pulsed-field gel electrophoresis show only linear mitochondrial DNA molecules from yeasts with linear-mapping and circular-mapping mitochondrial genomes. *Curr Genet* **30**: 3-11.
- Janska H. et Mackenzie S. A.** (1993). Unusual mitochondrial genome organization in cytoplasmic male sterile common bean and the nature of cytoplasmic reversion to fertility. *Genetics* **135**: 869-879.
- Janska H., Sarria R., Woloszynska M., Arrieta-Montiel M. et Mackenzie S. A.** (1998). Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. *Plant Cell* **10**: 1163-1180.
- Jensen D. E., Kelly R. C. et von Hippel P. H.** (1976). DNA "melting" proteins. II. Effects of bacteriophage T4 gene 32-protein binding on the conformation and stability of nucleic acid structures. *J Biol Chem* **251**: 7215-7228.
- Jing D., Beechem J. et Patton W.** (2004). The utility of a two-color electrophoretic mobility shift assay for the analysis of DNA replication complexes. *Electrophoresis* **25**: 2439-2446.
- Johnson S. K., Bhattacharyya S. et Griep M. A.** (2000). DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates. *Biochemistry* **39**: 736-744.
- Kaguni J. M. et Kornberg A.** (1984). Replication initiated at the origin (oriC) of the E. coli chromosome reconstituted with purified enzymes. *Cell* **38**: 183-190.
- Kanazawa A., Tsutsumi N. et Hirai A.** (1994). Reversible changes in the composition of the population of mtDNAs during dedifferentiation and regeneration in tobacco. *Genetics* **138**: 865-870.
- Kass-Eisler A. et Greider C. W.** (2000). Recombination in telomere-length maintenance. *Trends Biochem Sci* **25**: 200-204.
- Kauppi L., Jeffreys A. J. et Keeney S.** (2004). Where the crossovers are: recombination distributions in mammals. *Nat Rev Genet* **5**: 413-424.
- Kerr I. D., Wadsworth R. I., Cubeddu L., Blankenfeldt W., Naismith J. H. et White M. F.** (2003). Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein. *Embo J* **22**: 2561-2570.
- Khazi F. R., Edmondson A. C. et Nielsen B. L.** (2003). An Arabidopsis homologue of bacterial RecA that complements an E. coli recA deletion is targeted to plant mitochondria. *Mol Genet Genomics* **269**: 454-463.
- Knoop V.** (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr Genet* **46**: 123-139.
- Kolodner R., Fishel R. A. et Howard M.** (1985). Genetic recombination of bacterial plasmid DNA: effect of RecF pathway mutations on plasmid recombination in Escherichia coli. *J Bacteriol* **163**: 1060-1066.

- Komori K. et Ishino Y.** (2001). Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem* **276**: 25654-25660.
- Koths K. et Dressler D.** (1978). Analysis of the phiX DNA replication cycle by electron microscopy. *Proc Natl Acad Sci U S A* **75**: 605-609.
- Kowalczykowski S. C.** (2000). Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci* **25**: 156-165.
- Kowalczykowski S. C., Dixon D. A., Eggleston A. K., Lauder S. D. et Rehrauer W. M.** (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol Rev* **58**: 401-465.
- Kowalczykowski S. C. et Eggleston A. K.** (1994). Homologous pairing and DNA strand-exchange proteins. *Annu Rev Biochem* **63**: 991-1043.
- Krejci L. et Sung P.** (2002). RPA not that different from SSB. *Structure* **10**: 601-602.
- Kruft V., Eubel H., Jansch L., Werhahn W. et Braun H. P.** (2001). Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiol* **127**: 1694-1710.
- Kubo T., Nishizawa S., Sugawara A., Itchoda N., Estiati A. et Mikami T.** (2000). The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). *Nucleic Acids Res* **28**: 2571-2576.
- Kuzmin E. V., Duvick D. N. et Newton K. J.** (2005). A mitochondrial mutator system in maize. *Plant Physiol* **137**: 779-789.
- Laemmlli U. K., Beguin F. et Gujer-Kellenberger G.** (1970). A factor preventing the major head protein of bacteriophage T4 from random aggregation. *J Mol Biol* **47**: 69-85.
- Lang B. F., Burger G., O'Kelly C. J., Cedergren R., Golding G. B., Lemieux C., Sankoff D., Turmel M. et Gray M. W.** (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387**: 493-497.
- Lecrenier N. et Foury F.** (2000). New features of mitochondrial DNA replication system in yeast and man. *Gene* **246**: 37-48.
- Lei M., Podell E. R., Baumann P. et Cech T. R.** (2003). DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* **426**: 198-203.
- Lemieux C., Otis C. et Turmel M.** (2000). Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature* **403**: 649-652.
- Lewis L. K. et Resnick M. A.** (2000). Tying up loose ends: nonhomologous end-joining in *Saccharomyces cerevisiae*. *Mutat Res* **451**: 71-89.
- Li Y. et Sugiura M.** (1990). Three distinct ribonucleoproteins from tobacco chloroplasts: each contains a unique amino terminal acidic domain and two ribonucleoprotein consensus motifs. *Embo J* **9**: 3059-3066.
- Linke B. et Borner T.** (2005). Mitochondrial effects on flower and pollen development. *Mitochondrion* **5**: 389-402.
- Liu F. et Schnable P. S.** (2002). Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol* **130**: 1657-1674.
- Liu N., Lamerdin J. E., Tebbs R. S., Schild D., Tucker J. D., Shen M. R., Brookman K. W., Siciliano M. J., Walter C. A., Fan W., Narayana L. S., Zhou Z. Q., Adamson A. W., Sorensen K. J., Chen D. J., Jones N. J. et Thompson L. H.** (1998). XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. *Mol Cell* **1**: 783-793.
- Liu Y., Masson J. Y., Shah R., O'Regan P. et West S. C.** (2004). RAD51C is required for Holliday junction processing in mammalian cells. *Science* **303**: 243-246.
- Lloyd R. G. et Buckman C.** (1985). Identification and genetic analysis of *sbcC* mutations in commonly used *recBC sbcB* strains of *Escherichia coli* K-12. *J Bacteriol* **164**: 836-844.

- Lohman T. M. et Ferrari M. E.** (1994). Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu Rev Biochem* **63**: 527-570.
- Lohman T. M. et Overman L. B.** (1985). Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J Biol Chem* **260**: 3594-3603.
- Lovett S. T. et Kolodner R. D.** (1989). Identification and purification of a single-stranded-DNA-specific exonuclease encoded by the recJ gene of Escherichia coli. *Proc Natl Acad Sci U S A* **86**: 2627-2631.
- Lowry O. H., Rosebrough N. J., Farr A. L. et Randall R. J.** (1951). Protein measurement with the Folin phenol reagent. *J Biol Chem* **193**: 265-275.
- Lueking A., Holz C., Gotthold C., Lehrach H. et Cahill D.** (2000). A system for dual protein expression in Pichia pastoris and Escherichia coli. *Protein Expr Purif* **20**: 372-378.
- Mackenzie S. A. et Chase C. D.** (1990). Fertility Restoration Is Associated with Loss of a Portion of the Mitochondrial Genome in Cytoplasmic Male-Sterile Common Bean. *Plant Cell* **2**: 905-912.
- Malek O. et Knoop V.** (1998). Trans-splicing group II introns in plant mitochondria: the complete set of cis-arranged homologs in ferns, fern allies, and a hornwort. *Rna* **4**: 1599-1609.
- Maleszka R., Skelly P. J. et Clark-Walker G. D.** (1991). Rolling circle replication of DNA in yeast mitochondria. *Embo J* **10**: 3923-3929.
- Margulis L. et Bermudes D.** (1985). Symbiosis as a mechanism of evolution: status of cell symbiosis theory. *Symbiosis* **1**: 101-124.
- Marienfeld J., Unseld M. et Brennicke A.** (1999). The mitochondrial genome of Arabidopsis is composed of both native and immigrant information. *Trends Plant Sci* **4**: 495-502.
- Martinez-Zapater J. M., Gil P., Capel J. et Somerville C. R.** (1992). Mutations at the Arabidopsis CHM locus promote rearrangements of the mitochondrial genome. *Plant Cell* **4**: 889-899.
- Mazin A. V. et Kowalczykowski S. C.** (1998). The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *Embo J* **17**: 1161-1168.
- Mathur J. et Koncz C.** (1998). Establishment and maintenance of cell suspension cultures. *Methods Mol Biol* **82**: 27-30.
- Mazin A. V., Zaitseva E., Sung P. et Kowalczykowski S. C.** (2000). Tailed duplex DNA is the preferred substrate for Rad51 protein-mediated homologous pairing. *Embo J* **19**: 1148-1156.
- Menand B., Marechal-Drouard L., Sakamoto W., Dietrich A. et Wintz H.** (1998). A single gene of chloroplast origin codes for mitochondrial and chloroplastic methionyl-tRNA synthetase in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **95**: 11014-11019.
- Meyer R. R., Glassberg J. et Kornberg A.** (1979). An Escherichia coli mutant defective in single-strand binding protein is defective in DNA replication. *Proc Natl Acad Sci U S A* **76**: 1702-1705.
- Moriguchi T., Motomura T., Akihama T. et Omura M.** (1997). Analysis of mitochondrial genomes among citrus plants produced by the interspecific somatic fusion of 'Seminole' tangelo with rough lemon. *Plant Cell Rep* **16**: 397-400.
- Morimatsu K. et Kowalczykowski S. C.** (2003). RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell* **11**: 1337-1347.
- Mortensen U. H., Bendixen C., Sunjevaric I. et Rothstein R.** (1996). DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci U S A* **93**: 10729-10734.
- Mosig G.** (1987). The essential role of recombination in phage T4 growth. *Annu Rev Genet* **21**: 347-371.

- Murakami Y., Eki T. et Hurwitz J.** (1992). Studies on the initiation of simian virus 40 replication in vitro: RNA primer synthesis and its elongation. *Proc Natl Acad Sci U S A* **89**: 952-956.
- Murante R. S., Rumbaugh J. A., Barnes C. J., Norton J. R. et Bambara R. A.** (1996). Calf RTH-1 nuclease can remove the initiator RNAs of Okazaki fragments by endonuclease activity. *J Biol Chem* **271**: 25888-25897.
- Murzin A. G.** (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *Embo J* **12**: 861-867.
- Naktinis V., Onrust R., Fang L. et O'Donnell M.** (1995). Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. II. Intermediate complex between the clamp loader and its clamp. *J Biol Chem* **270**: 13358-13365.
- Narhi L. O., Arakawa T., Aoki K. H., Elmore R., Rohde M. F., Boone T. et Strickland T. W.** (1991). The effect of carbohydrate on the structure and stability of erythropoietin. *J Biol Chem* **266**: 23022-23026.
- Neuburger M., Journet E. P., Bligny R., Carde J. P. et Douce R.** (1982). Purification of plant mitochondria by isopycnic centrifugation in density gradients of Percoll. *Arch Biochem Biophys* **217**: 312-323.
- New J. H., Sugiyama T., Zaitseva E. et Kowalczykowski S. C.** (1998). Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391**: 407-410.
- Newton K. et Coe E.** (1986). Mitochondrial DNA changes in abnormal growth mutants of maize. *Proc Natl Acad Sci U S A* **83**: 7363-7366.
- Newton K. J., Knudsen C., Gabay-Laughnan S. et Laughnan J. R.** (1990). An abnormal growth mutant in maize has a defective mitochondrial cytochrome oxidase gene. *Plant Cell* **2**: 107-113.
- Notsu Y., Masood S., Nishikawa T., Kubo N., Akiduki G., Nakazono M., Hirai A. et Kadowaki K.** (2002). The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol Genet Genomics* **268**: 434-445.
- Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N., Akashi K., Kanegae T., Ogura Y., Kohchi T. et et al.** (1992). Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J Mol Biol* **223**: 1-7.
- Ogihara Y., Yamazaki Y., Murai K., Kanno A., Terachi T., Shiina T., Miyashita N., Nasuda S., Nakamura C., Mori N., Takumi S., Murata M., Futo S. et Tsunewaki K.** (2005). Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res* **33**: 6235-6250.
- Oldenburg D. J. et Bendich A. J.** (2001). Mitochondrial DNA from the liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatemers, and a 5' protein. *J Mol Biol* **310**: 549-562.
- Oldenburg D. J. et Bendich A. J.** (2004). Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. *J Mol Biol* **335**: 953-970.
- Oram M., Keeley A. et Tsaneva I.** (1998). Holliday junction resolvase in *Schizosaccharomyces pombe* has identical endonuclease activity to the CCE1 homologue YDC2. *Nucleic Acids Res* **26**: 594-601.
- Palmer J. D.** (1997). Organelle genomes: going, going, gone! *Science* **275**: 790-791.
- Palmer J. D. et Herbon L. A.** (1988). Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol* **28**: 87-97.
- Paques F. et Haber J. E.** (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**: 349-404.
- Parkinson C. L., Mower J. P., Qiu Y. L., Shirk A. J., Song K., Young N. D., Depamphilis C. W. et Palmer J. D.** (2005). Multiple major increases and

- decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evol Biol* **5**: 73.
- Peeters N. et Small I.** (2001). Dual targeting to mitochondria and chloroplasts. *Biochim Biophys Acta* **1541**: 54-63.
- Pellegrini L., Yu D. S., Lo T., Anand S., Lee M., Blundell T. L. et Venkitaraman A. R.** (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**: 287-293.
- Perales C., Cava F., Meijer W. J. et Berenguer J.** (2003). Enhancement of DNA, cDNA synthesis and fidelity at high temperatures by a dimeric single-stranded DNA-binding protein. *Nucleic Acids Res* **31**: 6473-6480.
- Petes T. D.** (2001). Meiotic recombination hot spots and cold spots. *Nat Rev Genet* **2**: 360-369.
- Pinto A. V., Mathieu A., Marsin S., Veaute X., Ielpi L., Labigne A. et Radicella J. P.** (2005). Suppression of homologous and homeologous recombination by the bacterial MutS2 protein. *Mol Cell* **17**: 113-120.
- Prevost C. et Takahashi M.** (2003). Geometry of the DNA strands within the RecA nucleofilament: role in homologous recombination. *Q Rev Biophys* **36**: 429-453.
- Prigodich R. V., Casas-Finet J., Williams K. R., Konigsberg W. et Coleman J. E.** (1984). ¹H NMR (500 MHz) of gene 32 protein--oligonucleotide complexes. *Biochemistry* **23**: 522-529.
- Raderschall E., Golub E. I. et Haaf T.** (1999). Nuclear foci of mammalian recombination proteins are located at single-stranded DNA regions formed after DNA damage. *Proc Natl Acad Sci U S A* **96**: 1921-1926.
- Raghunathan S., Kozlov A. G., Lohman T. M. et Waksman G.** (2000). Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nat Struct Biol* **7**: 648-652.
- Robbins J. B., Murphy M. C., White B. A., Mackie R. I., Ha T. et Cann I. K.** (2004). Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase BI. *J Biol Chem* **279**: 6315-6326.
- Sakamoto W., Kondo H., Murata M. et Motoyoshi F.** (1996). Altered mitochondrial gene expression in a maternal distorted leaf mutant of *Arabidopsis* induced by chloroplast mutator. *Plant Cell* **8**: 1377-1390.
- Sakamoto W., Spielewoy N., Bonnard G., Murata M. et Wintz H.** (2000). Mitochondrial localization of AtOXA1, an *Arabidopsis* homologue of yeast Oxa1p involved in the insertion and assembly of protein complexes in mitochondrial inner membrane. *Plant Cell Physiol* **41**: 1157-1163.
- Sarria R., Lyznik A., Vallejos C. E. et Mackenzie S. A.** (1998). A cytoplasmic male sterility-associated mitochondrial peptide in common bean is post-translationally regulated. *Plant Cell* **10**: 1217-1228.
- Schofield M. J. et Hsieh P.** (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* **57**: 579-608.
- Scully R., Chen J., Ochs R. L., Keegan K., Hoekstra M., Feunteun J. et Livingston D. M.** (1997). Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90**: 425-435.
- Scully R., Chen J., Plug A., Xiao Y., Weaver D., Feunteun J., Ashley T. et Livingston D. M.** (1997). Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**: 265-275.
- Seigneur M., Bidnenko V., Ehrlich S. D. et Michel B.** (1998). RuvAB acts at arrested replication forks. *Cell* **95**: 419-430.
- Seitz E. M. et Kowalczykowski S. C.** (2000). The DNA binding and pairing preferences of the archaeal RadA protein demonstrate a universal characteristic of DNA strand exchange proteins. *Mol Microbiol* **37**: 555-560.
- Seo Y. S. et Hurwitz J.** (1993). Isolation of helicase alpha, a DNA helicase from HeLa cells stimulated by a fork structure and signal-stranded DNA-binding proteins. *J Biol Chem* **268**: 10282-10295.

- Shamoo Y., Friedman A. M., Parsons M. R., Konigsberg W. H. et Steitz T. A.** (1995). Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* **376**: 362-366.
- Shaver J. M., Oldenburg D. J. et Bendich A. J.** (2005). Changes in chloroplast DNA during development in tobacco, *Medicago truncatula*, pea, and maize. *Planta*: 1-11.
- Sheahan M. B., McCurdy D. W. et Rose R. J.** (2005). Mitochondria as a connected population: ensuring continuity of the mitochondrial genome during plant cell dedifferentiation through massive mitochondrial fusion. *Plant J* **44**: 744-755.
- Sigurdsson S., Van Komen S., Petukhova G. et Sung P.** (2002). Homologous DNA pairing by human recombination factors Rad51 and Rad54. *J Biol Chem* **277**: 42790-42794.
- Silva-Filho M. C.** (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Curr Opin Plant Biol* **6**: 589-595.
- Skelly P. J. et Maleszka R.** (1989). Isolation of mitochondrial DNA using pulsed field gel electrophoresis. *Nucleic Acids Res* **17**: 7537.
- Small I., Isaac P. et Leaver C.** (1987). Stoichiometric differences in DNA molecules containing the *atp9* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. *Embo J* **6**: 865-869.
- Small I., Suffolk R. et Leaver C. J.** (1989). Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* **58**: 69-76.
- Smith G. R.** (1991). Conjugational recombination in *E. coli*: myths and mechanisms. *Cell* **64**: 19-27.
- Solinger J. A., Kiianitsa K. et Heyer W. D.** (2002). Rad54, a Swi2/Snf2-like recombinational repair protein, disassembles Rad51:dsDNA filaments. *Mol Cell* **10**: 1175-1188.
- Sourice S., Biaudet V., El Karoui M., Ehrlich S. D. et Gruss A.** (1998). Identification of the Chi site of *Haemophilus influenzae* as several sequences related to the *Escherichia coli* Chi site. *Mol Microbiol* **27**: 1021-1029.
- Speck C. et Messer W.** (2001). Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *Embo J* **20**: 1469-1476.
- Stillman B.** (2005). Origin recognition and the chromosome cycle. *FEBS Lett* **579**: 877-884.
- Struglics A., Fredlund K., Rassmusson A. et Moller I.** (1993). The presence of a short redox chain in the membrane of intact potato tuber peroxisome and the association of malate dehydrogenase with the peroxisomal membrane. *Physiologia Plantarum* **88**: 19-28.
- Studwell P. S. et O'Donnell M.** (1990). Processive replication is contingent on the exonuclease subunit of DNA polymerase III holoenzyme. *J Biol Chem* **265**: 1171-1178.
- Stukenberg P. T. et O'Donnell M.** (1995). Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. V. Four different polymerase-clamp complexes on DNA. *J Biol Chem* **270**: 13384-13391.
- Sugiyama T., Zaitseva E. M. et Kowalczykowski S. C.** (1997). A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the *Saccharomyces cerevisiae* Rad51 protein. *J Biol Chem* **272**: 7940-7945.
- Sugiyama Y., Watase Y., Nagase M., Makita N., Yagura S., Hirai A. et Sugiura M.** (2005). The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol Genet Genomics* **272**: 603-615.
- Sung P.** (1997). Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *J Biol Chem* **272**: 28194-28197.
- Sung P., Krejci L., Van Komen S. et Sehorn M. G.** (2003). Rad51 recombinase and recombination mediators. *J Biol Chem* **278**: 42729-42732.
- Szostak J. W., Orr-Weaver T. L., Rothstein R. J. et Stahl F. W.** (1983). The double-strand-break repair model for recombination. *Cell* **33**: 25-35.

- Timmis J. N., Ayliffe M. A., Huang C. Y. et Martin W.** (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5**: 123-135.
- Tougu K. et Marians K. J.** (1996). The extreme C terminus of primase is required for interaction with DnaB at the replication fork. *J Biol Chem* **271**: 21391-21397.
- Touzet P. et Budar F.** (2004). Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? *Trends Plant Sci* **9**: 568-570.
- Tseng Y. C., Hung J. L. et Wang T. C.** (1994). Involvement of RecF pathway recombination genes in postreplication repair in UV-irradiated *Escherichia coli* cells. *Mutat Res* **315**: 1-9.
- Tsurimoto T. et Stillman B.** (1991). Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. *J Biol Chem* **266**: 1961-1968.
- Tucker P. A., Tsernoglou D., Tucker A. D., Coenjaerts F. E., Leenders H. et van der Vliet P. C.** (1994). Crystal structure of the adenovirus DNA binding protein reveals a hook-on model for cooperative DNA binding. *Embo J* **13**: 2994-3002.
- Turmel M., Otis C. et Lemieux C.** (2002). The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. *Proc Natl Acad Sci U S A* **99**: 11275-11280.
- Turmel M., Otis C. et Lemieux C.** (2003). The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell* **15**: 1888-1903.
- Umez K., Nakayama K. et Nakayama H.** (1990). *Escherichia coli* RecQ protein is a DNA helicase. *Proc Natl Acad Sci U S A* **87**: 5363-5367.
- Unsel M., Marienfeld J. R., Brandt P. et Brennicke A.** (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat Genet* **15**: 57-61.
- van Engelen F. A., Molthoff J. W., Conner A. J., Nap J. P., Pereira A. et Stiekema W. J.** (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**: 288-290.
- Van Komen S., Petukhova G., Sigurdsson S., Stratton S. et Sung P.** (2000). Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. *Mol Cell* **6**: 563-572.
- Venkitaraman A. R.** (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**: 171-182.
- Vermel M., Guermann B., Delage L., Grienenberger J. M., Marechal-Drouard L. et Gualberto J. M.** (2002). A family of RRM-type RNA-binding proteins specific to plant mitochondria. *Proc Natl Acad Sci U S A* **99**: 5866-5871.
- Vitart V., De Paepe R., Mathieu C., Chetrit P. et Vedel F.** (1992). Amplification of substoichiometric recombinant mitochondrial DNA sequences in a nuclear, male sterile mutant regenerated from protoplast culture in *Nicotiana glauca*. *Mol Gen Genet* **233**: 193-200.
- Vogelauer M., Rubbi L., Lucas I., Brewer B. J. et Grunstein M.** (2002). Histone acetylation regulates the time of replication origin firing. *Mol Cell* **10**: 1223-1233.
- Voinnet O., Rivas S., Mestre P., Baulcombe D.** (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the P19 protein of tomato bushy stunt virus. *Plant J* **33**: 949-956.
- Vulic M., Dionisio F., Taddei F. et Radman M.** (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc Natl Acad Sci U S A* **94**: 9763-9767.
- Waga S. et Stillman B.** (1998). The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* **67**: 721-751.

- Ward B. L., Anderson R. S. et Bendich A. J.** (1981). The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). *Cell* **25**: 793-803.
- West S. C.** (1996). DNA helicases: new breeds of translocating motors and molecular pumps. *Cell* **86**: 177-180.
- West S. C.** (1996). The RuvABC proteins and Holliday junction processing in *Escherichia coli*. *J Bacteriol* **178**: 1237-1241.
- Whitby M. C., Ryder L. et Lloyd R. G.** (1993). Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* **75**: 341-350.
- Wischmann C. et Schuster W.** (1995). Transfer of rps10 from the mitochondrion to the nucleus in *Arabidopsis thaliana*: evidence for RNA-mediated transfer and exon shuffling at the integration site. *FEBS Lett* **374**: 152-156.
- Wold M. S.** (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**: 61-92.
- Yang H., Jeffrey P. D., Miller J., Kinnucan E., Sun Y., Thoma N. H., Zheng N., Chen P. L., Lee W. H. et Pavletich N. P.** (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* **297**: 1837-1848.
- Yanisch-Perron C., Vieira J. et Messing J.** (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.
- Yoda K., Yasuda H., Jiang X. W. et Okazaki T.** (1988). RNA-primed initiation sites of DNA replication in the origin region of bacteriophage lambda genome. *Nucleic Acids Res* **16**: 6531-6546.
- Youvan D. C. et Michel-Beyerle M. E.** (1996). Structure and fluorescence mechanism of GFP. *Nat Biotechnol* **14**: 1219-1220.
- Yu X., Jacobs S. A., West S. C., Ogawa T. et Egelman E. H.** (2001). Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. *Proc Natl Acad Sci U S A* **98**: 8419-8424.
- Zuo J., Niu Q. W. et Chua N. H.** (2000). An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* **24** : 265-273.

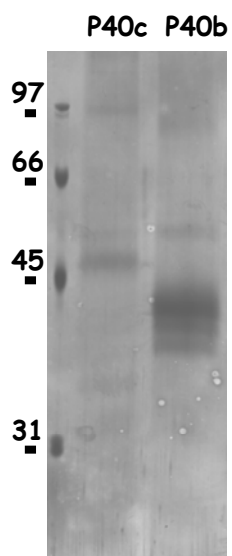


Figure annexe: Purification d'*At*-SBP40b et c à l'aide du système TAP

Les protéines *At*-SBP40b et c ont été exprimées dans des feuilles de tabac et purifiées par deux chromatographies d'affinité successives (voir paragraphe 6.6.2 de la partie Méthodes). On observe que nos protéines d'intérêt sont obtenues sous forme largement majoritaire et à la taille attendue.

ANNEXE

(LISTE DES OLIGONUCLEOTIDES UTILISES)

1 Sondes

R1 : Oligonucléotide sens utilisé lors du test de recombinaison *in vitro* (**Figure r3.12**)
5'AATAGAGAGAGATAGATTTGTAGAG3'

R2 : Oligonucléotide antisens utilisé lors du test de recombinaison *in vitro* (**Figure r3.12**)
5'ACTGAAAGCTTGCATGCCTGCAG3'

S1 : Oligonucléotide utilisé comme sonde dans des expériences de Selex et de Southwestern (**Figures r1.1, r3.5 et r3.20**)
5'GGCCACTAGTCGGATCCC(N)₂₅GGGCTGCAGGAATTCGACG3'

Oligonucléotides utilisés comme sondes en southwestern et en gel retard, afin de déterminer le type d'acide nucléique (ADNsb, ADNdb ou ARN) préférentiellement fixé par At-SBP40a et At-SBP40b (**Figures r3.6, r3.7, r3.9, r3.10 et r3.11**):

S2 : Sonde sb de 29 pb.
5'GGGACGCTGCCGAATTCTACCAGTGCCTT3'

S3 : Complémentaire à S2.
5'AAGGCACTGGTAGAATTCGGCAGCGTCCC3'

S4 : Oligonucléotide sens ayant servi à la synthèse d'une sonde ARN de même séquence que S2 ; il contient pour cela le promoteur T7.

5'TGTAATACGACTCACTATAGGGACGCTGCCGAATTCTACCAGTGCCTT3'

S5 : Oligonucléotide antisens ayant servi à la synthèse d'une sonde ARN de même séquence que S2 ; complémentaire à S4.

5'AAGGCACTGGTAGAATTCGGCAGCGTCCCTATAGTGAGTCGTATTACA3'

Oligonucléotides utilisés comme sondes dans des expériences de gel retard, afin de déterminer le nombre de base occupées par At-SBP40b sur son substrat (voir paragraphe 2.1.2 du Chapitre III de la partie Résultats):

S6 : 5'(AGTC)₄3'

S7 : 5'(AGTC)₈3'

S8 : 5'(AGTC)₈3' (**Figures r3.8, r3.15 et r3.16**)

S9 : 5'(AGTC)₁₀3'

S10 : 5'(AGTC)₁₂3'

S11 : 5'(AGTC)₁₄3'

S12 : 5'(AGTC)₁₆3'

Oligonucléotides utilisés comme sondes dans des expériences gel retard, afin de déterminer l'influence de différentes structures sur l'efficacité de fixation d'At-SBP40b à son substrat

(Figures r3.13 et r3.14) :

S13 : Sonde sb de 61 pb.

5'GACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTTTGCCACCTGCAGG
TTCACCC3'

S14 : Son extrémité 5' s'hybride avec l'extrémité 3' de S9 ; utilisé pour la synthèse de sondes mimant une fourche de réplication ou une jonction de Holliday.

5'TGGGTGAACCTGCAGGTGGGCAAAGATGTCCTAGCAATGTAATCGTCAAGCTTT
ATGCCGTT3'

S15 : Utilisé pour la synthèse d'une sonde mimant une jonction de Holliday.

5'CAACGGCATAAAGCTTGACGATTACATTGCTAGGACATGCTGTCTAGAGGATCC
GACTATCGA3'

S16 : Son extrémité 3' s'hybride avec l'extrémité 5' de S9 ; utilisé pour la synthèse de sondes mimant une fourche de réplication ou une jonction de Holliday.

5'ATCGATAGTCGGATCCTCTAGACAGCATGTCCTAGCAAGGCACTGGTAGAA
TTCGGCAGCGT3'

S17 : Complémentaire à S9 ; utilisé pour la synthèse d'une sonde d'ADNdb.

5'GGGTGAACCTGCAGGTGGGCAAAGATGTCCTAGCAAGGCACTGGTAGAATTCG
GCAGCGTC3'

S18 : Complémentaire à la partie 3' de S9 ; utilisé pour la synthèse d'une sonde dans laquelle l'extrémité 5' est sous forme de sb et l'extrémité 3' sous forme de db.

5'GGGTGAACCTGCAGGTGGGCAAAGATGTCC3'

S19 : Complémentaire à la partie 5' de S9 ; utilisé pour la synthèse d'une sonde dans laquelle l'extrémité 3' est sous forme de sb et l'extrémité 5' sous forme de db.

5'TAGCAAGGCACTGGTAGAATTCGGCAGCGTC3'

S20 : Complémentaire à S9 excepté pour les 6 bases médianes ; utilisé dans le but de vérifier si la présence de sb au sein d'une sonde de nature db influe sur l'efficacité de fixation.

5'GGGTGAACCTGCAGGTGGGCAAAGATGTAACCTAAAGGCACTGGTAGAATTCG
GCAGCGTC3'

2 Clonages

2.1 Clonage des séquences codant pour les précurseurs des protéines de la famille SBP40

L'ADNc de *St*-SBP40 a été synthétisé par RT-PCR à partir d'ARNm totaux de pomme de terre en utilisant des oligonucléotides dégénérés obtenus à l'aide de deux peptides identifiés par séquençage N-terminal (Pept1: IEIYDEAEDVSSWPKPSEI; Pep2: NFLLDENDNQHSDY). La séquence (450 pb) du fragment amplifié a ensuite été utilisée pour obtenir de nouvelles amorces permettant l'amplification complète de l'ADNc de *St*-SBP40 par 5' and 3' RACE (5' RACE System, Invitrogen). Les ADNc des protéines At-SBP40 ont été amplifiés à partir d'ARNm totaux d'*Arabidopsis* obtenus à l'aide d'oligo-dT, en employant des amorces complémentaires aux séquences des 5' et 3' UTR des gènes *At1g47720*,

At4g20010 et *At5g44785*, déterminées à partir des bases de données informatiques. Les oligonucléotides suivants ont été utilisés :

Pour *At-SBP40a* :

P3 : Oligonucléotide sens
5'TAAAAATCTCAAATTTCCCCAGTA3'

P4 : Oligonucléotide antisens
5'AACATACTTGTAATTTCCCAAGAA3'

Pour *At-SBP40b* :

P5 : Oligonucléotide sens
5'GGGATTTTTTCTCTCGTTTCAT3'

P6 : Oligonucléotide antisens
5'TGGAGCTCTTTACATGGAGATAA3'

Pour *At-SBP40c* :

P7 : Oligonucléotide sens
5'TCAGATTCTTTCTCTTTAGTATGAA3'

P8 : Oligonucléotide antisens
5'TCTCAACCAAGTTCTTCCAAATTT3'

2.2 Fusions avec l'étiquette TAP

Les séquences codant pour les précurseurs des protéines *At-SBP40* ont été agrémentées de l'étiquette TAP à leur extrémité C-terminale par clonage dans le vecteur pUCAP35S-TAP (voir paragraphe 3.4 de la partie Matériels). Elles ont pour cela été amplifiées par PCR en utilisant des amorces portant les sites *NcoI* (pour l'amorce sens) et *KpnI* (pour l'amorce antisens). Les produits de PCR ont ensuite été digérés à l'aide de ces enzymes de restriction et clonés dans les plasmides pUCAP35S-TAP puis pBINPLUS à l'aide des sites *PacI* et *AscI*. Les amorces ayant été utilisées sont les suivantes :

Pour *At-SBP40a* :

T1 : Oligonucléotide sens
5'TAGTGGGTACCAATATAGTCAGAAAGACGGTCACG3'

T2 : Oligonucléotide antisens
5'TAATCCATGGCCACGTTCTTCAAACCTCGGAAG3'

Pour *At-SBP40b* :

T3 : Oligonucléotide sens
5'TAATCCATGGCTTTGATTTCCAAATCCTTGGCAA3'

T4 : Oligonucléotide antisens
5'GCTGGTACCAAAGACAATACTTTCTTGTTTCTTC3'

Pour *At-SBP40c* :

T5 : Oligonucléotide sens
5'TAATCCATGGCTTTGATCTCAAGAACATTGACAA3'

T6 : Oligonucléotide antisens
5'ATTGGTACCGCCATGACAGGTCTTTC3'

2.3 Fusions avec l'eGFP

Les séquences codant pour les 43, 85 et 68 acides aminés N-terminaux des protéines *At*-SBP40a, b et c, respectivement, ont été clonées dans le plasmide pCK-GFP3 à l'aide des oligonucléotides suivants :

Pour *At*-SBP40a :

F1 : amorce sens ; contient le site *Bsp*HI.

5'CCCAGTTCATGAACACGTTCTTC3'

F2 : amorce antisens ; contient le site *Bam*HI.

5'CCTAGGATCCCGGTTTGGATTGAGAGG3'

Pour *At*-SBP40b :

F3 : amorce sens ; contient le site *Nco*I.

5'AACCATGGGTTTGATTTCCAAATCCT3'

F4 : amorce antisens ; contient le site *Bam*HI.

5'GCGGATCCAATTCATTCTCCGTCGT3'

Pour *At*-SBP40c :

F5 : amorce sens ; contient le site *Nco*I.

5'AACCATGGATTTGATCTCAAGAACATTGA3'

F6 : amorce antisens ; contient le site *Bam*HI.

5'GCGGATCCGATTCTTTCTCCGTCGCTACAT3'

2.4 Fusions avec le marqueur GUS

Les séquences comprenant environ 1000 nt situés en 5' des gènes codant pour les protéines *At*-SBP40 ont été fusionnées au gène de la β -glucuronidase (GUS) au moyen des oligonucléotides suivants :

Pour *At*-SBP40a :

G1 : amorce sens ; contient le site *Hind*III.

5'TATAAGCTTAATGGACTTTGAAATTCCTTTTCTT3'

G2 : amorce antisens ; contient le site *Xma*I.

5'ATACCCGGGCTACTGGGGAAATTTGAGATTTTTA3'

Pour *At*-SBP40b :

G3 : amorce sens ; contient le site *Hind*III.

5'TATAAGCTTGTTTTTAGTCATTAGATATTATGAA3'

G4 : amorce antisens ; situé dans la région codante ; le clonage a été réalisé à l'aide d'un site *Bam*HI situé dans le fragment amplifié.

5'CAATTCTTGCCAAGGATTTGGA3'

Pour *At*-SBP40c :

G5 : amorce sens ; contient le site *Hind*III.

5'TATAAGCTTAGTGAATTCCTCTAATGTGACTA3'

G6 : amorce antisens ; contient le site *Xma*I.

5'ATACCCGGGTTTCATACTAAAGAGAAAGAATCT3'