HAL
open science

Complexation d'un fragment de brin télomérique riche en C par des protéines de Saccharomyces cerevisiae. Identification génomique aux IMPDH. Possible repliement en motif i de l'ADN complexé.

Jean-François Cornuel

To cite this version:

Jean-François Cornuel. Complexation d'un fragment de brin télomérique riche en C par des protéines de Saccharomyces cerevisiae. Identification génomique aux IMPDH. Possible repliement en motif i de l'ADN complexé.. Biochimie [q-bio.BM]. Ecole Polytechnique X, 2000. Français. NNT : . tel-00003611

HAL Id: tel-00003611

https://pastel.hal.science/tel-00003611v1

Submitted on 21 Oct 2003

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

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

ÉCOLE POLYTECHNIQUE

THÈSE

pour obtenir le grade de

DOCTEUR de L'ÉCOLE POLYTECHNIQUE

Discipline : Biophysique

Présentée et soutenue publiquement par

Jean-François CORNUEL

le 24 novembre 2000

Complexation d'un fragment de brin télomérique riche en C par des protéines de *Saccharomyces cerevisiae*. Identification génomique aux IMPDH. Possible repliement en motif i de l'ADN complexé.

Directeur de thèse : M. Maurice Guéron

Jury :

M.	E. Gilson	Rapporteur
M.	J.-L. Mergny	Rapporteur
M.	M. Guéron	Examinateur
M.	M. Kochoyan	Examinateur
M.	M. Le Maire	Examinateur
M.	P. Plateau	Examinateur

A mes parents, ma sœur et mes amis

- R E M E R C I E M E N T S -

Je tiens à remercier M. E. Gilson et M. J.-L. Mergny pour l'examen attentif qu'ils ont fait de ce travail et de ce manuscrit, ainsi que M. M. Kochoyan, M. M. Le Maire et M. P. Plateau pour avoir accepté de le juger.

La découverte des thématiques de la biophysique, loin des rigueurs de la physique théorique, a constitué une expérience humaine enrichissante et gratifiante. Cela n'aurait été sans votre contribution à tous :

M. Maurice Guéron a dirigé ce projet durant toute sa durée. Nos nombreuses discussions sont autant de souvenirs marquants. A l'heure où s'achèvent les années passées en son laboratoire, je veux le remercier pour l'accueil qu'il me fît, pour sa confiance, pour la formation pluridisciplinaire qu'il m'a transmise, pour ses conseils et ses encouragements qui jalonnèrent mon quotidien.

M. le Professeur Mac Carroll m'a fait la confiance de me soutenir dans mon choix d'orientation au sortir de son DEA. J'ai toujours trouvé en lui un interlocuteur passionné. Je lui en suis reconnaissant.

Je remercie infiniment Anne Moraillon. Cela fut une collaboration étroite et enrichissante. Ce travail doit beaucoup à sa présence, à son investissement total, et surtout à sa bonne humeur. Mille fois merci ! Monique Leblond m'a initié aux « joies » de la paillasse et de la biochimie. Je ne peux oublier sa très grande disponibilité.

J'ai trouvé en chaque membre du Groupe de Biophysique un amical soutien. Je remercie Sylvie, l'amie complice attentive et toujours disponible, Kevin et Tuan, pour les heures tardives passées entre thésards, Karim pour sa contribution au sujet et de nombreuses discussions toujours enrichissantes, Laurent, parce qu'il est si rare de trouver quelqu'un avec qui parler du motif i, des protéines et de « La Force », Denis, Alvin et Renaud pour leur soutien, Janette, Eve, Annick et Mohamed pour leur aide et leur bonne humeur. Je remercie également Michel Kochoyan, à l'origine de ce projet, pour ses indications et la chaleur de son accent du sud et Jean Louis Leroy pour sa participation à cette étude.

Je remercie Jean-Pierre Le Caer et Valérie Labas pour l'aide apportée lors des identifications par spectrométrie de masse. Je remercie les membres du laboratoire de biophysique pour leur accueil et leur aide lors de mes passages au muséum. Je remercie également Olivier Gueldry pour son aide en informatique et biochimie.

Je n'ai la place de citer tous les membres des laboratoires PMC, BIOCH et BIOP2 de l'Ecole Polytechnique qui ont un jour ou l'autre croisé mon chemin. Je les remercie tous pour leur aide, leurs conseils et leurs encouragements au quotidien renouvelés. Je remercie tous les thésards de PMC pour leur confiance et leur dynamisme.

Sylvie et Fabrice Lecomte sont des amis d'exception. Je les remercie pour ces instants partagés où les tourments du thésard s'effaçaient, pour leur soutien sans faille et leur amitié sans lesquels je n'aurais pu durer. *L'aventure continue…*

Je remercie ma famille et mes amis pour leur chaleur, leurs marques de sympathie et leurs encouragements.

Parmi tous, mes parents et ma sœur m'ont apporté plus que quiconque. Ils sont ma force, ma confiance et ma richesse. Jamais je n'aurais réussi sans eux.

Résumé

Les régions terminales (les télomères) des chromosomes de la plupart des organismes sont essentielles au maintien de l'intégrité des chromosomes et sont impliquées dans la cancérogenèse et le vieillissement cellulaire. Les télomères sont formés d'un brin riche en guanines et d'un brin complémentaire riche en cytosines. *In vitro*, chaque brin s'organise en une structure repliée, le quadruplexe de guanines pour l'un, le motif i pour l'autre, ce qui pose la question du rôle biologique de ces structures. Plusieurs protéines s'associant au brin G et favorisant la formation de quadruplexes de guanines sont connues. Deux protéines reconnaissant spécifiquement le brin C de l'ADN télomérique humain viennent d'être identifiées (ASF/SF2 et hnRNP K).

Nous avons voulu contribuer à la compréhension des mécanismes de régulation des télomères en recherchant une protéine se liant spécifiquement à une séquence d'ADN télomérique riche en cytosines. Cette étude a été entreprise sur un organisme eucaryote, la levure *Saccharomyces cerevisiae*.

Nous avons commencé par montrer que le fragment d'ADN télomérique de levure, d[$(CCCACA)_3CCC$], pouvait former un motif i *in vitro*. Nous avons observé par migration électrophorétique non dénaturante, la formation d'un complexe ADN-protéine sur lequel nous avons focalisé nos efforts. La masse moléculaire de cette protéine (~ 250 kDa) a été déterminée par passage sur tamis moléculaire. L'association ADN-protéine a été observée pour une gamme de pH allant de 6 à 8. Un duplexe d'ADN B d(CGCGATCGCG) ou une petite suite de cytosines d(CCC) ne sont pas compétiteurs et le signal retardé est atténué de moitié pour un rapport monobrin d(T_{14})/sonde de 4000. L'ADN d'*E. coli*, coupé en fragments de 10 à 1000 bases, est compétiteur (50%) pour un rapport compétiteur/sonde égal à 1000. Mais toutes les séquences d'oligomères testées formant *in vitro* un motif i monomérique (séquences comptant 4 répétitions de 2, 3, 4 ou 5 de désoxycytidines, dont des séquences télomériques ou centromériques naturelles), entrent en compétition avec la séquence d[$(CCCACA)_3CCC$], notamment dans des conditions (pH, température, force ionique,..), où le motif i est observé *in vitro*.

Pour identifier les protéines impliquées dans l'association avec l'ADN télomérique, nous avons utilisé des billes de streptavidine greffées ou non avec la séquence d'ADN. Nous avons identifié les protéines liées à l'ADN par spectrométrie de masse MALDI-TOF et Q-TOF MS/MS. Il s'agit de trois protéines Yhr216p, Ylr432p et Yml056p, dont les séquences sont très proches. Cette identification a été confirmée par séquençage par dégradation d'Edman de peptides internes. Les séquences de ces protéines sont similaires (~ 78 %) aux deux inosine 5'-monophosphate déshydrogénases humaines (IMPDH 1 et 2 (masse de ~ 56 kDa)) connues pour leur implication dans le métabolisme des purines. Celles-ci sont habituellement observées sous forme tétramérique (masse > 200 kDa). Malgré la présence dans des extraits nucléaires humains HeLa d'une protéine ayant des caractéristiques de migration natives identiques à celles observées chez la levure, et différentes des protéines ASF/SF2 et hnRNP K déjà identifiées, nous avons montré que les protéines IMPDH 1 et 2 purifiées ne sont pas capables de fixer seules l'ADN télomérique humain.

Nous montrons également que les protéines identifiées (Yhr216p, Ylr432p et Yml056p) s'associent au brin complémentaire riche en G. Des migrations électrophorétiques en présence de compétiteurs montrent que cette association est spécifique. A notre connaissance, il s'agit des premières protéines de levures interagissant séparément avec chacun des deux brins télomériques dans des conditions d'incubation où l'on observe la formation de quadruplexes de guanines et de motif i.

Mots Clés : télomère, association ADN-protéine, motif i et quadruplexe de guanines.
Discipline : Biophysique.

Table des matières

Chapitre I : Introduction ***1***

1- Introduction aux acides nucléiques **2**

1-1- Les acides nucléiques 2

1-2- Structures de l'ADN 4

1-3- Les quadruplexes de guanines 5

1-4- Le motif i 6

2- Fonctions des acides nucléiques **7**

2-1- Transmission et expression de l'information génétique contenue dans les acides nucléiques 7

2-2- Mécanisme de réplication 8

2-3- Les télomères, extrémités des chromosomes 9

2-4- Mécanismes liés à la régulation des télomères 11

2-5- Protéines spécifiques des télomères 13

2-6- Région centromérique 15

3- Le sujet de thèse **15**

3-1- Les télomères de levure 16

3-2- Présentation du sujet de thèse 18

Chapitre II : Matériels et Méthodes ***19***

1- Biochimie des protéines et des acides nucléiques **19**

1-1- Culture de *Saccharomyces cerevisiae* et extraits de levure 19

1-2- Synthèse d'oligodésoxynucléotides 23

1-3- Marquage radioactif des oligodésoxynucléotides 23

1-4- Gel dénaturant (ADN marqué) 24

1-5- Electrophorèse en conditions non dénaturantes 24

1-6- Gel de protéines (électrophorèse dénaturante) 25

1-7- Séquençage d'Edman 28

1-8- Tamis moléculaire 29

1-9- Purification par affinité 30

1-10- Reconcentration, dessalage et changement de pH 31

1-11- Autres techniques utilisées 31

2- Biophysique **33**

2-1- Spectroscopie d'absorption 33

2-2- Spectrométrie de masse 35

2-3- Spectroscopie RMN 38

Chapitre III : Caractéristiques du complexe ADN-protéine ____ *41*

1- La séquence d'ADN télomérique de levure d[(CCCACA)$_3$CCC] ____ **42**

1-1- La séquence d[(CCCACA)$_3$CCC] ____ 42

1-2- Etude par spectroscopie d'absorption ____ 43

1-3- Etude par spectroscopie RMN du proton ____ 44

2- Caractéristiques du complexe ADN-protéines ____ **46**

2-1- Migration électrophorétique non dénaturante ____ 46

2-2- Protéolyse et hydrolyse des acides nucléiques des extraits de protéines ____ 53

2-3- Concentration en sites de fixation spécifiques et affinité ____ 57

2-4- Masse moléculaire apparente ____ 59

2-5- Spécificité de l'association ADN-protéines ____ 61

2-6- Comparaison de nos résultats avec des protéines connues ____ 68

Chapitre IV : Identification des protéines du complexe ____ *69*

1- Protéines identifiées par affinité (streptavidine immobilisée) ____ **69**

1-1- Choix de la séquence biotine-ADN ____ 69

1-2- Le complexe streptavidine-biotine-ADN-protéines ____ 71

1-3- Description du mode opératoire ____ 72

1-4- Essais de purification en conditions natives ____ 73

2- Identification différentielle de protéines ____ **74**

2-1- Principe de l'expérience ____ 74

2-2- Séquençage des protéines identifiées ____ 77

3- Reconnaissance du brin complémentaire riche en G ____ **80**

3-1- Caractérisation de l'association ____ 80

3-2- Identification par affinité ____ 83

3-3- Implication biologique ____ 85

4- Aspects biologiques des IMPDH ____ **87**

4-1- Caractéristiques des protéines identifiées ____ 87

4-2- Rôle biologique ____ 91

4-3- Comparaison avec les protéines humaines ____ 94

4-4- Conclusion de ces identifications et perspectives ____ 95

Chapitre V : Autres résultats ____ *99*

1- Etudes complémentaires par spectrométrie de masse ____ **99**

1-1- Les protéines Ku70p et Ku80p ____ 100

1-2- La protéine Hsp60p ____ 101

1-3- La protéine CBF3B ____ 102

2- Purifications partielles et éléments d'identification 103

2-1- Eléments de purification 103

2-2- Pontage ADN-protéines aux UV 107

2-3- Cohérence avec les résultats du chapitre IV 108

2-4- Protéines de choc thermique HSP90 109

CONCLUSION et PERSPECTIVES 113

1- Résultat principal 113

2- Autres contributions 114

3- Perspectives 115

Annexes 117

1- Liste des abréviations 119

2- Principales protéines impliquées au niveau des télomères de levures 121

3- Caractéristiques des différents tampons utilisés 123

4- Masses des peptides ionisés 124

5- Modes opératoires 125

5-1- Préparation d'extraits protéiques de levure 125

5-2- Préparation d'échantillons pour la spectrométrie de masse 126

6- Alignements de séquences 129

Bibliographie 133

Chapitre I

Introduction

Les acides nucléiques (l'ADN, acide désoxyribonucléique, et l'ARN, acide ribonucléique) sont les vecteurs de l'information génétique nécessaire au fonctionnement et à la réplication des cellules. Les molécules d'ADN constituant les chromosomes sont les macromolécules les plus longues de la cellule (jusqu'à 10^8 paires de bases).

Autres éléments actifs de la vie de la cellule, les protéines sont impliquées aussi bien dans l'organisation structurale des organismes eucaryotes et procaryotes que dans leur fonctionnement (enzyme). Les protéines sont des macromolécules résultant de la condensation d'acides aminés par formation de liaisons peptidiques. Chaque protéine se caractérise par sa composition en acides aminés et la séquence (provenant de l'information génétique inscrite dans les chromosomes) les agençant, sa longueur, son poids moléculaire et son point isoélectrique.

1- Introduction aux acides nucléiques

1-1- Les acides nucléiques

Les acides nucléiques sont constitués par l'enchaînement d'éléments chimiques, les nucléotides (figure 1.1.). Un nucléotide est composé d'une base aromatique hétérocyclique reliée par une liaison N-glycosidique à la position 1' d'un sucre pentose cyclique (β-D-2'-désoxyribose pour l'ADN et β-D-ribose pour l'ARN), phosphorylé en position 5'.

Les bases classiquement rencontrées sont classées en deux catégories :

- les purines : adénine (A) et guanine (G)
- les pyrimidines : cytosine (C) et thymine (T) (ou uracile (U) pour l'ARN).

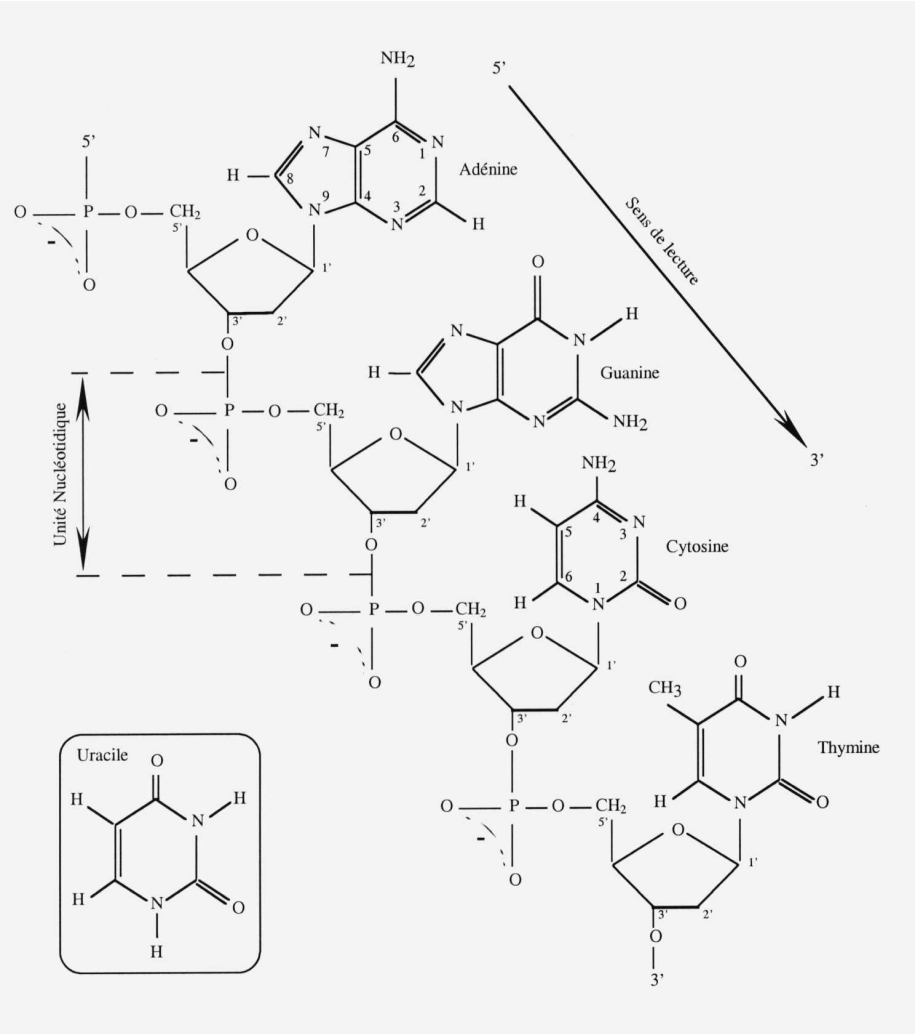


Figure 1.1. : Schéma d'un fragment d'ADN. Pour l'ARN, la base uracile et le sucre β-D-ribose remplacent la thymine et le sucre β-D-2'-désoxyribose de l'ADN.

Dans le cas des purines, c'est l'azote N_9 qui participe à la liaison N-glycosidique et dans le cas des pyrimidines, l'atome N_1.

Les sucres possèdent plusieurs degrés de liberté autour de différentes liaisons, et les angles dièdres définis par ces liaisons dépendent les uns des autres. Les conformations des sucres sont essentiellement C_2'-*endo* et C_3'-*endo* pour les doubles hélices d'ADN et d'ARN respectivement (figure 1.2.).

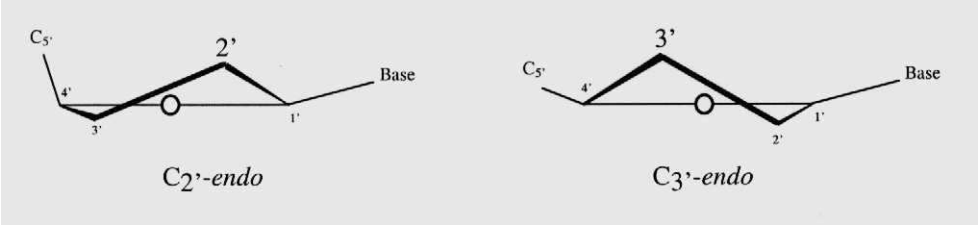


Figure 1.2. : Conformations C_2'-*endo* et C_3'-*endo* des sucres

La base peut tourner autour de la liaison N-glycosidique. L'espace conformationnel se réduit principalement à deux orientations, appelées *syn* et *anti* définies par l'angle χ, par suite de contrainte d'encombrement stérique (figure 1.3.).

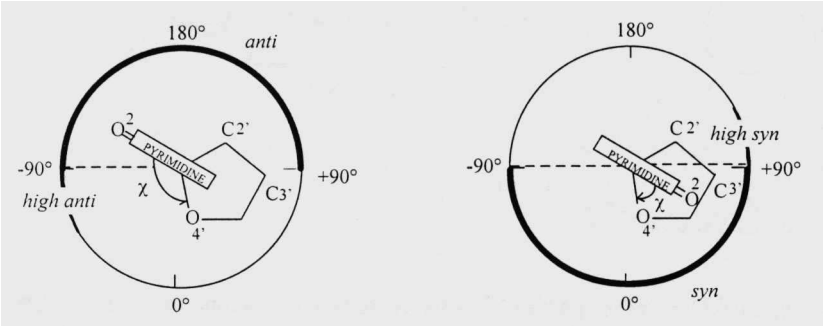
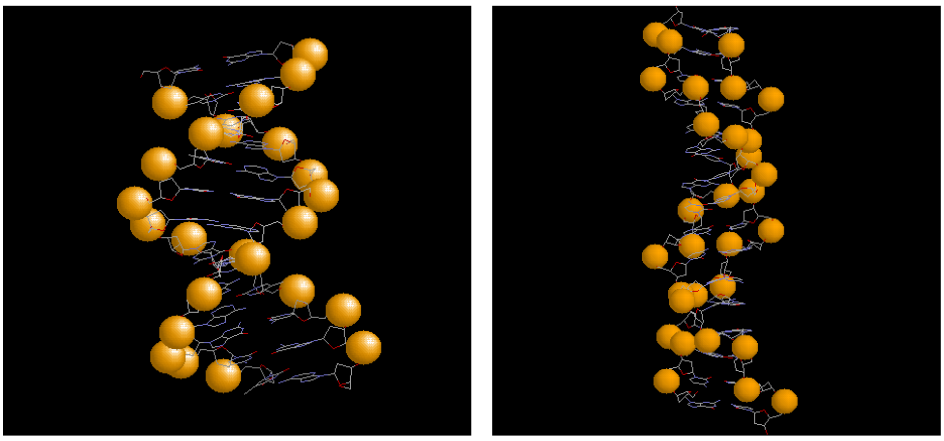


Figure 1.3. : Conformations *syn* et *anti* possibles autour de la liaison N-glycosidique pour une pyrimidine (Saenger, 1984)

Les acides nucléiques ont un caractère polyanionique dû à la présence de phosphates. Ceci permet des interactions avec des cations, modifie le pKa apparent des groupements acido-basiques et intervient dans les phénomènes de répulsion électrostatique.

1-2- Structures de l'ADN

Maximisant les interactions hydrophobes et réduisant au minimum le nombre de molécules d'eau en son cœur, la structure de l'ADN est constituée d'un empilement hélicoïdal de plateaux de bases (nature aromatique des bases, molécules planes, hydrophobes et rigides). Les bases s'égrainent donc le long de la chaîne sucre-phosphate, colonne vertébrale de l'ADN. Classiquement, le sens de lecture d'une séquence d'ADN s'effectue de l'extrémité 5' vers l'extrémité 3' de cette chaîne (figure 1.1.).



(a) ADN B *(b) ADN Z*

Figure 1.4. : Formes B et Z de l'ADN (*http://tome.cbs.univ-montp1.fr)*

La structure de l'ADN se caractérise par l'existence de deux chaînes polynucléotidiques appariées antiparallèlement l'une à l'autre en une structure hélicoïdale régulière. La forme dite « ADN B » découverte en 1953 par Watson et Crick (Watson & Crick, 1953 a et b) (figure 1.4.), est la structure stable la plus fréquemment rencontrée dans les conditions naturelles. Dans cette structure, une base pyrimidique est appariée à une purine (figure 1.5.).

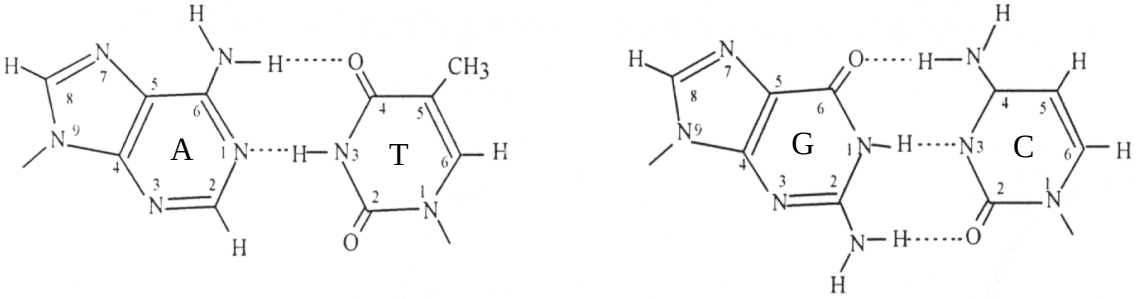


Figure 1.5. : Appariements Watson-Crick A·T et G·C

Deux autres configurations, l'ADN A et l'ADN Z ont été observées par cristallographie. Les conformations A et B montrent une forme hélicoïdale avec pas à droite, alors que l'hélice d'ADN Z (figure 1.4.) tourne dans l'autre sens. Des variations structurales dépendant de la séquence ont été découvertes, par exemple une courbure de l'hélice d'ADN (cas de l'ADN B').

Il existe encore d'autres structures dites non canoniques de l'ADN : ces structures ne s'appuient pas sur les paires de bases classiques Watson-Crick ou ne sont pas en double hélice anti-parallèle. Citons l'ADN π, un duplexe parallèle, les repliements en épingle à cheveux, et enfin les structures d'ADN à plusieurs brins telles que les hélices triples ou encore les quadruplexes de guanines et le motif i.

1-3- Les quadruplexes de guanines

Quatre guanines provenant de quatre brins distincts peuvent s'associer et former un plateau de guanines ou « tétrade de G » (figure 1.6.a.). La structure obtenue est appelée quadruplexe de guanines (figure 1.6.b.).

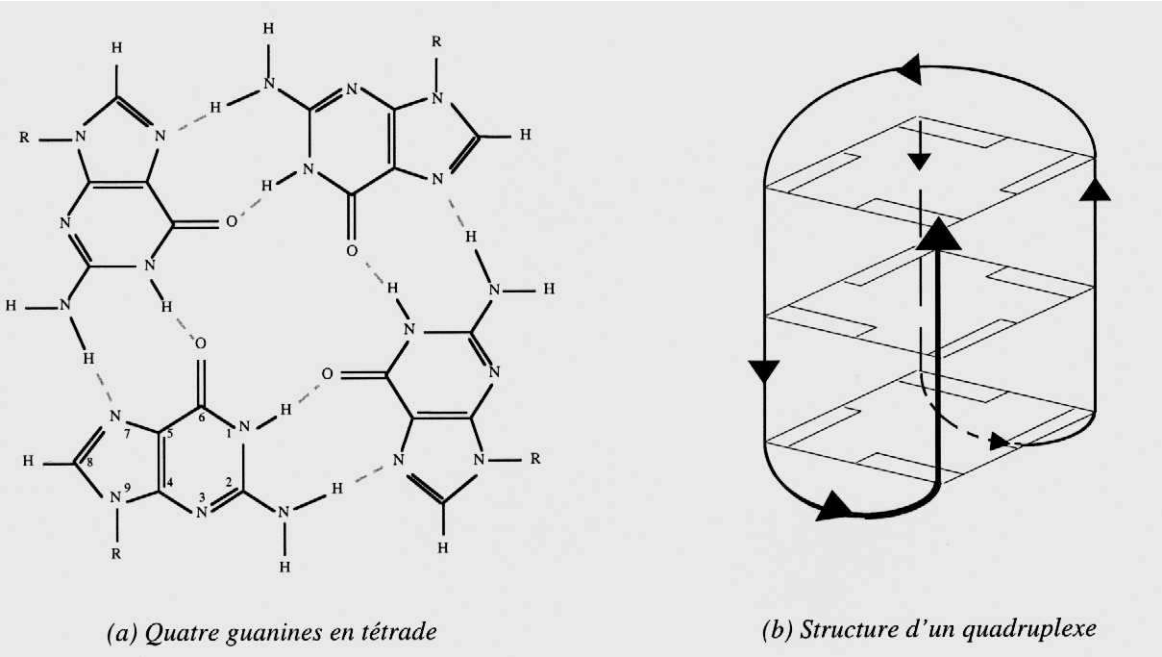


Figure 1.6. : *(a)* Tétrade de guanines. *(b)* Schéma du quadruplexe formé par le fragment du brin télomérique humain d[A(GGGTTA)$_3$GGG] (Wang & Patel, 1993).

L'empilement de tétrades est favorisé par la présence de cations monovalents s'insérant entre les plateaux (Patel et *al.*, 1999), entraînant une dépendance de l'architecture du quadruplexe en fonction de la nature du cation inséré (Sen & Gilbert, 1990; Miura et *al.*, 1999). Les structures les plus stables ont été observées quand des cations K^+ s'insèrent entre les tétrades de G.

Cette structure peut être formée de quatre brins pouvant provenir de fragments différents (quadruplexes dimériques et tétramériques) ou d'un seul (quadruplexe monomérique). Plusieurs orientations des brins sont possibles. Les brins d'un quadruplexe tétramérique sont orientés parallèlement, alors qu'il existe plusieurs conformations possibles pour les quadruplexes formés par repliement de un ou deux brins.

1-4- Le motif i

Première structure de l'ADN découverte par la RMN, le motif i est une structure reposant sur l'intercalation tête-bêche de deux duplexes parallèles formés de paires de bases $C{\cdot}CH^+$ (figure 1.7.a. et 1.7.b.) (Gehring et *al.*, 1993). Ses caractéristiques structurales principales peuvent se résumer ainsi :

- les 2 brins d'un même duplexe sont parallèles et reliés entre eux par des paires $C{\cdot}CH^+$
- la distance entre deux paires de bases intercalées est de ~ 0.31 nm, soit 0.62 nm entre deux bases consécutives d'un même brin (contre 0.34 nm dans le cas de l'ADN B)
- l'angle de rotation de l'hélice est compris entre 12° et 23° (contre 36° dans le cas de l'ADN B)
- le motif i présente deux sillons étroits et deux sillons larges
- l'angle de torsion glycosidique est toujours *anti* (figure 1.3.)
- la conformation des sucres est *C_3'-endo*

Deux topologies d'intercalation sont possibles selon que les désoxycytidines situées à l'extérieur du motif i sont localisées à l'extrémité 3' ou à l'extrémité 5'. Comme dans le cas des quadruplexes de guanines, le motif i peut être intermoléculaire (motif tétramérique ou dimérique) ou intramoléculaire (motif i monomérique), mais l'orientation des brins doit obéir à l'intercalation tête-bêche des deux duplexes. La stabilité du motif i dépend des conditions expérimentales : pH, température, concentration en brins (pour les formes multimériques) et force ionique.

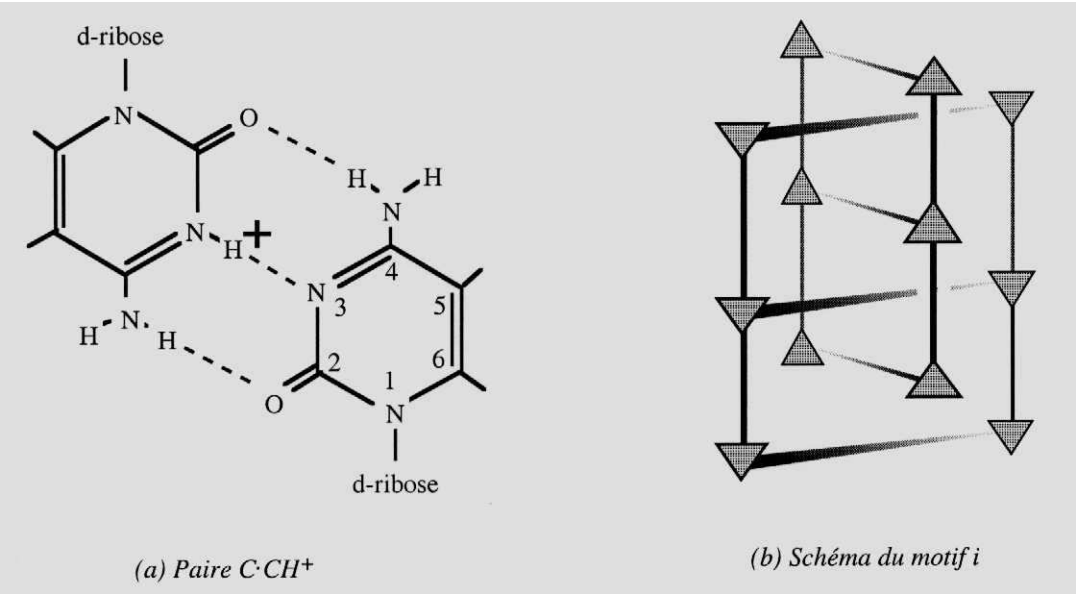


Figure 1.7. : Paire de bases C·CH$^+$ et schéma du motif i adopté par le tétramère d(^{m}CCC)$_4$, (Leroy et al., non publié)

Malgré l'obligation d'hémi-protonation des paires C·CH$^+$, le motif i intramoléculaire formé par la séquence naturelle télomérique humaine, d[(CCCTAA)$_3$CCCT] persiste à pH 7 (Phan et *al.*, 2000), et celui formé par la séquence d(CCTTTCCTTTACCTTTCC) persiste même à pH 7.4 (Han et *al.*, 1997 ; Guéron & Leroy, 2000 pour revue). Ces observations laissent supposer une existence possible du motif i *in vivo*.

Depuis la découverte du motif i, plusieurs études par cristallographie sont venues confirmer ces caractéristiques structurales (Berger et *al.*, 1995 ; Chen et *al.*, 1994 ; Kang et *al.*, 1994 ; Kang et *al.*, 1995).

2- Fonctions des acides nucléiques

2-1- Transmission et expression de l'information génétique contenue dans les acides nucléiques

L'information nécessaire à la synthèse des protéines est contenue dans des segments de chromosomes, les gènes. Pour les cellules eucaryotes, ces gènes peuvent être interrompus par des séquences non codantes, les introns. Les segments codants séparés par ces introns sont appelés exons. L'élaboration d'une protéine est réalisée à partir d'un gène lors de deux étapes distinctes : la transcription et la traduction.

La transcription est le transfert de l'information génétique contenue dans l'ADN d'un gène à un ARN messager (ARN_m). Cette transcription est initiée par l'association d'ARN polymérase avec un site promoteur de l'ADN et conduit à la synthèse de plusieurs copies d'ARN_m.

Lors de la traduction, le ribosome, un complexe ARN-protéines, traduit l'information contenue par l'ARN_m en polymères linéaires d'acides aminés, les protéines. Cette traduction fait intervenir les ARN de transfert (ARN_t). Ceux-ci reconnaissent des segments de trois bases consécutives, les codons, en y associant les acides aminés correspondants (pour chaque acide aminé, il existe au moins un codon), les insérant à la position adéquate dans la séquence polypeptidique. La traduction de l'ARN_m aboutira à la fabrication des protéines correspondantes.

2-2- Mécanisme de réplication

La réplication génère une copie fidèle des chromosomes d'une cellule et s'effectue dans les deux directions à partir d'un point origine. Chaque brin sert de modèle à un nouveau brin. L'ADN est synthétisé avec une grande fidélité par une ADN polymérase qui fonctionne uniquement dans le sens 5' $\rightarrow$ 3' (figure 1.8.). Au niveau de la fourche de réplication, la synthèse de la nouvelle molécule d'ADN procède de façon asymétrique : le brin directeur est synthétisé de façon continue dans le sens de progression de la fourche de réplication, alors que l'autre brin est synthétisé de façon discontinue, par séquences de 1000 à 2000 bases, appelées « fragments d'Okazaki » (Okazaki & Okazaki, 1969).

La synthèse du brin discontinu s'appuie sur des amorçages répétés orchestrés par une primase générant des amorces d'ARN. Après chaque amorçage, il y a élimination de l'amorce d'ARN et reprise du processus de synthèse par une polymérase d'ADN à partir de l'extrémité 3' de l'ADN issue du processus d'initiation suivant. Ce mécanisme permet la synthèse de l'ensemble du brin à

synthèse discontinue, à l'exception du dernier segment (long d'une quinzaine de nucléotides) pour lequel il n'y a pas de synthèse possible. En l'absence d'autre mécanisme, il en résulte que la réplication de l'ADN s'accompagne à chaque fois d'une perte à l'extrémité 5' du chromosome.

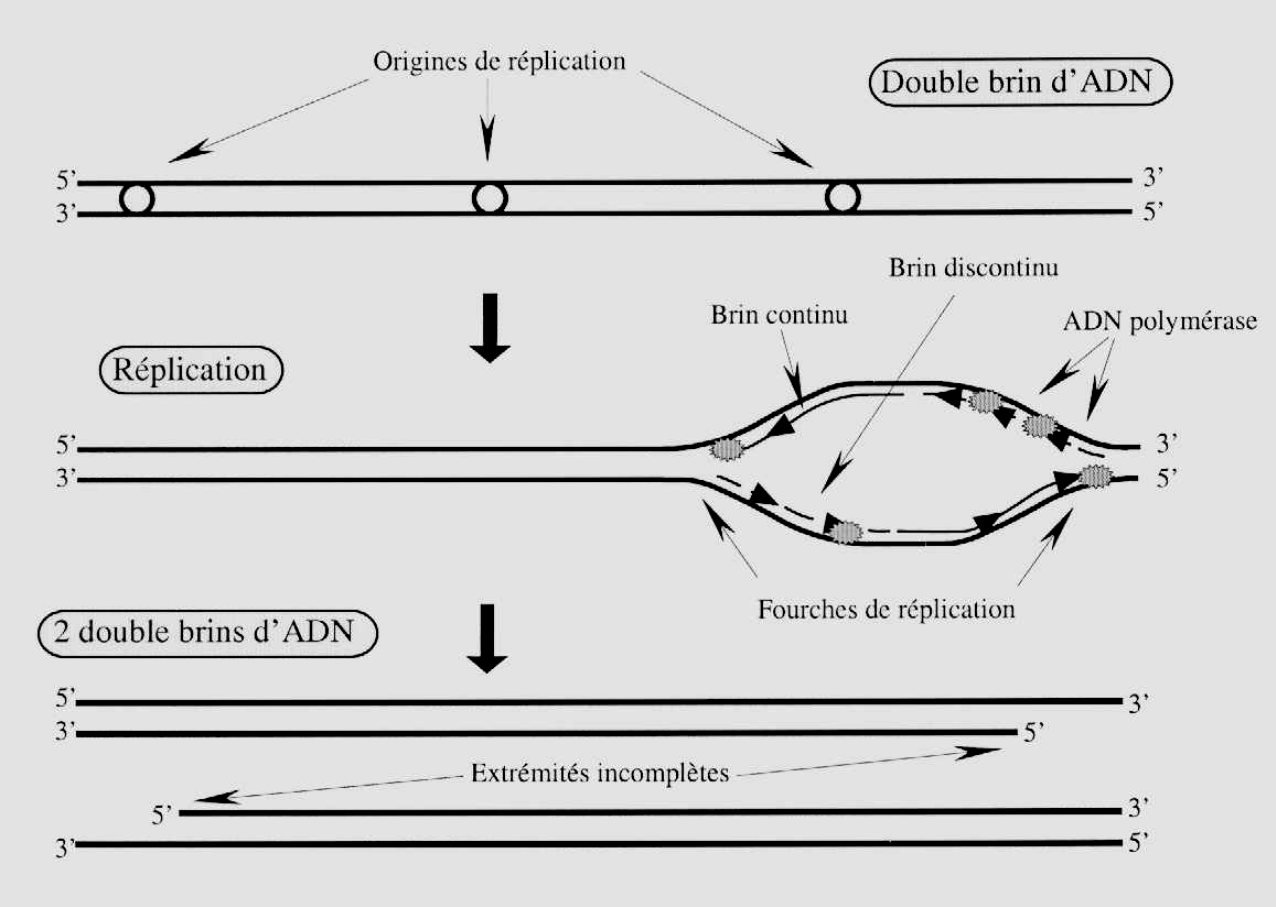
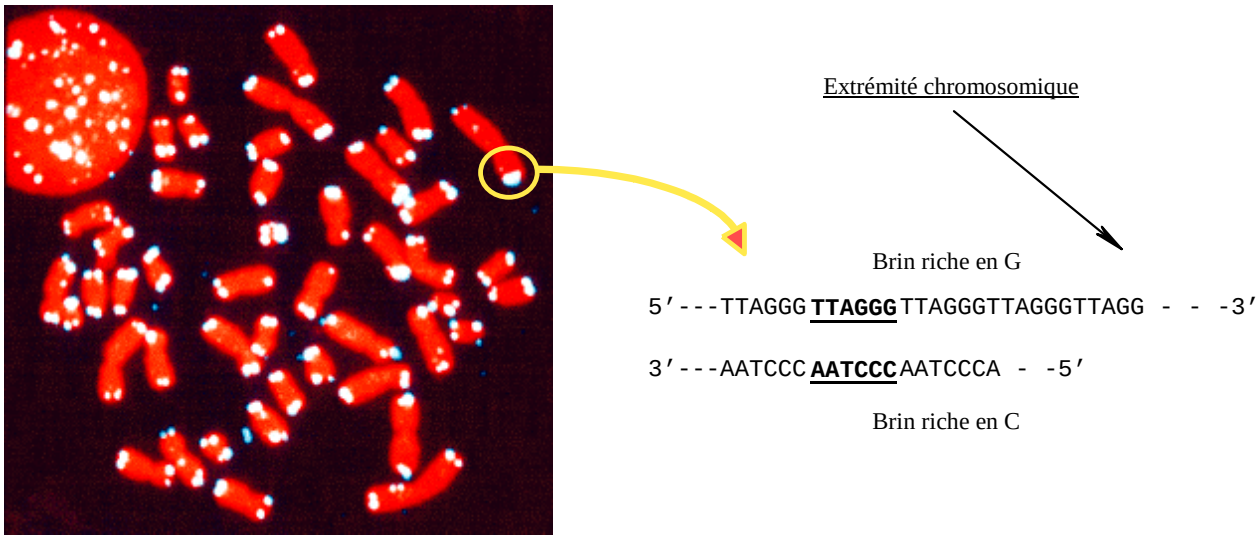


Figure 1.8. : Mécanisme de réplication; fonctionnement d'une fourche de réplication (Okazaki & Okazaki, 1969) et raccourcissement consécutif (Olovnikov, 1973).

2-3- Les télomères, extrémités des chromosomes

L'ADN codant ne représente que ~ 5% de l'ADN chromosomique humain. Le reste contient notamment des zones constituées d'un grand nombre de répétitions d'une séquence élémentaire. C'est le cas des extrémités des chromosomes, appelés télomères (figure 1.9.a.). Ces éléments structuraux, composés de répétitions d'une séquence d'ADN spécifique, associées à des protéines, ont un rôle de maintien de l'intégrité des chromosomes. Le schéma d'organisation de l'ADN télomérique de tous les organismes étudiés à ce jour présente une extension monocaténaire riche en

G à l'extrémité 3' (Greider, 1990 ; Blackburn, 1991 ; Greider, 1991 ; Blackburn, 1994 ; Fang & Cech, 1993 ; Price, 1990 ; Zakian, 1995). Par exemple, l'ADN télomérique des mammifères se compose de la séquence répétée, d[(CCCTAA)·(TTAGGG)] (figure 1.9.b.).



(a) Séquences télomériques détectées par fluorescence

(b) Répétitions d'ADN télomériques et décalage brin G / brin C

Figure 1.9. : *(a)* Chromosomes métaphasiques humain hybridés *in situ* avec la séquence humaine $d(TTAGGG)_n$ fluorescente (Kipling, 1995). *(b)* Schéma des deux brins télomériques des mammifères.

Les répétitions télomériques sont généralement régulières, à l'exception de celles de certains champignons et protozoaires. Leur longueur varie de quelques bases (cas des organismes ciliés) à plusieurs dizaines de kilobases (pour certains mammifères) (tableau 1.1.). A titre d'exemple, la longueur de l'ADN télomérique humain est supérieure à 10000 paires de bases (de Lange, 1994) et la longueur du monobrin riche en G est d'une centaine de bases (Blackburn, 1991).

Enfin, chacun des brins de l'ADN télomérique pris séparément peut former *in vitro*, soit un quadruplexe de guanines (Sen & Gilbert, 1988), soit un motif i (Ahmed et *al.*, 1994 ; Leroy et *al.*, 1994 ; Manzini et *al.*, 1994). Par exemple, les structures par RMN du quadruplexe de guanines d'un fragment d'ADN télomérique humain (Wang & Patel, 1993) et du motif i de la séquence complémentaire (Phan et *al.*, 2000) ont pu être déterminées.

Organismes	Longueurs (en paires de bases)	Séquences brin C 5'→3'	Séquences brin G 5'→3'
Protozoaires			
Tetrahymena	< 100 [a]	C_4A_2	T_2G_4
Oxytricha, Euplotes	< 100 [b,c]	C_4A_4	T_4G_4
Paramecium	< 100 [a]	$C_3[C/A]A_2$	$T_2[T/G]G_3$
Plasmodium	~ 1500	$C_3T[G/A]A_2$	$T_2[T/C]AG_3$
Champignons			
Neurospora		C_3TA_2	T_2AG_3
Cryptococcus		$C_{3-5}TA_2$	T_2AG_{3-5}
Saccharomyces cerevisiae	< 500 [d]	$C_{2-3}A(CA)_{1-6}$	$(TG)_{1-6}TG_{2-3}$
Schizosaccharomyces		$C_{1-6}G_{0-1}T_{0-1}GTA_{1-2}$	$T_{1-2}ACA_{0-1}C_{0-1}G_{1-6}$
Plantes			
Chlamydomonas	< 500 [e]	C_3TA_4	T_4AG_3
Arabidopsis	3000 [f]	C_3TA_3	T_3AG_3
Invertébrés			
Ascaris		GC_2TA_2	T_2AG_2C
Bombyx mori	> 10000 [g]	C_2TA_2	T_2AG_2
Vertébrés [h]		C_3TA_2	T_2AG_3
Homme	> 10000 [i]	C_3TA_2	T_2AG_3
Mus musculus	> 50000 [j]	C_3TA_2	T_2AG_3

Tableau 1.1. : Séquences télomériques de quelques organismes (d'après Wellinger & Sen, 1997). [a] Blackburn & Gall, 1978 ; [b] Oka et *al.*, 1980 ; [c] Wellinger et *al.*, 1993 ; [d] Shampay et *al.*, 1984 ; [e] Petracek et *al.*, 1990 ; [f] Richards & Ausubel, 1988 ; [g] Okazaki et *al.*, 1993 ; [h] Meyne et *al.*, 1989 ; [i] de Lange, 1994 ; [j] Kipling & Cooke, 1990.

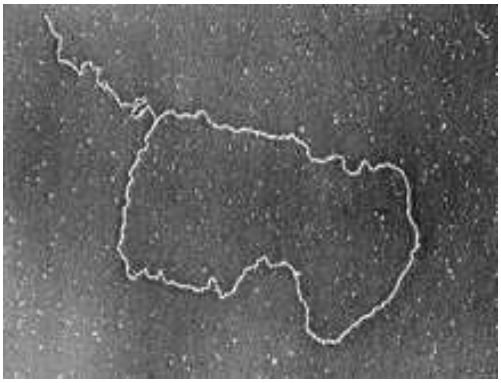
2-4- Mécanismes liés à la régulation des télomères

Lors de la réplication de la cellule, la longueur des télomères diminue (environ 50-200 bases par division pour les mammifères) selon le principe proposé par Olovnikov (Olovnikov, 1996 pour revue). En l'absence de mécanisme compensateur, ce phénomène limite le nombre possible de réplications cellulaires. A partir d'une longueur critique, un mécanisme conduit à la mort de la cellule. Mais le raccourcissement peut être corrigé par l'intermédiaire d'une transcriptase inverse spécialisée, la télomérase (Blackburn, 1992 ; Lingner et *al.*, 1995 ; Counter, 1996 ; Chiu & Harley, 1997 ; O'Reilly et *al.*, 1999 pour revue récente). Elle se compose d'une partie protéique et d'un ARN servant de matrice pour la synthèse de l'ADN télomérique. La région de l'ARN télomérase humain servant de modèle est le fragment r(CUAACCCUAAC) codant pour la répétition d'ADN télomérique d(TTAGGG).

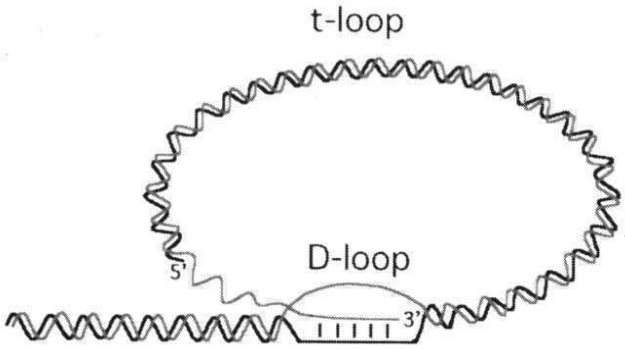
L'environnement nucléoprotéique d'un télomère régule l'activité de la télomérase. Des expériences ont montré *in vivo* que l'activité de la télomérase était régulée par la présence de protéines s'associant au duplexe (Bianchi et *al.*, 1997 ; Bilaud et *al.*, 1997 ; Van Steensel & de Lange, 1997 ; Ancelin et *al.*, 1998 ; Van Steensel et *al.*, 1998) et qu'elle nécessitait la présence de certaines protéines affines du brin G ainsi qu'une dégradation sélective du brin C (Wellinger et *al.*, 1996 ; Makarov et *al.*, 1997).

Un dérèglement de l'activité de la télomérase pourrait être responsable de l'apparition de certaines tumeurs (Shay, 1997 ; Holt & Shay, 1999). Ainsi, il a été récemment montré que l'inactivation de la télomérase jouait un rôle limitant dans la croissance des cellules tumorales, mais son activation seule ne suffit pas pour immortaliser toutes les cellules primaires ou les transformer en cellules tumorales (Hahn et *al.*, 1999a et b).

L'extrémité télomérique des chromosomes présente une caractéristique structurale récemment mise en évidence (Griffith et *al.*, 1999). Les télomères forment une boucle (observation par microscopie électronique (figure 1.10.a.) sur des extrémités télomériques de clône de cellule *HeLa*), la t-boucle. Au niveau de la fermeture de cette t-boucle, le monobrin G s'associerait avec la double hélice télomérique pour former une D-boucle (figure 1.10.b). Cette D-boucle pourrait se déplacer le long du télomère et sa structure serait induite par la protéine TRF2 (Greider, 1999).



(a) Photographie d'une extrémité télomérique d'un chromosome humain observée par microscopie électronique (Griffith et al., 1999)



(b) Schéma de la structure possible de l'extrémité des télomères (Lundblad, 2000)

Figure 1.10. : *(a)* Observation par microscopie électronique d'une t-boucle (ADN télomérique d'un clone de cellule *HeLa*) ; *(b)* Schéma proposé par Greider (Greider, 1999) de la structure du télomère.

Ces observations placent dans un contexte nouveau le suicide programmé d'une cellule (appelé apoptose). Ce phénomène reste mal compris et les mécanismes biologiques permettant de distinguer un télomère trop court, non élucidés. On pourrait imaginer que la t-boucle ne puisse plus se former si le télomère est trop court. La cellule pourrait le détecter alors comme un chromosome cassé et déclencher l'apoptose.

Plusieurs hypothèses structurales sont envisagées concernant les mécanismes associatifs intervenant au sein de la D-boucle :

- le monobrin G s'associe au brin C en s'insérant en lieu et place du brin G normalement impliqué dans le duplexe d'ADN télomérique. Celui-ci peut alors soit rester libre, soit se replier en quadruplexe de G ;
- les deux brins G forment chacun un quadruplexe ou s'associent en un quadruplexe multimérique et le brin C reste libre ou forme un motif i ;
- les trois brins s'associent et forment une triple hélice C·G×G selon un modèle similaire à celui proposé pour *Tetrahymena* (Veselkov et *al.*, 1993).

Actuellement, aucune des solutions proposées n'a été confirmée. La compréhension des interactions entre les brins télomériques et les facteurs extérieurs intervenant dans la régulation des télomères laissent envisager de nouveaux champs d'application thérapeutique concernant le vieillissement cellulaire et pourraient permettre de comprendre les étapes liées à la cancérogenèse. Différentes études sont en cours concernant, par exemple la possible formation d'une triple hélice C·G×G à partir de fragments d'ADN télomérique humain, ou encore, la stabilisation de quadruplexe de G par des ligands.

2-5- Protéines spécifiques des télomères

Plusieurs groupes de recherches (Blackburn, 1994 ; Zakian, 1995 ; Greider, 1996 ; Holt et *al.*, 1996 ; Olovnikov, 1996 ; Greider, 1998 ; Shore, 1998 ; Gottschling & Stoddard, 1999 ; O'Reilly et *al.*, 1999) ont dirigé leurs efforts dans la découverte et la caractérisation de protéines se liant spécifiquement aux séquences télomériques. Des protéines ont été identifiées comme reconnaissant le double brin, par exemple Rap1p pour la levure *Saccharomyces cerevisiae* (Giraldo & Rhodes, 1994, Giraldo et *al.*, 1994) (figure 1.12), PPT pour *Physarum* (Coren et *al.*, 1991) et Trf1p/Trf2p

dans les cellules des mammifères (Zhong et *al.*, 1992 ; Chong et *al.*, 1995 ; Bilaud et *al.*, 1997) (figure 1.11.).

Le simple brin télomérique riche en G de *Saccharomyces cerevisiae* est reconnu par les protéines Est1p (Virta-Pearlman et *al.*, 1996) et Cdc13p, qui assurent un rôle de protection du télomère (Porter et *al.*, 1996 ; Lin & Zakian, 1996). Le simple brin est aussi reconnu par les sous-unités protéiques α et β chez *Oxytricha* (Raghuraman & Cech, 1989), TBP chez *Euplotes crassus* (Price, 1990), TEP et TGP chez *Tetrahymena* (Sheng et *al.*, 1995 ; Schierer & Henderson, 1994), GBP chez *Chlamydomonas* (Petracek et *al.*, 1994), XTEF chez *Xenopus* (Cardenas et *al.*, 1993), MyoD dans certaines cellules de mammifères (Weisman-Shomer & Fry, 1993) et hnRNP A2/B1 dans les cellules *HeLa* (Ishikawa et *al.*, 1993). D'autres protéines interagissent spécifiquement avec les tétrades de G (Frantz & Gilbert, 1995 ; Liu et *al.*, 1995), ou encore induisent cette formation *in vitro* (Giraldo & Rhodes, 1994 ; Giraldo et *al.*, 1994).

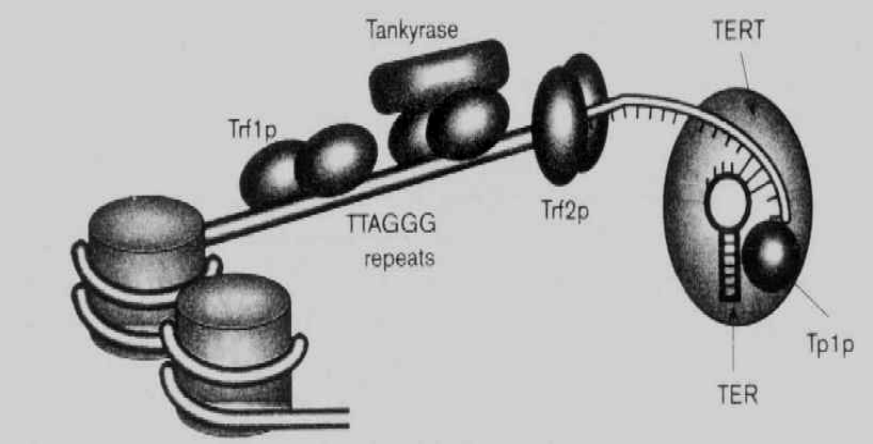


Figure 1.11. : Schéma d'une organisation fonctionnelle possible de l'extrémité télomérique humaine (O'Reilly et *al.*, 1999).

Concernant les séquences télomériques complémentaires riches en C, une protéine se liant spécifiquement au fragment d((CCCTAA)$_3$CCCT) a été mise en évidence pour les cellules humaines *HeLa* (Marsich et *al.*, 1996) et les cellules de mammifères (Marsich et *al.*, 1998). Deux autres protéines issues de cellules *HeLa*, ASF/SF2 et hnRNP K, ont été identifiées comme reconnaissant spécifiquement la séquence naturelle d((TAACCC)$_4$TAA) (Lacroix et *al.*, 2000). Ces protéines sont à même de s'associer avec d'autres séquences riches en C capables d'adopter le motif i *in vitro*, mais on ne connaît pas encore la structure de l'ADN dans le complexe.

2-6- Région centromérique

Chez l'homme, la zone centrale des chromosomes, appelée centromère, présente elle aussi des répétitions de séquences élémentaires riches en cytosines sur un brin et en guanines sur le brin complémentaire. Défini par les cytologistes comme la constriction primaire des chromosomes, le centromère intervient lors de la méiose et de la mitose. Le centromère humain représente environ 10% du génome, et se compose principalement de 5 satellites, dont les plus importants sont les satellites α et III. Le premier comporte des suites répétées de 171 paires de bases, contenant une sous-séquence de 17 nucléotides, d(TCCCGTTTCCAACGAAG), (site de l'association d'une protéine, la centromeric protein B ou CEN-P B), dont une partie, d(TCCCGTTTCCA) forme *in vitro* un motif i dimérique (Gallego et *al.*, 1997).

Le satellite III contient des millions de bases de répétitions presque parfaites de la séquence d((CCATT)·(TAAGG)) (Grady et *al.*, 1992). D'un point de vue structural, on sait que le brin riche en G peut adopter *in vitro* des structures 'tige-boucle' (Chou et *al.*, 1994 ;Zhu et *al.*, 1995) et que la séquence naturelle d[$(CCATT)_3$CC] peut former des structures en motif i monomériques ou dimériques (forme favorisée par la présence de KCl) (Nonin et *al.*, 1997 ; Nonin-Lecomte & Leroy, 2001).

A la différence des régions télomériques et de certaines protéines centromériques (CENP-A, B et C) remarquablement conservées entre espèces différentes, les ADN centromériques des différents mammifères ne présentent quasiment aucune homologie. Dans ce contexte, la variabilité des structures en motif i (mono- ou dimériques selon les séquences de boucles) adoptées par le brin C du satellite III *in vitro*, n'exclut pas un rôle biologique. Le brin complémentaire riche en G présente une variabilité comparable.

3- Le sujet de thèse

Les télomères, comme les centromères, présentent deux brins complémentaires pouvant adopter *in vitro* deux structures non canoniques, le quadruplexe de guanines et le motif i. Ce particularisme structural et l'importance biologique des régions chromosomiques impliquées renforcent la question

du rôle biologique de ces structures. Depuis la découverte du motif i, et de protéines associées au brin G initiant la formation de quadruplexes de guanines, les régions riches en répétitions de cytosines et de guanines, et plus particulièrement, les télomères, font l'objet d'un regain d'intérêt. Afin de caractériser une éventuelle fonction biologique du motif i, nous avons entrepris de rechercher des protéines capables de reconnaître cette structure. Cette étude a été entreprise chez une levure, *Saccharomyces cerevisiae*, dont le génome présentait l'avantage d'être totalement séquencé.

3-1- Les télomères de levure

Chez la levure *Saccharomyces cerevisiae*, la longueur des télomères est d'environ 300 ± 75 paires de bases. La séquence consensus du brin C des télomères est $C_{2\text{-}3}A(CA)_{1\text{-}6}$, mais il n'existe pas de répétitions régulières comme la séquence C_3TA_2 pour les mammifères. La matrice d'ARN de la télomérase de *Saccharomyces cerevisiae*, r ($^{3'}$ACACACACCCACACCAC$^{5'}$), plus longue que celle de la plupart des autres organismes [par exemple, r($^{3'}$CAATCCCAATC$^{5'}$) chez l'homme], n'est que partiellement copiée, selon un positionnement variable de l'amorce d'ADN non encore élucidé (Nugent & Lundblad, 1998 pour revue). Le gène *EST2* code la transcriptase inverse de la télomérase. Les protéines Est1p, Est3p et Est4p/Cdc13p composant la télomérase peuvent avoir un rôle de facteur de régulation. Un ensemble spécifique de protéines (Cdc13p, Est1p, Stn1p) est probablement directement impliqué dans la protection et la régulation de la longueur des télomères par la télomérase (Lin & Zakian, 1996).

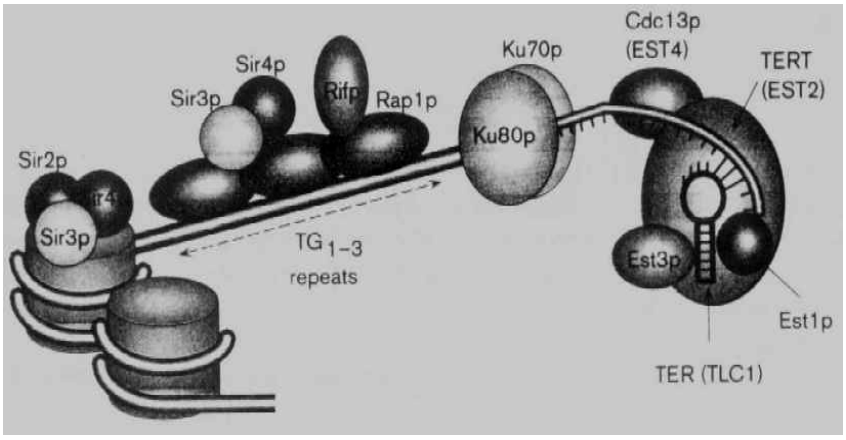


Figure 1.12. : Schéma d'une organisation fonctionnelle possible de l'extrémité télomérique de *Saccharomyces cerevisiae* (O'Reilly et *al.*, 1999).

L'élongation du brin G des télomères de *Saccharomyces cerevisiae* est étroitement liée à celle du brin opposé par les ADN polymérases α et δ (Diede & Gottschling, 1999). Plusieurs protéines, s'associant à l'ADN télomérique de *Saccharomyces cerevisiae*, coopèrent et participent aussi à la régulation de la longueur des télomères (figure 1.12., voir tableau non exhaustif en annexes) dont :

- les protéines Sir1p, Sir2p, Sir3p et Sir4p (Bourns et *al.*, 1998) par ailleurs impliquées dans un phénomène de régulation des gènes, le « silencing »;
- les protéines Rif1p et Rif2p (Bourns et *al.*, 1998) interagissant avec Rap1p (Rap1p interacting factor). Rif2p coopère avec Rif1p pour réguler la longueur des télomères (Wotton & Shore, 1997) ;
- la protéine Rap1p dont la surproduction altère la structure du télomère (Conrad et *al.*, 1990). Cette protéine favorise aussi la formation de quadruplexe de guanines (Giraldo et *al.*, 1994). Elle peut induire une distorsion de la double hélice d'ADN télomérique (Gilson et *al.*, 1993) et est également impliquée dans l'association des deux brins d'ADN (Gilson et *al.*, 1994).
- les protéines Ku70p et Ku80p reconnaissent la transition entre le double brin télomérique et l'extension monocaténaire riche en G, plus longue (Bianchi & de Lange, 1999) et interviennent en cas de dommage sur le double brin d'ADN (Martin et *al.*, 1999).
- la protéine Tel2p interagit avec le double brin d'ADN télomérique (Kota & Runge, 1998) et se fixe spécifiquement sur le simple brin G. Elle pourrait avoir un rôle dans la régulation des télomères (Kota & Runge, 1999).

Les gènes situés à proximité des télomères peuvent être inactivés par un phénomène appelé « silencing » (Loo & Rine, 1995), où l'on retrouve impliquée la protéine Rap1p (Cockell M. et *al.*, 1995), les protéines Ku70p et Ku80p (Martin et *al.*, 1999 ; Mishra & Shore, 1999) et les protéines Sir (Ivy et *al.*, 1986). Les télomères servent alors de réservoir pour des facteurs de transcription impliqués dans la structure de la chromatine réprimée (Marcand et *al.*, 1996), ce qui permet de faire un lien entre l'activité transcriptionnelle de la cellule et l'intégrité des télomères.

Un autre phénomène appelé « Telomere Position Effect » ou TPE (Gottschling et *al.*, 1990, Kyrion et *al.*, 1993) régule l'expression des gènes placés à proximité des télomères. Plusieurs protéines déjà citées jouent un rôle dans ce phénomène, parmi lesquelles, la protéine Tel2p (Runge & Zakian, 1996 ; Zakian, 1996) et les protéines Ku70p et Ku80p (Evans et *al.*, 1998 pour revue).

3-2- Présentation du sujet de thèse

L'étude développée dans le présent ouvrage a été entreprise à partir d'une séquence télomérique et d'extraits de protéines de *Saccharomyces cerevisiae*, respectivement synthétisés et produits au laboratoire.

Ce projet comporte plusieurs étapes :

- mettre en évidence une (ou plusieurs) protéine(s) de levure se liant spécifiquement à une séquence riche en cytosines ;
- déterminer les caractéristiques et la spécificité de cette association ;
- identifier les protéines impliquées ;
- purifier les protéines identifiées et vérifier qu'elles s'associent avec le brin C de l'ADN télomérique ;
- comprendre les mécanismes de cette association.

Chapitre II

Matériels et Méthodes

1- Biochimie des protéines et des acides nucléiques

L'organisme servant de base de travail à l'étude présentée dans ce manuscrit est une levure, *Saccharomyces cerevisiae*, communément appelé « levure du boulanger » dont le génome est entièrement séquencé.

1-1- Culture de *Saccharomyces cerevisiae* et extraits de levure

1-1-1- Souches utilisées

Plusieurs souches de levure *Saccharomyces cerevisiae* ont été utilisées durant cette étude (tableau 2.1. pour les deux principales).

Dénomination	Génotype
YDS 38	*MATα*, *ade*2-1, *trp*1-1, *can*1-100, *leu*2-3, 112, *his*3-11, 15, *ura*3-1, hmr :: TRP1
EG124 = GA59	*MAT**a***, *trp*1 *ura*3-52 *leu*2-3, 112 *prb*1-1122 *pep*4-3 *prc*1-407 *gal*2

Tableau 2.1. : Les deux principales souches de levure sources des extraits protéiques étudiés dans ce projet. Nous avons commencé par travailler avec la souche YDS 38. A partir de décembre 1998, tout le travail a été effectué avec la souche EG124 construite pour produire peu de protéases (Cockell et *al*, 1995).

1-1-2- Conditions de culture

La culture des levures *Saccharomyces cerevisiae* est réalisée en milieu liquide YPG (Yeast Peptone Glucose), sous agitation permanente (oxygénation des levures), à une température de 30°C. Pour permettre un bon démarrage de la culture, une colonie isolée provenant de levures souches (stockées sur boîtes de Petri à 4°C) est prélevée pour incubation pendant 24h dans un milieu de préculture d'une vingtaine de millilitres. Ce milieu est ensuite versé directement dans un fernbach de 1 litre de YPG pour 60 heures environ à 30°C sous agitation continue. La croissance des levures est suivie par la mesure de l'absorbance à 600 nm, lors de prélèvements de 1 ml du milieu de culture. Le développement de la culture est stoppé en fin de phase de croissance exponentielle.

1-1-3- Extraction des protéines

Plusieurs méthodes permettent de briser les levures en vue d'en extraire des protéines. Parmi les plus répandues, citons la méthode consistant à les écraser, dite « French press », ou celle consistant à déchirer les membranes par frottements contre des billes de verre. Nous avons commencé à travailler selon ces modes de préparation (cf. Annexes). Ces deux méthodes ne permettent pas de traiter des quantités conséquentes de levures (faible volume de la chambre d'écrasement dans le cas de la French press et viscosité croissante du milieu d'extraction dans le cas des billes de verre, diminuant d'autant l'efficacité de ces méthodes). Ne pouvant atteindre un objectif de production (plusieurs millilitres d'extraits de protéines à une concentration moyenne de 1 µg / 1µl, provenant de plusieurs dizaines de grammes de levures) nous permettant de travailler durablement avec des protéines venant d'une même préparation, nous avons développé une méthode d'extraction plus adaptée à nos besoins.

Cette méthode repose sur la cassure des levures par des congélations quasi-instantanées (azote liquide) et successives, avec entre chaque congélation des périodes de réchauffement en bain-marie à 25°C - 30°C. Pour faciliter ces opérations, les levures récupérées par centrifugation après culture, sont remises en suspension dans un faible volume d'eau (maximum 5% volume / masse des levures).

Nous avons pu contrôler l'efficacité de la méthode par inspection des cellules au microscope, lors des étapes de congélation/décongélation. En présence de SDS (1%), les levures cassées forment des agrégats alors que celles dont les membranes sont intactes restent isolées. La première congélation n'est efficace qu'à 30% alors qu'après la seconde congélation, plus de 95% des levures sont agrégées, donc effectivement cassées. Cependant, cette technique ne s'applique pas à la récupération de protéines membranaires, car les membranes sont simplement rompues. Elle ne donne pas non plus de résultats satisfaisants lorsque les levures sont collectées en phase stationnaire (membranes trop épaisses).

1-1-4- Protection contre les protéolyses

L'ensemble des étapes menant à l'obtention d'extraits de protéines de levure se fait en présence d'anti-protéases (voir modes opératoires en Annexes) afin d'éviter les protéolyses, principalement

lors des phases de réchauffement. Les anti-protéases utilisées au cours de la préparation de nos extraits protéiques, ainsi que les concentrations efficaces sont données dans le tableau 2.2..

Anti-protéases	Cibles	Concentrations habituellement utilisées
PMSF	Trypsine, Chymotrypsine, protéases à sérine	0.1 – 1 mM
Pefabloc®	Kallikrein, Thrombine, TPA, Subtilisine A, Trypsine, Chymotrypsine, Plasmine, protéases à sérine	0.4 – 4 mM
Benzamidine	Peptidases	0.1 µM
Aprotinine	Kallikrein, Trypsine, Chymotrypsine, Plasmine	0.01 – 0.3 µM
Inhibiteur de Trypsine (ovomucoïde)	Trypsine	0.5 – 5 µM
EDTA	Metalloprotéases	0.5 – 1.3 mM

Tableau 2.2. : Anti-protéases utilisées lors des étapes d'extraction des protéines de levure et de purification par chromatographie liquide. (sources : Sambrook et *al.*, 1989 ; Boehringer Mannheim : The complete guide for protease inhibition, 1998)

1-1-5- Préparation des extraits protéiques. Dosage colorimétrique des protéines.

Les levures cassées sont centrifugées et le surnageant récupéré. Les acides nucléiques présents sont éliminés par précipitation à la streptomycine (8mg /ml) (Spelsberg, 1983). Aucune amélioration à cette séparation n'est apportée par une précipitation complémentaire à l'aide de polyéthylèneimine (PEI) (Murray P., 1990). Nous avons vérifié qu'aucune co-précipitation importante de protéines du surnageant n'avait lieu lors de ces étapes en hydrolysant les acides nucléiques précipités à la streptomycine (utilisation d'une endonucléase, la benzonase, et gel de protéines).

Les protéines sont ensuite précipitées au sulfate d'ammonium à 55%, et le précipité obtenu est repris dans du tampon Tris (45 mM) borate (45 mM) EDTA (1 mM) (= TBE 0.5x) à pH 8, en présence de β-mercaptoéthanol (5mM), de détergent non ionique Nonidet P40 (20µl/l), de PMSF (0.5 mM), de Pefabloc® (4 mM), de benzamidine (6 µM), d'aprotinine (8.5 nM), d'inhibiteur de trypsine (3 nM) et de NaCl (50 mM). L'extrait de protéines obtenu est dialysé pour éliminer le sulfate d'ammonium, pendant 5 heures à 4°C contre du TBE 0.5x en présence de NaCl 10 mM, β-mercaptoéthanol 5mM, de $MgCl_2$ 2 mM et de Nonidet (20µl/l).

Le dosage des protéines présentes dans l'extrait est réalisé par colorimétrie (Bradford, 1976) à l'aide du kit Biorad Protein Micro Assay (1-20µg/l)). Une gamme étalon est réalisée à l'aide de BSA (sérumalbumine bovine). La concentration finale en protéines des extraits préparés varie de 1 à 2 µg protéines /µl d'extraits.

1-2- Synthèse d'oligodésoxynucléotides

La synthèse des acides désoxyribonucléiques est réalisée avec un synthétiseur Gene Assembler (Pharmacia). La séquence d'ADN est synthétisée de l'extrémité 3' vers l'extrémité 5' (Blackburn Gait, 1990), à partir des dérivés β-cyanoéthyle phosphoramidites des nucléosides. Le support solide utilisé est constitué de billes de silice sur lesquelles est greffé le premier nucléotide. La réaction de couplage est conduite en milieu anhydre dans l'acétonitrile, en présence de tétrazole comme activateur. Les synthèses sont réalisées à l'échelle de 2 µmol ou de 15 µmol. L'insertion d'une biotine dans la chaîne d'ADN est possible en utilisant une biotine amidite ou biodite® (Pharmacia) comme base modifiée. La partie biotine de la molécule est séparée de la partie amidite par un lien d'une dizaine d'atomes à partir du groupement carboxyle de la chaîne latérale de la biotine. Nous avons généralement procédé à ce type d'insertion à l'extrémité 5' de la séquence synthétisée.

Les oligodésoxynucléotides synthétisés sont purifiés par passage sur colonne anionique Q-Sepharose (Pharmacia), dans un tampon à 60 mM NaOH (pH 12). L'élution est réalisée à l'aide d'un gradient de NaCl optimisé selon la séquence de l'oligodésoxynucléotide à purifier, et suivie par mesure de l'absorbance à 280 nm (longueur d'onde choisie pour ne pas saturer le détecteur). Après neutralisation, l'échantillon est dessalé par dialyse. La concentration est déterminée à partir du coefficient d'extinction molaire établi en fonction des premiers voisins (Cantor et *al.*, 1970) et de la mesure de l'absorbance à 265 nm. La biotine n'absorbant pas à 265 nm, elle n'affecte pas la mesure d'absorbance.

1-3- Marquage radioactif des oligodésoxynucléotides

Les oligodésoxynucléotides sont marqués radioactivement à l'aide de γATP-^{33}P. La phosphorylation en 5' est réalisée à l'aide de la T4-polynucléotide kinase (Roche). La réaction a

lieu à 37°C pendant 30 minutes, selon les procédures communément utilisées (Sambrook et *al.*, 1989). L'ADN phosphorylé est séparé de la polynucléotide kinase par précipitation dans l'éthanol. Le précipité est lavé avec de l'éthanol à 70% puis repris dans 30 à 50 µl de Tris (90 mM) citrate (90 mM) EDTA (2 mM) (= TCE 1x) à pH 7.

L'efficacité de la réaction de phosphorylation est visualisée par migration sur couche mince d'un échantillon du milieu réactionnel prélevé avant l'incubation à 37°C, et d'un échantillon de même volume prélevé en fin d'incubation. Le γATP-^{33}P non utilisé est visualisé au front de migration. La différence de signal relevée entre les deux échantillons permet d'établir le rendement de kination. Les séquences radioactivement marquées seront annotées d'un « * » par la suite.

1-4- Gel dénaturant (ADN marqué)

La qualité des échantillons d'acides nucléiques purifiés est contrôlée par migration électrophorétique dénaturante. L'échantillon préalablement marqué radioactivement est dénaturé à 90°C en présence d'urée (6M) pendant cinq minutes avant d'être déposé sur gel (acrylamide/bisacrylamide 12%, urée 6M). La migration se fait sur système d'électrophorèse non réfrigéré, à 30W (pour un gel de 20 cm x 20 cm x 0.8 mm). Après migration, les acides nucléiques sont fixés et l'urée éliminée dans un mélange acide acétique 10% - éthanol 20%. Le gel est séché sur papier Whatman. Les résultats sont obtenus après exposition nocturne et développement d'un film Kodak Biomax MR. Le rendement final de la synthèse d'ADN est généralement de 97 % pour une séquence de 20-25 bases, 75% si une biodite™ est insérée en fin de synthèse.

1-5- Electrophorèse en conditions non dénaturantes

Les techniques de migration électrophorétique non dénaturante ont servi à la mise en évidence des interactions entre ADN* (radioactivement marqué) et protéines, nécessaires pour déterminer la spécificité de l'association entre ADN et protéines, et pour suivre les étapes de purification menant à l'identification des protéines recherchées.

Nous avons utilisé pour cela des gels de polyacrylamide de 18 cm x 20 cm x 0.8 mm généralement à 7% d'acrylamide (acrylamide / bisacrylamide, 29 :1, sous forme liquide, Biorad). Le tampon d'électrophorèse est du TBE 0.5x, à pH 8-8.3. Le dépôt des échantillons (volume ~15 µl) est facilité par l'ajout de glycérol (5% v/v) après incubation. La vitesse de migration est visualisée par la progression de deux marqueurs colorés, le xylène cyanol et le bleu de bromophénol (Sambrook et *al.*, 1989).

Pour éviter un échauffement des protéines durant la migration, nous avons travaillé en chambre froide et nous avons eu recours à un système de doubles cuves d'électrophorèse à réfrigération centrale. Dans ces conditions, la migration a lieu à 15W constant (~ 25V/cm). Après migration, les gels sont séchés et autoradiographiés.

Les autoradiographies sont numérisées à l'aide d'un scanner. Le bruit de fond est éliminé par traitement de l'image à l'aide des programmes Scion Image (PC Windows) ou NIH Image 1.62 (Macintosh). Ces mêmes programmes permettent de mesurer l'intensité des bandes retardées.

1-6- Gel de protéines (électrophorèse dénaturante)

1-6-1- Conditions de migration

La migration électrophorétique en conditions dénaturantes des extraits protéiques partiellement purifiés a été utilisée à la fois comme outil de contrôle de l'avancement des différentes étapes de purification et comme étape de purification. Il s'agit d'un passage obligé avant le séquençage de protéines.

Les gels sont préparés selon la méthode développée par Laemmli (1970). Un gel de concentration (pH 6.8, 5.1 % d'acrylamide) est coulé au-dessus d'un gel de résolution. La zone de concentration mesure 2 cm pour une profondeur des puits de dépôts de 1 cm. Le pourcentage d'acrylamide du gel de résolution dépend de la masse des protéines que l'on veut observer. Les gels sont de mêmes dimensions que les gels natifs détaillés précédemment. Dans ces conditions, la distance de migration possible dans le gel de résolution est de 16 cm. La tension au début de la migration est de 500V, à puissance constante de 60W dans le gel de concentration. Après passage des protéines dans

le gel de résolution, la puissance maximale est ramenée à 40-50W. L'ensemble de la migration est thermiquement régulé entre 60 et 70°C (système de double cuves d'électrophorèse à réfrigération centrale).

Les échantillons d'extraits protéiques en phase liquide sont dénaturés par incubation à 80°C pendant 5 minutes en présence de SDS (1%) et de β-mercaptoéthanol (100 mM). La dénaturation de protéines piégées sur un support (gel retard natif ou billes d'agarose) a lieu dans le même milieu et les mêmes conditions d'incubation, en présence d'urée, 2.5 M.

1-6-2- Coloration des gels de protéines

La visualisation d'une bande de protéine repose sur l'utilisation de colorants se fixant spécifiquement sur les peptides. La résolution et la sensibilité dépendent du colorant utilisé, de la concentration, de la masse de la protéine et de son affinité pour les colorants. A titre indicatif, la sensibilité des colorants utilisés pour l'albumine bovine (Bovine Serum Albumine Fraction V, Sigma) et le temps moyen d'obtention du résultat sont donnés dans le tableau 2.3.

Différents types de coloration utilisés	Sensibilité (/BSA)	Durée moyenne avant obtention de résultats	Remarques
Rouge Ponceau	10-100 µg	10 minutes	Coloration réversible
Noir Amido	1-10 µg	1 heure	
Bleu de Coomassie	0.1-1 µg	12 heures	
Sulfate de zinc	10-100 ng	10 minutes	Coloration en négatif et réversible
Nitrate d'argent	0.1-10 ng	2 heures	

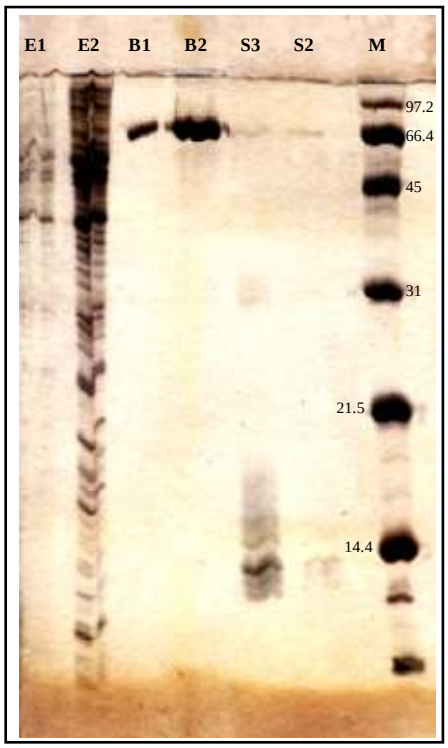
Tableau 2.3. : Sensibilités par rapport à la BSA des différentes colorations utilisées lors de la coloration de gels de protéines.

La coloration au nitrate d'argent est effectuée en utilisant le kit « Silver Stain Plus » de BioRad. Cette méthode ne nécessite pas de fixation préalable des protéines à l'aide de glutaraldéhyde, fixation qui empêcherait l'exploitation d'échantillons par spectrométrie de masse.

➢ Remarques concernant les différents types de coloration :

Le temps total de coloration / décoloration des protéines au bleu de Coomassie est réduit en chauffant le gel pendant l'étape de coloration / fixation. On constate que la décoloration est accélérée et le bruit de fond diminué.

La sensibilité et l'affinité de la coloration au nitrate d'argent sont augmentées si le gel est préalablement coloré au bleu de Coomassie, puis quasi-totalement décoloré avec du méthanol à 40%. La figure 2.1. donne une idée des gains de sensibilité obtenus. De plus la sélectivité de la coloration est fortement diminuée. Mais nous n'avons pas pu exploiter en spectrométrie de masse d'échantillons obtenu par double coloration, contrairement à l'une ou l'autre des deux colorations.



Coloration au NiAg

Double coloration Coomassie / NiAg

Figure 2.1. : Comparaison entre coloration au nitrate d'argent et double coloration Coomassie/$AgNO_3$. B1 = BSA 0.1 µg ; B2 = BSA 1 µg; E1 = 0.1 µl d'extraits de protéines de levure; E2 = 1 µl d'extraits de protéines de levure; E3 = 10 µl d'extraits de protéines de levure; M = marqueurs de poids moléculaire (en kDa); S1 = 0.01 µg de streptavidine; S2 = 0.1 µg de streptavidine; S3 = 1 µg de streptavidine

Dans le cadre de la préparation d'échantillons destinés à la spectrométrie de masse (par découpage du gel dénaturant dans les zones correspondant aux protéines à analyser), nous avons systématiquement procédé à une identification des protéines par coloration au sulfate de zinc. Cette coloration est réversible mais reste moins sensible que la double coloration bleu de Coomassie – nitrate d'argent.

La coloration au noir amido est mal adaptée à la coloration d'un gel de protéines faiblement concentrées. En revanche, elle se justifie pour les gels riches en protéines à partir desquels sont préparés les échantillons pour le séquençage par dégradation d'Edman. Nous avons constaté une augmentation du contraste entre protéines et bruit de fond de la coloration à l'amido black, par coloration préalable au sulfate de zinc, suivie d'une décoloration.

1-6-3- Transfert sur membrane

Les protéines sont transférées à partir du gel d'acrylamide sur une membrane en fluorure de polyvinylidène (ou PVDF) Immobilon-P (Millipore) à l'aide d'un système de transfert humide de type Trans-Blot Cell (BioRad). La membrane est préalablement trempée dans du méthanol (99.8%), puis rincée avec de l'eau type Milli Q, avant d'être équilibrée dans le tampon Tris-Borate 50 mM (~ pH 9) servant lors des transferts.

Le gel et la membrane sont maintenus entre des carrés de papier Whatman, puis immergés dans le système. Le transfert s'effectue sur la nuit. On règle la tension de transfert en fonction de la masse des protéines à transférer, de l'épaisseur et de la concentration en acrylamide du gel, de façon à éviter un transfert incomplet ou un sur-transfert. Par exemple, une protéine de 50 kDa séparée sur un gel à 7% d'acrylamide et de 0.8 mm d'épaisseur, est électro-transférée sous une tension maintenue constante de 28V. Le courant est de 120 mA à 140 mA pour un gel de 10 cm^2 et 0.8 mm d'épaisseur.

1-7- Séquençage d'Edman

Les micro-séquençages séquentiels ont été réalisés par dégradation d'Edman automatique avec un séquenceur de protéines en phase liquide Procise (Perkin Elmer, Applied Biosystems division) au Laboratoire de micro-séquençage des protéines de l'Institut Pasteur. L'identification est obtenue à partir de 1 à 10 µg de protéines. Le gel a une épaisseur de 1.5 mm, de façon à permettre le dépôt d'un volume plus conséquent de protéines. Les conditions de migration, les temps de transfert sur membrane et de coloration, sont ajustés par rapport aux conditions précédemment décrites pour un gel de 0.8 mm.

Les protéines sont directement analysées après transfert sur membrane PVDF. En cas de blocage de la partie N-terminale de la protéine, une digestion enzymatique en présence d'endolysine-C est pratiquée. Les peptides obtenus sont séparés par HPLC sur colonne DEAE – C_{18} de 1 mm de diamètre, éluée par un gradient eau-acétonitrile (de 2 à 70 % acétonitrile) en présence d'acide trifluoroacétique (TFA) 0.1%, avant d'être analysés.

1-8- Tamis moléculaire

En première approximation, la masse moléculaire des protéines natives est déterminée par filtration sur tamis moléculaire. Cette séparation en fonction de la masse a également servi d'étape de purification.

Le support utilisé dans notre étude est une résine Superdex (Pharmacia), matrice composée d'agarose et de dextran. Les protéines sont séparées sur une colonne de 90 cm de long et de diamètre 1 cm, préparée à partir de résine Superdex 200 (résolutive de 100 kDa à 600 kDa). La colonne est préalablement calibrée à l'aide des protéines des kits de calibration Pharmacia (poids moléculaires de 67 à 749 kDa).

La colonne est éluée dans un tampon TBE 0.5x à pH 8 et NaCl 175 mM (le sel évite les interactions ioniques entre matrice et protéines), à un débit de 1 ml/min, en présence d'antiprotéases. Les échantillons collectés sont analysés par migration électrophorétique native, après incubation avec l'ADN*.

1-9- Purification par affinité

La purification par affinité repose sur les interactions spécifiques existant entre deux molécules. La streptavidine est une glycoprotéine d'origine bactérienne formée de 4 sous-unités (masse moléculaire = 16.4 kDa chacune). Elle est neutre à pH 7.6, non glycosylée, et donne une très faible fixation non-spécifique. Chaque sous-unité est capable de lier la biotine, un coenzyme de transfert de CO_2, avec une très forte affinité ($K_d = 10^{-15}$ M). La biotine ne présente pas d'absorption significative dans l'ultraviolet ou dans le visible.

L'identification de protéines se liant spécifiquement à l'ADN télomérique peut donc être entreprise en mettant à profit les propriétés associatives de la streptavidine pour la biotine et la spécificité de l'association entre la séquence « biotine-ADN » et les protéines (figure 2.2.). Pour cela, nous avons utilisé de la streptavidine immobilisée sur des billes d'agarose de diamètre 1µm (Sigma) comme support.

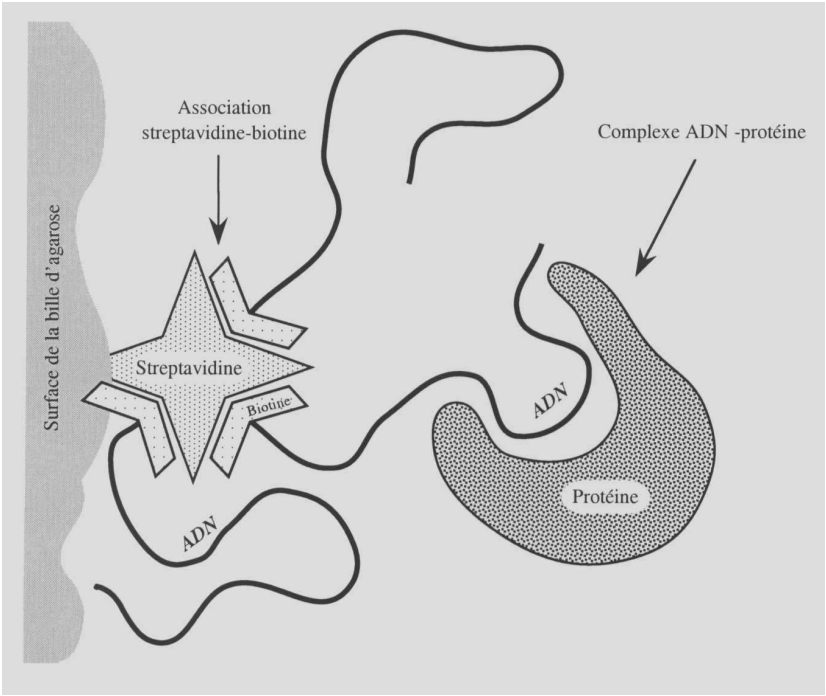


Figure 2.2. : Associations spécifiques streptavidine - biotine et ADN - protéine.

Pour préparer la résine d'affinité, la 'résine-streptavidine' (20-50 µl) est équilibrée dans du phosphate de potassium (20 mM), pH 7.5, en présence de NaCl (150 mM). La résine est mise en contact avec l'ADN biotinylé (~ 400 pmol), 30 minutes à température ambiante, sous agitation

ménagée, avant d'être finalement équilibrée en TBE 0.5x, pH 7.4, KCl (150 mM). L'ADN biotinylé est amené en excès sur la résine et la quantité d'ADN liée est calculée en mesurant l'absorbance du surnageant après incubation et sédimentation de la résine (correspondant au ligand en excès).

Le contact entre la résine d'affinité et les protéines se fait selon une procédure de lots, sous agitation ménagée. Après incubation, une centrifugation légère (une impulsion à 2000 tr/mn) permet de séparer les deux phases. Cette procédure a été préférée à une chromatographie classique à l'aide d'une colonne pour éviter de se retrouver dans une situation où le nombre de ligands serait très élevé par rapport à l'estimation du nombre de protéines à fixer.

1-10- Reconcentration, dessalage et changement de pH

La dilution inhérente aux expériences de séparation (chromatographie, lot, …) nécessite la reconcentration des échantillons. A cet effet, nous avons utilisé un dispositif d'ultrafiltration Vivaspin (Vivascience) pour centrifugeuse de laboratoire. La membrane en polyéthersulfone (PES) est disposée longitudinalement, ce qui assure une grande surface de filtration et élimine le risque de reconcentration à sec.

Le changement de milieu des échantillons protéiques a été obtenu principalement par dilutions successives et reconcentrations, moyens mieux adaptés pour de petits volumes que la dialyse. Plusieurs seuils de filtration ont été testés (5 kDa, 30 kDa et 50 kDa) sans modification apparente des résultats. Nous avons choisi de travailler avec un seuil de coupure de 5kDa.

1-11- Autres techniques utilisées

1-11-1- Chromatographie anionique

Nous avons recouru à la chromatographie anionique pour deux applications distinctes : la purification des oligodésoxynucléotides synthétisés au laboratoire (*cf.* paragraphe 2-) et la séparation de protéines en fonction de leur charge. Dans les deux cas, le support utilisé est la

matrice Q Sepharose (Pharmacia). Le groupement échangeur d'ions est une amine quaternaire qui reste chargée entre pH 2 et pH 12.

La séparation des protéines s'effectue par passage sur une colonne de 2 cm de long par 1 cm de section. Le tampon d'élution est du TBE 0.5X à pH 7.4, en présence d'anti-protéases. Nous avons utilisé du chlorure de potassium (KCl) comme sel d'élution. La colonne est préparée à l'aide d'un gradient (de 50 mM à 2 M de KCl puis retour à 50 mM). Pour obtenir une meilleure séparation des protéines, nous avons procédé à une élution par plateaux de KCl.

1-11-2- Autres types d'élution par chromatographie

D'autres modes de séparation ont été essayés, dont la séparation par pseudo - affinité à l'aide d'héparine (Pharmacia) ou de résines colorées (Sigma) (interactions hydrophobes, échangeuse d'ions). Aucune de ces résines n'a permis d'isoler les protéines recherchées.

1-11-3- Combinaison d'électrophorèses non dénaturante / dénaturante

La migration de protéines au cours d'une électrophorèse non dénaturante est un moyen de séparation en fonction de la charge et de la forme. Il est possible de visualiser les protéines présentes à une distance de migration donnée, en les électrotransférant dans un gel dénaturant.

Sur le gel natif, la zone correspondant aux protéines que l'on souhaite analyser est repérée à l'aide de marqueurs colorés (xylène cyanol et bleu de bromophénol) et confirmée par autoradiographie du gel (les protéines migrant sous la forme complexée à l'ADN*). On prélève le fragment de gel natif (bande de gel d'un millimètre de hauteur) que l'on insère sur le dessus d'un gel dénaturant. La dénaturation des protéines est réalisée par l'immersion de la bande de gel natif dans du tampon Laemmli porté à 90°C en présence d'urée (2.5 M). La migration en conditions dénaturantes a lieu selon le mode opératoire défini précédemment. L'ADN* est dissocié des protéines et migre au front de migration. Une autoradiographie de ce front de migration permet de vérifier que la zone découpée correspond à la position du complexe ADN*-protéines du gel natif.

1-11-4- Pontage ADN*-protéines par irradiation aux UV

L'irradiation aux UV de complexes ADN*-protéines peut donner lieu à des pontages entre bases de l'ADN* et acides aminés aromatiques. Le complexe covalent formé peut être visualisé aisément par autoradiographie.

La préparation des échantillons est la même que celle définie dans le cadre de migrations électrophorétiques natives. Les échantillons sont ensuite irradiés sur lit de glace pendant 1 h, avec une lampe UV de paillasse (254 nm, 6 W), avant d'être dénaturés et analysés par gel SDS dénaturant. Le gel est coloré puis séché sur papier Whatman. Après autoradiographie, la superposition du gel de protéines et du signal radioactif indique l'emplacement des protéines associées à l'ADN.

2- Biophysique

2-1- Spectroscopie d'absorption

Les oligodésoxynucléotides possèdent une bande d'absorption dans l'ultraviolet, avec un maximum compris entre 255 et 270 nm. La spectroscopie d'absorption permet d'établir la concentration d'un échantillon d'ADN.

On peut également déterminer si un brin d'ADN est structuré. En effet, le coefficient d'absorption optique d'un oligodésoxynucléotide est différent selon qu'il est structuré ou non (effet hypochromique). L'observation de l'absorbance d'un échantillon en fonction de la température permet de déterminer la température de dénaturation, T_m, point de demi-transition de la dénaturation thermique, lors du passage d'une structure ordonnée à une structure désordonnée. Des observations de ce type ont été faites à l'aide d'un spectrophotomètre Kontron-Uvikon 940 au Laboratoire de Biophysique du Muséum National d'Histoire Naturelle.

Les échantillons sont introduits dans des cellules en quartz (1 cm de chemin optique), recouvertes d'une mince couche de paraffine. La température des cellules est régulée par circulation de liquide thermostaté. L'évolution du profil d'absorbance en fonction de la température est observée à 265 et

295 nm. A 265 nm, le coefficient d'absorption ne dépend pas de la protonation de la cytosine (point isobestique) et l'absorbance reflète directement le changement de structure (Mergny et *al.*, 1995). Les profils sont enregistrés en fonction de la température. Pour s'assurer de la réversibilité des phénomènes observés, l'enregistrement a lieu durant un cycle de descente-montée en température (variation de 0.2°C / minute) (figure 2.3.). Si les courbes correspondant à la montée et à la descente sont superposables, on peut conclure que chaque point de ces courbes correspond à une valeur à l'équilibre.

Indépendamment de la température, on peut faire varier la concentration (15 à 50 µM selon la longueur des séquences d'ADN traitées), le pH (pH 4 et 5, tampon acétate de sodium 10 mM ; pH 6, 7 et 7.5, tampon cacodylate de sodium 10 mM), ou la force ionique des échantillons (175 mM NaCl et 2 mM $MgCl_2$) que l'on observe. Ces observations ont permis de déterminer les conditions dans lesquelles certaines séquences d'acides nucléiques étaient structurées.

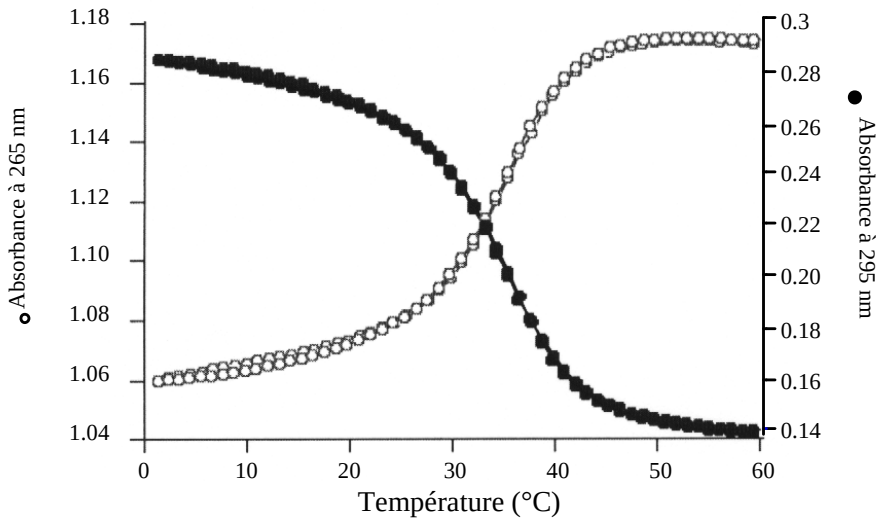


Figure 2.3. : Profil de dénaturation thermique de la séquence d(CCCACACCCACACCCACACCC) à 265 nm et 295 nm [→ T_m=34°C], à pH 6, pour une concentration de 5 µM.

Une autre application de la spectroscopie d'absorption consiste à suivre l'avancement d'une purification en disposant une cellule de détection en série au sortir d'une colonne de chromatographie. Pour les protéines, deux acides aminés, le tryptophane et la tyrosine, absorbent à 280 nm.

2-2- Spectrométrie de masse

2-2-1- Introduction et principe

La spectrométrie de masse a pris de plus en plus d'importance dans l'étude et la caractérisation des biomolécules. La cartographie par mesure des masses moléculaires de mélanges peptidiques obtenus après digestion enzymatique permet d'identifier une protéine. L'analyse par spectrométrie de masse peut être couplée à des digestions enzymatiques (aminopeptidase ou carboxypeptidase), pour déterminer des séquences d'acides aminés. Technique en plein essor, la spectrométrie de masse présente l'avantage de nécessiter très peu de matériel biologique (quelques dizaines de femtomol) en comparaison des méthodes de séquençage habituelles.

Un spectromètre de masse permet de mesurer le rapport masse/charge (m/z) des ions formés à partir de l'échantillon analysé. Il est constitué d'une source d'ions (passage en phase gazeuse [vaporisation / sublimation / désorption], ionisation et décomposition des ions de l'échantillon à analyser), d'un analyseur triant les ions en fonction de leur rapport m/z, d'un enregistreur et d'un système de calibration (figure 2.4.). Deux approches sont possibles pour l'ionisation de l'échantillon : le MALDI (Matrix-Assisted Laser Desorption Ionization), et l'électrospray. Ces deux techniques sont brièvement décrites ci-après.

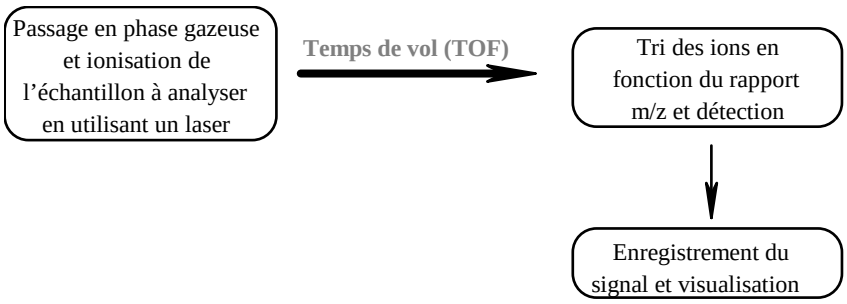


Figure 2.4. : Principe schématisé de l'expérience de spectrométrie de masse MALDI.

L'ensemble des expériences de spectrométrie de masse a été réalisé en collaboration avec M. Jean-Pierre Le Caer du Laboratoire de Biologie de l'ESPCI.

2-2-2- Préparation des échantillons

Les échantillons à analyser sont sélectionnés après identification par coloration au sulfate de zinc sur gel de protéines dénaturant. Les fragments de gel sont lavés (élimination du SDS), puis les protéines sont réduites avec du DTT et de l'iodoacétamide avant d'être digérées à la trypsine. Les peptides obtenus sont extraits du gel SDS par traitement à l'acétonitrile et à l'acide formique, puis dessalés à l'aide de micro colonnes C_{18} (voir mode opératoire en annexes).

2-2-3- Spectrométrie de masse MALDI

La technique MALDI utilise un laser UV pulsé pour désorber et ioniser le mélange de matrice/échantillon co-cristallisé. La matrice permet de minimiser la dégradation de l'échantillon provoquée par le faisceau laser incident. L'énergie absorbée par la matrice provoque son expansion en phase gazeuse en entraînant les molécules d'échantillon sous forme d'un mélange d'ions mono- ou multi- chargés notés $[M + nH]^{n+}$, où M correspond à la masse moléculaire de la molécule analysée, M + nH à la masse moléculaire de molécule ionisée (masse moléculaire + masse de n protons). Les ions mono chargés sont généralement majoritaires. La masse M peut être déterminée très simplement par déconvolution du spectre et transformation de ce dernier en une courbe masse-intensité. La résolution pour la masse d'un peptide lors de ce genre d'expérience, est de ± 0.1 Dalton.

Dans les expériences réalisées, l'échantillon peptidique est associé à une matrice DHB (acide gentisique ou 2,5-hydroxybenzoïque) en proportions 1:1, pour un volume final de 1µl. Le dépôt s'effectue sur une plaque métallique en acier inoxydable, pouvant supporter jusqu'à 100 dépôts. L'analyse a lieu après cristallisation du dépôt. L'appareil utilisé est un Voyager Elite (PerSeptive Biosystems).

2-2-4- Spectrométrie de masse ESI et fonctionnement en tandem MS/MS

Le principe de l'électrospray est relativement simple : une solution est introduite dans un capillaire porté à un haut potentiel électrique. Le champ électrique intense appliqué à la sortie du capillaire provoque la formation d'un nuage de gouttelettes chargées qu'un gradient de champ électrique et un

gradient de pression (entre la pression atmosphérique et le vide poussé à 10^{-6} bar) transforment en ions désolvatés $[M + nH]^{n+}$.

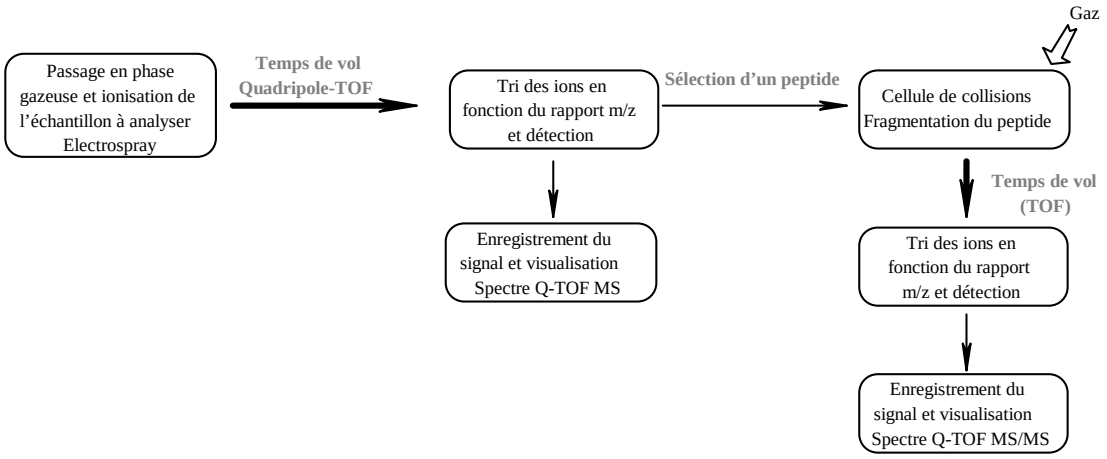


Figure 2.5. : Principe de l'expérience de spectrométrie de masse en tandem (MS /MS).

Un analyseur quadripolaire, constitué de quatre barreaux de section hyperbolique ou circulaire, disposés parallèlement le long de la trajectoire des ions, permet de les filtrer en fonction de leur rapport m/z. Chaque paire de barreaux diamétralement opposés est portée à un potentiel électrique opposé à celui de l'autre paire. On peut alors sélectionner un ion précurseur, lequel passera ensuite dans une cellule de collisions. Lors de ces collisions, son énergie interne augmente et certaines liaisons chimiques sont rompues, ce qui provoque sa décomposition en ions fragments. Un deuxième analyseur en temps de vol, à accélération par impulsion électrique, orthogonal au premier analyseur, sert alors à l'étude de ces fragments (Dawson & Guilhaus, 1989) permettant d'accéder à la séquence d'un peptide.

Un spectre Q-TOF MS/MS d'un peptide de n acides aminés correspond à la superposition du signal de l'ion précurseur, des pics correspondants aux différents fragments de tailles $(n-1)-i_{0\to n-1}$ et des pics de tailles $i_{0\to n-1}$ correspondant aux fragments complémentaires, généralement de plus faibles intensités.

L'ion précurseur apparaît comme ayant une masse apparente correspondant à (masse réelle + 2) / 2. A partir de cette masse initiale (correspondant à la masse des n acides aminés du peptide étudié), on procède de proche en proche [fragment n-1, n-2, etc..] (cf. table des masses des acides aminés en

annexes) et on obtient la séquence du peptide sélectionné. A ce stade, l'indétermination demeure entre leucine et isoleucine.

L'échantillon peptidique à analyser doit être dissous à une concentration de 0.01 à 10 µM dans un solvant de type eau-méthanol avec 5% d'acide formique. Cette méthode nécessite très peu de matériel (~ 10^{-17} mol).

2-2-5- Analyse des résultats

L'analyse des résultats obtenus avec ces deux méthodes s'effectue grâce à des logiciels (Profound et Peptident) accessibles sur internet, utilisant toutes les banques de données disponibles. Pour les expériences MALDI, l'identification est réalisée à partir des données expérimentales suivantes : masses des peptides, erreurs sur les mesures, et masse moléculaire déterminée sur gel SDS dénaturant.

2-3- Spectroscopie RMN

L'accès à la structure haute définition de macromolécules biologiques peut être mené selon deux approches complémentaires : l'analyse par diffraction de rayons X dans un cristal et la résonance magnétique nucléaire (RMN). La RMN permet l'étude dynamique et structurale en solution, ce qui autorise des modifications importantes du milieu, comme le pH, la température, la force ionique ou la concentration.

Nous avons étudié certains des acides nucléiques utilisés lors de ce travail par RMN ^{1}H à l'aide des spectromètres RMN à 360 MHz et à 500 MHz disponibles au Groupe de Biophysique de l'Ecole Polytechnique. L'expérience acquise par le laboratoire concernant la RMN des acides nucléiques et particulièrement les études sur la structure en motif i adoptée par les fragments d'ADN synthétiques riches en cytosines, a servi de référence dans la détermination du type d'association pouvant exister entre des séquences d'ADN télomériques et des protéines.

2-3-1- Principe

Les noyaux atomiques possèdent un moment cinétique intrinsèque, le spin **I**, auquel est associé un moment magnétique colinéaire et caractéristique de la particule étudiée, **μ**, tel que :

$$\boldsymbol{\mu} = \gamma \hbar \mathbf{I}$$

où le rapport gyromagnétique γ est une constante propre à la particule étudiée.

L'expérience de RMN consiste à modifier l'aimantation sous l'action d'un champ magnétique radiofréquence $\mathbf{B_1}$ en présence d'un champ magnétique fort statique $\mathbf{B_0}$ et à observer son évolution. Deux effets contribuent au retour à l'équilibre : un couple, $\mathbf{C} = \mathbf{M} \wedge \mathbf{B}$, provenant de l'interaction avec le champ magnétique, et la relaxation due à des champs magnétiques microscopiques. Le mouvement résultant sera une précession autour du champ $\mathbf{B_0}$ à une fréquence $\boldsymbol{\omega_0}$, la fréquence de Larmor, telle que $\boldsymbol{\omega_0} = \gamma\, \mathbf{B_0}$. On enregistre la précession libre avec une bobine perpendiculaire à $\mathbf{B_0}$ et le spectre en fréquence est obtenu par transformation de Fourier du signal de précession. L'environnement spatial et le type de groupement auxquels appartient un proton modifient sa fréquence de résonance. Bien qu'ayant un même rapport gyromagnétique, chaque proton aura donc sa propre fréquence de résonance.

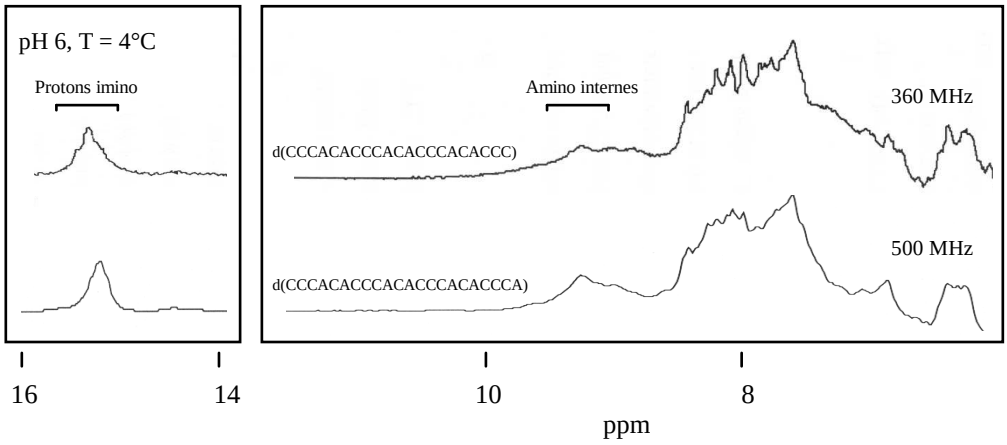


Figure 2.6. : Comparaison des spectres RMN ^{1}H de deux fragments télomériques de levure (*Saccharomyces cerevisiae*) [les concentrations en brin d'ADN sont de 1 mM].

Les interactions entre deux moments magnétiques nucléaires sont liées à la distance les séparant. L'observation de ces interactions permet de remonter à certaines distances internucléaires et de là, d'accéder à la structure de la molécule étudiée.

2-3-2- Préparation des échantillons et expériences réalisées

L'échantillon d'ADN synthétique (concentration 1 mM) est dissous dans un mélange H_2O / D_2O (90 / 10), en présence d'EDTA (1 mM) et de DSS (0.5 mM, servant de marqueur de déplacement chimique), dans un volume final de 450 µl.

Les expériences RMN à une dimension ont été réalisées en utilisant une séquence d'impulsion JR pour annuler le signal de l'eau (Guéron et *al.*, 1991). L'expérience RMN à deux dimensions est une expérience de type JR-NOESY.

Chapitre III

Caractéristiques du complexe ADN-protéine

Nous avons mis en évidence un complexe constitué d'une séquence télomérique et de protéines de levure *Saccharomyces cerevisiae*. L'étude des propriétés physico-chimiques régissant cette association forme la base du travail de compréhension et d'analyse menant à l'isolement et à l'identification des protéines impliquées. Ces caractéristiques reposent à la fois sur les propriétés propres à chacun des deux partenaires en fonction des conditions du milieu (pH, température, force ionique,…) et sur les interactions entre parties, générées au moment de l'association. Tels sont les sujets abordés dans ce chapitre.

1- La séquence d'ADN télomérique de levure d[(CCCACA)$_3$CCC]

Nous avons volontairement choisi un fragment d'ADN synthétique pouvant adopter la structure en motif i par repliement intramoléculaire. Le choix de l'organisme, *Saccharomyces cerevisiae*, sur lequel ce travail s'est effectué et les informations concernant la stabilité du motif i intramoléculaire, nous ont conduit à choisir pour séquence de référence, le fragment d[(CCCACA)$_3$CCC], appelé par la suite **Tlev**.

1-1- La séquence d[(CCCACA)$_3$CCC]

Cette séquence est formée de 4 répétitions de 3 désoxycytidines pouvant permettre un arrangement en motif i intramoléculaire composé d'un cœur de 6 paires de bases C·CH$^+$ coiffé par des boucles ACA, selon les modèles structuraux proposés par X. Han (Han et *al.*, 1998) et A.T. Phan (Phan et *al.*, 2000). On ne trouve ce fragment d'ADN qu'aux extrémités télomériques des chromosomes I, VIII, XII et XIII (données provenant de *Blast alignments* (site : http://www.ncbi.nlm.nih.gov/BLAST/), cf. tableau 3.1.) de *Saccharomyces cerevisiae*.

Chromosome	Longueur du chromosome (en bases)	Position de la séquence (en bases)	Région
I	230203	230127 → 230159 [a]	Télomérique
VIII	562639	556950 → 556970	Sub-télomérique
		562500 → 562520	Télomérique
XII	1078172	1065115 → 1065141 [b]	Sub-télomérique
XIII	924430	28 → 48	Télomérique

Tableau 3.1. : Emplacements où l'on retrouve la séquence d[(CCCACA)$_3$CCC] dans les chromosomes de *Saccharomyces cerevisiae* (*Blast alignments).* [a] Tlev s'inscrit 3 fois dans la séquence ^{230127}[(CCCACA)$_5$CCC]^{230159} ; [b] Tlev s'inscrit 2 fois dans la séquence ^{1065115}[(CCCACA)$_4$CCC]^{1065141}

Nous allons montrer par des mesures de dénaturation thermique (UV) et des expériences de RMN ^{1}H à 1 et 2 dimensions, que sous certaines conditions, le fragment Tlev peut adopter la structure en motif i.

1-2- Etude par spectroscopie d'absorption

La mesure de la température de dénaturation du fragment d'ADN Tlev en fonction de sa concentration (en monobrin) (figure 3.1.) et en fonction du pH (figure 3.1. en bas à droite) permet de vérifier l'existence d'une structure et la stœchiométrie de celle-ci.

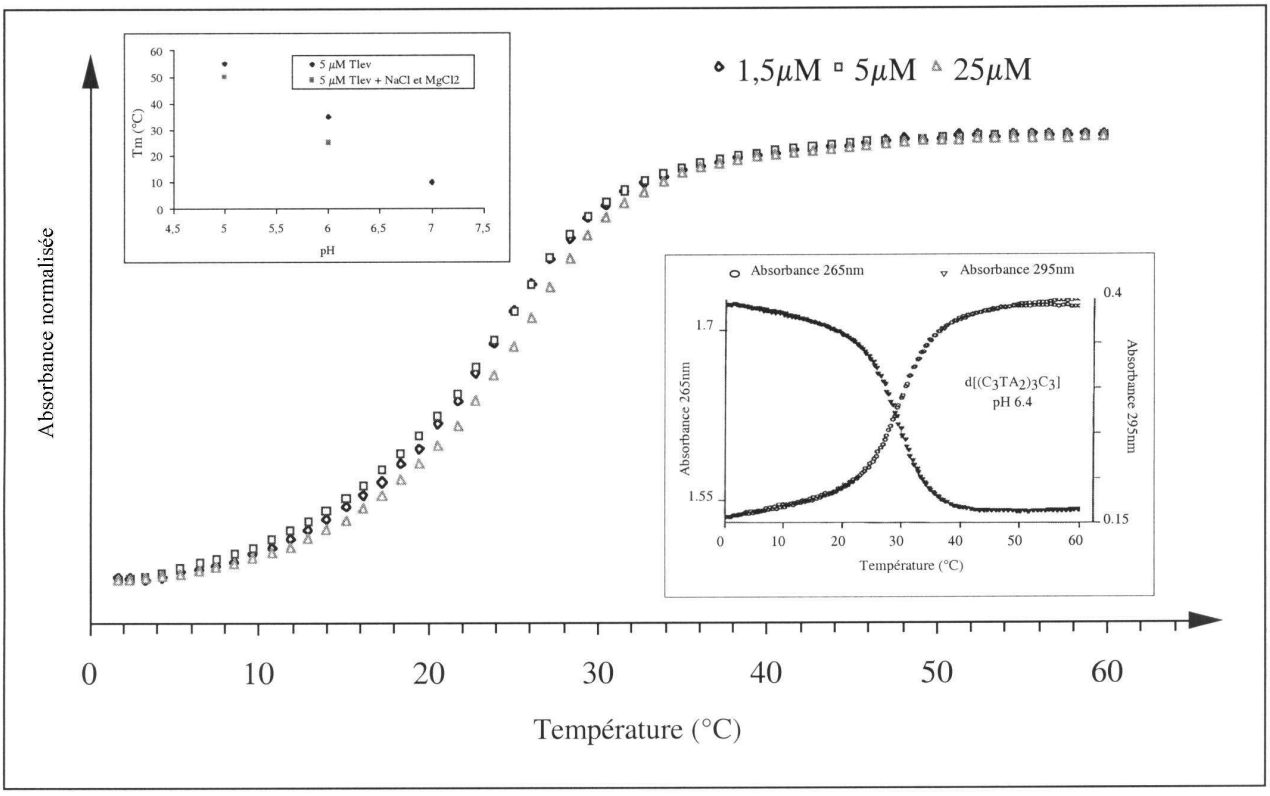
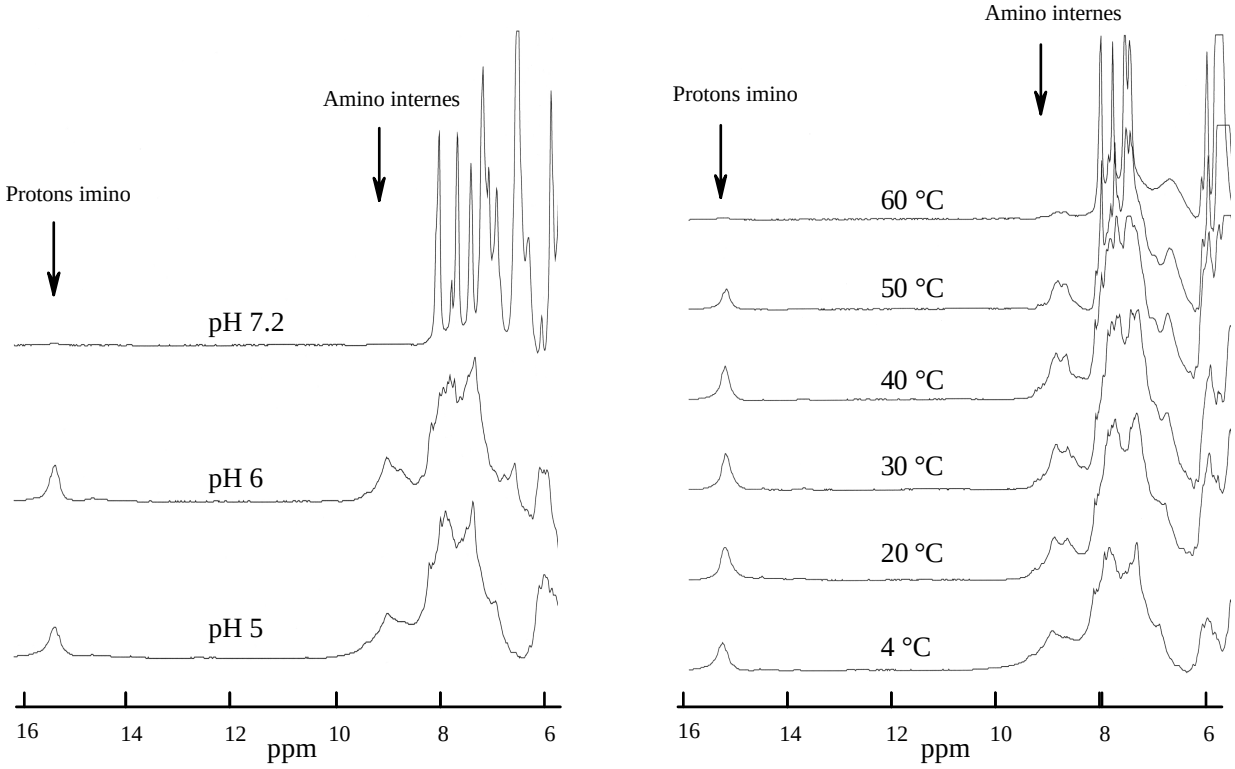


Figure 3.1. : Profil de dénaturation thermique à 265 nm de la séquence Tlev à pH 6 en fonction de la concentration en brin, en présence de 175 mM NaCl et 2 mM $MgCl_2$. En haut à gauche : Evolution de la température de demi-dissociation de la séquence Tlev en fonction du pH, sans sel ou en présence de 175 mM NaCl et 2 mM $MgCl_2$. En bas à droite : Profil de dénaturation thermique de la séquence télomérique humaine à pH 6 (Leroy et *al.*, 1994).

A pH 6, en présence de NaCl 175 mM et $MgCl_2$ 2 mM, le profil de dénaturation thermique à 25 µM diffère légèrement de ceux obtenus à 1.5 et 5 µM. La température de dénaturation thermique T_m varie de ~ 24°C (à 1.5 et 5 µM) et ~ 26°C (à 25 µM). A titre de comparaison, la température de dénaturation thermique de la séquence télomérique d[$(CCCTAA)_3CCC$] des mammifères est de 30°C à pH 6.4 en l'absence de sel (Leroy et *al.*, 1994) (figure 3.1. en bas à droite). L'évolution de T_m en fonction du pH est très semblable à celle observée (figure 3.1 en haut à gauche).

1-3- Etude par spectroscopie RMN du proton

Les expériences de spectroscopie RMN ont été conduites sur la séquence d[$(C_3ACA)_3C_3A$], dont le spectre à 1 dimension, très proche de celui de la séquence Tlev, présente dans la zone des protons amino un profil un peu mieux résolu (expériences à 1 dimension sur spectromètre RMN à 360 MHz, (figure 2.6.)). La concentration en monobrin de l'échantillon utilisé est de ~ 1 mM.



(a) Evolution en fonction du pH *(b) Evolution en fonction de la température*

Figure 3.2. : Spectres RMN à 500 MHz du fragment d'ADN télomérique de levure d[$(C_3ACA)_3C_3A$]. *(a)* Evolution en fonction du pH à T = 4 °C. *(b)* Evolution en fonction de la température à pH 5.

La présence d'un pic large centré sur 15.5 ppm est caractéristique de protons imino provenant de bases CH^+ appariées. Ce signal persiste à 50 °C, pH 5, et on en observe encore la trace à pH 7.2, 4 °C et à pH 5, 60 °C (figure 3.2.).

Un spectre NOESY dans H_2O présente des pics croisés entre protons amino et protons $H_{2'}$ / $H_{2''}$, caractéristiques du motif i (figure 3.3.).

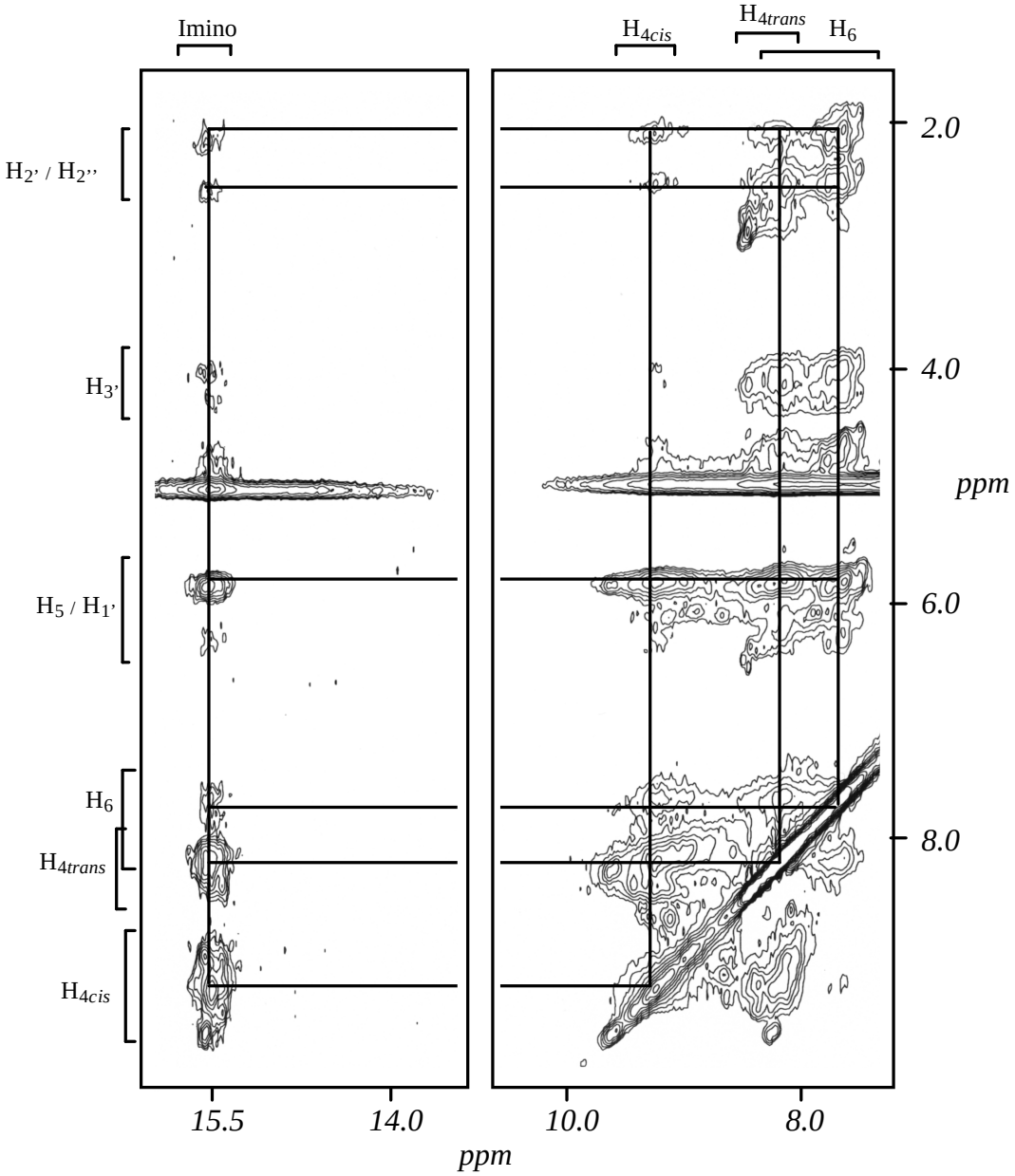


Figure 3.3. : Spectre NOESY à 500 MHz du brin d'ADN télomérique de levure d[(CCCACA)$_3$CCCA] dans H_2O, à 125 ms de temps de mélange, à T = 4 °C et pH 5.

Nous pouvons déduire des résultats obtenus avec le fragment d'ADN télomérique de levure d[(CCCACA)$_3$CCCA], que la séquence Tlev forme, comme dans le cas de la séquence télomérique humaine, un motif i intramoléculaire *in vitro*, et donc potentiellement *in vivo*.

2- Caractéristiques du complexe ADN-protéines

La visualisation du complexe ADN-protéines se fait uniquement par l'intermédiaire du fragment d'ADN synthétique marqué radioactivement à l'extrémité 5' au ^{33}P. Cette méthode ne permet pas une observation directe du complexe et apporte une observation différée de la présence de la protéine, notamment au cours des expériences de purification. La visualisation du complexe ADN-protéines dépend de la capacité des protéines à reconnaître et fixer l'ADN, donc de leur intégrité.

2-1- Migration électrophorétique non dénaturante

Le principe des expériences qui vont être maintenant décrites, consiste à incuber des protéines et un fragment d'ADN marqué radioactivement, puis à réaliser une migration électrophorétique native permettant de visualiser d'éventuels complexes ADN*/protéines. Les conditions d'incubation d'une part, les conditions de migration, d'autre part, sont les paramètres ajustables de ces expériences.

Nous avons généralement constaté l'apparition d'un précipité (ou agrégat) dans l'extrait de protéines, à partir de pH 7, phénomène s'amplifiant à pH 6. Afin d'éviter de perdre les protéines que nous recherchions lors de nos expériences, nous avons volontairement travaillé le plus souvent entre pH 7.4 et pH 8.

2-1-1- Conditions d'incubation

Nous avons défini les paramètres d'une migration électrophorétique servant de référence pour l'ensemble des tests présentés :

- ~ 5 µg de protéines provenant des extraits protéiques de levure (préparés comme décrit au chapitre II et en annexes) sont mélangés avec 0.65 pmol de Tlev partiellement marquée (3%) dans 10µl (soit une concentration de 65 nM)
- le mélange est incubé à température ambiante durant 15 min, dans TBE 0.5x à pH 8
- dans certains cas, la cellule d'incubation contient aussi 0.5 nmol d'ADN d(T_{14}) simple brin en tant que compétiteur non spécifique. Cependant, nous avons constaté que malgré son abondance relative (1000 fois celle de Tlev), ce compétiteur est sans effet sur la formation du complexe. C'est pourquoi la plupart des expériences de migrations électrophorétiques natives décrites par la suite ont été réalisées sans ADN compétiteur.

La migration dure ~ 45 min (voir conditions d'électrophorèse décrites au chapitre II). Dans ces conditions, nous observons principalement une bande retardée intense sur laquelle nous avons focalisé nos efforts de caractérisation et d'identification. Nous observons au moins trois autres bandes d'intensité relative beaucoup plus faible (figure 3.4.) que nous n'étudierons pas.

La quantité d'ADN radioactif est saturante comme l'atteste la présence de sonde libre Tlev* et l'absence de changement de l'intensité obtenue pour la bande retardée principale si on augmente d'un facteur 2 la quantité de Tlev*.

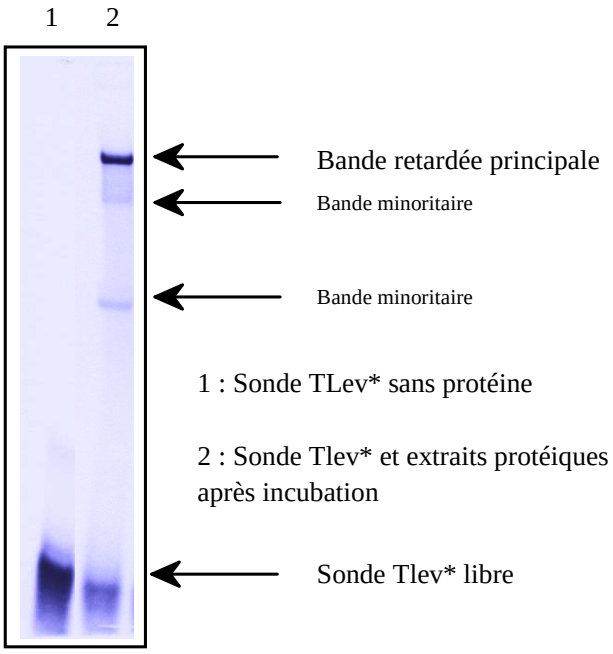


Figure 3.4. : Expérience de référence. Autoradiographie d'un gel non dénaturant avec la sonde d'ADN Tlev* et avec le mélange ADN Tlev*- extraits de protéines de levure.

Nous avons vérifié que nous obtenions des résultats identiques (distance de migration de la bande retardée la plus intense) en utilisant des protéines provenant d'extraits nucléaires préparés dans le laboratoire du Professeur Suzanne Gasser à partir de la même souche Ga59 que celle que nous utilisions (figure 3.4.bis). Cela nous permet également de considérer que l'activité retardée est due à une protéine localisée au moins partiellement dans le noyau. On remarque encore que l'abondance en protéines spécifiques des extraits nucléaires est le 10ème de celle de nos extraits (on a un signal retardé identique pour 10µgr de protéines nucléaires et pour 1µgr de nos extraits protéiques). Ceci nous laisse penser que les 9/10ème des protéines impliquées dans la reconnaissance de l'ADN Tlev pourrait être d'origine cytoplasmique.

EP : Extraits de protéines préparés selon la méthode de congélation-décongélation

EN : Extraits de protéines nucléaires (Pr. S. Gasser)

Protéines s'associant à Tlev*:~0.1 pmole
Tlev* : ~0.5 pmole

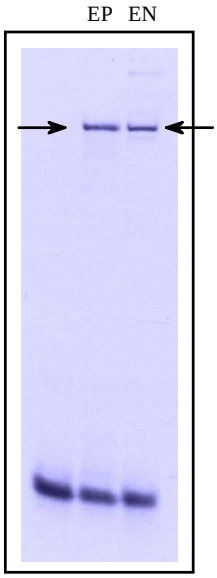


Figure 3.5. : Expérience de référence : Comparaison entre nos extraits de protéines préparés selon la méthode de congélation-décongélation et les extraits de protéines nucléaires (Pr. S. Gasser)

2-1-2- Conditions de migration

Pendant une migration électrophorétique à 4°C, l'intensité du signal retardé reste constante pour une durée de migration totale allant de 30 minutes à 8 heures. Dans ces conditions, le complexe ADN-protéines migre sous une forme stable et n'évolue pas au cours de la migration.

Pour des durées de migration identiques mais à température supérieure (20-30°C), l'intensité de la bande retardée diminue et s'étale, indiquant que le complexe ADN-protéines se dissocie en cours de migration. L'ADN une fois séparé des protéines sort de la zone de migration des protéines et la ré-association avec les protéines devient impossible.

2-1-3- Cinétique de l'association

Nous avons étudié la vitesse d'association de l'ADN avec les protéines à température ambiante. L'intensité du signal observé reste identique à celle de l'expérience de référence pour des temps d'incubation allant de 1 à 60 minutes (conditions expérimentales par ailleurs identiques à l'expérience référence), révélant une association rapide.

Une autre expérience confirme la cinétique rapide. Nous avons commencé par faire migrer les protéines seules (15 minutes), avant de faire migrer l'ADN*. Celui-ci rattrape les protéines et s'y associe dans des proportions identiques à celles de l'expérience de référence, montrant que cette association reste rapide (inférieure au temps de passage de l'ADN (1 min.) dans la zone où migrent les protéines).

En complément de cette expérience, nous avons commencé par faire migrer des protéines préalablement incubées avec le fragment Tlev non marqué radioactivement, puis après 15 min, nous avons fait migrer Tlev marqué radioactivement. Après autoradiographie nous visualisons une bande retardée moins intense que celle observée habituellement, indiquant qu'au passage de l'ADN Tlev*, celui-ci entre en compétition avec l'ADN Tlev non marqué radioactivement.

	Températures d'incubation				
	4°C	25°C	37°C	50°C	60°C
Intensité de la bande retardée (par rapport au signal maximum) après **1** min. d'incubation	50%	100%	100%	100%	-
Intensité de la bande retardée (par rapport au signal maximum) après **15** min. d'incubation	100%	100%	100%	60%	0%

Tableau 3.2. : Evolution de l'intensité de la bande retardée principale en fonction de la température d'incubation.

Enfin la cinétique d'association dépend de la durée mais aussi de la température d'incubation. A 25 et 37°C on n'observe pas de différence de l'intensité de la bande retardée. A 4°C, l'association est ralentie. Après 1 minute, l'intensité du signal retardé est moitié plus faible à 4°C qu'à 25 °C et nous ne discernons pas de différence entre 25 et 37 °C. Après 15 minutes à 4, 25 ou 37 °C, l'intensité du signal retardé est équivalente. A 50°C, le signal est atténué de 60% pour complètement disparaître à

partir de 60°C, ce que l'on peut expliquer par la dégradation des protéines (dénaturation, agrégat par exemple).

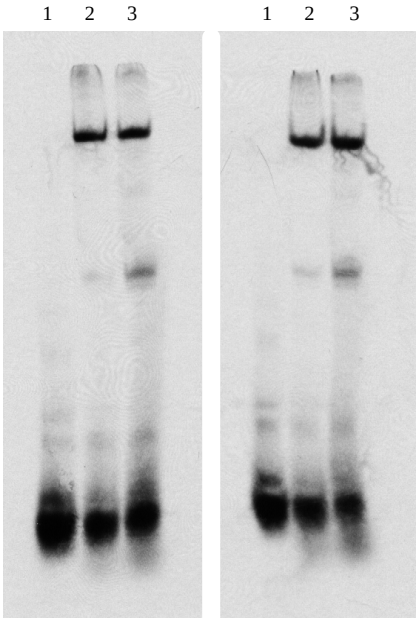
2-1-4- Stabilité des protéines

➢ *en fonction de la force ionique*

Après que la force ionique dans les extraits de protéines ait été augmentée en KCl 2M ou en NaCl 4M (1h à T°C), puis ramenée à des conditions permettant de réaliser des migrations électrophorétiques exploitables (50 mM final) (à l'aide d'étapes de dilution - reconcentration dans un tube reconcentrateur avec un tampon sans sel), nous conservons l'activité retardée habituelle.

➢ *en fonction du pH*

Nous avons réalisé deux expériences de migration électrophorétique non dénaturante en fonction du pH de migration (figure 3.6., pH 6.5 et figure 3.7., pH 8) .



Incubation à pH 6.5 Incubation à pH 8

Figure 3.6. : Migration électrophorétique à pH 6.5 en conditions non dénaturantes après incubation à pH 6.5 ou pH 8, à température ambiante, durant 15 minutes, de 1.1 pmol de Tlev* :

1 : en absence de protéines ;

2 : en présence de 5 µg d'extraits de protéines de levure et de 500 pmol de d(T_{14}) ;

3 : en présence de 5 µg d'extraits de protéines de levure.

Dans le cas de la migration à pH 6.5 (TCE 0.5x), l'incubation a eu lieu à deux pH (tampons TBE pH 8 et TCE pH 6.5), avec ou sans une suite de désoxythymidines d(T_{14}) pour compétiteur non

spécifique (500 fois plus que de Tlev). On observe une bande d'intensité comparable à celle observée pour une migration à pH 8. Une bande d'intensité plus faible et de mobilité électrophorétique plus grande apparaît sensible à la présence de d(T_{14}).

La migration à pH 8 (TBE 0.5x) a lieu après incubation à différents pH (tampons TBE pH 8, HEPES pH 7.2, TCE pH 6.8 et MES pH 6 à 50mM). Aucune variation de l'intensité de la bande retardée intense n'est observée.

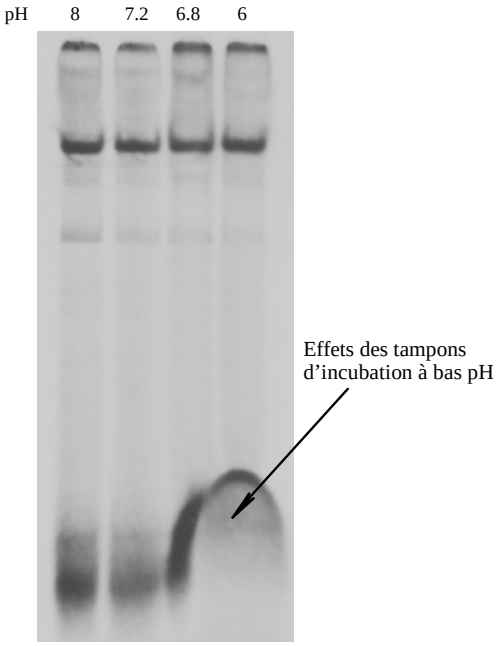


Figure 3.7. : Migration électrophorétique à pH 8.0 en conditions non dénaturantes, après incubation à différents pH de 5 µg d'extraits de protéines de levure et 0.65pmol de Tlev*, à température ambiante durant 15 minutes. On remarque les effets dus aux différents tampons d'incubation sur le front de migration.

Cependant, nous ne pouvons établir avec certitude de correspondance directe entre les bandes intenses observées à pH 6.5 et pH 8. Par ailleurs, la cinétique d'association étant rapide, l'effet d'une incubation à un pH différent de celui de la migration montre seulement que les possibilités de fixation des protéines étudiées ne sont pas altérées (figure 3.7.).

Aussi, pour montrer que l'association peut avoir lieu à pH 6 (MES à pH 6 50mM), nous avons déplété presque totalement en protéines spécifiques, un extrait de protéines de levure et comparé les résultats avec un extrait complet. Pour cela, nous avons utilisé une résine d'affinité porteuse d'un ligand comportant la séquence Tlev. Cette expérience sera ultérieurement commentée (cf. Chapitre IV). A partir de ces deux extraits, nous montrons qu'après une incubation à pH 6 pendant 15 min à 4°C et une migration en conditions non dénaturantes à pH 6, les protéines s'associent à la séquence Tlev (figure 3.8.). On observe à pH 6 pour des extraits nucléaires de protéines de levure, une bande

persistante en présence de d(T_{14}) (figure 3.8., Bande 1) et une autre pour laquelle la séquence d(T_{14}) est compétitrice de Tlev (figure 3.8., Bande 2).

L'incubation des protéines à pH 6 s'accompagnant de l'apparition d'un précipité, pour éviter toute dénaturation ou dégradation des protéines, nous avons réalisé toutes nos expériences à pH 8.

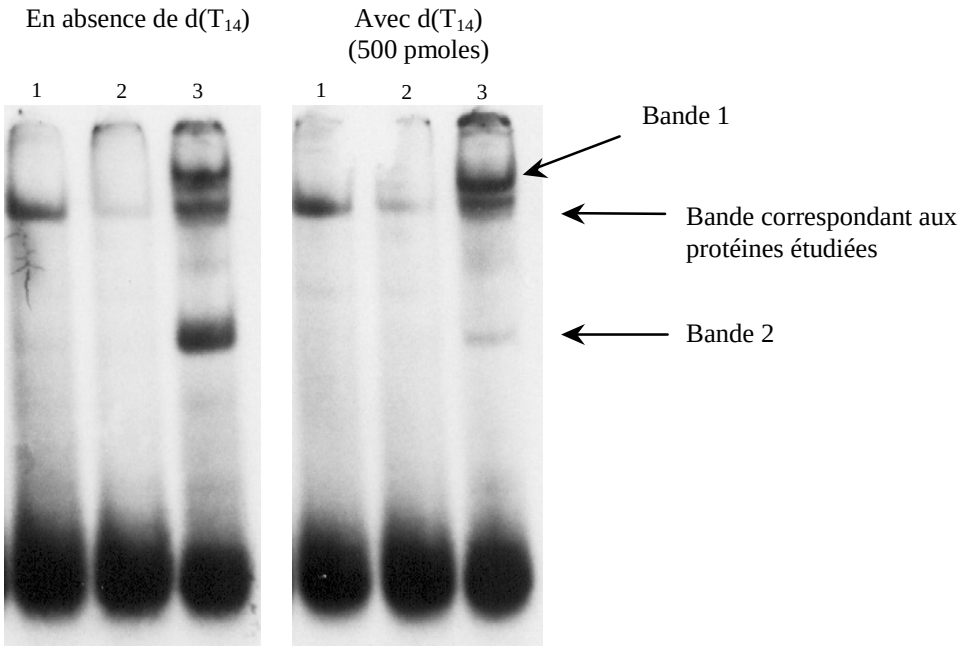


Figure 3.8. : Expérience d'incubation et de migration électrophorétique non dénaturante à pH 6. Incubation 15 minutes à 4°C de 1 pmol de Tlev avec : **1** : 3 µg d'extraits complets de protéines ; **2** : 3 µg d'extraits déplétés en protéines spécifiques de Tlev ; **3** : 24 µg d'extraits nucléaires de protéines de levure.

La formation du complexe Tlev-protéines à pH 6, dans des conditions compatibles avec la formation du motif i, renforce, sans en apporter la preuve, l'idée qu'une association entre Tlev structurée en motif i et des protéines soit possible.

➢ *en fonction de la température*

Indépendamment de toute cinétique d'association avec l'ADN*, la stabilité des extraits protéiques varie en fonction de la température. Nous gardons une activité retardée normale sur des échantillons restés 48 h à 4°C, 15 h à 10° C et 2 h à 37°C. Pour des périodes d'incubation plus importantes à ces températures, l'intensité de la bande retardée diminue. La congélation lente (congélateur à -20 °C)

dégrade totalement les protéines et nous avons systématiquement procédé par congélation à l'azote liquide pour assurer la conservation à -20 °C de nos échantillons. Toutes ces observations sur la stabilité des protéines sont valables pour les extraits protéiques dont la préparation a été décrite préalablement (Chapitre II).

2-2- Protéolyse et hydrolyse des acides nucléiques des extraits de protéines

Nous avons étudié la nature des constituants (acides nucléiques et protéines) du complexe multimérique s'associant à Tlev.

2-2-1- Protéolyse ménagée

Nous avons soumis les extraits de protéines de levure à une protéolyse ménagée, à l'aide d'une protéase alcaline, la subtilisine ou subtilopeptidase A (de masse moléculaire 30 kDa). Cette protéase est active à 37 °C en présence de 10 mM d'acétate de sodium et 5 mM d'acétate de calcium, et est efficace pour 7 < pH < 11. Nous avons utilisé 3 µg d'extraits de protéines de levure par essai en testant une gamme allant de 3 ng à 3 µg de protéase. Le volume réactionnel est constant (11 µl). La période d'incubation à 37°C entre protéase et protéines est de 5 minutes. Ensuite, on incube l'échantillon avec l'ADN* pendant 15 minutes à température ambiante (figure 3.9.).

En parallèle, une expérience en l'absence de protéase est menée pour montrer que l'association ADN-protéine reste possible à 37 °C, pH 8, en présence d'acétate de sodium 10 mM et d'acétate de calcium 5 mM.

On ne distingue pas d'effet particulier en présence de 3 ng de protéase. Pour 30 ng de protéase, on observe un épaississement de la bande retardée, traduisant des différences de vitesses de migration, et donc de tailles, dues à l'action de la protéase, dans l'ensemble des protéines qui s'associent à l'ADN. A partir de 300 ng de protéase, nous ne voyons plus trace d'activité retardée. Le complexe retardé observé dépend donc de la présence de protéines.

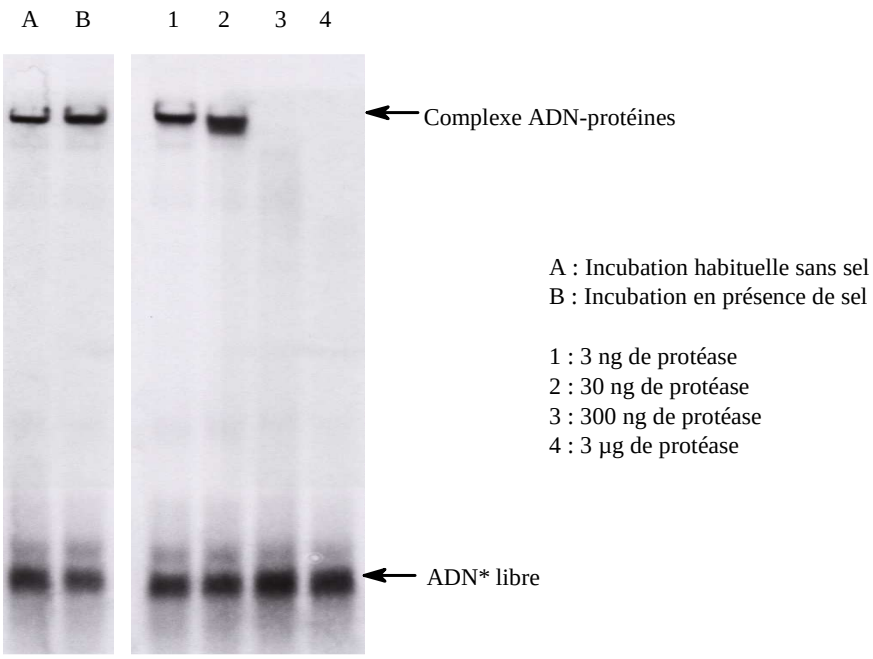


Figure 3.9. : Hydrolyse des acides nucléiques des extraits protéiques de levure. Incubation : 5 min. à 37 °C, en présence de 10 mM de CH_3COONa et 5 mM de $(CH3COO)_2Ca$ (réalisée par K. Snoussi).

2-2-2- Hydrolyse des acides nucléiques de l'extrait de protéines

Nous ne pouvons exclure la présence d'acides nucléiques dans les extraits protéiques de levure. Pour s'assurer que ceux-ci ne participent pas activement à l'association du fragment d'ADN Tlev avec une ou des protéines de levure, nous avons hydrolysé les acides nucléiques présents dans les extraits protéiques en utilisant une endonucléase, la benzonase. Cependant, cette endonucléase doit être par la suite séparée des protéines pour que la séquence Tlev* demeure intacte et soit reconnue durant l'incubation préalable à toute visualisation par migration électrophorétique native du complexe ADN*-protéine. Pour cela, nous avons séparé les protéines fixant Tlev de la benzonase (30 kDa) par passage sur une mini-colonne (5 cm) de Superdex 200.

La benzonase dégrade les ADN et ARN simples et double brins, linéaires ou circulaires, en oligonucléotides n'excédant pas 3 à 5 bases de longueur. Elle est active même en présence de détergents ou d'agents dénaturants. Elle fonctionne entre 0 et 42 °C (maximum d'efficacité à 37 °C) entre pH 6 et pH 10 et nécessite 1 à 2 mM d'ions Mg^{2+} pour une activité maximale. L'efficacité de la benzonase a été vérifiée sur la séquence Tlev et sur un fragment d'ARN d'une dizaine de bases

marquées radioactivement. Les fragments radioactifs d'acides nucléiques simple brin obtenus font moins de 4 bases de longueur.

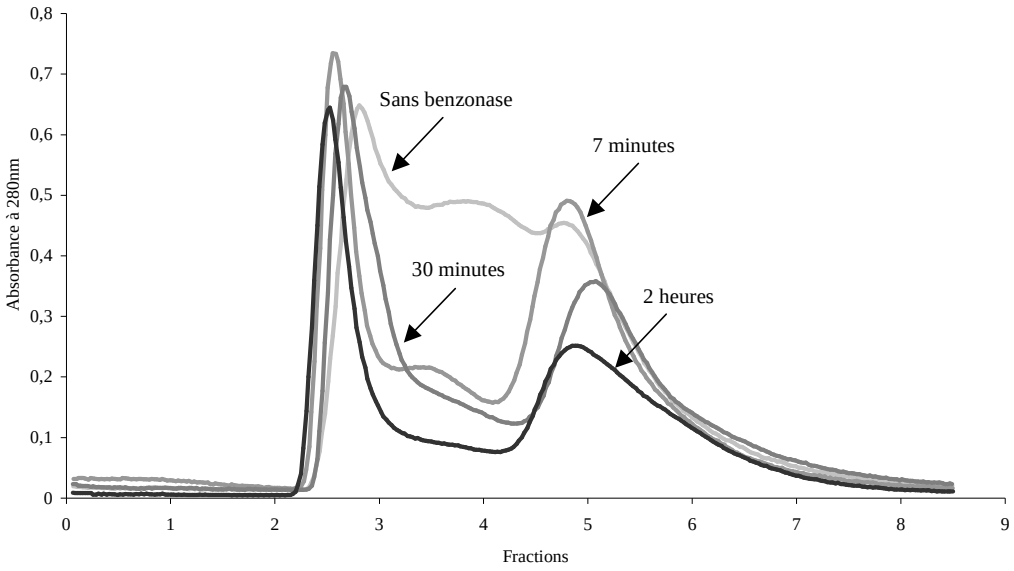


Figure 3.10. : Profils d'élution des essais d'hydrolyse par la benzonase des acides nucléiques présents dans les extraits, en fonction du temps d'incubation.

Nous avons incubé 340 unités de benzonase avec 15 µg de protéines à 37 °C, pendant 7 min, 30 min et 2h, avant de séparer par chromatographie d'exclusion, la benzonase des protéines fixant l'ADN. L'observation des profils d'élution nous permet de constater que l'absorbance à 280 nm est due en partie aux acides nucléiques, lesquels sont bien coupés par la benzonase, comme le montre le décalage de leur sortie vers la fin du profil d'élution avec les poids moléculaires faibles (figure 3.10.).

Les fractions recueillies sont incubées avec l'oligonucléotide Tlev, puis une migration électrophorétique non dénaturante montre que l'intensité de la bande retardée ne varie pas en fonction du temps d'incubation avec la benzonase (figure 3.11.). La masse moléculaire correspondant au signal retardé reste comprise entre 200 et 300 kDa.

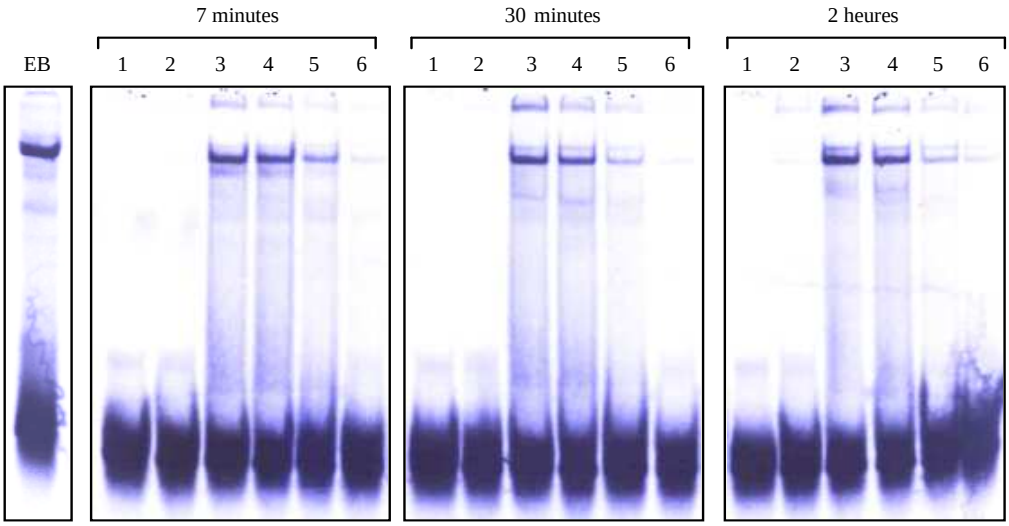


Figure 3.11. : Migrations électrophorétiques natives réalisées à partir des profils d'élution des chromatographies d'exclusion traités à la benzonase en fonction du temps d'incubation des protéines avec la benzonase. EB : extraits protéiques non hydrolysés avant colonne. 1 à 6 : fractions du profil d'élution.

Une expérience complémentaire en fonction de la concentration en benzonase (incubation de 30 minutes à 37 °C), sans élimination de la benzonase par passage sur colonne d'exclusion a été réalisée (figure 3.12.).

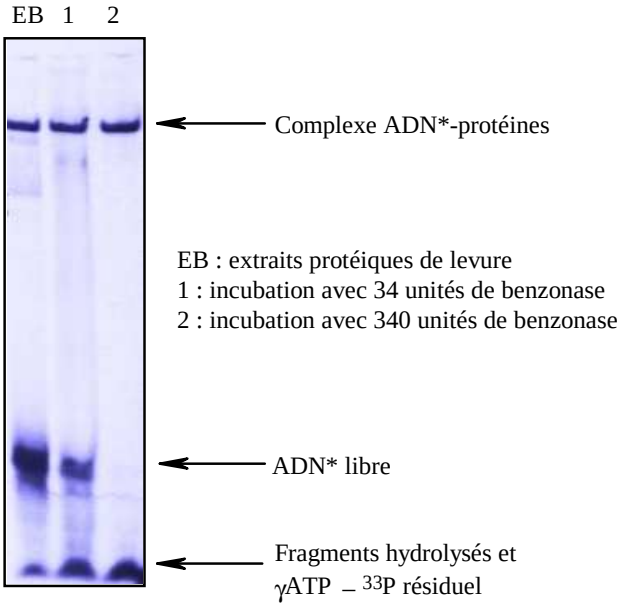


Figure 3.12. : Résultats après incubation avec l'endonucléase sans élimination de celle-ci par passage sur colonne d'exclusion.

Après incubation (15 minutes à température ambiante), l'ADN* libre non fixé par les protéines est dégradé, comme le montre la tache radioactive présente au front de migration. Cependant l'intensité du signal retardé ne varie pas. Ces résultats semblent indiquer qu'après hydrolyse partielle des acides nucléiques présents dans les extraits protéiques (cf. figure 3.10.), la benzonase n'a pas hydrolysé l'ADN Tlev* intervenant dans l'association ADN-protéines ou encore n'a pas hydrolysé l'ADN Tlev* dans des proportions telles que celui-ci ne puisse être reconnu par les protéines alors que la séquence Tlev libre l'était. Cela semble indiquer que Tlev associé aux protéines est relativement bien protégé de l'action de la benzonase.

2-3- Concentration en sites de fixation spécifiques et affinité

La concentration en protéines spécifiques de Tlev et la constante de dissociation du complexe Tlev-protéines formé, peuvent être estimées à partir de la loi d'action de masse. Si l'on considère n sites de fixation de Tlev, équivalents et indépendants, par protéine, la réaction d'association s'écrit à l'équilibre :

$$[\text{ADN*}_{\text{libres}}] + [\text{sites de fixation}_{\text{libres}}] \leftrightarrows [\text{sites de fixation}_{\text{liés}}*]$$

ou encore :

$$[\text{ADN*}_{\text{libres}}] + ([\text{sites de fixation}_{\text{total}}] - [\text{sites de fixation}_{\text{liés}}*]) \leftrightarrows [\text{sites de fixation}_{\text{liés}}*]$$

avec $[\text{sites de fixation}_{\text{liés}}*] = [\text{ADN*}_{\text{liés}}] = x$.

La loi d'action de masse s'écrit alors :

$$\begin{aligned} K_d &= \frac{[\text{ADN*}_{\text{libre}}].[\text{sites de fixation}_{\text{libres}}]}{[\text{sites de fixation}_{\text{liés}}*]} \\ &= \frac{[\text{ADN*}_{\text{libre}}].([\text{sites de fixation}_{\text{total}}] - [\text{sites de fixation}_{\text{liés}}*])}{[\text{sites de fixation}_{\text{liés}}*]} \\ &= \frac{[\text{ADN*}_{\text{libre}}].([\text{sites de fixation}_{\text{total}}] - [\text{ADN}_{\text{liés}}*])}{[\text{ADN}_{\text{liés}}*]} \end{aligned}$$

En posant : $y = [\text{ADN*}_{\text{liés}}] / [\text{ADN*}_{\text{libres}}]$, on obtient : $\boxed{y = \frac{[\text{sites de fixation}_{\text{total}}] - x}{K_d}}$

En faisant varier la concentration en Tlev*, pour un volume d'incubation constant et une concentration en protéines donnée (figure 3.13), on obtient un diagramme de Scatchard (figure 3.14.) qui présente une droite dont la pente correspond à la valeur de la constante de dissociation et l'extrapolation sur l'axe des abscisses permet d'obtenir la concentration totale en sites spécifiques. Cette expérience est réalisée en présence d'un excès de compétiteur non spécifique d(T_{14}).

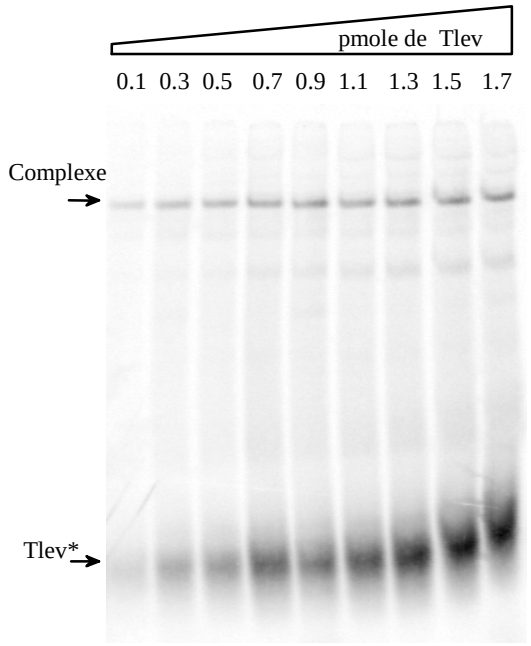


Figure 3.13. : Titration des sites de fixation de la séquence Tlev* : Migration électrophorétique réalisée pour un volume d'incubation constant (15 µl) et 1 µl (=1 µg) d'extraits de protéines de levure en présence de 500 pmol de d(T_{14}). La concentration en Tlev* varie de 6.6 nM à 113 nM.

A partir de ces données, nous obtenons une concentration en sites de fixation de ~ 6.1 . 10^{-8} M pour une constante de dissociation K_d de ~ 3.2 . 10^{-8} M.

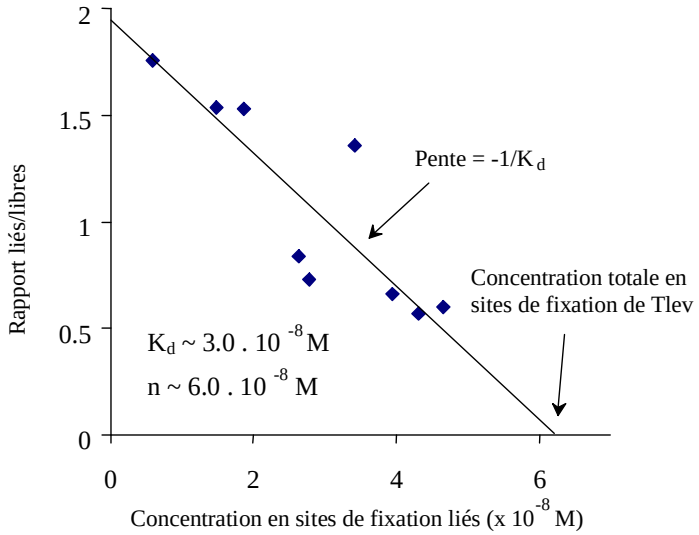


Figure 3.14. : Diagramme de Scatchard : la droite obtenue a une pente proportionnelle à l'inverse de la valeur de K_d et l'intersection avec l'axe des abscisses donne la concentration totale en sites de fixation.

2-4- Masse moléculaire apparente

La masse moléculaire apparente des protéines s'associant à l'ADN est déterminée par chromatographie d'exclusion, sur résine Superdex200 (longueur de la colonne : 90 cm, diamètre : 1cm) préalablement étalonnée. Les fractions du profil d'élution sont récupérées et testées par électrophorèse non dénaturante après incubation avec Tlev*. La quantité de protéines spécifiques de l'ADN est conservée (comparaison entre l'intensité du signal retardé avant et après passage sur colonne). Dans les conditions décrites au chapitre II, nous trouvons une masse moléculaire comprise entre 200 et 300 kDa.

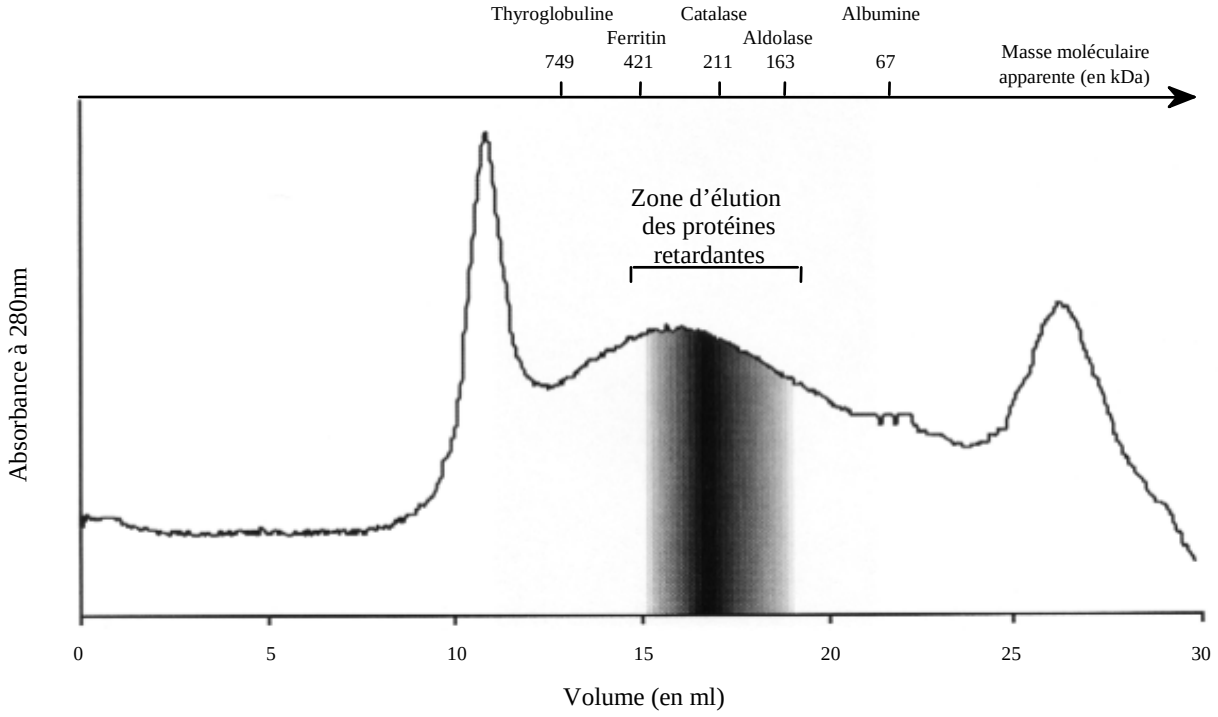


Figure 3.15. : Détermination de la masse des protéines par passage sur tamis moléculaire en conditions natives. Les fractions retardant l'ADN Tlev* sont centrées sur 250 kDa.

La concentration en protéines des extraits étant de 1 g/l, en fonction des résultats obtenus par titration et de la masse moléculaire apparente du complexe protéique (250 kDa), nous pouvons dire que si n = 1, la masse en protéines spécifiques représente $6.1.10^{-8}$ (mol/l) x 250000 (g/mol) / 1 (g/l) = 0.015 g par gramme de protéines de l'extrait , soit 1.5 % de la masse totale de l'extrait de protéines étudié.

Nous avons dénaturé la fraction du profil d'élution donnant un signal retardé et nous l'avons comparé avec un extrait protéique dosé pour donner une bande retardée d'intensité identique, lors d'une expérience de migration électrophorétique en conditions dénaturantes. Nous constatons un appauvrissement significatif de l'extrait et nous n'observons aucune protéine abondante ayant une masse supérieure à 150 kDa (figure 3.16.). Il peut donc s'agir d'un complexe multimérique dont nous ne pouvons rien dire de son homogénéité.

Les protéines s'associant à Tlev pouvant se présenter sous la forme d'un complexe multimérique, nous avons étudié la stabilité de ce complexe. En effet, sous certaines conditions (concentration, force ionique, pH, température), un tel complexe pourrait se dissocier. Si un catalyseur (ou un cofacteur) nécessaire à l'association est absent, la dissociation pourrait être irréversible de même que si l'un des composants du complexe est dégradé après dissociation.

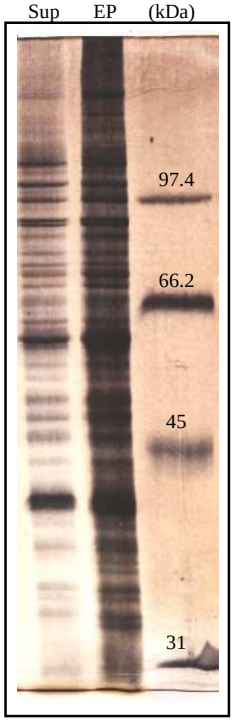


Figure 3.16. : Gel de protéines dénaturant permettant de comparer la fraction du profil de chromatographie d'exclusion centrée sur 250 kDa et l'extrait protéique de levure pour un signal retardé d'intensité identique lors d'une électrophorèse non dénaturante.

En diluant (ou concentrant) d'un facteur 10 les extraits de protéines de levure, nous obtenons un signal retardé proportionnellement plus faible (ou plus fort) que normalement. De plus, en diluant un extrait protéique 800 fois (pendant 1h à 4°C, correspondant à la durée moyenne d'une chromatographie anionique), nous conservons la capacité à retarder habituelle, sans constater de perte de l'intensité du signal. Cela semble indiquer que le complexe protéique, s'il se dissocie, peut se réassocier de façon réversible et rapide.

2-5- Spécificité de l'association ADN-protéines

On peut avoir une idée de la spécificité de l'association entre l'ADN Tlev et les protéines en réalisant des expériences de compétition. En ajoutant au milieu d'incubation un acide nucléique dit 'compétiteur' en plus de la séquence Tlev (l'une des deux séquences étant radioactivement marquée), on observe à partir de quelle concentration ce compétiteur entre en compétition avec Tlev (disparition du signal retardé). Il existe plusieurs catégories de compétiteurs, définies en fonction du rapport [compétiteur]/[Tlev] à partir duquel on observe une compétition (diminution supérieure à 50% de l'intensité du signal retardé) :

- les séquences non spécifiques : [compétiteur]/[Tlev] > 1000
- les mauvais compétiteurs : 10 < [compétiteur]/[Tlev] < 1000
- les compétiteurs spécifiques : [compétiteur]/[Tlev] < 10
- les compétiteurs interagissant avec l'ADN Tlev

Pour vérifier qu'une séquence s'associe avec une mauvaise affinité ou ne peut pas s'associer aux protéines ciblées, nous avons entrepris des migrations électrophorétiques non dénaturantes pour lesquelles les séquences à tester étaient marquées et l'ADN Tlev était le compétiteur froid.

Tous les essais sont réalisés avec 5 µg de protéines d'extraits de levure, soit environ ~ 0.5 pmol de sites de fixation spécifiques.

➢ *Séquences non spécifiques*

Des fragments d'ADN d'*E. coli* de longueurs variant de 100 à plus de 1000 bases et la molécule d'ATP sont classés dans la catégorie des compétiteurs non spécifiques. Autre compétiteur classé dans cette catégorie, le fragment d(CCC) pour lequel nous n'avons vu aucun effet. Ce fragment peut se structurer en motif i tétramérique à pH < 6, pour une concentration de l'ordre du millimolaire. Il ne peut former cette structure dans les conditions de l'incubation.

Pour le simple brin 21mix (Lacroix et *al.*, 2000), on observe un effet de compétition à partir d'un facteur 1000 de même que pour la séquence T_{14}. Le double brin Igo n'entre pas en compétition avec Tlev*, même pour un rapport [compétiteur] / [Tlev] = 3000 (figure 3.12).

Noms	Séquences non spécifiques (5'→3')	Remarques
Igo	d(CGCGATCGCG)	Double brin d'ADN B [a]
21mix	d(CTCACACTCACACTCACACCC)	Simple brin [b]
T_{14}	d(TTTTTTTTTTTTTT)	Simple brin
C_3	d(CCC)	Simple brin ou motif i [c]

Tableau 3.3. : Séquences non spécifiques. [a] Leroy et *al.*, 1988, [b] Lacroix et *al.*, 2000, [c] Leroy et *al.*, non publié

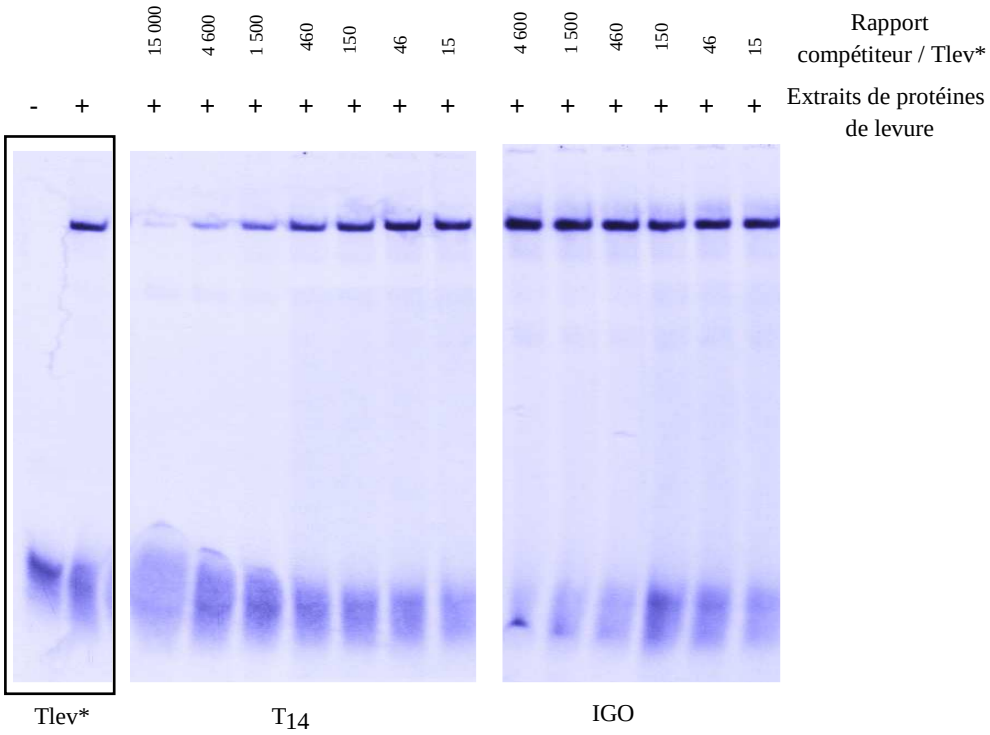


Figure 3.17. : Gel retard non dénaturant obtenu après incubation de Tlev* et avec les séquences T_{14} et Igo et ~ 0.5 pmol de sites de fixation spécifiques. La concentration en Tlev est de 50 nM et le volume d'incubation de 16.8 µl.

➢ *Mauvais compétiteurs*

Lors d'une rapide étude par spectroscopie d'absorption de la séquence d(CCCACACCC), nous avons observé une température de dénaturation thermique de 32°C à 50 µM, pH 5, en présence de NaCl 175 mM et $MgCl_2$ 2 mM. Nous avons vu qu'à 3 µM, cette séquence ne semblait pas adopter de structure.

Noms	Mauvais compétiteurs (5'→3')	Remarques
Crypt1.5	d(CCCCTAACCCC)	Motif i dimérique [a]
Sacc1.5	d(CCCACACCC)	Motif i dimérique [b]
27ccg	d(CCGCCGCCGCCGCCGCCGCCGCCGCCG)	Répétition trinucléotidique[c]
T_{26}	d(TTTTTTTTTTTTTTTTTTTTTTTTTT)	Simple brin

Tableau 3.4. : Mauvais compétiteurs. [a] Nonin-Lecomte S., communication personnelle, [b] ce travail, [c] (Moore et *al.*, 1999)

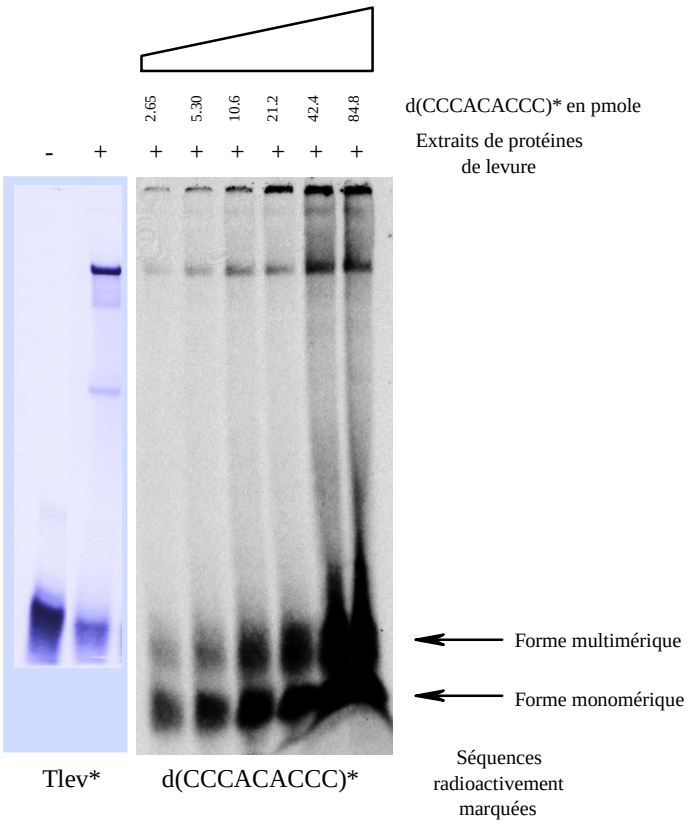


Figure 3.18. : Gel retard non dénaturant obtenu après incubation de ~ 0.5 pmol de sites de fixation spécifiques de levure et différentes concentrations de la séquence d(CCCACACCC)*. A gauche rappel des résultats obtenus avec la séquence Tlev* pour une même quantité de protéines de levure (temps d'exposition différents) (cf. figure 3.4.).

Nous pensons que cette séquence, de même que dans le cas de la séquence Crypt1.5, peut former un motif i dimèrique. Nous observons un faible signal retardé lorsque nous utilisons la séquence Sacc1.5 comme sonde radioactive (figure 3.18). L'amplitude de ce signal augmente en fonction de la concentration en brin, Sacc1.5 ou Crypt1.5, de même que le signal observé d'une forme

multimérique. Ceci peut indiquer que la protéine se fixe à un multimère. La répétition trinucléotidique 27ccg est capable de former une structure tige-boucle *in vivo* (Moore et *al.*, 1999). Les résultats obtenus avec cette séquence indiquent qu'elle est un mauvais compétiteur de la séquence Tlev.

Lors d'une incubation dans un rapport [T_{26}] / [Tlev] = 100, nous observons que le signal retardé est fortement atténué (figure 3.19.). Après incubation avec les extraits de protéines de levure, la séquence T_{26}* marquée radioactivement présente un retard diffus dans la zone d'observation habituelle du signal caractéristique des protéines de levure s'associant à Tlev (figure 3.19.).

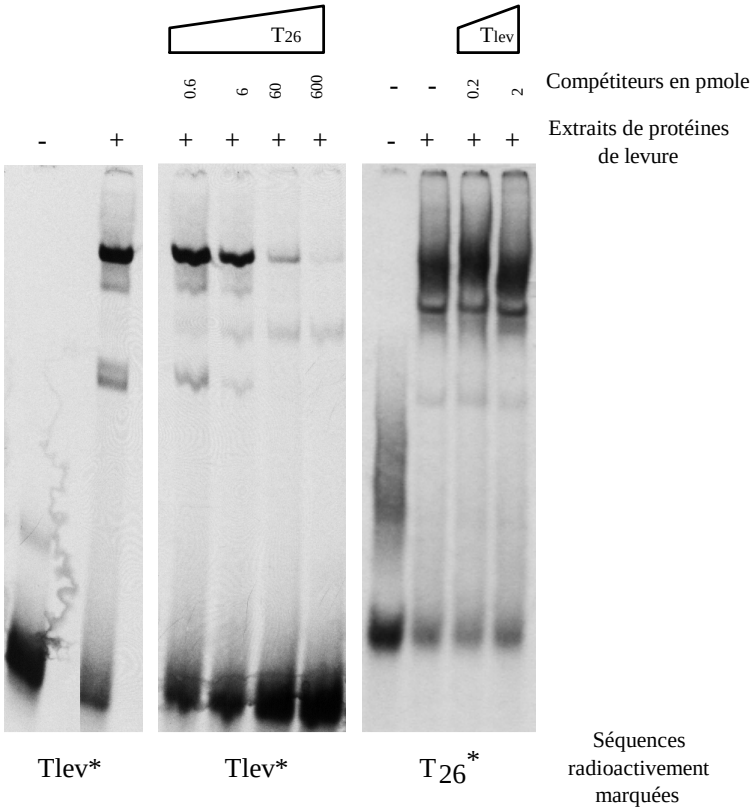


Figure 3.19. : Gel retard non dénaturant obtenu après compétition entre les séquences Tlev et T_{26}, et incubation avec ~ 0.5 pmol de sites de fixation spécifiques de levure. La quantité de Tlev* est constante (0.65pmol) de même que pour T_{26}* (0.2pmol).

➢ *Compétiteurs spécifiques*

Les séquences testées retenues dans cette catégorie peuvent toutes adopter un motif i intramoléculaire. A l'exception de la séquence C_{26}, ces séquences présentent toutes 4 blocs de 2 à 5 désoxycytidines entre lesquels sont intercalées des variations de 2 à 3 bases A et / ou T. Aucune différence n'a pu être faite sur l'efficacité des compétiteurs testés en fonction de la nature de ces variations de séquences.

Les séquences comportant des répétitions à 2 désoxycytidines ont une affinité moins bonne que celles présentant des répétitions de 3 à 5 désoxycytidines. Parmi les différents oligodésoxynucléotides testés, nous ne mentionnerons dans le tableau suivant que ceux présentant des caractéristiques significativement différentes les unes des autres.

Noms	Compétiteurs spécifiques (5'→3')	Compétitions	Remarques
Tlev	d(**CCCACACCCACACCCACACCC**)	+++	Possible répétition télomérique de levure / motif i
C_{26}	d(CCCCCCCCCCCCCCCCCCCCCCCCCC)	++	Motif i / simple brin [a]
20sc	d(CCACACCACACCACACCACA)	+	Possible répétition télomérique de levure / motif i
24sc	d(CCACACCCACACCCACACCCACAC)	++	Possible répétition télomérique de levure / motif i
50sc	d[$C_{2-3}A(CA)_{1-6}$] (50 bases)	+*	Séquence télomérique naturelle provenant de levure (50 bases)
Centmod	d(CCTTTCCTTTACCTTTCC)	+	Motif i intramoléculaire [b]
Mcent-u	d(CCATTCCAUTCCUTTCC)	+	Centromère modifié / motif i intramoléculaire [c]
Cena	d(CCATTCCATTCCATTCCA)	+	Centromère (satellites II et III humain, satellite CDE III de levure) / motif i [d]
Centro6	d(ATTCCATTCCATTCCATTCCATTCCATTCC)	+*	Centromère (satellites II et III humain, satellite CDE III de levure) / motif i [d]
22hm	d(CCCTAACCCTAACCCTAACCCT)	+++	Possible répétition télomérique humaine / motif i intramoléculaire [e]
27hm	d(TAACCCTAACCCTAACCCTAACCCTAA)	+++	Possible répétition télomérique humaine / motif i intramoléculaire [a]
45hm'	d(CCCTAACCCTAACCCTAACCCTAACCCTAAC CCTAACCCTAACCC)	++*	Possible répétition télomérique humaine / motif i intramoléculaire [a]
BTlev	B-d(TTTT**CCCACACCCACACCCACACCC**)	+++	Séquence Tlev biotinylée en 5'
T_4Tlev	d(TTTT**CCCACACCCACACCCACACCC**)	+++	Séquence BTlev sans biotine
Crypt25	d(CCCCTAACCCCTAACCCCTAACCCC)	++*	Possible répétition télomérique de *Cryptococcus* / motif i [d]
Tetra22	d(CCCCAACCCCAACCCCAACCCC)	+++*	Possible répétition télomérique de *Tétrahymena* / motif i [f]
29h	d(CCCCCTAACCCCCTAACCCCCTAACCCCC)	+++*	Motif i [a]
29i	d(CCCCCTTTCCCCCTTTCCCCCTTTCCCCC)	+++*	Motif i [a]

Tableau 3.5. : Compétiteurs spécifiques. +, ++, +++ : diminution de moitié du signal retardé pour un rapport compétiteur / Tlev* de 100, 10 et 1 respectivement ; * séquences pour lesquelles on observe la formation de complexes ADN*-protéines autres que celui sur lequel porte cette étude. [a] Lacroix et *al.*, 2000, [b] Han et *al.*, 1998, [c] Nonin-Lecomte & Leroy, 2001, [d] Nonin-Lecomte, communication personnelle, [e] Phan et *al.*, 2000, [f] Leroy J.L., non publié

Dans le cas des séquences Centmod, Mcent-u et 22hm, nous avons observé qu'elles s'associaient aux protéines lors d'incubation à pH 6.8 suivi de migration à pH 6.8 (tampon d'électrophorèse TCE 0.5x, pH 6.8), dans des conditions où elles ont été observées structurées en motif i par RMN. Les séquences 29h et 29i (structurées en motif i à pH 6.8) s'associent elles aussi avec les protéines dans ces conditions.

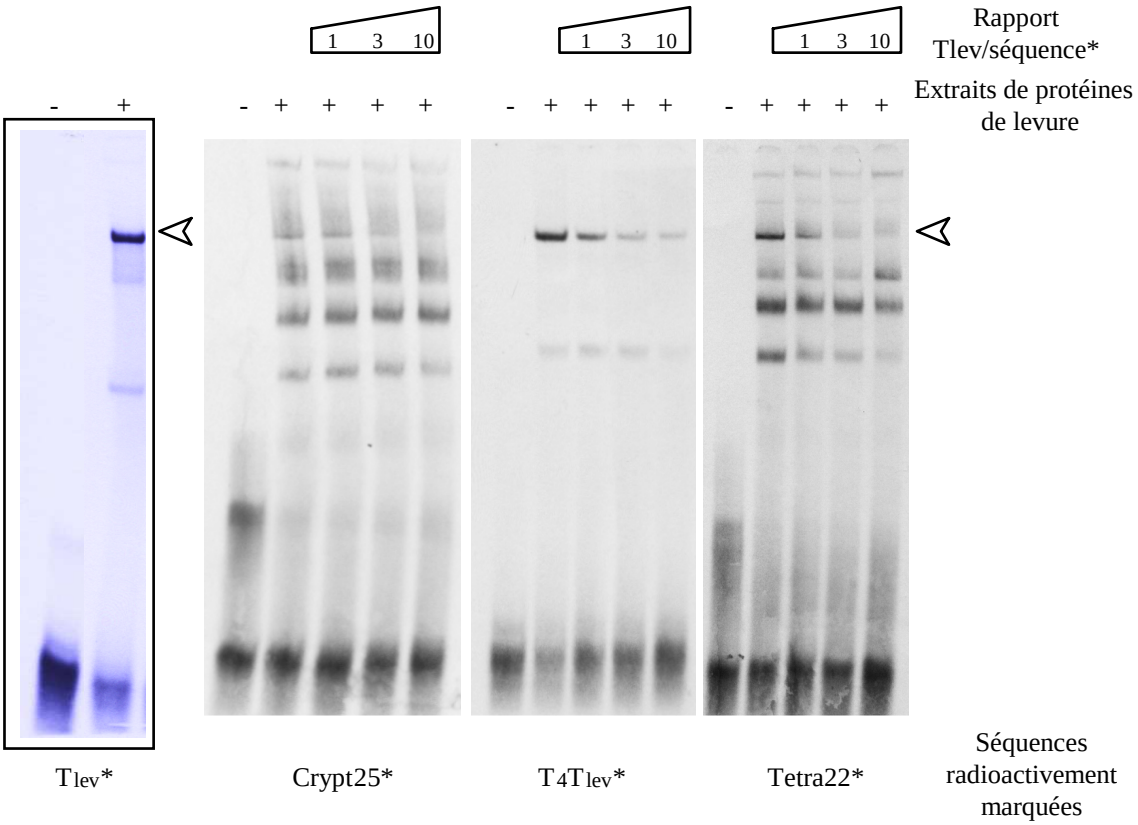


Figure 3.20. : Gels retard non dénaturant obtenus pour différents compétiteurs spécifiques. La séquence Tlev est utilisée comme compétiteur froid et les séquences testées (Crypt25, T_4Tlev et Tetra22 sont radioactivement marquées).

Comme on peut le remarquer sur les exemples donnés en figure 3.20., certaines séquences testées peuvent s'associer à des protéines pour lesquelles la séquence Tlev n'est pas compétitrice. Ces séquences sont indexées d'un * dans la colonne 'compétitions' du tableau 3.5. .

➢ *Compétiteurs en interaction avec Tlev*

Dans cette catégorie, nous regroupons les séquences pour lesquelles on observe une compétition due vraisemblablement à l'interaction avec la séquence Tlev. Les fragments d'ADN Ghum et

GTlev, partiellement ou totalement complémentaires de la séquence Tlev, peuvent interagir avec cette séquence (double brin). Nous décrirons ultérieurement les phénomènes intervenant en présence de ces séquences riches en G.

Noms	Compétiteurs interagissant avec la séquence Tlev (5'→3')	Remarques
27cgg	d(CGGCGGCGGCGGCGGCGGCGGCGG)	Tige-boucle [a]
21gc	d(CCCTAACCCTAAGGGTTAGGG)	Tige-boucle [b]
Ghum	d(AGGGTTAGGGTTAGGGTTAGGG)	Possible répétition télomérique humaine / tétrade de G [c]
GTlev	d(GGGTGTGGGTGTGGGTGTGGG)	Brin complémentaire de Tlev

Tableau 3.6. : Séquences interagissant avec Tlev. [a] Moore et *al.*, 1999 ; [b] Mergny, communication personnelle ; [c] Thèse Phan Anh Tuan

➢ *Conclusion*

Les protéines recherchées sont spécifiques de séquences d'ADN riches en cytosines. Cependant, elles ne s'associent pas au fragment élémentaire d(CCC) dont nous avons montré qu'il formait un motif i intermoléculaire sous des conditions de concentration et de température différentes de celles dans lesquelles peuvent avoir lieu le contact avec les protéines (Leroy et *al.*, en préparation).

Par ailleurs, l'association avec les fragments Crypt1.5 et Sacc1.5 est fonction de la concentration en ADN, ce qui peut traduire l'éventuelle association entre les protéines et une forme multimérique adoptée par ces séquences.

Il existe une différence d'affinité entre les séquences de type centromérique à 4 répétitions de deux désoxycytidines et les séquences télomériques à 4 répétitions de 3 désoxycytidines et plus. Toutes les séquences à 4 répétitions de désoxycytidines testées sont capables d'adopter un motif i intramoléculaire. Cependant, la stabilité des séquences de type centromérique est moins grande que celle des séquences télomériques. Ainsi dans le cas du fragment centromérique humain, Mcent-u, deux intercalations sont observées en fonction du pH (Nonin-Lecomte & Leroy, 2001). Toutes ces constatations nous orientent vers l'hypothèse d'une association spécifique entre les protéines et un motif i intramoléculaire, sans cependant en apporter la moindre preuve.

2-6- Comparaison de nos résultats avec des protéines connues

Nous avons étudié la participation éventuelle dans le complexe protéique associé au fragment télomérique Tlev de protéines connues pour interagir avec les télomères de levure. Nous avons regardé ce qu'il en était des protéines Rap1p, Sir1p, Sir3p et Sir4p.

Nous avons disposé de sérum anti-Rap1p et d'une souche d'*E. coli* portant le plasmide de la protéine Rap1p, fournis par M. E. Gilson (Laboratoire de biologie moléculaire et cellulaire de l'Ecole Normale Supérieure de Lyon).

Après incubation des extraits protéiques et de Tlev* des anticorps de Rap1p, nous n'avons pas observé d'effet de supershift ou de disparition partielle ou totale de la bande retardée habituelle lors d'une électrophorèse non dénaturante. De même, nous n'avons pas observé de bande retardée après migration des protéines produites à partir du plasmide de Rap1p [Expériences réalisées par Mme M. Leblond (Gilson et *al.*, 1993)].

Pour les protéines Sir1p, Sir3p et Sir4p, nous disposions pour chacune, de souches de levure où les protéines n'étaient pas exprimées (tableau 3.7). Nous avons produit des extraits de protéines à partir de ces souches selon le mode opératoire défini précédemment. Après incubation entre ces différents extraits et Tlev*, nous avons procédé à des migrations électrophorétiques natives. La bande retardée habituelle est apparue sans changement.

Dénomination	Génotype
YLS 126	*MATα, ade2-1, trp1-1, can1-100, leu2-3, 112, his3-11, 15, ura3-1,* hmr :: TRP1 Δsir3::LEU2
YLS 134	*MATα, ade2-1, trp1-1, can1-100, leu2-3, 112, his3-11, 15, ura3-1,* hmr :: TRP1 Δsir4::LEU2
YLS 396	*MATα, ade2-1, trp1-1, can1-100, leu2-3, 112, his3-11, 15, ura3-1,* hmr :: TRP1 Δsir1::LEU2

Tableau 3.7. : Souches où les protéines Sir1p, Sir3p et Sir4p n'étaient pas exprimées.

Chapitre IV

Identification des protéines du complexe

1- Protéines identifiées par affinité (streptavidine immobilisée)

1-1- Choix de la séquence biotine-ADN

Nous avons développé un protocole d'identification par affinité en utilisant de la streptavidine immobilisée. A cette fin, nous avons synthétisé la séquence BTlev (cf. tableau 4.1.) sur la base de la séquence Tlev, en prenant soin d'éloigner par une suite de thymines, la biotine de la partie Tlev. Si les protéines s'associent à la séquence Tlev sous une forme structurale en motif i ou sous une autre, l'ajout d'une biotine à l'extrémité 5' de la séquence ne devrait pas perturber la structure adoptée par l'ADN.

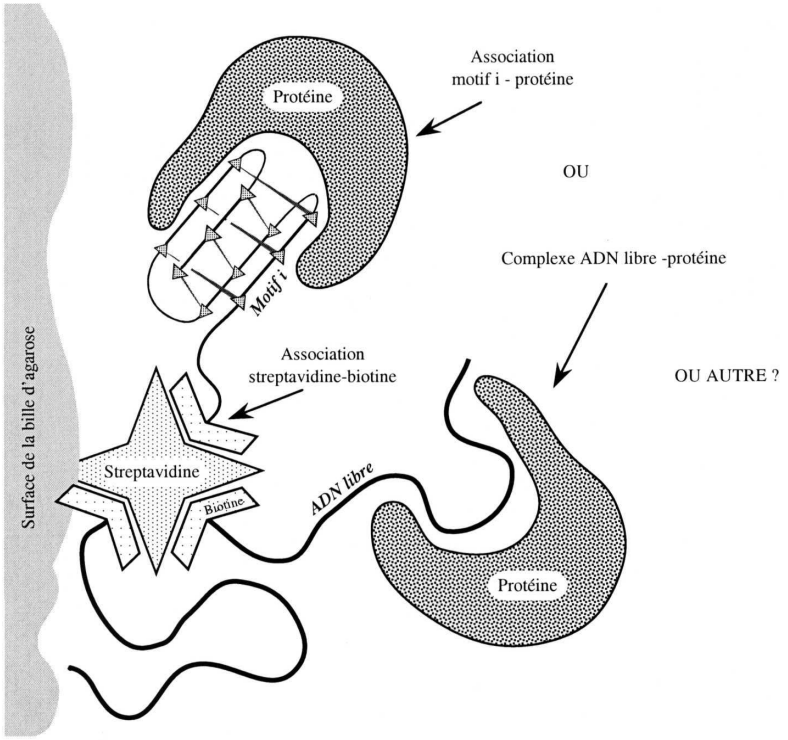


Figure 4.1. : Séparation par affinité. Quelle que soit la structure adoptée par la séquence Tlev et reconnue par les protéines, la formation du complexe ADN-protéines doit rester possible malgré la présence d'une biotine à l'extrémité 5' et la proximité immédiate de la streptavidine et de son support.

Pour nous assurer que la séquence modifiée n'entraînera pas l'association artéfactuelle d'autres protéines, et vérifier la longueur du segment de thymines nécessaire et suffisant pour ne pas perturber la structure adoptée par l'ADN, nous avons réalisé plusieurs essais de migrations électrophorétiques en conditions natives, à partir de protéines non purifiées et de séquences présentant des variations de la longueur du bras de thymines (tableau 4.1.).

Noms	Séquences (5'→3')
Tlev	**d(CCCACACCCACACCCACACCC)**
TBT_2Tlev	dT-B-d(TT**CCCACACCCACACCCACACCC**)
T_4Tlev	d(TTTT**CCCACACCCACACCCACACCC**)
BTlev	B-d(TTTT**CCCACACCCACACCCACACCC**)
TBTlev	dT-B-d(TTTT**CCCACACCCACACCCACACCC**)
TBT_7Tlev	dT-B-d(TTTTTTT**CCCACACCCACACCCACACCC**)

Tableau 4.1. : Séquences testées en prévision de la séparation par affinité.

On constate que l'insertion d'une biotine (séquences BTlev* et TBTlev*, figure 4.2.), de même que l'allongement de la séquence Tlev par une suite de 4 thymines (séquence T_4Tlev*), ne permet pas

de visualiser de nouvelle bande de protéines intense. Il en va de même lors de tests réalisés pour des suites de 2 ou 7 thymines (résultats non montrés).

On remarque également que le rendement de kination est moins bon pour ces séquences qu'avec la séquence Tlev* (notamment pour la séquence BTlev* où la biotine est située à l'extrémité 5' du fragment d'ADN).

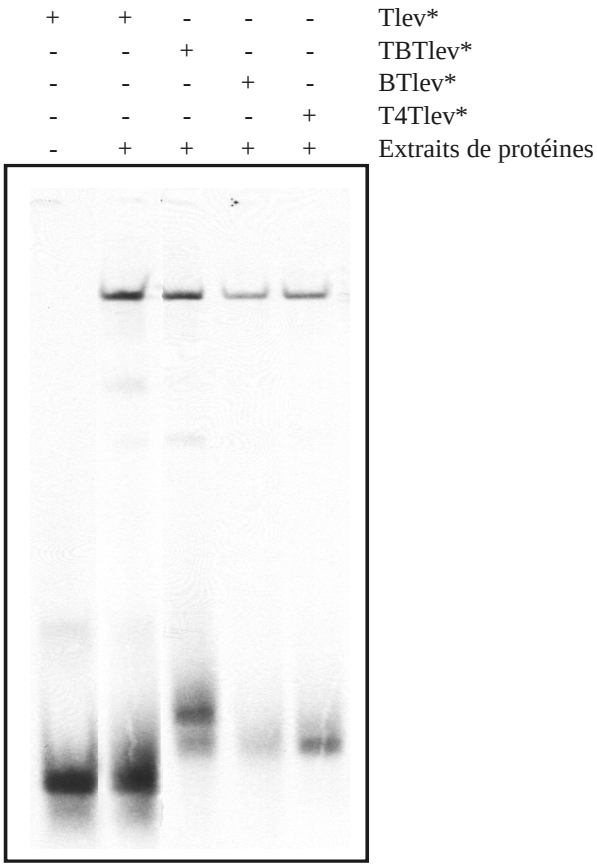


Figure 4.2. : Comparaison entre les migrations électrophorétiques observées pour la séquence Tlev* et pour les séquences T_4Tlev*, BTlev* et TBTlev*. Les concentrations en ADN* sont identiques (~ 65 nM). Sites de fixation spécifiques ~ 0.5pmol

En fonction de ces expériences, le choix du ligand utilisé pour les expériences de séparation par affinité s'est porté sur la séquence BTlev, dont les rendements de synthèse étaient supérieurs à ceux de séquences plus longues.

1-2- Le complexe streptavidine-biotine-ADN-protéines

Nous avons vérifié la formation d'un complexe streptavidine-biotine-ADN-protéines unique par migration électrophorétique en conditions natives. Pour cela, nous avons utilisé la séquence TBTlev* comme sonde radioactive. Dans un premier essai, nous avons vérifié que cette séquence

s'associait avec de la streptavidine à pH 8, température ambiante, en présence de KCl 50 mM durant 15 minutes dans un volume d'incubation de 25 µl (figure 4.3.a.). Nous avons constaté la présence d'un complexe streptavidine-biotine-ADN-protéines (sous la forme d'une bande retardée supplémentaire) après incubation dans des conditions d'incubation non dénaturantes (figure 4.3.b.). De plus, l'intensité de la bande retardée habituellement observée a diminué.

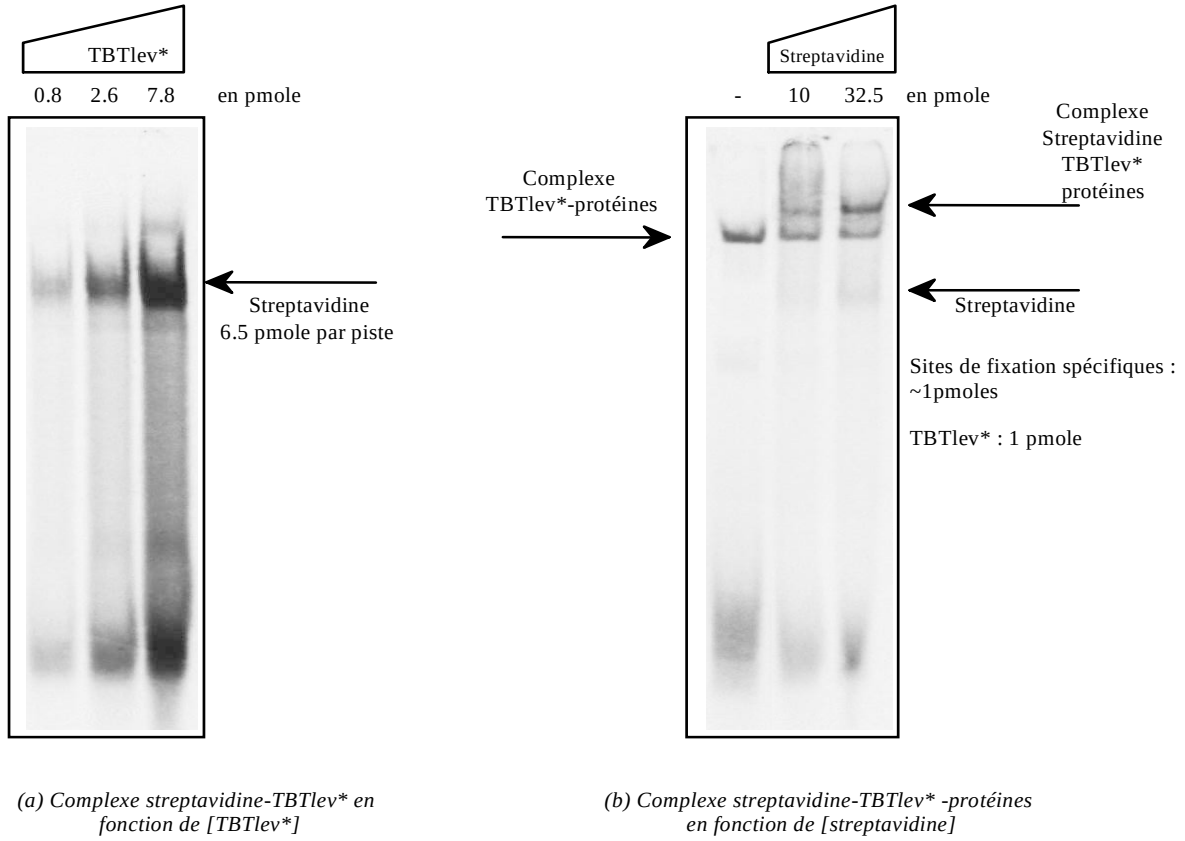


Figure 4.3. : *(a)* Formation du complexe streptavidine-biotine-ADN. *(b)* Formation du complexe streptavidine-biotine-ADN-protéines unique.

1-3- Description du mode opératoire

Afin de vérifier qu'il n'existait pas d'interaction entre la résine-streptavidine et les protéines, nous avons procédé à des essais d'incubation par lots entre la résine (sans le ligand d'ADN) et les protéines. Pour un rapport égal à 1/5 et une incubation de 18h à 4°C dans TBE 0.5x, pH 7.4, 150 mM KCl, l'activité retardante n'est pas affectée. Mais elle est perdue, de même que toutes les autres protéines de l'extrait, si le rapport (volume de billes en suspension (1.4 mg de streptavidine / ml de

résine en suspension) / volume de protéines (1mg de protéines /ml d'extraits)) est supérieur à 1 (vraisemblablement suite à une agrégation ou un amalgame entre résine et protéines).

Lors de nos essais de purification par affinité, nous avons travaillé avec 20 µl de résine-streptavidine en suspension (~ 2 nmol de sites de fixation de biotine), pour 100 - 200µl d'extraits protéiques (~ 8 - 24 pmol de sites de fixation spécifiques), à pH 7.4 et température ambiante.

1-4- Essais de purification en conditions natives

Selon le principe de la séparation par affinité, après incubation des protéines avec la résine-streptavidine préalablement saturée en ADN ligand BTlev, nous avons procédé à des lavages successifs des billes avec le même tampon que celui servant à l'incubation (6 x 1 ml de TBE 0.5x, pH 7.4, 150 mM KCl), avant d'essayer de séparer les protéines du ligand BTlev par élution.

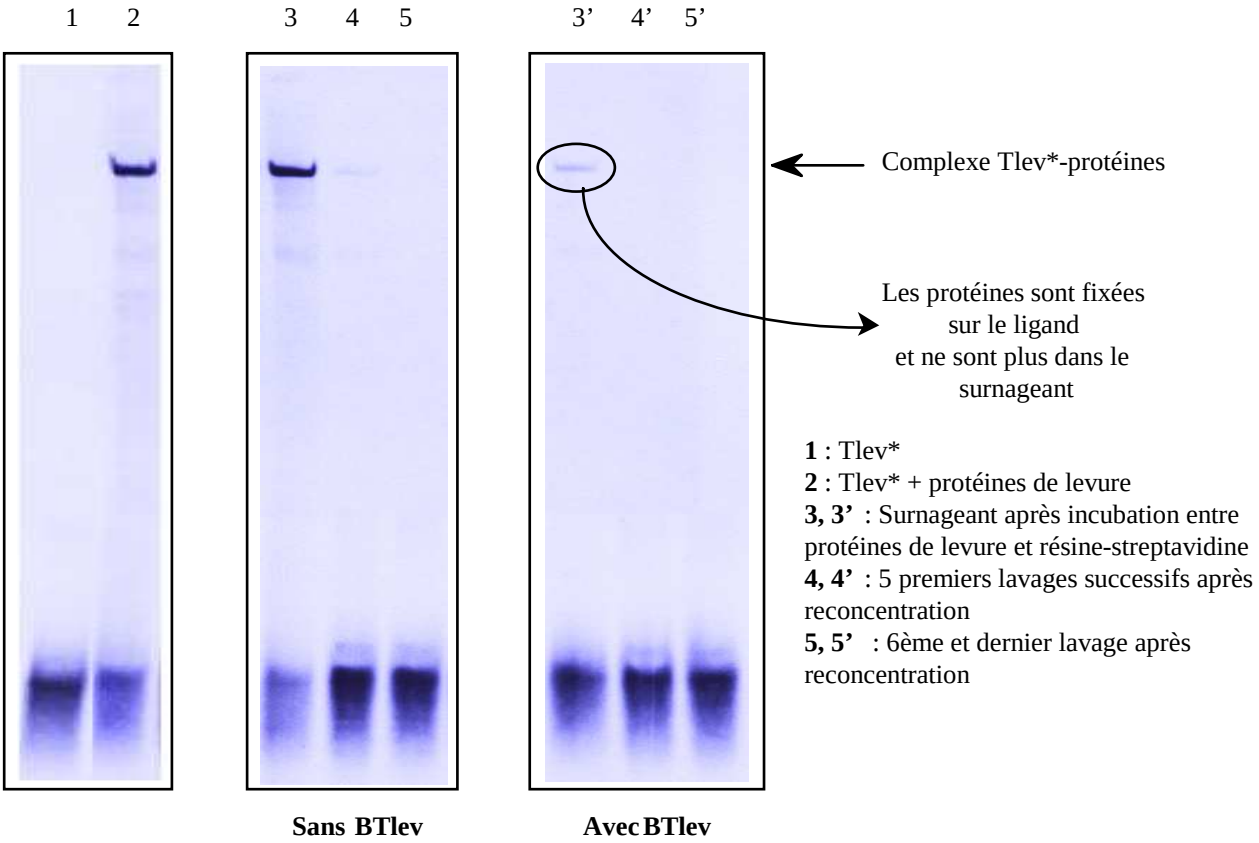


Figure 4.4. : Comparaison entre les surnageants obtenus après contact entre les protéines de levure et la résine-streptavidine porteuse ou non du ligand BTlev

Pour contrôler l'efficacité de l'association spécifique des protéines avec le ligand, nous avons observé systématiquement par migration électrophorétique non dénaturante, les différents surnageants obtenus lors des lavages (figure 4.4.). Aucune bande retardée n'étant relevée pour ces lavages, nous en avons conclu que les protéines sont effectivement fixées sur la matrice d'affinité.

Pour procéder à la séparation entre protéines et ligands, nous avons testé différentes conditions d'élution :

- variation du pH (de pH 7 à pH 8)
- augmentation de la force ionique (de 150 mM à 2 M KCl)
- passage en NaCl (variation de la concentration de 150 mM à 2 M)
- passage en urée (de 2 à 6 M, puis élimination de l'urée par filtration à l'aide des tubes reconcentrateurs).

Aucun de ces procédés n'a permis d'obtenir une activité retardante. Deux types d'explications sont possibles :

- les protéines sont restées liées à l'ADN sur les billes;
- les protéines recherchées sont partiellement éluées, mais le complexe protéique s'est dissocié de manière irréversible pendant ou après l'élution.

2- Identification différentielle de protéines

2-1- Principe de l'expérience

Ayant mis en évidence la fixation des protéines spécifiques sur la résine d'affinité, (figure 4.4.), nous avons entrepris d'identifier ces protéines par différenciation entre l'ensemble des protéines provenant de résines porteuses du ligand BTlev et celles provenant de résines sans ligand.

Pour éliminer les protéines s'associant à la résine d'agarose, nous procédons à une incubation préalable entre les protéines et de la résine-streptavidine sans ligand BTlev. Ce premier surnageant est partiellement débarrassé des protéines spécifiques de la résine et contient cependant les

protéines recherchées, comme en atteste une migration électrophorétique non dénaturante après incubation avec Tlev*.

Afin d'éliminer des associations artéfactuelles entre des protéines non spécifiques et la séquence BTlev, l'incubation entre les protéines du premier surnageant et les résines porteuses ou non de BTlev, a lieu en présence d'un excès de T_{14} dont on a vu au chapitre III paragraphe 2-5 qu'il ne s'associait pas trop avec les protéines recherchées.

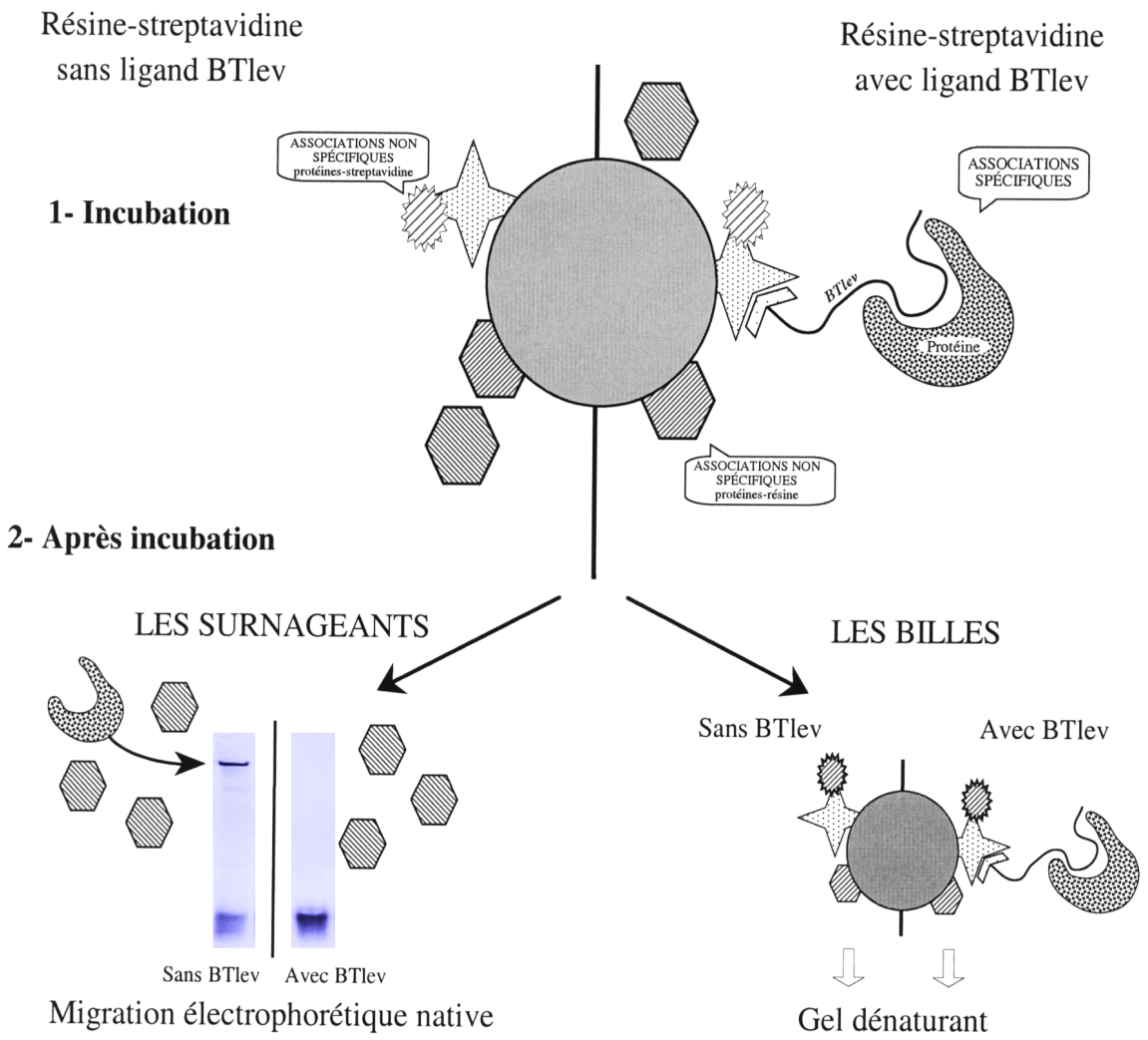


Figure 4.5. : Principe permettant d'identifier les protéines recherchées par différenciation entre l'ensemble des protéines provenant des billes de résine porteuses du ligand BTlev et celles provenant de billes ne portant pas la séquence BTlev.

On étudie dans les deux cas, avec et sans ligand, les protéines restant sur les billes, en les dénaturant et en faisant une migration électrophorétique dénaturante (cf. chapitre II). Nous nous assurons de la

réussite de l'association entre les protéines et le ligand BTlev en réalisant un gel non dénaturant avec les différents surnageants.

En l'absence de T_{14}, on visualise 12 bandes de protéines, absentes dans le cas des billes sans ligand BTlev et 9 bandes quand l'incubation a lieu en présence de 100 fois plus de T_{14} que de ligand (figure 4.6.). En présence de 1000 fois plus de T_{14} on ne visualise plus que trois bandes différenciées, dont un doublet intense (agrandissement de la figure 4.6.). C'est donc en présence de 1000 fois plus de T_{14} que de ligand lors de l'incubation que nous avons procédé aux découpages des bandes de protéines devant servir aux identifications par spectrométrie de masse. On proposera une brève étude de la bande de plus faible intensité au chapitre V.

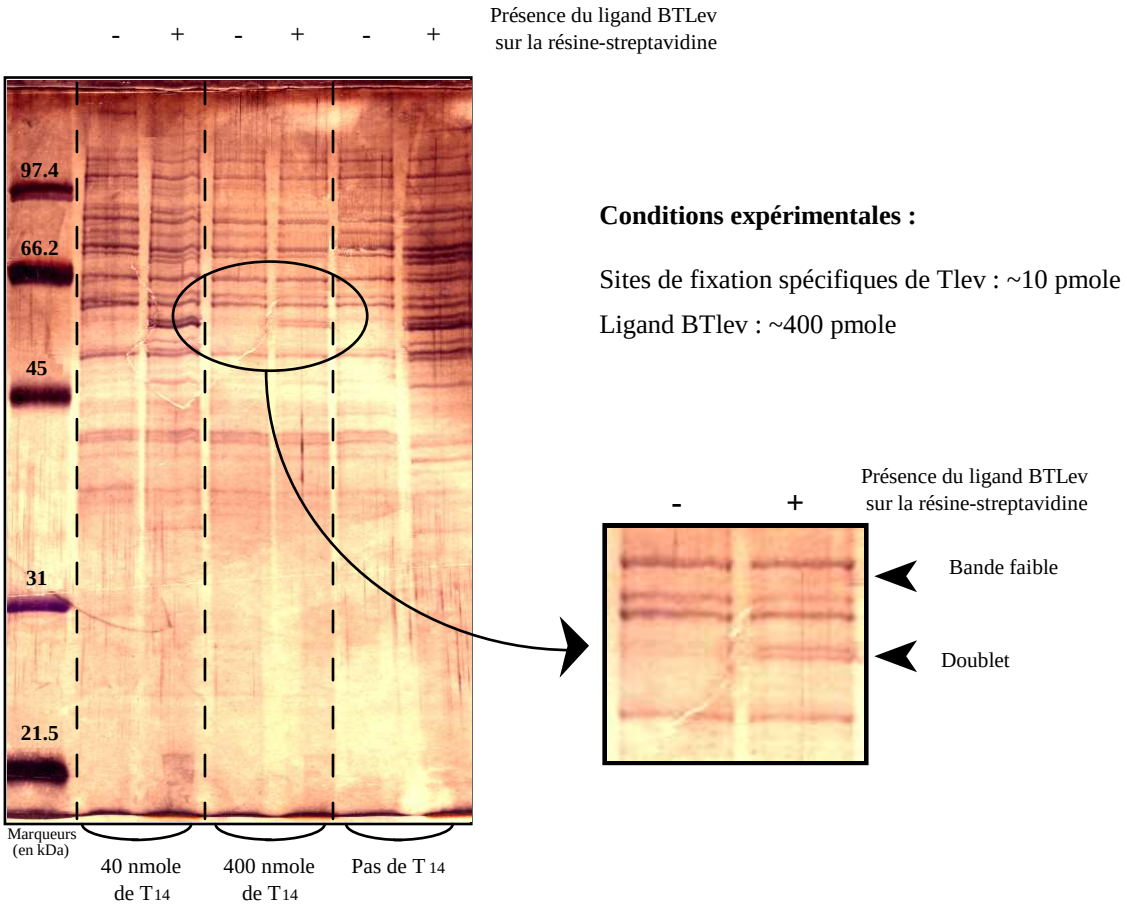


Figure 4.6. : Migration électrophorétique dénaturante réalisée à partir de résine-streptavidine porteuse ou non du ligand BTlev, en fonction d'un excès de T_{14} par rapport à BTlev. La coloration des protéines est obtenue par double coloration (Bleu de Coomassie / nitrate d'argent).

2-2- Séquençage des protéines identifiées

Lors de l'étude des protéines du doublet, nous identifions trois protéines Yhr216p, Ylr432p et Yml056p, s'associant à la séquence Tlev. Ces trois protéines ont des séquences extrêmement proches (voir figure 4.7. et alignements de séquences en annexes). Aussi, afin de lever les incertitudes inhérentes à ces similitudes, nous avons dû procéder selon trois voies d'identification distinctes :

- la spectrométrie de masse MALDI
- la spectrométrie de masse en tandem MSMS
- le séquençage N-terminal par dégradation d'Edman

```
             1          11         21         31         41         51
             |          |          |          |          |          |
YHR216W    1 MAAIRdykta ldftkslprp dglsvqelmd skirggltyn dflilpglvd fassevslqt    60
YLR432W    1 maavrDYKTA LEFAKSLPRL DGLSVQELMD SKtrggltyn dflvlpglvd fpssevslqt    60
YML056C    0 MSAAPLDYKKA LEHLKTYSSK DGLSVQELMD STTRGGLTYN DFLVLPGLVN FPSSAVSLQT    60

YHR216W   61 kltrnitlni plvsspmdtv tesematfma llggigfihh nctpedqadm vrrVKNYENG   120
YLR432W   61 kltrnitlnt pfvsspmdtv tesemaifma llggigfihh nctpedqadm vrrVKNYENG   120
YML056C   61 Kltkkitlnt pfvsspmdtv teadmaiyma llggigfihh nctpkEQASM VKKVKMFENG   120

YHR216W  121 FINNPIVISP TTTVGEAKSM KEKYGFAGFP VTTDGKRNAK LVGVITSRDI QFVEDNSLLV   180
YLR432W  121 FINNPIVISP TTTVGEAKsm kERFGFSGFP VTEDGKRNGK LMGIVTSRDI QFVEDNSLLV   180
YML056C  121 FINSPIVISP TTTVGEVKvm krKFGFSGFP VTEDGKCPGK LVGLVTSRDI QFLEDDSLVV   180

YHR216W  181 QDVMTKNPVT GAQGITLSEG NEILKKikkg rllvvdekgn lvsmlsrtdl mkNQNYPLAS   240
YLR432W  181 QDVMTKNPVT GAQGITLSEG NEILKKikkg kLLIVDDNGN LVSMLSRTDL MKNQNYPLAS   240
YML056C  181 SEVMTKnpvt gikgitlkeg neilkQTKKG KLLIVDDNGN LVSMLSRADL MKNQNYPLAS   240

YHR216W  241 Ksantkqllc gasigtmdad kerLRLLVKa gldvvildss qgnsifelnm lkwvkesfpg   300
YLR432W  241 KSATTKQLLC GAAIGTIDAD KERLRllvea gldvvildss qgnsifqlnm ikWIKETFPD   300
YML056C  241 KSATTKQLLC GAAIGTIEAD KERlrllvea gldvvildss qgnsvfqlnm ikWIKETFPD   300

YHR216W  301 leviagnvvt rEQAANLIAA GADGLRigmg tgsicitqev macgrpqgta vynvcefanq   360
YLR432W  301 LEIIAGNVAT REQAANLIAA GADGLRigmg sgsicitqev macgrpqgta vynvcefanq   360
YML056C  301 LEIIAGNVAT REQAANLIAA GADGLRigmg sgsicitqev macgrpqgta vynvcqfanq   360

YHR216W  361 fgvpcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfyqdgk rlkAYRGMGS   420
YLR432W  361 fgipcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfyqdgk rlkAYRGMGS   420
YML056C  361 fgvpcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfykdgk rlkayrGMGS   420

YHR216W  421 IDAMQKTGTK GNASTSRYFS ESDSVLVAQG VSGAVVDKGS IKKfipylyn glqhscqdig   480
YLR432W  421 IDAMQKTGTK GNASTSRYFS ESDSVLVAQG VSGAVVDKGS IKKfipylyn glqhscqdig   480
YML056C  421 IDAMQKTGNK GNASTSRYFS ESDSVLVAQG VSGAVVDKGS IKKfipylyn glqhscqdig   480

YHR216W  481 crSLTLLKNN VQRgkVRFEF RTASAQLEGG VHNLHSYEKR LHN
YLR432W  481 ykSLTLLKEN VQSGKVRFEF RTASAQLEGG VHNLHSYEKR LHN
YML056C  481 cesltslken vqngevrFEF RTASAQLEGG VHNLHSYEKR LYN
```

Figure 4.7. : Identifications obtenues. En rouge, les peptides identifiés par spectrométrie de masse MALDI; en bleu, les peptides identifiés par séquençage N-terminal par dégradation d'Edman ; en encadré, le peptide déterminé par spectrométrie de masse en tandem, spécifique de Ylr432p

Pour nous assurer que nous avions bien identifié des protéines spécifiques de Tlev, nous avons procédé à une expérience parallèle où le ligand d'ADN était la séquence T_{26} biotinylée à l'extrémité 5'. Nous ne retrouvons aucune bande correspondant aux protéines Yhr216p, Ylr432p et Yml056p

dans cette expérience parallèle. Ce résultat conforte nos identifications et relativise la compétition observée entre T_{26} et Tlev.

2-2-1- Spectroscopie MALDI

Les résidus peptidiques de la trypsine ont servi pour la calibration interne des expériences de spectrométrie de masse MALDI (figure 4.8. pour exemple). La plupart des peptides identifiés sont communs à deux ou trois des protéines (cf. alignements de séquences en annexes). Cependant la présence d'au moins trois peptides réellement spécifiques de chaque protéine a permis une identification sûre. En spectrométrie de masse, l'intensité d'un pic de spectrométrie de masse n'est pas directement reliée à la concentration du peptide correspondant dans l'échantillon. En effet, l'efficacité de la désorption de la matrice de DHB est différente selon les peptides. De plus, certains peptides peuvent être perdus par adsorption sur les surfaces des récipients de préparation, etc.

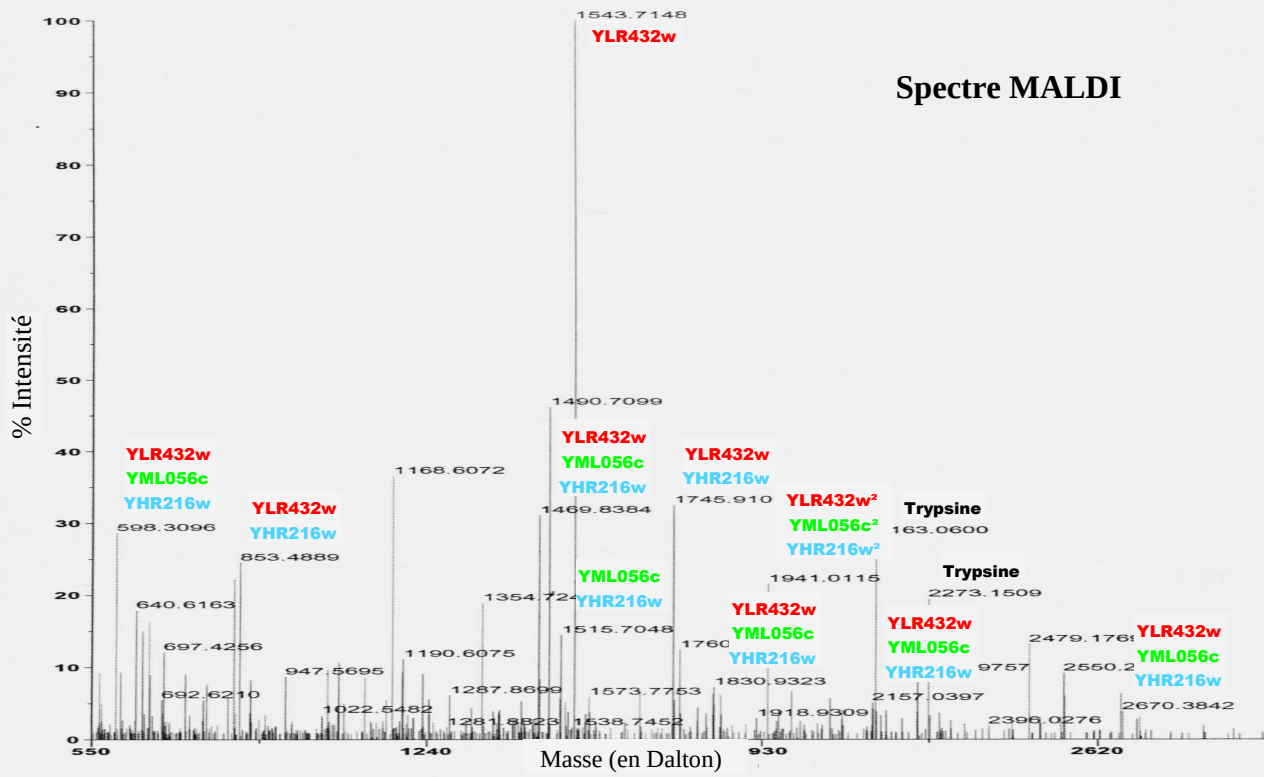


Figure 4.8. : Exemple de spectre MALDI-TOF. Seuls les pics les plus intenses correspondants à des peptides des protéines Yhr216p, Yml056p et Ylr432p ou de la trypsine sont indexés. Nous avons identifié pour cette expérience, 24 peptides pour la protéine Ylr432p, 19 peptides pour la protéine Yml056p et 18 peptides pour la protéine Yhr216p, représentant des taux de recouvrement des séquences de 45.3%, 40.1% et 36.5% respectivement. ² Le pic à 1941.0115 Da correspond à deux peptides communs aux trois protéines (masses de ces peptides : 1940.946 Da et 1941.029 Da).

D'autre part, les spectres de la bande inférieure du doublet de protéines (figure 4.6.), ne contiennent aucun peptide provenant des parties N-terminales des protéines Ylr432p, Yml056p et Yhr216p. Ceci suggère que cette bande correspond à une dégradation N-terminale.

Gènes	Acides aminés identifiés	Taux de recouvrement	Nombre de fragments peptidiques identifiés
YHR216W	217 / 523	41.5 %	23
YLR432W	296 / 523	56.6 %	35
YML056C	304 / 524	58.0 %	36

Tableau 4.2. : Résultats obtenus lors des identifications MALDI réalisées avec la séquence BTlev

2-3-2- Spectroscopie en tandem ESI Q-TOF MSMS

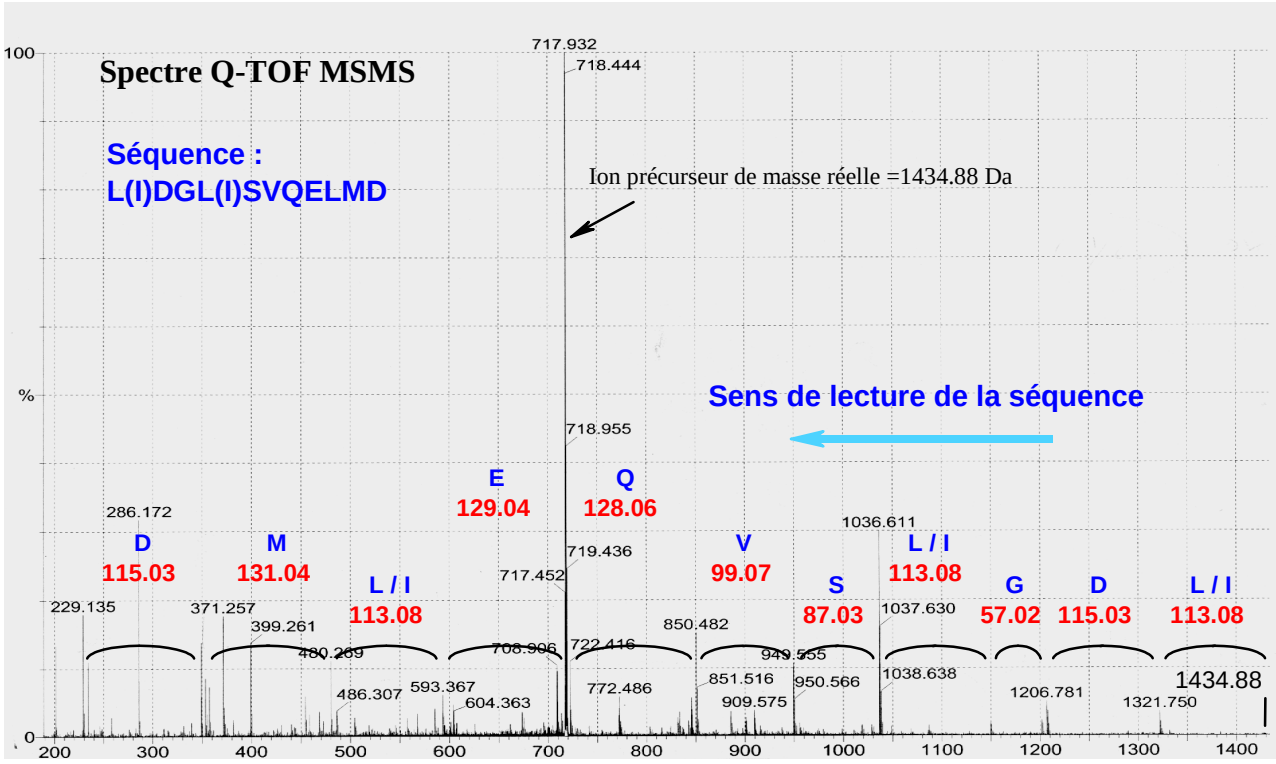


Figure 4.9. : Exemple de spectre obtenu par spectroscopie de masse en tandem MS/MS. Le peptide précurseur sélectionné a une masse de 1434.8 Da. La séquence déterminée ici est spécifique d'un peptide de la protéine Ylr432p.

Après avoir sélectionné un peptide (choisi par analyse des résultats obtenus en spectrométrie de masse MALDI) pouvant correspondre spécifiquement à une seule des trois protéines, nous avons procédé à des expériences de spectrométrie de masse en tandem visant à confirmer l'identification et la présence des trois protéines dans les bandes de protéines sélectionnées. Ces expériences ont permis de formellement identifier un peptide (figure 4.9.) spécifique de la protéine Ylr432p dont la masse réelle de l'ion précurseur est : $718.44 \times 2 - 2 = 1434.88$ Da. Ces expériences ont confirmé plusieurs autres peptides communs à deux des trois protéines identifiées.

2-3-3-Séquençage par dégradation d'Edman

Aucun séquençage N-terminal n'a pu être pratiqué directement à partir de protéines séparées préalablement lors d'une migration électrophorétique dénaturante, transférées sur membrane de PVDF, et identifiées après coloration au noir amido. La cause probable de ce blocage est l'acétylation prédite dans l'étude réalisée par Huang (Huang et *al.*, 1987).

Cependant, après digestion enzymatique (endolysine-C) réalisée sur les protéines du doublet étudié par spectrométrie de masse, on obtient l'identification d'un peptide spécifique de la protéine Yml056p (MFENGFINSPIVISP) et d'un peptide commun aux protéines Yhr216p, Ylr432p et Yar073p (NPVTGAQGITLSEGN). Cette analyse confirme définitivement la présence de la protéine Yml056p dans les bandes de protéines analysées.

3- Reconnaissance du brin complémentaire riche en G

3-1- Caractérisation de l'association

Après incubation du brin complémentaire de la séquence Tlev, d[$(GGGTGT)_3GGG$] (appelé par la suite GTlev) avec les extraits de protéines de levure non purifiés, on procède à une migration électrophorétique en conditions non dénaturantes. On observe trois bandes retardées, nommées A, G et B (piste 2, figure 4.10.). La bande retardée G est au même niveau et présente une intensité comparable à la bande retardée (notée C) observée dans le cas d'une incubation avec comme sonde radioactive la séquence Tlev* (piste 6).

Lorsque l'on mélange les deux brins complémentaires marqués radioactivement, de façon à avoir un excès de brin Tlev*, on observe en piste 3 la formation du double brin, et l'excès de brin Tlev*. Après incubation avec les extraits de protéines, nous constatons l'apparition de trois nouvelles bandes retardées correspondant vraisemblablement à des protéines fixant le double brin d'ADN. La majeure partie du signal Tlev* est compléxée soit avec les protéines, soit avec le brin complémentaire GTlev*. La résolution du gel ne permet pas de résoudre les bandes C et G. Les quantités de brin GTlev* utilisées lors de ces expériences étant constantes, la séquence radioactive GTlev*, qui se répartissait totalement entre les bandes A, G et B (piste 2), se répartit maintenant entre les bandes A, db3, db, G, db2, B et dans l'association en double hélice avec le brin complémentaire Tlev*.

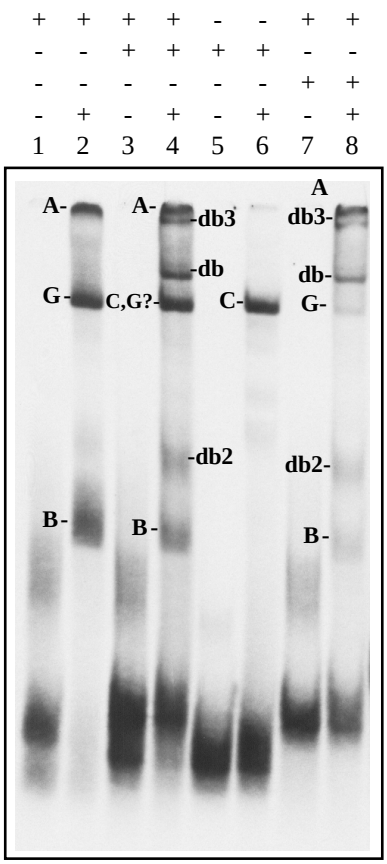


Figure 4.10. : Migration électrophorétique en conditions non dénaturantes réalisées après incubation des protéines de levure avec les fragments télomériques Tlev*, Tlev et GTlev*.

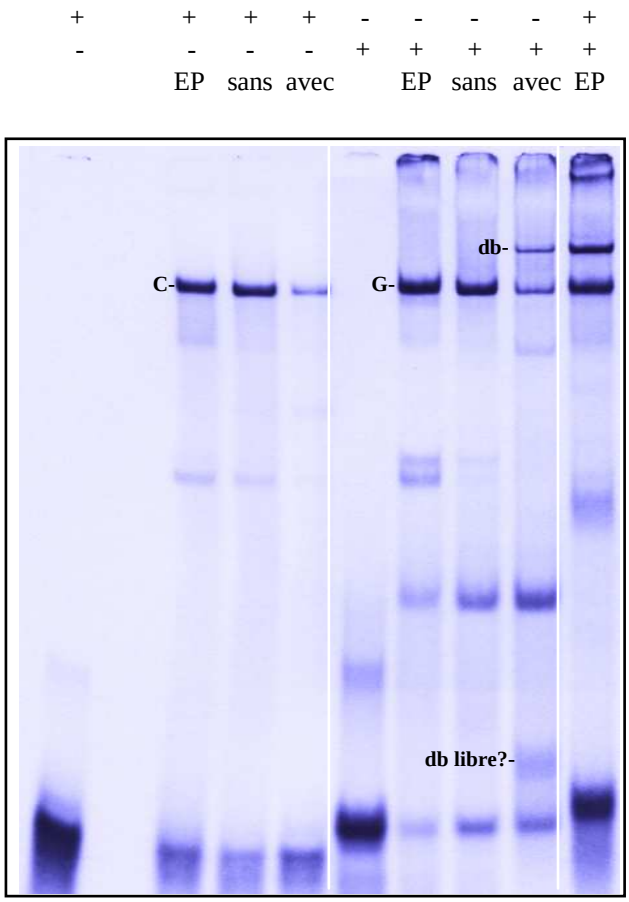
Brins Tlev* et Tlev : ~ 1pmol
Brin GTlev* : ~ 0.5 pmol

Pour les pistes 7 et 8, le brin Tlev est en quantité identique à celles de Tlev* utilisées dans les autres pistes et n'est pas marqué radioactivement. On confirme que tout le brin GTlev* est recruté dans la formation du double brin d'ADN. En piste 8, on note la présence des bandes db, db2 et db3, d'intensités plus faibles qu'en piste 4 (où les deux brins étaient radioactivement marqués), indiquant

que ces protéines sont spécifiques du double brin. On remarque aussi la présence des bandes retardées A, G et B, ayant dans le cas des bandes G et B une intensité nettement plus faible que celles observées en piste 2.

Toutes ces observations ne permettent pas de conclure que la bande G et la bande C correspondent à une seule et même protéine, car tout le brin GTlev* est associé en double brin d'ADN, pour lequel la protéine de la bande G ne semble pas avoir d'affinité. Mais on peut penser qu'il s'agit de la même protéine, saturée par la séquence Tlev (cf. piste 6) plus affine que le double brin ou GTlev*.

Pour vérifier que les bandes G et C étaient vraisemblablement dues à une même protéine, nous avons profité des propriétés des surnageants récupérés après incubation entre les résines greffées ou non avec le ligand BTlev. Dans un cas, nous obtenions un signal retard avec la séquence Tlev*, dans l'autre cas, un faible signal (figure 4.4).



EP : Extraits de protéines

Sans : surnageant récupéré après incubation avec la résine sans ligand BTlev

Avec : surnageant récupéré après incubation avec la résine avec ligand BTlev

Figure 4.11. : Migrations électrophorétiques natives obtenues à partir des surnageants de l'expérience d'identification par affinité avec ou sans le ligand BTlev, pour les sondes radioactives Tlev* et GTlev*.

Lors de l'expérience présentée en figure 4.11., on note que l'intensité de la bande G décroît comme celle de la bande C, pour le surnageant correspondant aux protéines ayant incubé en présence du ligand BTlev. Cela semble confirmer que les protéines responsables de la bande retardée C obtenue avec la séquence Tlev* ont une mobilité électrophorétique identique à la bande retardée G.

Par ailleurs, on visualise à nouveau la bande référencée db dans la figure 4.10. On peut supposer que les protéines de la bande C sont initialement partiellement liées à de l'ADN riche en cytosines présents dans les extraits protéiques. Lors de l'incubation avec la résine porteuse du ligand BTlev, la compétition entre la séquence BTlev et l'ADN riche en C libère celui-ci. Présent dans le surnageant, il peut alors former avec GTlev* un double brin, dont la bande diffuse notée 'db libre?' pourrait correspondre à la forme libre, et s'associer avec les protéines spécifiques du double brin.

Sans réaliser une étude aussi complète que celle entreprise pour déterminer la spécificité de l'association de Tlev à des protéines, nous avons vérifié ce qui se passait avec les compétiteurs T_{14} et ds10 précédemment définis. Nous n'observons aucune compétition avec ces séquences (rapport compétiteur / GTlev* ~ 10000).

Par contre, l'affinité des protéines pour les séquences GTlev et d[A(GGGTTA)$_3$GGG] semble être la même.

3-2- Identification par affinité

Il était intéressant de vérifier l'identification par affinité des protéines Yhr216p, Ylr432p et Yml056p avec le ligand BGTlev (=d[BTTTT(GGGTGT)$_3$GGG]). Pour cela, nous avons utilisé le même mode opératoire que dans le cas de l'identification avec la séquence BTlev.

Comme dans le cas de l'expérience d'identification différentielle réalisée avec le ligand BTlev, nous procédons à une première incubation entre les protéines et de la résine-streptavidine non porteuse de ligand. Puis nous avons procédé à une incubation entre les protéines et les résines greffées ou non avec le ligand BGTlev, en présence d'un excès de T_{14}.

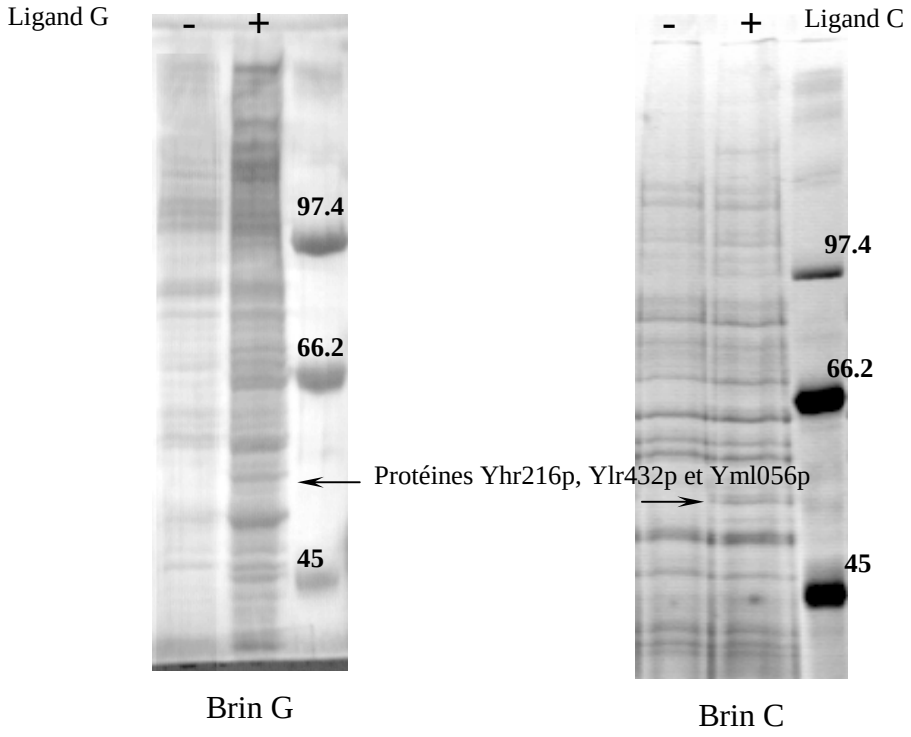


Figure 4.12. : Comparaison entre les migrations électrophorétiques dénaturantes réalisées à partir de résine-streptavidine porteuse ou non du ligand BGTlev (gel de résolution à 8% d'acrylamide) ou du ligand BTlev (gel de résolution à 10% d'acrylamide). Les photographies présentées correspondent aux gels colorés selon le protocole de coloration au sulfate de zinc avant découpage de bandes de protéines pour analyse par spectrométrie de masse.

Nous différencions plusieurs bandes de protéines (migration électrophorétique dénaturante colorée au sulfate de zinc de la figure 4.12.) d'intensités inégales, dont une bande intense de masse ~ 50 kDa. A nouveau, nous obtenons une identification par spectrométrie de masse MALDI des protéines Yhr216p, Ylr432p et Yml056p. Les résultats obtenus (figure 4.13.) sont présentés et comparés aux résultats obtenus précédemment dans le tableau 4.2. Comme dans le cas des identifications réalisées avec le brin BTlev, certains peptides identifiés sont spécifiques d'une seule des trois protéines, ce qui nous assure de la présence de ces trois protéines lors de l'association avec le ligand BGTlev.

Aucune des autres bandes ne permet d'obtenir d'identification formelle de protéines (principalement par manque de matériel peptidique à analyser). Nous n'avons pas noté d'autres protéines communes entre les expériences réalisées avec l'ADN télomérique riche en cytosines et l'ADN télomérique riche en guanines, ni de protéines déjà connues pour se lier au brin G des télomères de levure.

```
1           11          21          31          41          51
              |           |           |           |           |           |
YHR216W   1  maairdykTA LDFTKslprp dglsvqelmd skirggltyn dflilpglvd fassevslqt    60
YLR432W   1  maavrdykta lefakslprl dglsvqelmd sktrggltyn dflvlpglvd fpssevslqt    60
YML056C   0 msaapldykka lehlktyssk dglsvqelmd sttrggltyn dflvlpglvn fpssavslqt    60

YHR216W   61 kltrnitlni plvsspmdtv tesematfma llggigfihh nctpedqadm vrrvknyeng   120
YLR432W   61 kltrnitlnt pfvsspmdtv tesemaifma llggigfihh nctpedqadm vrrvknyeng   120
YML056C   61 kltkkitlnt pfvsspmdtv teadmaiyma llggigfihh nctpkeqasm vkkvkMFENG   120

YHR216W  121 finnpivisp tttvgeaksm kekygfagfp vttdgkrnak LVGVITSRdi qfvednsllv   180
YLR432W  121 finnpivisp tttvgeaksm kerfgfsgfp vtedgkrNGK LMGIVTSRdi qfvednsllv   180
YML056C  121 FINSPIVISP TTTVGEVKvm krkfgfsgfp vtedgkcpgk LVGLVTSRdi qfleddslvv   180

YHR216W  181 qdvmtkNPVT GAQGITLSEG NEILKkikkg rllvvdekgn lvsmlsrtdl mknqnyplas   240
YLR432W  181 qdvmtkNPVT GAQGITLSEG NEILKkikkg kllivddngn lvsmlsrtdl mknqnyplas   240
YML056C  181 sevmtknpvt gikgitlkeg neilkqtkkg kllivddngn lvsmlsrADL MKnqnyplas   240

YHR216W  241 ksantkqllc gasigtmdad kerlrllvka gldvvildss qgnsifelnm lkwvkesfpg   300
YLR432W  241 kSATTKQLLC GAAIGTIDAD KERlrllvea gldvvildss qgnsifqlnm ikwikETFPD   300
YML056C  241 ksattkqllc gaaigtiead kerlrllvea gldvvildss qgnsvfqlnm ikwikETFPD   300

YHR216W 301 leviagnvvt rEQAANLIAA GADGLRigmg tgsicitqev macgrpqgta vynvcefanq   360
YLR432W 301 LEIIAGNVAT REQAANLIAA GADGLRigmg sgsicitqev macgrpqgta vynvcefanq   360
YML056C 301 LEIIAGNVAT REQAANLIAA GADGLRigmg sgsicitqev macgrpqgta vynvcqfanq   360

YHR216W 361 fgvpcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfyqdgk rlkayrGMGS   420
YLR432W 361 fgipcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfyqdgk rlkayrGMGS   420
YML056C 361 fgvpcmadgg vqnighitka lalgsstvmm ggmlagttes pgeyfykdgk rlkayrGMGI   420

YHR216W 421 IDAMQKtgtk gnastsrYFS ESDSVLVAQG VSGAVVDKgs ikkFIPYLYN GLQHSCQDIG   480
YLR432W 421 IDAMQKtgtk gnastsrYFS ESDSVLVAQG VSGAVVDKgs ikkFIPYLYN GLQHSCQDIG   480
YML056C 421 IDAMQKtgnk gnastsrYFS ESDSVLVAQG VSGAVVDKgs ikkfipylyn glqhscqdig   480

YHR216W 481 CRsltllknn vqrgkvrFEF RTASAQLEGG VHNLHSYEKr lhn
YLR432W 481 YKsltllken vqsgkvrFEF RTASAQLEGG VHNLHSYEKr lhn
YML056C 481 cesltslken vqngevrFEF RTASAQLEGG VHNLHSYEKr lyn
```

Figure 4.13. : Résultats obtenus par identification MALDI avec le ligand BGTlev.

Gènes	Acides aminés identifiés	Taux de recouvrement
YHR216W	121 (217) / 523	23.1 % (41.5 %)
YLR432W	154 (296) / 523	29.6 % (56.6 %)
YML056C	120 (304) / 524	22.9 % (58.0 %)

Tableau 4.3. : Résultats obtenus lors des identifications MALDI réalisées avec la séquence BGTlev et comparaisons avec ceux obtenus avec le ligand BTlev entre parenthèses.

3-3- Implication biologique

Les trois protéines identifiées pourraient fixer séparément le simple brin télomérique riche en C et le simple brin riche en G. Or sur le modèle des fragments télomériques humains, ces deux séquences sont capables de former *in vitro*, des structures canoniques particulières. (le motif i (Phan, 1999 et Phan et *al.*, 2000) pour la séquence d[(CCCTAA)$_3$CCCT] et un quadruplexe de guanines (Wang et Patel, 1993) pour la séquence d[A(GGGATT)$_3$GGG]). Rappelons que les conditions dans lesquelles a lieu l'association entre les protéines et le brin GTlev sont favorables à la formation d'un quadruplexe de guanines et que l'association avec le brin Tlev reste possible à pH 6.

La compétition existant entre la formation d'un duplexe d'une part, et celles d'un motif i et d'un quadruplexe de guanines, d'autre part, pourrait être modulée par la présence de protéines s'associant à l'une ou l'autre de ces structures non canoniques. Dans ce contexte, une protéine pouvant s'associer à chacun des brins de l'ADN télomérique et pouvant éventuellement stabiliser le motif i et le quadruplexe de guanines, place la compréhension des mécanismes de formation éventuelle et concertée de ces structures, sous un nouveau jour (figure 4.14.).

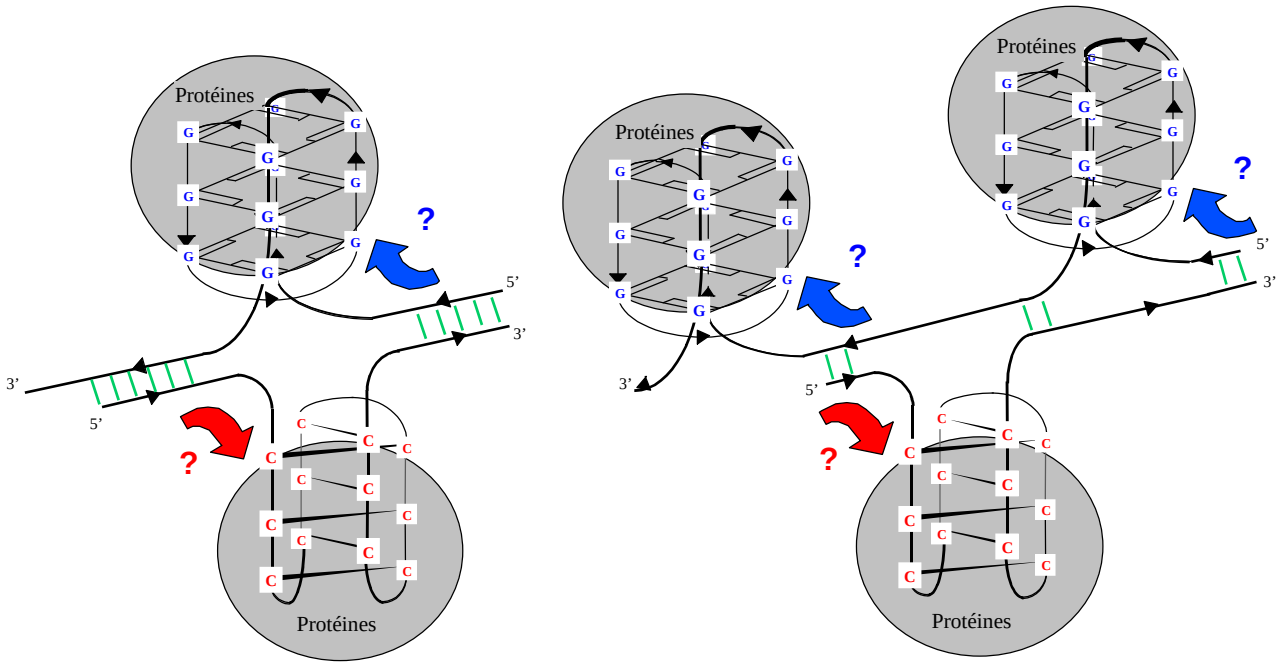


Figure 4.14. : Différentes possibilités de coexistence du motif i et de quadruplexe de guanines en présence de protéines stabilisant ces deux structures. L'ouverture du duplexe pourrait permettre la formation concertée de ces structures (schéma de gauche).

La formation de ces deux structures non canoniques nécessitant l'ouverture du double brin d'ADN, nous avons essayé de savoir si ces protéines reconnaîtraient également le duplexe d'ADN télomérique, en variant par exemple, l'ordre et les temps d'incubation des deux brins d'ADN Tlev*

et GTlev avec l'extrait protéique. Les difficultés d'interprétation inhérentes au mode de visualisation utilisé ne nous permettent pas de conclure sur la possible reconnaissance du double brin d'ADN télomérique par les protéines Yhr216p, Ylr432p et Yml056p. Un moyen de vérifier si l'association est possible avec un double brin d'ADN, serait d'utiliser, par exemple, la séquence GTlev et la séquence d(CCCACACCC) (dont on a vu qu'elle ne s'associait aux protéines qu'à concentration élevée), pouvant former partiellement un double brin.

4- Aspects biologiques des IMPDH

4-1- Caractéristiques des protéines identifiées

4-1-1- La protéine Yhr216p

Le gène *YHR216W* (ou IMD2) est situé à l'extrémité du chromosome VIII (figure 4.15.). Cette position le place sous une possible régulation par le phénomène de Telomere Position Effect. Par homologie de séquence avec *Candida albicans* (Kohler et *al.*, 1997), cette protéine aurait un rôle dans le métabolisme des purines. D'une masse calculée de 56 kDa, sa mobilité électrophorétique en conditions dénaturantes correspond à une masse de 54 kDa (Perrot et *al.*, 1999).

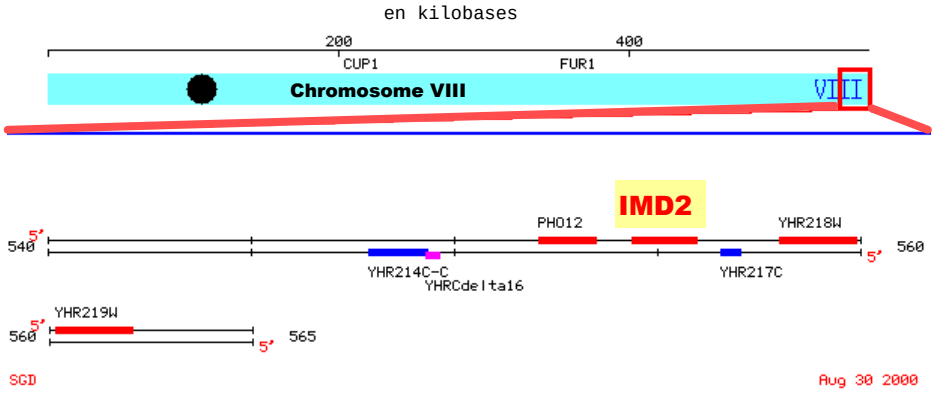


Figure 4.15. : Représentation de la zone du chromosome VIII où se situe le gène *YHR216W*.

Cette protéine est en interaction avec la protéine Lhp1p (protéine s'associant et stabilisant le pré t-RNA) (Rosenblum et *al.*, 1998).

Il existe à l'extrémité du chromosome I une région de 25 kilobases que l'on retrouve quasiment à l'identique sur le chromosome VIII (Barton et *al.*, 1997). On trouve ainsi un gène, *YAR073W* (ou IMD1, figure 4.15), quasiment identique à l'extrémité N-terminale du gène *YHR216W* (voir les alignements de séquences en annexes).

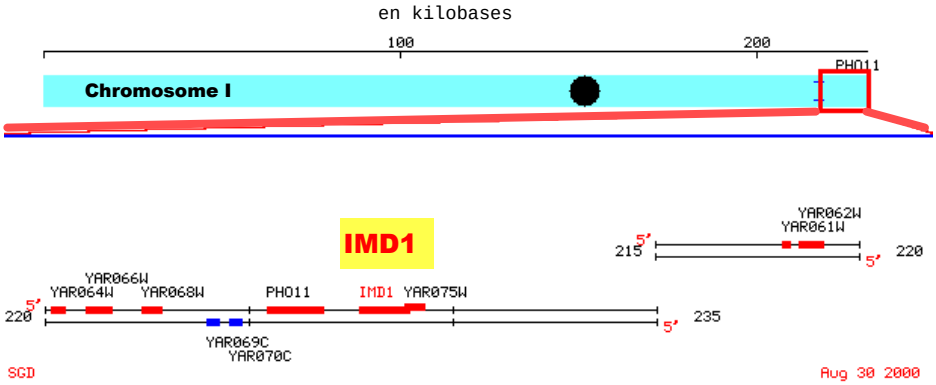


Figure 4.16. : Représentation de la zone du chromosome I où se situe le gène *YAR073W*.

Les gènes situés sur le chromosome I sont faiblement transcrits par rapport aux gènes existants sur le chromosome VIII. Barton et *al.* suggèrent qu'ils seraient réprimés par un effet naturel de TPE et qu'ils auraient un rôle structural pour les chromosomes.

La protéine Yar073p aurait une masse de 44.4 kDa, et pourrait être confondue avec les protéines Yhr216p, Ylr432p et Yml056p, dégradées à l'extrémité N-terminale de la bande inférieure (masse de ~ 47 kDa) du doublet étudié lors des expériences réalisées avec le ligand BTlev. L'identité de séquence avec la protéine Yhr216p est telle que nous n'avons aucun moyen de vérifier cette hypothèse. Les identifications de peptides pouvant correspondre à Yar073p sont reprises dans la figure 4.17. et représentent 25.6 % de recouvrement de séquence.

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  1 maairdykta ldltkslprp dglsvqelmd skirgglayn dflilpglvd fassevslqt    60
 61 kltrnitlni plvsspmdtv tesematfma lldgigfihh nctpedqadm vrrVKNYENG   120
121 FINNPIVISP TTTVGEAKsm kekygfagfp vtadgkrnak lvgaitsrDI QFVEDNSLLV   180
181 QDVMTKNPVT GAQGITLSEG NEILKKikkg rllvvdekgn lvsmlsrtdl mknqkyplas   240
241 ksantkqllw gasigtmdad kerLRLLVKa gldvvildss qgnsifqlnm ikWIKETFPD   300
301 LEIIAGNVVT KEQAANLIAA GADGLRigmg tgsicitqkv macgrpqgta vynvcefanq   360
361 fgvpcmadgg vqkhwsyyyq sfgswffycy dgwyvgryyr itr
```

Figure 4.17. : Reprise des résultats obtenus avec le ligand Tlev pour la protéine Yar073p.
On peut noter l'absence de peptides dans la région N-terminale.

Cependant, dans le cadre des expériences réalisées à partir du ligand d'ADN BGTlev, en l'absence de bande pouvant correspondre aux protéines dégradées, l'écart de masse semble trop important pour que l'on puisse prétendre que Yar073p fixe le brin G.

4-1-2- La protéine Ylr432p

Le gène *YLR432W* (ou IMD3) est situé à l'extrémité du chromosome XII (figure 4.18.), non loin du gène *SIR3* (Silent Information Regulator (Ivy et *al.*, 1986)). Il est suffisamment près du télomère pour être sensible à une régulation par le Telomere Position Effect.

La protéine est en interaction avec les protéines Ydr469p (rôle et localisation inconnus), Gcn3p (synthèse des protéines), Gdh2p (glutamate déshydrogénase (NAD)) et Taf25p (transcription de Pol II (Poon et *al.*, 1995 ; Klebanow et *al.*, 1996), localisée dans le noyau) (Uetz et *al.*, 2000). La protéine Ylr432p serait donc partiellement présente dans le noyau.

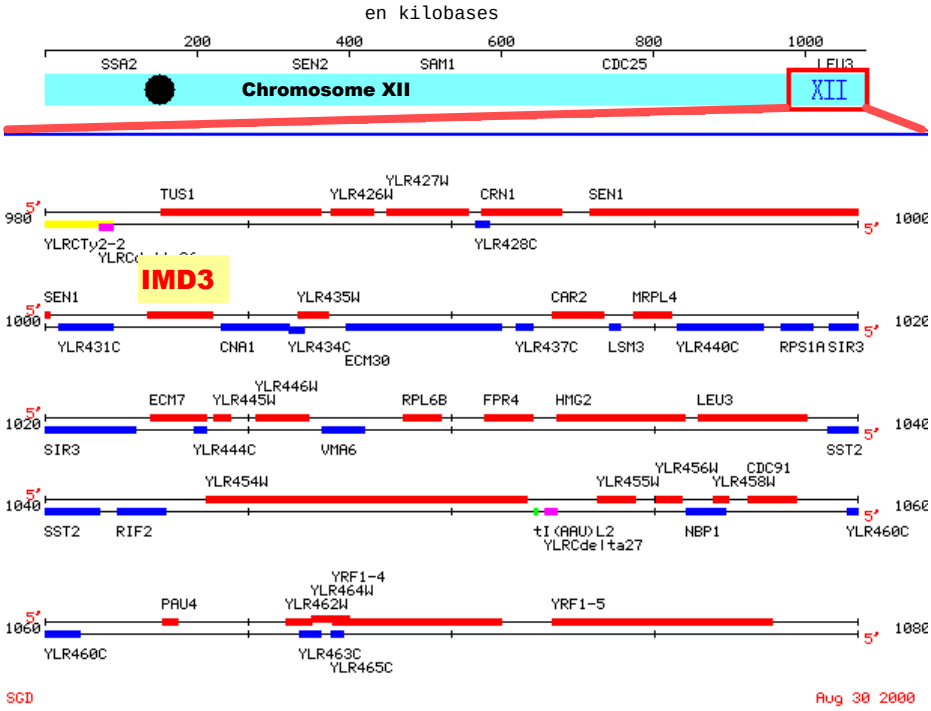


Figure 4.18. : Représentation de la zone du chromosome XII où se situe le gène *YLR432W*.

4-1-3- La protéine Yml056p

Le gène *YML056C* (ou IMD4) est situé sur le chromosome XIII (figure 4.19.). Des trois protéines identifiées, Yml056p est celle qui se différencie le plus (84% et 89% d'identité de séquence par rapport à Yhr216p et Ylr432p respectivement).

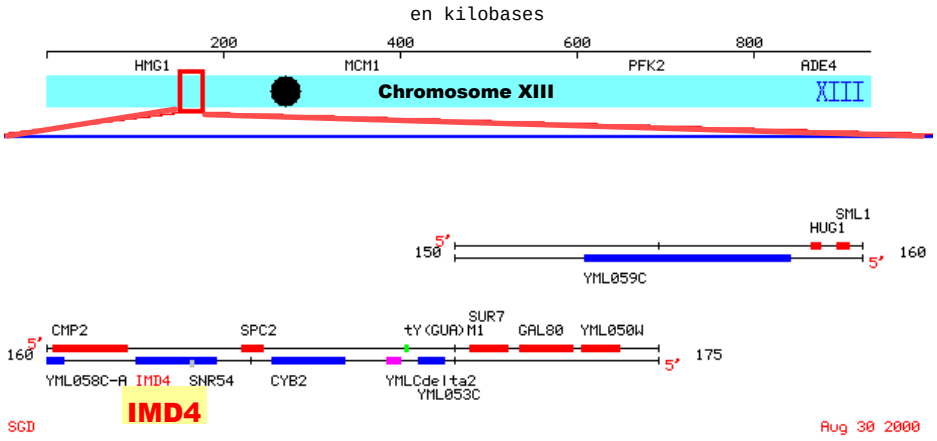


Figure 4.19. : Représentation de la zone du chromosome XIII où se situe le gène *YML056C*.

4-1-4- Conclusions

Du point de vue de leurs séquences respectives, les trois protéines Yhr216p, Ylr432p et Yml056p sont très proches. La viabilité des mutants *yar073w yhr216w* (Barton et *al.*, 1997), *ylr432w* et *yml056c*, montre qu'elles peuvent avoir une (ou plusieurs) fonctions biologiques communes. Cependant il existe des interactions spécifiques avec d'autres protéines, propres à chacune des trois, permettant de les différencier.

Il existe sur le chromosome I un gène, *YAR073W*, qui est la fidèle copie d'une partie du gène *YHR216W*. Sa transcription pourrait être régulée par un effet naturel de TPE.

Les 4 gènes *YAR073W*, *YHR216W*, *YLR432W* et *YML056C* sont situés sur les chromosomes I, VIII, XII et XIII. Dans le séquençage disponible des chromosomes de levure *S. cerevisiae* (*Blasts alignments*), il s'agit des 4 chromosomes dont on retrouve la séquence Tlev inscrite dans les télomères et dont, par ailleurs, les régions télomériques et subtélomériques sont assez proches de

Tlev alors que l'ADN télomérique des autres chromosomes apparaissent singulièrement différents. Coïncidence fortuite ou possible particularisme structurale de ces chromosomes ? Nous n'avons à ce jour aucun élément nous permettant d'en parler.

	Chromosome	Nombre d'acides aminés	pI	Masse moléculaire	% iden[a]	% sim[a]	Interaction protéine-protéine
YHR216W	VIII	523 aa	8.79	56534 Da	93%	97%	Lhp1p
YLR432W	XII	523 aa	7.31	56452 Da	-	-	Taf25p [b,c], Gdh2p[b], Ydr463p[b], Gnc3p[b]
YML056C	XIII	524 aa	8.46	56396 Da	89%	95%	
IMPDH1[e]	7q31.3-q32	506 aa			61%	78%	
IMPDH2[f]	3p21.2	505 aa			62%	78%	

Tableau 4.4. : Tableau résumant les principales caractéristiques des protéines 5'-inosine monophosphate déshydrogénase de *Saccharomyces cerevisiae* et de leurs homologues chez l'homme [a] Comparaison par rapport à *YLR432W*, [b] Uetz et *al.*, 2000 ; [c] ARN polymérase et localisation nucléaire, [d] Rosenblum et *al.*, 1998 ; [e,f] 5'-inosine monophosphate déshydrogénase humaines

4-2- Rôle biologique

Les protéines Yhr216p, Ylr432p et Yml056p sont les homologues, chez *Saccharomyces cerevisiae*, de protéines qui catalysent la conversion de l'inosine monophosphate (IMP) en xanthosine 5' monophosphate (XMP) avec réduction de la nicotinamide adénine dinucléotide (NAD) en NADH (figure 4.20.). Ces protéines, appelées inosine monophosphate déshydrogénase (IMPDH), interviennent dans le métabolisme des guanines.

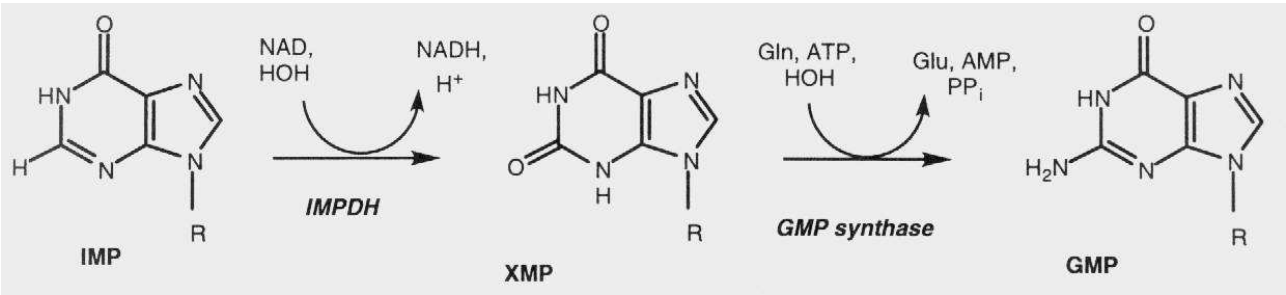


Figure 4.20. : Schéma des étapes menant à la synthèse de la guanosine monophosphate (GMP).

Elles font depuis longtemps l'objet d'études chez différents organismes et présentent d'importantes homologies de séquences entre espèces (voir en Annexes 2.2. les alignements de séquences). Elles

ont une grande importance dans différents phénomènes biologiques. La concentration en IMPDH dans la cellule décroît en présence de l'anti-oncogène p53 et les inhibiteurs de IMPDH induisent une différenciation dans les cellules tumorales (Sokoloski et *al.*, 1986 ; Sherley, 1991 ; Liu et *al.*, 1998). Aussi, ces protéines sont depuis longtemps considérées comme cibles de thérapies immunosuppressives et de chimiothérapies anticancéreuse et antivirale (Weber, 1983 ; Robins, 1982 ; Malinovski & Stollar, 1981).

A notre connaissance, aucune étude n'a été réalisée sur les protéines IMPDH de *Saccharomyces cerevisiae*, malgré l'importante homologie de séquence entre espèces.

Plusieurs structures, obtenues par cristallographie, dont celle de l'IMPDH de type II humaine, sont aujourd'hui disponibles. Chacune de ces structures cristallise sous forme d'un tétramère homogène offrant 4 sites de fixation à l'IMP et au NAD (figure 4.21.).

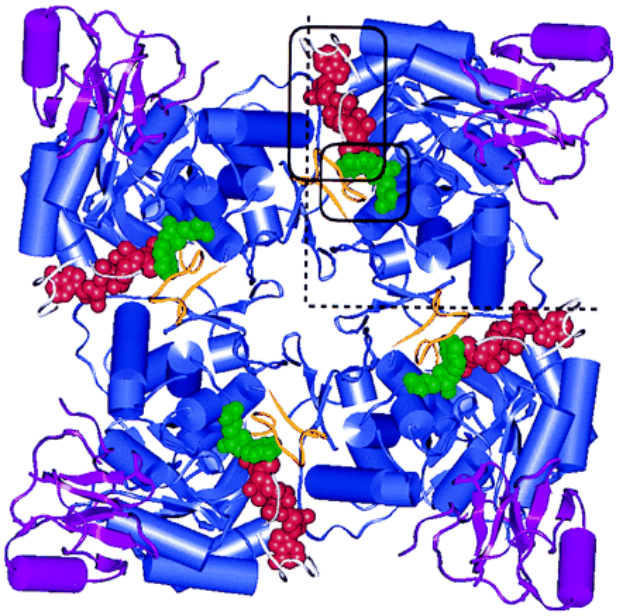


Figure 4.21. (tirée de Colby et *al.*, 1999) : IMPDH tétramérique humaine de type II et sites de fixation de SAD (Selenazole-4-carboxamide adénine dinucléotide), analogue de NAD (cercles rouges) et de 6-Cl-IMP (6-chloropurine riboside 5'-monophosphate), analogue de l'IMP (cercles verts). Le dinucléotide se fixe à l'interface monomère/monomère.

Nous avons incubé des extraits de protéines de levure et Tlev* en présence d'IMP et/ou de NAD. Pour cela, nous avons réalisé différentes incubations en présence de 20 mM IMP, 20 mM NAD en présence de 20 mM KCl (correspondant à une activité maximale de l'IMPDH humaine de type II (Xiang et *al.*, 1996)). La migration électrophorétique après incubation des protéines avec Tlev* (figure 4.22.) ou GTlev* (non montrés) est identique et ne montre aucune variation d'intensité du signal retardé.

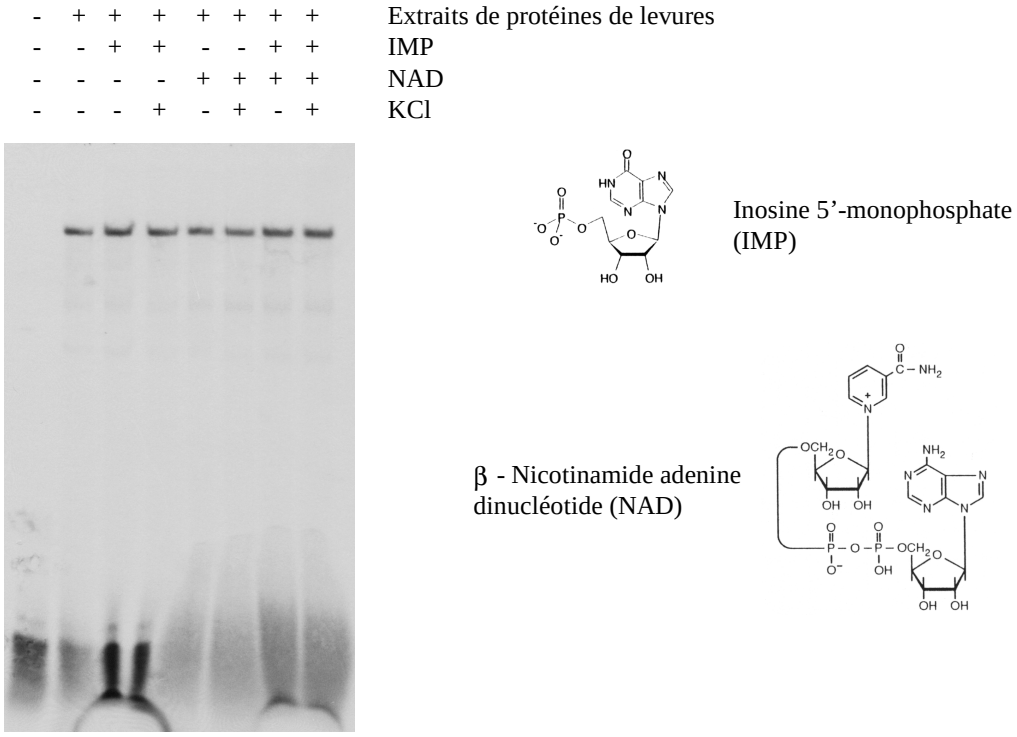


Figure 4.22. : Migrations électrophorétiques obtenues en présence d' IMP (20 mM), NAD (20 mM) et KCl (20 mM) et rappels des représentations chimiques de l' IMP et du NAD.

Plusieurs inhibiteurs de IMPDH sont bien connus et utilisés cliniquement, entre autre, dans les traitements anti-cancéreux et anti-viraux (Robins, 1982 ; Allison et *al.*, 1993a ; Allison et *al.*, 1993b ; DeClerq, 1993). Parmi ceux-ci, l'acide mycophénolique (ou MPA) est considéré comme un puissant inhibiteur spécifique des IMPDH humaines.

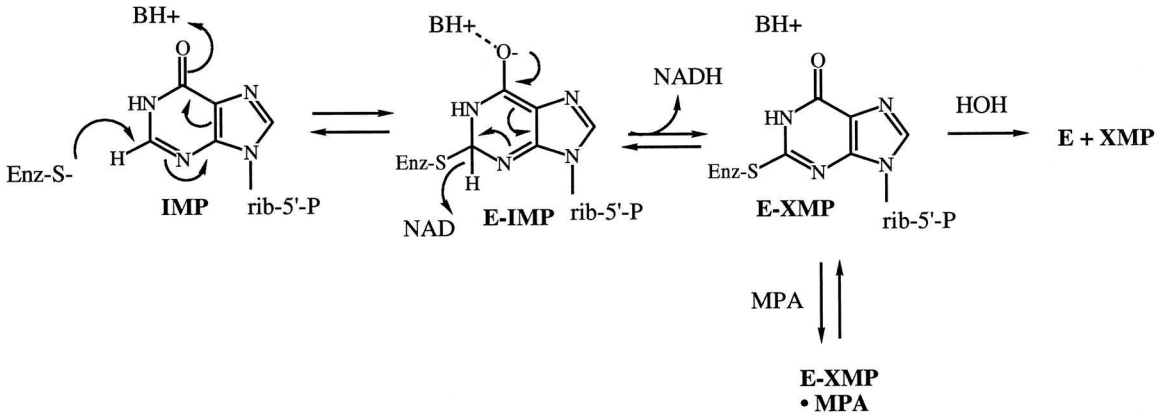


Figure 4.23. : Mécanisme réactionnel et inhibition de l'IMPDH par l'acide mycophénolique.

Le MPA piège l'intermédiaire E-XMP lors du passage de IMP à XMP (Farazi et *al.*, 1997) (figure 4.23.). L'inhibition de l'activité des IMPDH par l'acide mycophénolique varie grandement selon les

organismes. Par exemple, la constante de dissociation K_d de MPA, est de 22 nM pour les IMPDH humaines, 50 nM pour celle de *Bacillus subtilis*, et 20 µM pour celle d'*Escherichia coli*.

En ajoutant de l'acide mycophénolique lors de l'incubation entre les protéines et les séquences Tlev* (figure 4.24.) et GTlev* (résultats non montrés), nous n'avons pas constaté de changement dans l'intensité du signal retardé habituellement observé.

Nous avons donc montré que la présence d'un inhibiteur des protéines IMPDH pendant l'incubation entre les protéines de levure et les fragments télomériques Tlev* et GTlev*, n'entraînait pas de modification dans la migration électrophorétique du complexe ADN*-protéines.

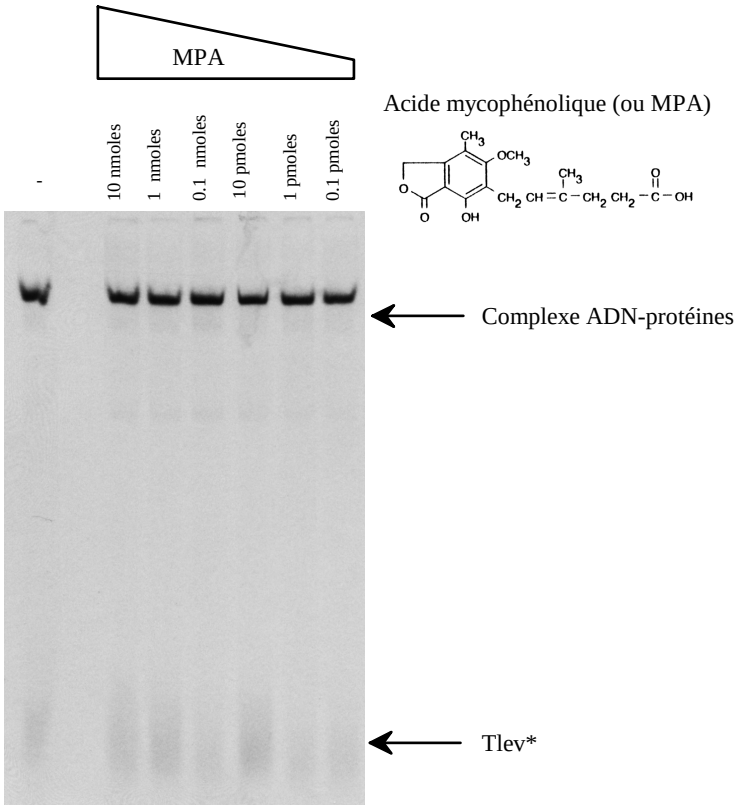


Figure 4.24. : Migration électrophorétique en fonction de la concentration en acide mycophénolique présente pendant l'incubation entre les protéines de levure et l'ADN Tlev* et représentation chimique de MPA.

4-3- Comparaison avec les protéines humaines

Le Professeur G.D. Markham (Institute for Cancer Research, Philadelphie, USA) nous a fourni de l'IMPDH humaine de type II purifiée. De plus, le Professeur B.S. Mitchell (Internal Medicine and Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, USA nous a fourni des IMPDH humaines de types I et II purifiées. Nous avons réalisé des migrations

électrophorétiques en conditions non dénaturantes avec ces protéines humaines, après incubation avec les fragments d'ADN télomériques humain d[(CCCTAA)$_3$CCCT]* ou de levure Tlev* (incubation en présence ou en absence de 20 mM KCl). Aucun retard n'a été observé lors de ces expériences, indiquant qu'aucune des deux homologues humaines des IMPDH de levure, ne s'associait avec une séquence d'ADN télomérique riche en cytosines.

Cependant, en l'absence de compétiteur non spécifique, nous observons qu'une protéine provenant d'extraits nucléaires de cellules HeLa, s'associe avec les séquences télomériques Tlev, 22hm et 27hm. Cette protéine a une mobilité électrophorétique identique à celles des protéines IMPDH des extraits de levure, mais migre à des positions différentes des protéines hnRNP K et ASF/SF2 (identifiées comme fixant la séquence 27hm (Lacroix et *al.*, 2000)). Nous remarquons encore que cette protéine humaine fixe également la séquence centromérique Mcent-u, comme les protéines IMPDH de levure.

Par contre, l'association ne se fait plus en présence du compétiteur non spécifique T_{26}. Nous envisageons d'entreprendre une électrophorèse en conditions non dénaturantes après incubation entre les protéines humaines et la séquence T_{26}* radioactivement marquée pour vérifier la spécificité de l'association observée avec les séquences Tlev et 22hm.

4-4- Conclusion de ces identifications et perspectives

Si on identifie bien l'association entre les séquences télomériques de levure Tlev et GTlev, et les protéines IMPDH, celle-ci peut-elle avoir lieu avec les seules protéines IMPDH ?

En regardant la structure habituelle des IMPDH, on observe qu'elle se présente toujours sous la forme d'un tétramère homogène possédant un axe de symétrie d'ordre 4 (figures 4.22. et 4.26.A.). En regardant la structure sur la tranche (figure 4.25.B), on observe que la partie N-terminale de chaque IMPDH se trouve sur le dessus et qu'il existe en dessous du plateau formé, 4 doigts flexibles (Zhang et *al.*, 1999). Ces doigts sont constitués de deux feuillets β anti-parallèles et forment une région structurale conservée d'une espèce à l'autre. Ils n'interviennent pas dans la fonction catalytique de IMPDH, mais des mutations dans ces domaines sont associées à des conséquences pathologiques (Zhang et *al.*, 1999). Il est possible (Bateman, 1997) que ce domaine puisse être

impliqué dans des fonctions de régulation. Dans tous les cas, la dépense énergétique requise pour la conservation de ces structures suggère un rôle fonctionnel sous-jacent (Zhang et *al.*, 1999).

Dans le cas où il n'y aurait qu'un site spécifique de Tlev par tétramère d'IMPDH, on pourrait supposer que l'association entre les protéines et l'ADN se fasse soit par le dessus, soit par le dessous, ou encore en passant par le centre, du plateau constitué par le tétramère d'IMPDH.

On peut encore penser que l'association se fasse à d'autres endroits du tétramère. Dans ce cas, la symétrie axiale des structures obtenues par cristallographie, laisse penser qu'il pourrait y avoir 4 sites de fixation de Tlev par tétramère.

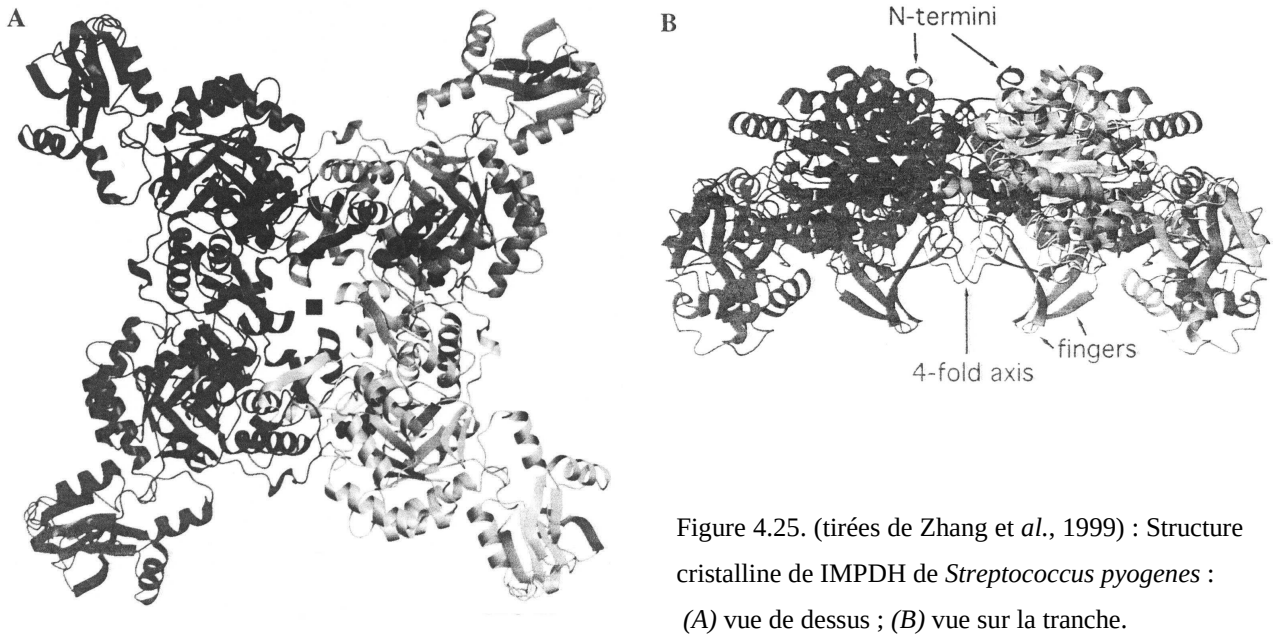


Figure 4.25. (tirées de Zhang et *al.*, 1999) : Structure cristalline de IMPDH de *Streptococcus pyogenes* : *(A)* vue de dessus ; *(B)* vue sur la tranche.

Toutes ces remarques sont à confronter avec les observations précédentes. L'absence de la partie N-terminale des protéines IMPDH ne semble pas influencer l'association avec l'ADN télomérique (puisque l'on montre lors des expériences MALDI que des protéines IMPDH dont la partie N-terminale est absente se sont associées avec le ligand BTlev). Donc celle-ci se ferait soit par le cœur, soit par le dessous du plateau formé par les quatre IMPDH.

On sait aussi que la masse observée du complexe protéique en conditions non dénaturantes est supérieure à 200 kDa. Or le tétramère d'IMPDH a une masse d'environ 220-240 kDa. L'imprécision sur la masse du complexe protéique s'associant à Tlev ne permet pas de déterminer si les IMPDH seules s'associent à ce fragment d'ADN. D'autre part, si les tétramères habituellement observés sont

homogènes, *Saccharomyces cerevisiae* est la seule espèce possédant plusieurs IMPDH quasiment identiques, pouvant laisser penser qu'elles pourraient s'associer de façon 'pseudo'-hétérogène.

Enfin, de récents travaux montrent que la protéine Sir2p utilise le NAD comme substrat lors d'une réaction de désacétylation requérant la présence de lysines acétylées (telles que celles présentent à l'extrémité N-terminale des histones). Le recrutement de NAD par Sir2p lors de cette réaction suggère que cette protéine pourrait être un senseur de l'état énergétique ou de l'oxydation des cellules (Imai et *al.*, 2000 ; Landry et *al.*, 2000 ; Smith et *al.*, 2000). S'associant à ces travaux, les propriétés associatives des protéines IMPDH pour l'ADN télomérique renforcent l'idée suggérée par Imai et *al.* d'interactions possibles entre métabolisme, 'silencing' génomique et sénescence chez la levure.

Chapitre V

Autres résultats

1- Etudes complémentaires par spectrométrie de masse

Nous avons procédé à une étude de la bande faible observée à ~ 60 kDa lors de l'identification différentielle par affinité (figure 4.6.). Nous avons des éléments d'identification par spectrométrie de masse pour plusieurs protéines : CBF3B, Hsp60p, Ku70p et Ku80p. Aucune présence de ces protéines n'a été retrouvée dans la bande de gel découpée dans la piste où la résine ne portait pas le ligand BTlev (piste servant de témoin lors de cette expérience). A la différence des IMPDH, nous ne retrouvons aucune de ces protéines lors de l'expérience réalisée avec le ligand BGTlev.

Ces identifications ont été réalisées à partir de peu de matériel peptidique par rapport à l'identification des protéines Yhr216p, Yml056p et Ylr432p. Dans le cas des protéines Ku70p, Ku80p et CBF3B, il ne s'agit pas d'identification formelle obtenue par le séquençage d'un peptide, mais d'une identification de peptides par spectrométrie de masse MALDI. Nous pensons à une implication éventuelle de ces protéines dans une association spécifique avec la séquence d'ADN télomérique de levure Tlev. Cela pourrait correspondre avec les bandes retardées d'intensités mineures devant la bande retardée due aux protéines IMPDH, observées lors de migrations électrophorétiques en conditions non dénaturantes. Dans le cas de la protéine Hsp60p, une expérience de spectrométrie de masse en tandem confirme l'identification.

1-1- Les protéines Ku70p et Ku80p

Nous avons identifié la présence probable des protéines yKu70p (ou Ku70p) et yKu80p (ou Ku80p). Dans le cas de la protéine Ku70p, les résultats obtenus par spectrométrie de masse MALDI montrent un recouvrement par les peptides identifiés de 21.0 % de la séquence de cette protéine (figure 5.1.).

```
     1          11         21         31         41         51
     |          |          |          |          |          |
   1 mrsvtnafgn sgelndqvde tgyrkfdihe gilfcielse tmfkessdle ykspllleile    60
  61 sldelmsqlv itrpgtaigc yfyycnreda kEGIYELFPL RDINATFMKK lndlledlss   120
 121 grislydyfm fqqtgsekqv rlsvlftfml dtfleeipgq kqlsnkrvfl ftdidkpqea   180
 181 qdiderarlr RLTIDLFDNK vnfatffigy adkpfdnefy sdilqlgsht nentgldsef   240
 241 dgpstkpida kyiksrilrk kevkrimfqc plildektnf ivgvkgytmy thekagvryk   300
 301 LVYEHEDIRQ EAYSKRkfln pitgedvtgk tvkvypygdl dinlsdsqdq ivmeaytqkd   360
 361 aflkiigfrS SSKSIHYFNN IDKssfivpd eakyegsirt lasllkilrk KDKIAILWGK   420
 421 LKsnshpsly tlspssvkDY NEGFYLYRVP FLDEIRkfps llsyddgseh kldydnmkkv   480
 481 tqsimgyfnl rdgynpsdfk npllqkhykv lhdyllqiet tfdenetpnt kkdrMMREDD   540
 541 SLRKLYYIrn kileseksed piiqrlnkYV KIWNMFYKKF NDDNISIKEE Kkpfdkkpkf   600
 601 ni
```

Figure 5.1. : Répartition des peptides identifiés par spectrométrie de masse MALDI correspondant à la protéine Ku70p.

Dans le cas de la protéine Ku80p, nous avons obtenu pour cette protéine un recouvrement de 20 % de la séquence (figure 5.2.).

Indépendamment d'un taux de recouvrement important, trouver des éléments de séquence de ces deux protéines agissant habituellement de concert, accrédite l'hypothèse d'une association avec l'ADN Tlev. Jusqu'à présent, le complexe hétérogène Ku70p/Ku80p était connu pour s'associer au

double brin et au brin G de l'ADN télomérique de *Saccharomyces cerevisiae*. Dans ces conditions, il n'est pas surprenant de retrouver ces protéines dans une association apparemment avec le brin C.

```
      1          11         21         31         41         51
      |          |          |          |          |          |
    1 mssesttfiv dvspsmmknn nvskSMAYLE YTLLNKSKks rktdwiscyl ancpvsensq    60
   61 eipnvfqiqs flapvtttat igfikrlkqy cdqhshdssn eglqsmiqcl lvvsldikqq   120
  121 fqarkilkqi vvftdnlddl ditdeeidll teelstriil idcgkDTQEE RKKsnwlklv   180
  181 eaipnsriyn mnellveits patsvvkpvr VFSGELRlga dilstqtsnp sgsmqdencl   240
  241 cikVEAFPAT Kavsglnrkt avevedsqkk eryvgvksii eyeihnegnk knvseddqsg   300
  301 ssyipvtisk dsvtkayryg adyvvlpsvl vdqtvyesfp gldlrgflnr EALPRyflts   360
  361 essfitadtr lgcqsdlmaf salvdvmlen rKIAVARYVS KKdsevnmca lcpvliehsn   420
  421 insekkfvkS LTLCRlpfae dervtdfpkl ldrtttsgvp lkketdghqi delmeqfvds   480
  481 mdtdelpeip lgnyyqpige vttdttlplp slnkdqeenk kDPLRIPTVJ VYRqqqvlle   540
  541 wihqlminds rEFEIPELPD SLKNKISPYT HKKFDSTKlv evlgikkvdk lkLDSELKTE   600
  601 LEREKIPDLE TLLKrGEQHS Rgspnnsnn
```

Figure 5.2.: Répartition des peptides identifiés par spectrométrie de masse MALDI correspondant à la protéine Ku80p.

Pour permettre de nous assurer de cette association avec la séquence Tlev, nous pourrions dépléter les extraits de levure en protéines Ku70p et Ku80p, par immunoprécipitation à l'aide d'anticorps spécifiques. Cette immunoprécipitation devrait, dans l'hypothèse d'une association de Tlev avec les protéines Ku70p et Ku80p, entraîner la disparition de l'une des bandes retardées d'intensité mineure observées après une migration électrophorétique non dénaturante.

1-2- La protéine Hsp60p

Nous obtenons une identification MALDI de la protéine Hsp60p, correspondant à un recouvrement de 37.8 % de la séquence. Nous en avons eu confirmation par identification par spectrométrie de masse en tandem d'un peptide spécifique de cette séquence (figure 5.3. en bleu).

```
      1          11         21         31         41         51
      |          |          |          |          |          |
    1 MLRSSVVRSR atlrpllrrA YSSHkELKFG VEGRasllkG VETLAEAVAA TLGPKgrNVL 60
   61 IEQPFGPPKi tkdgvtvakS IVLKDKFENM GAKllqevas ktNEAAGDGT TSATVLGRAI 120
  121 FTESVKnvaa gcnpmdlrRG SQVAVEKvie flsankkeit tseeiaqvat isangdshvg 180
  181 kLLASAMEKv gkEGVITIRE GRTLEDELEV TEGMRfdrGF ISPYFITDPK sskVEFEKPL 240
  241 LLLSEKkISS IQDILPALEI SNQSRrplli iaedvdgeal aacilnklrg qvkvcavkAP 300
  301 GFGDNRknti gdiavltggt vfteeldlkp eqctienlgs cdsitvtked tvilngsgpk 360
  361 eaiqerieqi kgsiditttn syekeklqer laklsggvav irvggaseve vgekkdrydd 420
  421 alnatrAAVE EGILPGGGTA LVKasrVLDE VVVDNFDQKl gvdiirkaIT RPAKQIIENA 480
  481 GEEGSVIIGK lideygddfa kgydasksey tdmlatgiid pfkvvrsglv dasgvaslla 540
  541 ttevaivdap eppaaagagg mpggmpgmpg mm
```

Figure 5.3. : Répartition des peptides identifiés par spectrométrie de masse MALDI (en rouge) et spectrométrie de masse en tandem (en bleu) correspondant à la protéine Hsp60p.

Cette protéine est localisée dans les mitochondries (Lithgow et *al.*, 1995) où elle joue un rôle au niveau du repliement des protéines (Koll et *al.*, 1992). Ses principales fonctions biochimiques connues sont celles de chaperone (Koll et *al.*, 1992) et de protéine de choc thermique (Reading et *al.*, 1989), mais à notre connaissance, aucune association de cette protéine avec de l'ADN n'a jusqu'à présent été mentionnée.

1-3- La protéine CBF3B

Nous obtenons une identification incomplète (11.3 % de séquence couverte) de la protéine CBF3B, sans confirmation par un autre moyen de séquençage. Cette protéine nucléaire fait partie du complexe protéique Cbf3 se liant à l'ADN centromérique (Lechner, 1994) et intervenant dans la mitose (Leopardi et *al.*, 1996) et le contrôle du cycle cellulaire (Sorger et *al.*, 1995).

```
    1          11         21         31         41         51
    |          |          |          |          |          |
  1 MFNRttqlks khpcsvctrr KVKCDRMIPC GNCRKRgqds ecmkstklit assskeylpd    60
 61 lllfwqnyey witniglykt kqrdltrtpa nldtdteecm fwmnylqkdq sfqlmnfame   120
121 nlgalyfgsi gdiselylrV EQYWDRradk NHSVDGKywd aliwsvftmc iyympvekla   180
181 eifsvyplhe ylgsnkrlnw edgmqlvmcq nfarcslfql kqcdfmahpd irlvqaylil   240
241 atttfpydep llanslltqc ihtfknfhvd dfrpllnddp vesiakvtlg rIFYRlcgcd   300
301 ylqsgprkpi alhtevssll qhaaylqdlp nvdvyreens tevlywkiis ldrdldqyln   360
361 kSSKPPLKtl dairreldif qykVDSLEED FRsnnsrfqk fialfqistv swklfkmyli   420
421 yydtadsllk vihyskviis livnnfhaks effnrhpmvm qtitrvvsfi sfyqifvesa   480
481 avkqllvdlt eltanlptif gskldklvyl terlsklkll wdkvqlldsg dsfyhpvfki   540
541 lqndikIIEL Kndemfslik GLGSLVPLNK lrqeslleee denntepsdf rtiveefqse   600
601 ynisdils
```

Figure 5.4. : Répartition des peptides identifiés correspondant à la protéine CBF3B

La répartition aléatoire des peptides recensés (figure 5.4.) et la masse native de cette protéine (71.349 Da) sont de bons indices de la présence de cette protéine dans les échantillons analysés. Nous n'avons pas pu déterminer la séquence d'un peptide par spectroscopie de masse en tandem et ces résultats sont insuffisants pour affirmer que cette protéine s'associe spécifiquement à la séquence Tlev. On sait cependant qu'elle se fixe à l'ADN centromérique du satellite CDEIII de *Saccharomyces cerevisiae* (Grady et *al.*, 1992), qui est homologue au satellite III du centromère humain. La séquence de ce dernier est constitué de répétitions de CC (Fitzgerald-Hayes et *al.*, 1982) et peut adopter une structure en motif i (Nonin-Lecomte & Leroy, 2001).

2- Purifications partielles et éléments d'identification

Dans le cadre de la recherche de protéines qui s'associeraient à la séquence télomérique de levure d[(CCCACA)$_3$CCC], nous avons entrepris de développer un protocole de purification menant à l'isolement en conditions natives des protéines impliquées. Les difficultés rencontrées au cours de ces travaux sont en partie dues à l'instabilité constatée du complexe protéique en fonction de l'avancement des étapes de purification.

2-1- Eléments de purification

Le travail présenté décrit les résultats obtenus lors de différentes phases de purifications partielles (pour chacune de ces étapes de purification, une étude préalable du comportement des protéines et de l'efficacité de la séparation a été conduite).

2-1-1- Chromatographie d'exclusion et chromatographie anionique

Une première étape de séparation consiste à récupérer la fraction de chromatographie d'exclusion à 250 kDa (chapitre III-2-3-). La stabilité du complexe protéique est comparable à ce qu'elle était dans les extraits de protéines de levure (rappel : concentration en protéines = 1µg/µl). Le tampon d'élution est TBE 0.5x, pH 8 et NaCl 175 mM (utilisé pour neutraliser les interactions électrostatiques avec la résine. Le remplacement de NaCl par KCl a conduit à la perte de l'activité retardante.) .

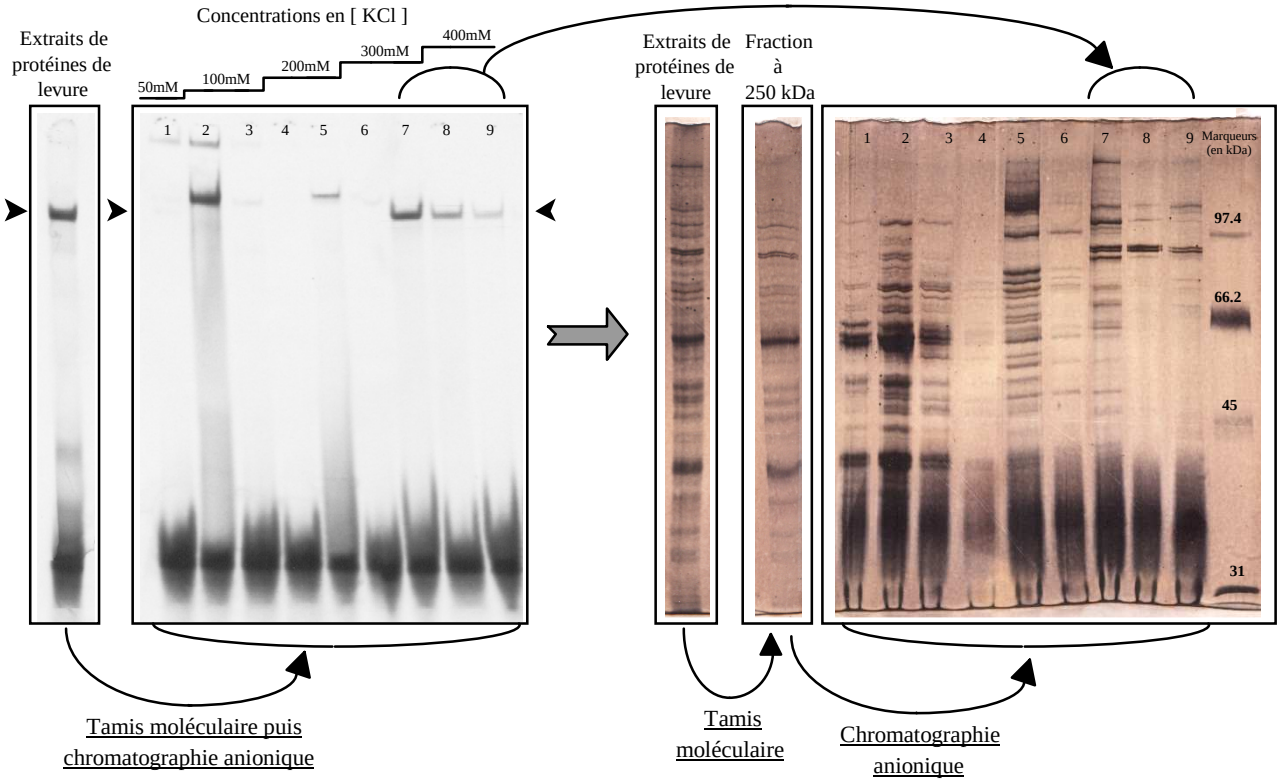
Nous observons un effet de sel de même genre lors de nos essais de séparation par chromatographie anionique : l'élution doit être réalisée en KCl, pour un pH compris entre 7.2 et 7.6. A l'opposé de ce qui se passe lors de la chromatographie d'exclusion, nous perdons l'activité retardée si nous utilisons du chlorure de sodium dans le tampon d'élution.

Lors des expériences de chromatographie anionique, nous avons aussi déterminé que nous perdions l'activité retardée si la résine utilisée avait déjà préalablement servi. Après traitement dans des conditions d'élution dénaturantes (passage à pH 2 puis pH 10), nous observons le décrochage de

produits que les conditions de chromatographie « natives » (gradient jusqu'à 2M KCl) ne permettent pas d'éluer. Ceux-ci pourraient être responsable de la perte de matériel constaté.

Enfin, la chromatographie anionique consiste en une élution par pas de 50 et 100 mM de KCl plutôt que par gradient, afin d'obtenir une meilleure séparation entre les protéines présentes dans les extraits de levure. Dans ces conditions, on observe 90% du complexe protéique recherché dans la fraction correspondant au plateau à 300 mM KCL (figure 5.5.a.).

Les protéines provenant de la fraction à 250 kDa de la chromatographie d'exclusion sont amenées aux conditions d'élution de la chromatographie anionique durant une étape de dilution - reconcentration (utilisation de tubes reconcentrateurs, chapitre II-10-) dans le tampon approprié (TBE 0.5x, pH 7.4, 50 mM KCl), sans perte de l'activité retardée. A la sortie de la colonne anionique, les fractions sont reconcentrées (volume 1 ml, volume final 10 µl) et dessalées avant d'en observer le contenu par migration électrophorétique native (figure 5.5.a) et par gel dénaturant de protéines (figure 5.5.b).



(a) Migration électrophorétique non dénaturante

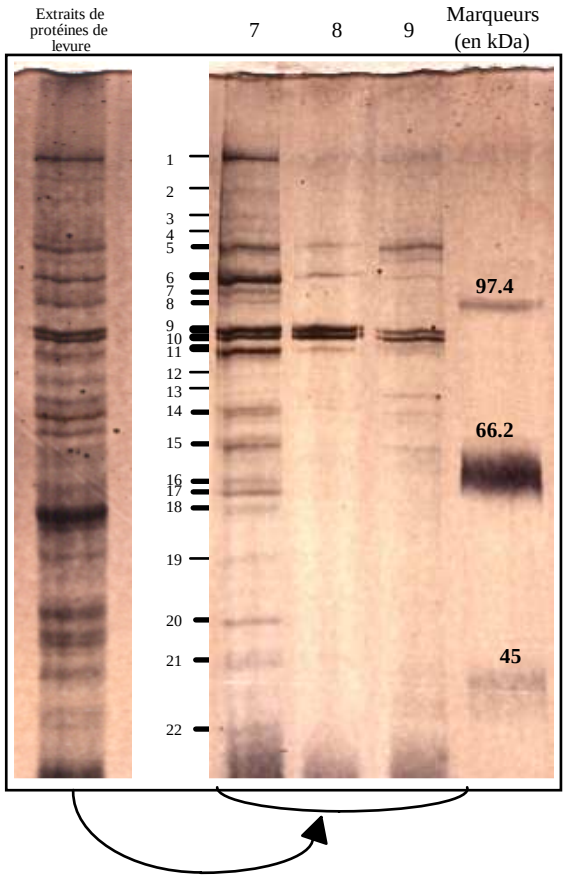
(b) Migration électrophorétique dénaturante

Figure 5.5. : Résultats obtenus après chromatographie d'exclusion puis chromatographie anionique. Les protéines du gel dénaturant *(b)* sont colorées au $AgNO_3$. Les pistes 1 à 9 correspondent aux fractions provenant de la chromatographie anionique.

Les résultats obtenus après ces deux étapes de purifications sont présentés dans la figure 5.5. . Pour la migration électrophorétique en conditions natives (figure 5.5.a.), les concentrations en protéines des différents échantillons sont établies de façon à pouvoir comparer directement l'intensité de la bande retardée obtenue avec les extraits de protéines (~ 0.5 pmol de sites de fixation spécifiques) et celle correspondant à la somme des intensités de l'ensemble des fractions du profil d'élution.

Lors de cette expérience, on perd 20% de l'intensité du signal retardé observé pour les extraits de protéines de levure, les 80% restant se trouvant dans les fractions 7, 8 et 9 (300mM KCl et passage à 400mM). On observe en piste 2 de l'autoradiographie une bande retardée d'intensité comparable à celle présente dans les extraits de protéines de levure ainsi qu'une bande de faible mobilité. En piste 5, on relève la présence d'une bande située au même niveau, que la bande principale observée en piste 2 mais d'intensité plus faible. Les interactions avec la résine anionique pourraient être la cause de modifications rendant possible l'association de protéines avec l'ADN Tlev*.

Figure 5.6. : Agrandissement du gel de protéines présenté en figure 5.5. (fractions 7, 8 et 9, contraste renforcé)



Les concentrations en protéines des pistes ‘Extraits de protéines de levure’ et ‘Fraction à 250 kDa’ correspondent au 1/5ème des concentrations des pistes 1 à 9. On observe entre 31 et 45 kDa des effets de coloration résultants du tampon d’élution de la chromatographie anionique (figure 5.5.b).

A partir de cette expérience (agrandissement du gel de protéines en figure 5.6), on dénombre pour la fraction 7, 22 bandes de protéines distinctes. On note la présence dans les pistes 7, 8 et 9 d’un doublet de protéines (9 - 10) persistant à environ 80-85 kDa, comptant parmi les protéines les plus contrastées observées dans les extraits de protéines non purifiées et dans la fraction à 250 kDa de la chromatographie d’exclusion.

Les bandes 13, 14 et 15 ne sont probablement pas impliquées dans la reconnaissance de la séquence Tlev* puisqu’elles sont absentes de la piste 8, ce qui va à l’encontre des intensités des bandes retardées observées pour les fractions 8 et 9 (figure 5.5.a.).

2-1-2- Développements ultérieurs

Devant l’instabilité du complexe protéique obtenu après ces deux étapes de purification (50% de perte du signal retardé après 2 h à 4°C et pas de bande retardée observée lors de migration électrophorétique native survenant soit après congélation à l’azote, soit après 1 nuit à 4°C), nous avons essayé d’identifier les protéines recherchées en profitant de la séparation complémentaire obtenue lors de la migration électrophorétique native.

Pour cela, après une migration non dénaturante (réalisée à partir des fractions de la chromatographie anionique) le gel est séché sur une feuille de cellophane. Les bandes correspondant à l’activité retardée peuvent être découpées selon l’autoradiographie du gel, avant d’être insérées dans les puits pratiqués en haut d’un gel de protéines (figure 5.7.).

Le découpage est tel que l’on retrouve 10% de l’activité retardée dans la piste de gauche du gel dénaturant, 85% dans la piste centrale et une fraction dans la piste de droite (estimations effectuées en mesurant la radioactivité au front de migration du gel de protéines). La bande de protéines de forte intensité (coloration saturée) observée dans la piste de gauche et la piste centrale correspond

au doublet 9-10 de la figure 5.6. D'autres bandes sont également observées dans la piste centrale présentant une intensité relative beaucoup plus faible.

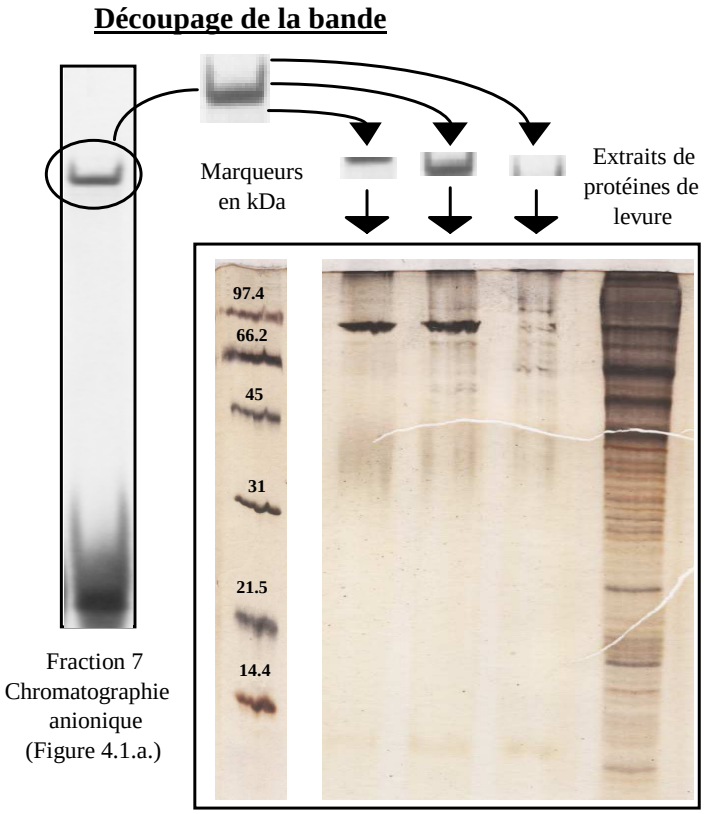


Figure 5.7. : Visualisation des protéines restant après purification partielle par chromatographie d'exclusion puis chromatographie anionique et séparation en fonction de la mobilité électrophorétique. Coloration au $AgNO_3$.

Lors d'expériences de contrôle, nous avons constaté que ce protocole de transfert de protéines d'un gel non dénaturant vers un gel dénaturant ne permettait pas de traiter l'ensemble des protéines (nous estimons transférer 75% des protéines). Nous pensons que la dénaturation des protéines, avant le début de la migration, reste imparfaite. Les agents dénaturants doivent pénétrer dans le gel natif avant d'être efficaces, ce qui signifie préalablement que celui-ci soit réhydraté. Une légère amélioration (n'entraînant pas de résultats marquants nouveaux) est apportée si le gel natif n'est pas séché. Cette amélioration peut s'expliquer d'une part par une meilleure dénaturation générale des protéines, d'autre part, par l'impossibilité de dégradation intervenant lors du séchage du gel sur cellophane. Cependant, les protéines ne pouvant être fixées dans le gel (puisque leur mobilité ne doit pas être entravée ultérieurement lors de la migration en conditions dénaturantes), l'autoradiographie doit être aussi rapide que possible, malgré la présence de molécules d'eau constitutives du gel non séché atténuant le signal radioactif.

2-2- Pontage ADN-protéines aux UV

Le pontage ADN-protéines permet d'associer le fragment d'ADN Tlev* marqué radioactivement avec les protéines spécifiques de cet oligodésoxynucléotide. Lors d'une migration électrophorétique dénaturante, la position de la radioactivité correspond au complexe ADN*-protéines.

L'expérience présentée (figure 5.8.) a été réalisée à partir d'extraits de protéines de levure non purifiés. Nous observons une bande intense (A) entre 85-92 kDa pouvant correspondre au doublet de protéines observé lors des étapes de purification partielle (figures 4.2. et 4.3.), et trois autres bandes d'intensités plus faibles (B, C et D), respectivement vers 47, 50 et 55 kDa. Ces masses moléculaires donnent une idée relative des masses réelles des protéines sur lesquelles l'ADN se sera fixé lors de cette expérience de pontage aux UV.

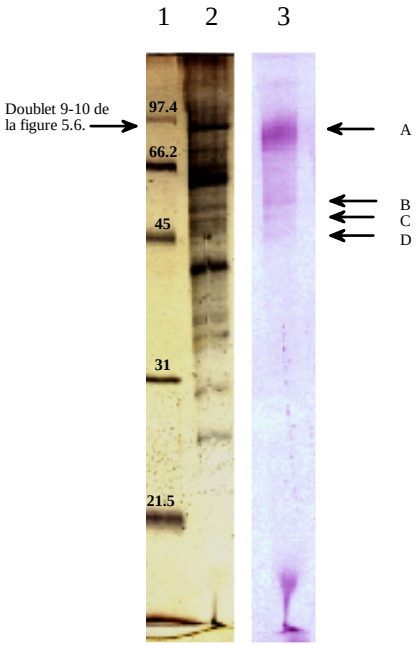


Figure 5.8. : Expérience de pontage ADN-protéines aux UV.
1 : Marqueurs de poids moléculaire ;
2 : migration dénaturante ;
3 : autoradiographie du gel.

On remarque que la mobilité du doublet de protéines 9-10 de la figure 5.6. semble légèrement différente (~ 85 kDa) de celle observée précédemment (~ 80-85 kDa).

2-3- Cohérence avec les résultats du chapitre IV

Le doublet intense de protéines repéré en figure 4.6., dont l'analyse par spectrométrie de masse conduit à l'identification des protéines Yhr216p, Ylr 432p et Yml056p, fait partie des bandes de protéines présentées dans les résultats obtenus lors de purifications partielles. Ainsi, on peut établir

une correspondance entre ce doublet (figure 4.6.) et la bande de protéines observées en figure 5.6. (bande n°19). On le retrouve en figure 5.7. parmi les protéines indicées et en figure 5.8. (bande B).

2-4- Protéines de choc thermique HSP90

Indépendamment des protéines IMPDH, nous retrouvons de façon récurrente dans ce chapitre un doublet de protéines à partir de purifications partielles et de pontage aux UV. Celles-ci ont été formellement identifiées par séquençage par dégradation d'Edman. Il s'agit des protéines Hsc82p et Hsp82p.

Ces deux protéines font partie d'une famille de protéines, les Heat Shock Protein 90. Il s'agit d'une famille de protéines chaperonnes rapidement induites dans les cellules lors de chocs thermiques ou d'un stress environnemental. Les protéines Hsp90p présentent d'importantes homologies d'une espèce eucaryote à l'autre. Elles sont connues pour former de nombreux complexes avec plusieurs autres protéines (Csermely et *al.*, 1998 pour revue).

Récemment, il a été montré que l'activité de la télomérase nécessitait la présence d'une association de protéines, p23 et Hsp90p (Holt et *al.*, 1999). Inhibiteur spécifique de la protéine Hsp90p, la geldanamycine semble par contrecoups ralentir l'activité de la télomérase (Masutomi et *al.*, 2000). C'est la première contribution observée d'un complexe protéique à l'activité de la télomérase.

Dénomination	Génotype
S150	*MATa, leu2-3, 112, his3Δ1 trp1-289, ura3-52,* HSP82 HSC82
PP501	*MATa, leu2-3, 112, his3Δ1 trp1-289, ura3-52,* HSC82Δhsp82 : :KANMX4
PP502	*MATa, leu2-3, 112, his3Δ1 trp1-289, ura3-52,* HSP82Δhsc82 : :KANMX4

Tableau 4.1. : Souches ne produisant pas de Hsc82p et Hsp82c. Les souches PP501 et PP502 sont construites à partir de la souche S150.

Nous avons obtenu de la protéine Hsp90p pure de rat (Pr. P. Csermely, Department of Medical Chemistry, Semmelweis University Budapest) et de la protéine Hsp90p pure de poulet (A. Chadli, Laboratoire des communications hormonales, INSERM U. 33, Kremlin-Bicêtre). Avec ces

protéines, nous avons procédé à différents tests de migration électrophorétique non dénaturante. Nous n'avons obtenu aucun retard de la séquence Tlev* utilisée lors de ces essais. Cela montre que Hsp90p seule ne peut associer des séquences comportant des répétitions de désoxycytidines.

Nous avons également réalisé différents essais de migration non dénaturante, à partir d'extraits de protéines provenant de souches construites pour ne produire que l'une des deux protéines Hsp82c ou Hsp82p (tableau 5.1., Pr. P. Piper, Department of Biochemistry and Molecular Biology, University College London).

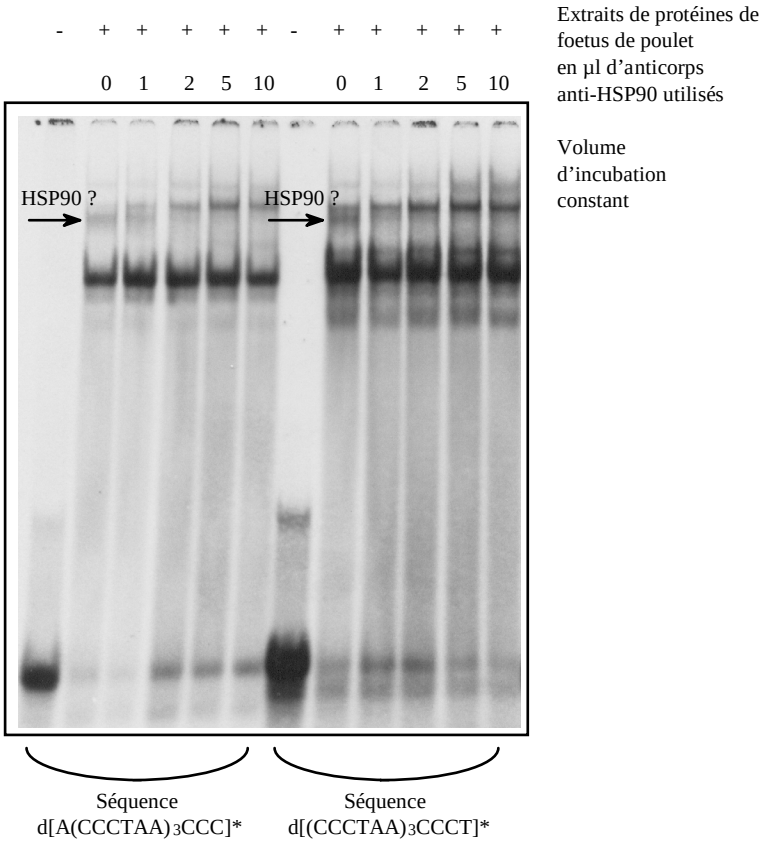


Figure 5.9. : Effets d'anticorps spécifiques de HSP90 de poulet en fonction de la concentration en anticorps.

Nous n'avons pas vu d'effet dû à l'absence de l'une des deux protéines de chocs thermiques. Cependant, ces deux protéines ont des séquences très proches l'une de l'autre (701 acides aminés communs sur 710), et une délétion simultanée des deux gènes étant impossible, on peut imaginer que chacune des protéines pourrait contribuer de la même façon dans un complexe s'associant au brin C de l'ADN télomérique de levure.

En travaillant sur des extraits de protéines de poulet et d'anticorps spécifiques des deux formes de la Hsp90p de poulet (fournis par A. Chadli), nous avons mis en évidence un effet des anticorps anti-

Hsp90p, sur une des bandes de protéines observées par migration électrophorétique après incubation des protéines de poulet et de la sonde télomérique (figure 5.9.). Nous n'observons pas d'effet similaire avec un anticorps anti-Hsp82p de levure.

Les protéines de choc thermique pourraient être impliquées dans un complexe protéique capable de reconnaître le brin C de l'ADN télomérique. Mais, dans toutes les expériences d'identification différentielles par affinité pour les séquences BTlev et BGTlev, nous avons pu observer ce doublet en concentration apparemment identique selon que la résine-streptavidine était porteuse ou non de ligand d'ADN.

CONCLUSION

et

PERSPECTIVES

1- Résultat principal

Nous avons montré que le fragment télomérique de levure d[(CCCACA)$_3$CCC] peut former *in vitro* un motif i par repliement intramoléculaire. L'identification des protéines de levure, Yhr216p, Ylr432p et Yml056p, s'associant spécifiquement à cette séquence d[(CCCACA)$_3$CCC], constitue le principal résultat de l'étude présentée. Cette association peut avoir lieu en milieu acide (pH 6), dans des conditions d'incubation compatibles avec l'observation *in vitro* du motif i. Jusqu'à présent,

seules quelques protéines avaient été identifiées comme s'associant à des séquences d'ADN riches en cytosines. Ainsi une protéine issue de cellules tumorales de souris, p70, qui reconnaît la répétition $(CCT)_n$ (Muraiso et *al.*, 1992), et surtout les protéines hnRNP K et ASF/SF2, issues de cellules HeLa, qui fixent le fragment télomérique humain d[$(CCCTAA)_3CCC$] (Lacroix et *al.*, 2000), pourraient s'associer au motif i, sans que cela ne soit établi pour l'instant. Les résultats obtenus concernant la spécificité des protéines IMPDH pour toutes les séquences formant un motif i intramoléculaire nous oriente également dans l'idée d'une reconnaissance de cette structure par ces protéines.

Cependant malgré la présence d'une bande de protéines (provenant d'extraits nucléaires HeLa) ayant une mobilité électrophorétique comparable, des essais réalisés à partir de protéines IMPDH pures de types I et II montrent que celles-ci ne s'associent pas au fragment télomérique humain.

L'association avec le brin G complémentaire des télomères de levure est un résultat inattendu obtenu lors de ce travail. Très récemment, les propriétés associatives de hnRNP D pour chacun des deux brins télomériques vient d'être démontrées *in vitro* (Eversole & Maizels, 2000). A notre connaissance, il s'agit du seul exemple d'association spécifique d'une protéine pour chacun des deux brins télomériques. Le rôle joué par de telles protéines dans la régulation des télomères reste encore à démontrer.

Dans le cas des IMPDH de levure, la localisation des gènes identifiés, sur des chromosomes où l'on retrouve fréquemment la séquence d[$(CCCACA)_3CCC$] ou des séquences consensus très proches, et leur possible régulation par le phénomène de TPE, renforce l'idée que ces protéines pourraient jouer un rôle structural important dans le maintien de l'intégrité de ces chromosomes.

2- Autres contributions

Nous avons obtenu au cours de notre étude, différents autres résultats. Nous avons montré que les protéines Ku70p et Ku80p, dont on connaissait déjà l'association à l'ADN télomérique double brin et au brin G, peuvent aussi s'associer avec le brin riche en C. On montre aussi qu'une protéine du complexe Cbf3, dont on connaissait l'association au satellite CDEIII du centromère de

Saccharomyces cerevisiae, et que la protéine Hsp60p, peuvent s'associer à un fragment télomérique riche en cytosines.

Nous avons montré que la protéine Hsp90p de poulet peut s'associer avec des séquences riches en cytosines. Cette caractéristique a été obtenue en se basant sur de nombreux indices impliquant cette protéine chez *Saccharomyces cerevisiae*, mais l'existence de deux protéines homologues, Hsc82p et Hsp82p, distinctes, ne nous a pas permis jusqu'à présent de confirmer cette hypothèse chez la levure. On sait que Hsp90p est impliquée dans un rôle de régulation de l'activité de la télomérase des mammifères et des études montrent que l'on pourrait inhiber l'activité de la télomérase en inhibant Hsp90p. On ne sait pas pour l'instant de quelle façon elle agit. L'association au brin C de l'ADN télomérique pourrait permettre d'expliquer le mode de fonctionnement de cette régulation.

3- Perspectives

Un premier objectif sera de confirmer que l'association des protéines IMPDH avec la séquence Tlev, ne fait pas intervenir d'autres facteurs que nous n'aurions pas appréhender par notre mode d'identification. Pour cela, nous avons déjà entrepris de produire l'une des protéines, Ylr432p. Nous pourrons alors entreprendre de définir le domaine d'interaction avec l'ADN, puis essayer de caractériser les mécanismes de cette association. Nous pouvons aussi envisager de vérifier qu'elles s'associent au motif i, en s'appuyant par exemple, sur des expériences de transfert d'énergie de fluorescence (Mergny, 1999).

La caractérisation approfondie de l'association avec le brin G complémentaire constitue une autre étape des développements envisageables. Notamment, il sera intéressant de regarder sous quelle forme le brin G s'associe avec ces protéines, et quelle est l'affinité des protéines pour ce brin.

Annexes

1- Liste des abréviations

Abréviations	Définitions
AMP	Adénosine monophosphate
ATP	Adénosine triphosphate
ade	Adénine (nomenclature des souches de levures)
BSA	Albumine bovine (Bovine Serum Albumine)
can	Canavanine (nomenclature des souches de levures)
Da	Dalton (défini comme 1/12ème de la masse d'un atome de carbone ^{12}C), approximativement égal à la masse d'un atome ^{1}H
DEAE	Diéthylaminoéthyl
DHB	Acide gentisique ou 2,5-hydroxybenzoïque
DO	Densité optique
EDTA	Acide éthylènediaminetétraacétique
ESI	Ionisation électrospray (ElectroSpray Ionization)
HEPES	Acide (N-[2-hydroxyéthil]piperazine-N'-[2-éthanesulfonique]), pKa = 7.5 à 25°C
his	Histidine (nomenclature des souches de levures)
IMP	Inosine monophosphate
IMPDH	Inosine monophosphate déshydrogénase
KDa	KiloDalton
leu	Leucine (nomenclature des souches de levures)
MALDI	Désorption / ionisation de la matrice assistée par laser (Matrix-Assisted Laser Desorption/Ionisation)
MATα et MAT**a**	Type de souches de levures (nomenclature des souches de levures) (mating type)
MES	Acide 2-[N-Morpholino]éthanesulfonique
MPA	Acide mycophénolique
MS	Spectrométrie de masse (Mass Spectrometry)
MS / MS	Spectrométrie de masse en tandem
NAD	Nicotinamide Adénine Dinucléotide
NADH	NAD sous forme réduite
NOESY	Nuclear Overhauser Enhancement SpectroscopY

Abréviations	Définitions
PEI	Polyéthylèneimine
pep	Pepstatine (nomenclature des souches de levures)
PMSF	Fluorure de Phénylméthylsulfonyl
PVDF	Fluorure de Polyvinylidène
RMN	Résonance Magnétique Nucléaire
SDS	Dodécyl Sulfate de Sodium
TCE	Tris-base Citrate EDTA
TBE	Tris-base Borate EDTA
TFA	Acide trifluoroacétique
Tlev	Séquence télomérique de levure d[$(CCCACA)_3CCC$]
TOF	Temps de vol (Time Of Flight)
trp ou TRP	Tryptophane (nomenclature des souches de levures)
ura	Uracile (nomenclature des souches de levures)
XMP	Xanthosine 5’-monophosphate
YPG	Yeast Peptone Glucose

2- Principales protéines impliquées au niveau des télomères de levures

Protéines	*Noms*	*Rôle / fonctions*
Est1p	Ever Shorter Telomere 1	Régulation de l'activité de la télomérase
Est2p	Ever Shorter Telomere 2	Composante catalytique de la télomérase
Est3p	Ever Shorter Telomere 3	Protéine requise *in vivo* pour le fonctionnement de la télomérase
Est4p/Cdc13p	Ever Shorter Telomere 4	Protection du télomère. Nécessaire pour que la télomérase est accès à l'extrémité du chromosome
Ku70p		Reconnaissance de la transition entre le double brin d'ADN télomérique et le simple brin. Activité de réparation du double brin. Implication dans les phénomènes de silencing trascriptionnel et de TPE. Action concertée avec Ku80p
Ku80p		Action concertée avec Ku70p
Mre11p		Maintien de la longueur du télomère
Pif1p ou Tst1p	Petite Integration Frequency ou Telomere STability	Régulation de la longueur des télomères
Pol1p	DNA polymerase 1	Allongement du brin G si protéine non exprimée
Pol2p	DNA polymerase 2	Allongement de la longueur des télomères si protéine non exprimée
Rap1p	Repressor Activator Protein 1	Favorise la formation de quadruplexe de guanines. S'associe au double brin d'ADN télomérique. Régulation de la longueur des télomères.
Rad50p		Maintien de la longueur des télomères en l'absence de la télomérase
Rad51p		Maintien de la longueur des télomères en l'absence de la télomérase
Rad52p		Protéine requise pour le maintien de télomère long
Rfc1p		Allongement de la longueur des télomères si protéine non exprimée
Rif1p	Rap1p Interaction Factor 1	Régulation de la longueur des télomères et

		silencing transcriptionnel. Interactions avec Rap1p
Rif2p	Rap1p Interaction Factor 2	Régulation de la longueur des télomères et silencing transcriptionnel. Interactions avec Rap1p
Sir1p	Silent information regulator 1	Régulation de la longueur des télomères. Silent information regulator.
Sir2p	Silent information regulator 2	Régulation de la longueur des télomères. Silent information regulator.
Sir3p	Silent information regulator 3	Régulation de la longueur des télomères. Silent information regulator.
Sir4p	Silent information regulator 4	Régulation de la longueur des télomères. Silent information regulator.
Stn1p		Protéine régulant la longueur des télomères en association avec Cdc13p
Tel1p		Régulation de la longueur des télomères
Tel2p		Régulation de la longueur des télomères et TPE
Xrs2p		Maintien de la longueur des télomères

3- Caractéristiques des différents tampons utilisés

Appellations	Eléments constituants	Concentrations
TBE 10x pH 8.3	Tris	0.89 M
	H_3BO_3	0.89 M
	EDTA	20 mM
TCE 10x pH 6.8	Tris	0.89 M
	Acide Citrique	0.89 M
	EDTA	20 mM
TES pH 7.5	Tris	50 mM
	Sorbitol	0.6 M
	EDTA	1 mM
Chromatographie anionique	TBE 0.5x	
pH 7.4 Sel d'élution : KCl	ßmercaptoéthanol	5 mM
	Nonidet P40	20 µl/l
+	PMSF	0.2 mM
	Benzamidine	0.65 mM
Expériences d'affinité	Aprotinin	0.9 µM
(175 mM 50mM KCl)	Inhibiteur de trypsine	0.3 µM
Chromatographie d'exclusion	TBE 0.5x	
pH 8	ßmercaptoéthanol	5 mM
Sel d'élution : NaCl (175 mM)	Nonidet P40	20 µl/l
	PMSF	0.2 mM
	Benzamidine	0.65 mM
	Aprotinin	0.9 µM
	Inhibiteur de trypsine	0.3 µM
Tampon de dénaturation de la résine d'affinité (pour migration électrophorétique dénaturante)	SDS 10%	25µl
	ßmercaptoéthanol	5µl
	Tris-HCl (1M pH 8)	10µl
	H_2O	40µl
	Urée	60mg
Tampon d'électrophorèse dénaturante (protéines)	Tris	25 mM
	Glycine	250 mM
	SDS	0.1 %
Tampon de fixation du ligand biotine	KH_2PO_4	20 mM
	NaCl	150 mM
Tampon PBS 1x	NaCl	135 mM
	KCl	2.68 mM
pH 7.4	Na_2HPO_4	8.1 mM
	KH_2PO_4	1.76 mM

4- Masses des peptides ionisés

Acides Aminés	Codes à 1 lettre	Codes à 3 lettres	Masses des résidus monoisotopique (en Da)
Glycine	G	Gly	57.02
Alanine	A	Ala	71.04
Sérine	S	Ser	87.03
Proline	P	Pro	97.05
Valine	V	Val	99.07
Thréonine	T	Thr	101.05
Cystéine	C	Cys	103.01
Leucine	L	Leu	113.08
Isoleucine	I	Ile	113.08
Asparagine	N	Asn	114.04
Acide aspartique	D	Asp	115.03
Lysine	K	Lys	128.09
Glutamine	Q	Gln	128.06
Acide glutamique	E	Glu	129.04
Méthionine	M	Met	131.04
Histidine	H	His	137.06
Phenylalanine	F	Phe	147.07
Arginine	R	Arg	156.1
Tyrosine	Y	Tyr	163.06
Tryptophane	W	Trp	186.08

5- Modes opératoires

5-1- Préparation d'extraits protéiques de levure

5-1-1- Extraits protéiques natifs obtenus après éclatement des levures à l'azote liquide

- Récupérer les levures dans des pots à centrifuger en Nalgène
- Centrifuger à 8000 tr/min., 4°C, pendant 15 min.
- Eliminer le surnageant sans assécher les levures
- Eclatement des levures :
  - Ajouter des anti-protéases :
    100 µl de PMSF 0.2 M pour 10 grammes de levures, 100 µl minimum, du Pefabloc (4 mM), de la benzamidine (65 µM), de l'aprotinin (8.5 nM), de l'inhibiteur de trypsine (3 nM).
  - Congeler en versant directement de l'azote liquide sur les levures
  - Décongeler en trempant les pots dans une cuvette d'eau à 25-35°C
  - Répéter ces étapes une deuxième fois en ajoutant des anti-protéases une troisième fois
- Centrifuger 15 min. à 8000 tr/min à 4°C
- Récupérer le surnageant et ajouter des anti-protéases
- Ajouter de la streptomycine (8 mg/ml)
- Laisser sur lit de glace avec agitation ponctuelle pendant 15 min.
- Centrifuger 20 min. à 10000 tr/min à 4°C
- Précipiter le surnageant au sulfate d'ammonium 55 % (326 g/l)
- Laisser, après dissolution complète des cristaux de sulfate d'ammonium, sur lit de glace pendant 15 min.
- Centrifuger 20 min. à 10000 tr/min à 4°C
- Reprendre le précipité dans un volume minimum de TBE 0.5x + βmercaptoéthanol (5 mM) + Nonidet 0.4% + PMSF 0.2 mM + NaCl (175 mM) + $MgCl_2$ (2 mM)
- Dialyser 4 – 5 h contre TBE 0.5x + NaCl (10 mM) + βmercaptoéthanol (5 mM) + $MgCl_2$ (2 mM) + Nonidet (20 µl pour 4 litres)
- Aliquoter en fractions de 500 µl le dialysat

5-1-2- Extraits protéiques natifs obtenus selon la méthode dite « des billes de verre »

- Récupérer les levures et ajouter des anti-protéases (cf. précédemment)
- Reprendre un gramme de levures dans 2 ml de TES (Tris (50 mM), EDTA (1 mM), Sorbitol (0.6 M)), pH 7.5
- Ajouter un volume équivalent de billes de verre, préalablement lavées à l'acide chlorhydrique concentré et équilibrées dans du TES, pH 7.5, puis agiter 8 minutes en vortexant fortement
- Ajouter des anti-protéases (cf. précédemment)
- Centrifuger 15 min. à 8000 tr/min, 4°C
- Récupérer le surnageant et ajouter des anti-protéases puis reprendre la suite du protocole précédent

5-1-3- Extraits protéiques obtenus selon la méthode dite « des billes de verre » dénaturante

- Récupérer dans un eppendorf de 2 ml, 10 DO_{600nm} de cellules (~ 10^8 cellules) par essai
- Rincer avec 1 ml de NaN_3 (10 mM) froid
- Centrifuger et reprendre dans 0.4 ml de NaN_3 (10 mM), 2% TCA froid
- Ajouter un volume de billes (lavées comme précédemment décrit) nécessaire et suffisant pour atteindre le bas du ménisque (~ 1.1 ml)
- Vortexer à pleine vitesse pendant 4 min. à température ambiante
- Sédimenter les billes (sans centrifuger) et prélever le premier surnageant
- Laver trois fois les billes avec 500 µl de TCA 2% et prélever le surnageant à chaque fois en lassant sédimenter sans centrifuger
- Réunir les trois surnageants avec le premier et laisser précipiter les protéines pendant 15 min. dans la glace.
- Centrifuger 15 min. à 12000 tr/min. à 4°C
- Eliminer le surnageant
- Remettre le culot en suspension avec 100 µl de Tris-HCl (0.1M) pH 8
- Ajouter 100 µl de tampon Laemmli puis chauffer 1 min. à ~ 100°C
- Vortexer et centrifuger quelques secondes
- Déposer 10 µl de surnageant (~ 25-50 µg de protéines) par piste de gel SDS dénaturant

5-2- Préparation d'échantillons pour la spectrométrie de masse

5-2-1- Remarques préliminaires

La spectrométrie de masse présente l'avantage de pouvoir travailler avec de très faibles quantités de matériel (de l'ordre de la femtomole). Dans ces conditions, il faut particulièrement faire attention à différents types de contamination possibles. Aussi, il faut :

- Utiliser des tubes de marque Treff de façon à éviter des réactions chimiques possibles pendant les périodes d'incubation en présence d'acétonitrile (contamination des échantillons par des polymères)
- Pour limiter les contaminations par les kératines humaines et ovines, manipuler avec des gants, non poudré et une blouse, et éviter la présence à proximité de la paillasse de vêtements en laine.

Autres remarques :

- Rincer les échantillons (fragments de gel de protéines) pour éviter de les contaminer avec du polyacrylamide
- Ne pas fixer les protéines par précipitation à l'acide acétique

5-2-2- Excision et rinçage des spots

- Prélever les échantillons du gel de protéines en découpant les spots au scalpel
- Couper les échantillons en bouts de 1 mm / 1 mm
- Laver le gel deux fois dans de l'eau (100 à 200 µl de H_2O milliQ selon le volume de gel à laver), 10 min sous agitation douce

5-2-3- Lavage des spots (élimination du SDS qui inhibe les protéases)

- Eliminer le liquide restant
- Couvrir avec de l'acétonitrile 100% (3 à 4 min.). Les bouts de gel s'agglutinent (par élimination de l'eau) et blanchissent
- Eliminer le surnageant
- Hydrater avec 0.1 M de carbonate d'ammonium (NH_4HCO_3, 2 g / 250 ml)
- Après 5 min., ajouter 1.2 volume d'acétonitrile
- Incuber 15 min et attendre l'agglutination et le blanchiment
- Enlever le surnageant et sécher au speedvac (~ 30 min.)

5-2-4- Réduction et Alkylation

- Laisser le gel gonfler dans 10 mM dithiotréitol / 0.1 M NH_4HCO_3 = 1.5 mg/ml (préparer cette solution au dernier moment)
- Incuber 45 min. à 56°C
- Ramener à T° ambiante et enlever l'excès de solution
- Ajouter 1 volume d'acétonitrile pour baigner le culot
- Remplacer rapidement par 55 mM iodoacétamide / 0.1 M NH_4HCO_3 = 10.2 mg/ml (préparer cette solution au dernier moment)
- Incuber 30 min. à T° ambiante dans le noir
- Eliminer le surnageant et laver avec du carbonate d'ammonium NH_4HCO_3 (0.1 M)
- Après 5 min, ajouter 1.2 volume d'acétonitrile
- Incuber 15 min, attendre que le gel blanchisse, enlever le surnageant, sécher au speedvac

5-2-5- Digestion

- Réhydrater avec 50 mM NH_4HCO_3, 5mM $CaCl_2$ et 12.5 ng/µL de trypsine (Trypsin Sequencing Grade 1418475 - 85739320, Boehringer Mannheim) à 4°C sur la glace avec 100µl. Réajuster si toute la solution est absorbée (la trypsine rentre dans le gel).
- Après 45 min, éliminer l'excès et ajouter du tampon 50 mM NH_4HCO_3 sans trypsine pour maintenir le gel immergé à 37°C pour la nuit.

5-2-6- Extraction des peptides

- Prélever le surnageant après centrifugation
- Reprendre les bouts de gel avec 50 µl 1% d'acide formique et agiter 10 min.
- Pooler avec le surnageant précédent
- Reprendre les bouts de gel avec 50 µl 1% d'acide formique et agiter 5 min.
- Ajouter 60 µl d'acétonitrile et agiter 10 min
- Pooler avec le surnageant précédent
- Sécher au speedvac (env. 1 h)
- Redissoudre dans 20 µl d' acide formique 1%

5-2-7- Dessalage et préparation des échantillons

On utilise pour le dessalage des micro colonnes $ZipTip_{C18}$ (Size P10, ZTC 185960, Millipore).

- Préparation des $ZipTip_{C18}$ (ne pas les laisser sécher une fois hydraté) :
  - 3x aspiration/refoulement avec 50 : 50 (acétonitrile : acide formique 1%) (3 x 10µl)
  - 3x aspiration/refoulement avec acide formique 1% (3 x 10µl)

- Fixation des peptides sur la résine C18. Procéder dans l'ordre sans temps d'attente pour chaque échantillon.
  - Redissoudre le digest dans 20 µl d'acide formique à 1% et soniquer 5 min.
  - 10x aspiration/refoulement avec l'échantillon
  - 3x aspiration/refoulement avec Ac. formique 1% (3 x 10µl)
  - Eluer avec 4 µL de mélange avec 50 : 50 (Acétonitrile : Ac. formique 1%).
  - Récupérer l'éluat
  - Eluer avec 4 µL de mélange avec 80 : 20 (Acétonitrile : Ac. formique 1%).
  - Récupérer l'éluat.
  - Evaporer l'acétonitrile au speedvac

5-2-8- Analyse par spectrométrie de masse :

- Spectroscopie MALDI :
  - redissoudre le digest dans 3 µl d'acide formique à 1% et soniquer 3 min.
  - Déposer 0.5 µl de DHB et 0.5 µl de l'échantillon, soit 1 / 6ème de l'échantillon total, sur la plaque d'analyse du spectromètre.
  - Attendre la cristallisation avant de placer la plaque dans l'appareil

- Spectroscopie Q-TOF MS/MS :
  - Dissoudre le digest dans un solvant de type méthanol / H_2O 1 : 1 avec 5% d'acide formique
  - Introduire la solution dans un capillaire pour analyse

6- Alignements de séquences

Les données présentées dans le tableau ci-dessous correspondent aux résultats obtenus lors de l'alignement de gènes (Blast alignments) avec la séquence de Ylr432p.

Gènes	Appellations	Organismes	Nombres d'acides aminés	% identités	% similarités
YHR216W	IMD2, IMH1 ou PUR5	*Saccharomyces cerevisiae*	523	93 % (484/523)	97 % (507/523)
YML056C	IMH3	*S. cerevisiae*	522	89 % (465/522)	95 % (495/522)
YAR073W	IMD1, FUN63, (possible pseudo-gène)	*S. cerevisiae*	372	90 % (336/372)	95 % (354/372)
YAR075W	(possible pseudo-gène)	*S. cerevisiae*	153	96 % (147/153)	98 % (150/153)
IMPDH1	IMP déshydrogénase 1	*Homo sapiens*	506	61 % (309/506)	78 % (394/506)
IMPDH2	IMP déshydrogénase 2	*H. sapiens*	505	62 % (315/505)	78 % (396/505)
Ras	(IMP déshydrogénase)	*Drosophila melanogaster*	502	61 % (308/502)	77 % (385/502)
T22D1.3	(IMP déshydrogénase)	*Caenorhabditis elegans*	521	49 % (257/521)	67% (351/521)

6-2-1- Alignement de séquences entre le gène *YLR432W* et le gène *YHR216W*

```
YLR432W  1    MAAVRDYKTALEFAKSLPRLDGLSVQELMDSKTRGGLTYNDFLVLPGLVDFPSSEVSLQT 60
              MAA+RDYKTAL+F KSLPR DGLSVQELMDSK RGGLTYNDFL+LPGLVDF SSEVSLQT
YHR216W  1    MAAIRDYKTALDFTKSLPRPDGLSVQELMDSKIRGGLTYNDFLILPGLVDFASSEVSLQT 60

YLR432W  61   KLTRNITLNTPFVSSPMDTVTESEMAIFMALLGGIGFIHHNCTPEDQADMVRRVKNYENG 120
              KLTRNITLN P VSSPMDTVTESEMA FMALLGGIGFIHHNCTPEDQADMVRRVKNYENG
YHR216W  61   KLTRNITLNIPLVSSPMDTVTESEMATFMALLGGIGFIHHNCTPEDQADMVRRVKNYENG 120

YLR432W  121  FINNPIVISPTTTVGEAKSMKERFGFSGFPVTEDGKRNGKLMGIVTSRDIQFVEDNSLLV 180
              FINNPIVISPTTTVGEAKSMKE++GF+GFPVT DGKRN KL+G++TSRDIQFVEDNSLLV
YHR216W  121  FINNPIVISPTTTVGEAKSMKEKYGFAGFPVTTDGKRNAKLVGVITSRDIQFVEDNSLLV 180

YLR432W  181  QDVMTKNPVTGAQGITLSEGNEILKKIKKGKLLIVDDNGNLVSMLSRTDLMKNQNYPLAS 240
              QDVMTKNPVTGAQGITLSEGNEILKKIKKG+LL+VD+ GNLVSMLSRTDLMKNQNYPLAS
YHR216W  181  QDVMTKNPVTGAQGITLSEGNEILKKIKKGRLLVVDEKGNLVSMLSRTDLMKNQNYPLAS 240

YLR432W  241  KSATTKQLLCGAAIGTIDADKERLRLLVEAGLDVVILDSSQGNSIFQLNMIKWIKETFPD 300
              KSA TKQLLCGA+IGT+DADKERLRLLV+AGLDVVILDSSQGNSIF+LNM+KW+KE+FP
YHR216W  241  KSANTKQLLCGASIGTMDADKERLRLLVKAGLDVVILDSSQGNSIFELNMLKWVKESFPG 300

YLR432W  301  LEIIAGNVATREQAANLIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVCEFANQ 360
              LE+IAGNV TREQAANLIAAGADGLRIGMG+GSICITQEVMACGRPQGTAVYNVCEFANQ
```

```
YHR216W  301  LEVIAGNVVTREQAANLIAAGADGLRIGMGTGSICITQEVMACGRPQGTAVYNVCEFANQ 360

YLR432W  361  FGIPCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGS 420
              FG+PCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGS
YHR216W  361  FGVPCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGS 420

YLR432W  421  IDAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIG 480
              IDAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIG
YHR216W  421  IDAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIG 480

YLR432W  481  YKSLTLLKENVQSGKVRFEFRTASAQLEGGVHNLHSYEKRLHN 523
               +SLTLLK NVQ GKVRFEFRTASAQLEGGVHNLHSYEKRLHN
YHR216W  481  CRSLTLLKNNVQRGKVRFEFRTASAQLEGGVHNLHSYEKRLHN 523
```

6-2-2- Alignement de séquences entre le gène *YLR432W* et le gène *YML056C*

```
YLR432W  2    AAVRDYKTALEFAKSLPRLDGLSVQELMDSKTRGGLTYNDFLVLPGLVDFPSSEVSLQTK 61
              AA  DYK ALE  K+    DGLSVQELMDS TRGGLTYNDFLVLPGLV+FPSS VSLQTK
YML056C  3    AAPLDYKKALEHLKTYSSKDGLSVQELMDSTTRGGLTYNDFLVLPGLVNFPSSAVSLQTK 62

YLR432W  62   LTRNITLNTPFVSSPMDTVTESEMAIFMALLGGIGFIHHNCTPEDQADMVRRVKNYENGF 121
              LT+ ITLNTPFVSSPMDTVTE++MAI+MALLGGIGFIHHNCTP++QA MV++VK +ENGF
YML056C  63   LTKKITLNTPFVSSPMDTVTEADMAIYMALLGGIGFIHHNCTPKEQASMVKKVKMFENGF 122

YLR432W  122  INNPIVISPTTTVGEAKSMKERFGFSGFPVTEDGKRNGKLMGIVTSRDIQFVEDNSLLVQ 181
              IN+PIVISPTTTVGE K MK +FGFSGFPVTEDGK  GKL+G+VTSRDIQF+ED+SL+V
YML056C  123  INSPIVISPTTTVGEVKVMKRKFGFSGFPVTEDGKCPGKLVGLVTSRDIQFLEDDSLVVS 182

YLR432W  182  DVMTKNPVTGAQGITLSEGNEILKKIKKGKLLIVDDNGNLVSMLSRTDLMKNQNYPLASK 241
              +VMTKNPVTG +GITL EGNEILK+ KKGKLLIVDDNGNLVSMLSR DLMKNQNYPLASK
YML056C  183  EVMTKNPVTGIKGITLKEGNEILKQTKKGKLLIVDDNGNLVSMLSRADLMKNQNYPLASK 242

YLR432W  242  SATTKQLLCGAAIGTIDADKERLRLLVEAGLDVVILDSSQGNSIFQLNMIKWIKETFPDL 301
              SATTKQLLCGAAIGTI+ADKERLRLLVEAGLDVVILDSSQGNS+FQLNMIKWIKETFPDL
YML056C  243  SATTKQLLCGAAIGTIEADKERLRLLVEAGLDVVILDSSQGNSVFQLNMIKWIKETFPDL 302

YLR432W  302  EIIAGNVATREQAANLIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVCEFANQF 361
              EIIAGNVATREQAANLIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVC+FANQF
YML056C  303  EIIAGNVATREQAANLIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVCQFANQF 362

YLR432W  362  GIPCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGSI 421
              G+PCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFY+DGKRLKAYRGMGSI
YML056C  363  GVPCMADGGVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYKDGKRLKAYRGMGSI 422

YLR432W  422  DAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGY 481
              DAMQKTG KGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIG
YML056C  423  DAMQKTGNKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGC 482

YLR432W  482  KSLTLLKENVQSGKVRFEFRTASAQLEGGVHNLHSYEKRLHN 523
              +SLT LKENVQ+G+VRFEFRTASAQLEGGVHNLHSYEKRL+N
YML056C  483  ESLTSLKENVQNGEVRFEFRTASAQLEGGVHNLHSYEKRLYN 524
```

6-2-3- Alignement de séquences entre le gène *YHR216W* et le gène *YAR073W*

```
YHR216W    MAAIRDYKTALDFTKSLPRPDGLSVQELMDSKIRGGLTYNDFLILPGLVDFASSEVSLQT 60
           MAAIRDYKTALD TKSLPRPDGLSVQELMDSKIRGGL YNDFLILPGLVDFASSEVSLQT
YAR073W    MAAIRDYKTALDLTKSLPRPDGLSVQELMDSKIRGGLAYNDFLILPGLVDFASSEVSLQT 60

YHR216W    KLTRNITLNIPLVSSPMDTVTESEMATFMALLGGIGFIHHNCTPEDQADMVRRVKNYENG 120
           KLTRNITLNIPLVSSPMDTVTESEMATFMALL GIGFIHHNCTPEDQADMVRRVKNYENG
YAR073W    KLTRNITLNIPLVSSPMDTVTESEMATFMALLDGIGFIHHNCTPEDQADMVRRVKNYENG 120

YHR216W    FINNPIVISPTTTVGEAKSMKEKYGFAGFPVTTDGKRNAKLVGVITSRDIQFVEDNSLLV 180
           FINNPIVISPTTTVGEAKSMKEKYGFAGFPVT DGKRNAKLVG ITSRDIQFVEDNSLLV
YAR073W    FINNPIVISPTTTVGEAKSMKEKYGFAGFPVTADGKRNAKLVGAITSRDIQFVEDNSLLV 180

YHR216W    QDVMTKNPVTGAQGITLSEGNEILKKIKKGRLLVVDEKGNLVSMLSRTDLMKNQNYPLAS 240
           QDVMTKNPVTGAQGITLSEGNEILKKIKKGRLLVVDEKGNLVSMLSRTDLMKNQ YPLAS
```

```
YAR073W    QDVMTKNPVTGAQGITLSEGNEILKKIKKGRLLVVDEKGNLVSMLSRTDLMKNQKYPLAS 240

YHR216W    KSANTKQLLCGASIGTMDADKERLRLLVKAGLDVVILDSSQGNSIFELNMLKWVKESFPG 300
           KSANTKQLL GASIGTMDADKERLRLLVKAGLDVVILDSSQGNSIF+LNM+KW+KE+FP
YAR073W    KSANTKQLLWGASIGTMDADKERLRLLVKAGLDVVILDSSQGNSIFQLNMIKWIKETFPD 300

YHR216W    LEVIAGNVVTREQAANLIAAGADGLRIGMGTGSICITQEVMACGRPQGTAVYNVCEFANQ 360
           LE+IAGNVVT+EQAANLIAAGADGLRIGMGTGSICITQ+VMACGRPQGTAVYNVCEFANQ
YAR073W    LEIIAGNVVTKEQAANLIAAGADGLRIGMGTGSICITQKVMACGRPQGTAVYNVCEFANQ 360

YHR216W    FGVPCMADGGVQ 372
           FGVPCMADGGVQ
YAR073W    FGVPCMADGGVQ 372
```

6-2-4- Alignement de séquences entre le gène *YLR432W* et le gène *YAR075W*

```
YLR432W  372  QNIGH-ITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGSIDAMQKTGTK 430
              +NIGH ITKALALGSSTVMMGGMLAGTTESPGEY YQDGKRLKAYRGMGSIDAMQKTGTK
YAR075W  5    KNIGHIITKALALGSSTVMMGGMLAGTTESPGEYLYQDGKRLKAYRGMGSIDAMQKTGTK 64

YLR432W  431  GNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGYKSLTLLKEN 490
              GNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIG +SLTLLKEN
YAR075W  65   GNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGCRSLTLLKEN 124

YLR432W  491  VQSGKVRFEFRTASAQLEGGVHNLHSYEKRLHN 523
              VQSGKVRFEFRTASAQLEGGV+NLHSYEKRLHN
YAR075W  125  VQSGKVRFEFRTASAQLEGGVNNLHSYEKRLHN 157
```

6-2-5- Alignement de séquences entre le gène *YLR432W* et le gène IMPDH1

```
YLR432W 21   DGLSVQELMDSKTRGGLTYNDFLVLPGLVDFPSSEVSLQTKLTRNITLNTPFVSSPMDTV 80
             DGL+ Q+L  S     LTYNDFL+LPG +DF + EV L + LTR ITL TP +SSPMDTV
IMPDH1  16   DGLTAQQLFAS--ADDLTYNDFLILPGFIDFIADEVDLTSALTRKITLKTPLISSPMDTV 73

YLR432W 81   TESEMAIFMALLGGIGFIHHNCTPEDQADMVRRVKNYENGFINNPIVISPTTTVGEAKSM 140
             TE++MAI MAL+GGIGFIHHNCTPE QA+ VR+VKN+E GFI +P+V+SP+ TVG+
IMPDH1  74   TEADMAIAMALMGGIGFIHHNCTPEFQANEVRKVKNFEQGFITDPVVLSPSHTVGDVLEA 133

YLR432W 141  KERFGFSGFPVTEDGKRNGKLMGIVTSRDIQFV--EDNSLLVQDVMTK--NPVTGAQGIT 196
             K R GFSG P+TE G    KL+GIVTSRDI F+  +D++ L+ +VMT     V    G+T
IMPDH1  134  KMRHGFSGIPITETGTMGSKLVGIVTSRDIDFLAEKDHTTLLSEVMTPRIELVVAPAGVT 193

YLR432W 197  LSEGNEILKKIKKGKLLIVDDNGNLVSMLSRTDLMKNQNYPLASKSATTKQLLCGAAIGT 256
             L E NEIL++ KKGKL IV+D   LV++++RTDL KN++YPLASK  + KQLLCGAA+GT
IMPDH1  194  LKEANEILQRSKKGKLPIVNDCDELVAIIARTDLKKNRDYPLASKD-SQKQLLCGAAVGT 252

YLR432W 257  IDADKERLRLLVEAGLDVVILDSSQGNSIFQLNMIKWIKETFPDLEIIAGNVATREQAAN 316
              + DK RL LL +AG+DV++  SSQGNS++Q+ M+ +IK+ +P L++I GNV T  QA N
IMPDH1  253  REDDKYRLDLLTQAGVDVIVFHSSQGNSVYQIAMVHYIKQKYPHLQVIGGNVVTAAQAKN 312

YLR432W 317  LIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVCEFANQFGIPCMADGGVQNIGH 376
             LI AG DGLR+GMG GSICITQEVMACGRPQGTAVY V E+A +FG+P +ADGG+Q +GH
IMPDH1  313  LIDAGVDGLRVGMGCGSICITQEVMACGRPQGTAVYKVAEYARRFGVPIIADGGIQTVGH 372

YLR432W 377  ITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGSIDAMQKTGTKGNASTS 436
             + KALALG+STVMMG +LA TTE+PGEYF+ DG RLK YRGMGS+D M+K+    ++S
IMPDH1  373  VVKALALGASTVMMGSLLAATTEAPGEYFFSDGVRLKKYRGMGSLDPMEKS----SSSQK 428

YLR432W 437  RYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGYKSLTLLKENVQSGKV 496
             RYFSE D V +AQGVSG++ DKGSI+KF+PYL  G+QH CQDIG +SL++L+  + SG++
IMPDH1  429  RYFSEGDKVKIAQGVSGSIQDKGSIQKFVPYLIAGIQHGCQDIGARSLSVLRSMMYSGEL 488

YLR432W 497  RFEFRTASAQLEGGVHNLHSYEKRLH 522
             +FE RT S Q+EGGVH LHSYEKRL+
IMPDH1  489  KFEKRTMSPQIEGGVHGLHSYEKRLY 514
```

6-2-6- Alignement de séquences entre le gène *YLR432W* et le gène IMPDH2

```
YLR432W 21   DGLSVQELMDSKTRGGLTYNDFLVLPGLVDFPSSEVSLQTKLTRNITLNTPFVSSPMDTV 80
             DGL+ Q+L +     GLTYNDFL+LPG +DF + +V L + LT+ ITL TP VSSPMDTV
IMPDH2  16   DGLTAQQLFN--CGDGLTYNDFLILPGYIDFTADQVDLTSALTKKITLKTPLVSSPMDTV 73

YLR432W 81   TESEMAIFMALLGGIGFIHHNCTPEDQADMVRRVKNYENGFINNPIVISPTTTVGEAKSM 140
             TE+ MAI MAL GGIGFIHHNCTPE QA+ VR+VK YE GFI +P+V+SP   V +
IMPDH2  74   TEAGMAIAMALTGGIGFIHHNCTPEFQANEVRKVKKYEQGFITDPVVLSPKDRVRDVFEA 133

YLR432W 141  KERFGFSGFPVTEDGKRNGKLMGIVTSRDIQFV--EDNSLLVQDVMTKNP--VTGAQGIT 196
             K R GF G P+T+ G+   +L+GI++SRDI F+  E++   ++++MTK    V   + IT
IMPDH2  134  KARHGFCGIPITDTGRMGSRLVGIISSRDIDFLKEEEHDCFLEEIMTKREDLVVAPRSIT 193

YLR432W 197  LSEGNEILKKIKKGKLLIVDDNGNLVSMLSRTDLMKNQNYPLASKSATTKQLLCGAAIGT 256
             L E NEIL++ KKGKL IV+++  LV++++RTDL KN++YPLASK A  KQLLCGAAIGT
IMPDH2  194  LKEANEILQRSKKGKLPIVNEDDELVAIIARTDLKKNRDYPLASKDA-KKQLLCGAAIGT 252

YLR432W 257  IDADKERLRLLVEAGLDVVILDSSQGNSIFQLNMIKWIKETFPDLEIIAGNVATREQAAN 316
             + DK RL LL +AG+DVV+LDSSQGNSIFQ+NMIK+IK+ +P+L++I GNV T  QA N
IMPDH2  253  HEDDKYRLDLLAQAGVDVVVLDSSQGNSIFQINMIKYIKDKYPNLQVIGGNVVTAAQAKN 312

YLR432W 317  LIAAGADGLRIGMGSGSICITQEVMACGRPQGTAVYNVCEFANQFGIPCMADGGVQNIGH 376
             LI AG D LR+GMGSGSICITQEV+ACGRPQ TAVY V E+A +FG+P +ADGG+QN+GH
IMPDH2  313  LIDAGVDALRVGMGSGSICITQEVLACGRPQATAVYKVSEYARRFGVPVIADGGIQNVGH 372

YLR432W 377  ITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGSIDAMQKTGTKGNASTS 436
             I KALALG+STVMMG +LA TTE+PGEYF+ DG RLK YRGMGS+DAM     K  +S +
IMPDH2  373  IAKALALGASTVMMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDAMD----KHLSSQN 428

YLR432W 437  RYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQHSCQDIGYKSLTLLKENVQSGKV 496
             RYFSE+D + VAQGVSGAV DKGSI KF+PYL  G+QHSCQDIG KSLT ++  + SG++
IMPDH2  429  RYFSEADKIKVAQGVSGAVQDKGSIHKFVPYLIAGIQHSCQDIGAKSLTQVRAMMYSGEL 488

YLR432W 497  RFEFRTASAQLEGGVHNLHSYEKRL 521
             +FE RT+SAQ+EGGVH+LHSYEKRL
IMPDH2  489  KFEKRTSSAQVEGGVHSLHSYEKRL 513
```

Bibliographie

Adams, A.K. & Holm, C. (1994). Specific DNA replication mutations affect telomere length in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**(9), 4614-20.

Ahmed, S., Kintanar, A. & Henderson, E. (1994). Human telomeric C-strand tetraplexes. *Nat. Struct. Biol.* **1**, 83-88.

Allison, A.C., Kowalski, W.J., Muller, C.J., Waters, R.V. & Eugui, E.M. (1993a). Mycophenolic acid and brequinar, inhibitors of purine and pyrimidine synthesis, block the glycosylation of adhesion molecules. *Transplant Proc.* **25**(3 Suppl 2), 67-70.

Allison, A.C., Kowalski, W.J., Muller, C.D. & Eugui, E.M. (1993b). Mechanisms of action of mycophenolic acid. *Ann. N. Y. Acad. Sci*. **696**, 63-87.

Allsopp, R. C., Chang, E., Kashefiaazam, M., Rogaev, E. I., Piatyszek, M. A., Shay, J. W. & Harley, C. B. (1995). Telomere shortening is associated with cell division *in vitr*o and *in vivo*. *Exp. Cell Res.* **220**(1), 194-200.

Ancelin, K., Brun, C. & Gilson, E. (1998). Role of the telomeric DNA-binding protein TRF2 in the stability of human chromosome ends. *Bioessays* **20**(11), 879-883.

Autexier, C. & Greider, C. W. (1996). Telomerase and cancer : revisiting the telomere hypothesis. *Trends Biochem. Sci.* **21**(10), 387-91.

Barton, A. B., Bussey, H., Storms, R. K. & Kaback, D. B. (1997). Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae* : Characterization of the 54 kb right terminal CDC15-FLO1-PHO11 region. *Yeast* **13**, 1251-1263.

Bateman, A. (1997). The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends Biochem. Sci.* **22**(1), 12-3.

Berger, I., Kang, C., Fredian, A., Ratliff, R., Moyzis, R. & Rich, A. (1995). Extension of the four-stranded intercalated cytosine motif by adenine.adenine base pairing in the crystal structure of d(CCCAAT). *Nat. Struct. Biol.* **2**, 416-425.

Bianchi, A., Smith, S., Chong, L., Elias, P. & de Lange, T. (1997). TRF1 is a dimer and bends telomeric DNA. *EMBO J.* **16**(7), 1785-94.

Bianchi, A. & de Lange, T. (1999). Ku binds telomeric DNA in vitro. *J. Biol. Chem.* **274**(30), 21223-21227.

Bilaud, T., Koering, C. E., Binetbrasselet, E., Ancelin, K., Pollice, A., Gasser, S. M. & Gilson, E. (1996). The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res.* **24**(7), 1294-1303.

Bilaud, T., Brun, C., Ancelin, K., Koering, C. E., Laroche, T. & Gilson, E. (1997). Telomeric localization of TRF2, a novel human telobox protein. *Nat. Genet.* **17**(2), 236-239.

Blackburn, E.H. & Gall, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J. Mol. Biol.* **120**(1), 33-53.

Blackburn, E. H. (1991). Structure and function of telomeres. *Nature* **350**, 569-573.

Blackburn, E.H. (1992). Telomerases. *Annu. Rev. Biochem.* **61**, 113-29.

Blackburn, E.H. (1994). Telomeres: no end in sight. *Cell* **77**, 621- 623.

Blackburn, G.M. & Gait, M.J. Editors (1990). *Nucleic Acids in Chemistry and Biology.* Oxford University Press. Oxford.

Bourns, B. D., Alexander, M. K., Smith, A. M. & Zakian, V. A. (1998). Sir proteins, rif proteins, and cdc13p bind saccharomyces telomeres In vivo. *Mol. Cell. Biol.* **18**, 5600-8.

Bradford, 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**, 248-54.

Cantor, C. R., Warshaw, M. M. & Shapiro, H. (1970). Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* **9**(9), 1059-1077.

Cardenas M.E., Bianchi A. & de Lange T. (1993). A Xenopus egg factor with DNA-binding properties characteristic of terminus-specific telomeric proteins. *Genes & Development* **7**, 883-894.

Chiu, C.P. & Harley, C.B. (1997). Replicative senescence and cell immortality: the role of telomeres and telomerase. *Proc. Soc. Exp. Biol. Med.* **214**(2), 99-106.

Chen, L., Cai, L., Zhang, X. H. & Rich, A. (1994). Crystal structure of a four-stranded intercalated DNA: d(C4). *Biochemistry* **33**, 13540-13546.

Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P. & de Lange, T. (1995). A human telomeric protein. *Science* **270**(5242), 1663-7.

Chou, S.H., Zhu, L. & Reid, B.R. (1994). The unusual structure of the human centromere $(GGA)^2$ motif. Unpaired guanosine residues stacked between sheared G·A pairs. *J. Mol. Biol.* **244**(3), 259-68.

Cockell M., Palladino F., Laroche T., Kyrion G., Liu C., Lustig A.J. & Gasser S.M. (1995). The carboxy termini of Sir4 and Rap1 affect Sir3 localization: evidence for a multicomponent complex required for yeast telomeric silencing. *J. Cell. Biol.* **129**(4), 909-24.

Colby, T.D., Vanderveen, K., Strickler, M.D., Markham, G.D., Goldstein, B.M. (1999). Crystal

structure of human type II inosine monophosphate dehydrogenase: implications for ligand binding and drug design. *Proc. Natl. Acad. Sci. USA* **96**(7), 3531-6.

Cong, Y. S., Wen, J. P. & Bacchetti, S. (1999). The human telomerase catalytic subunit hTERT: organization of the gene and characterization of the promoter. *Hum. Mol. Genet.* **8**(1), 137-142.

Conrad, M. N., Wright, J. H., Wolf, A. J. & Zakian, V. A. (1990). RAP1 protein interacts with yeast telomeres in vivo: overproduction alters telomere structure and decreases chromosome stability. *Cell* **63**(4), 739-50.

Coren J.S., Epstein E.M. & Vogt V.M. (1991). Characterization of a Telomere-Binding Protein from Physarum polycephalum. *Molecular & Cellular Biology* **11**, 2282-2290.

Counter, C.M. (1996). The roles of telomeres and telomerase in cell life span. *Mutat. Res.* **366**(1), 45-63.

Counter, C. M., Meyerson, M., Eaton, E. N. & Weinberg, R. A. (1997). The catalytic subunit of yeast telomerase. *Proc. Natl. Acad. Sci. USA* **94**(17), 9202-9207.

Csermely, P., Schnaider, T., Soti, C., Prohaszka, Z. & Nardai, G. (1998). The 90-kDa molecular chaperone family: structure, function, and clinical applications. A comprehensive review. *Pharmacol. Ther.* **79**(2), 129-68.

Dawson, J.H.J. & Guilhaus, M. (1989). *Rapid Commun. Mass Spectrom.* **3**, 155.

Diede, S. J. & Gottschling, D. E. (1999). Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. *Cell* **99**, 723-33.

Dionne, I. & Wellinger, R. J. (1996). Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase. *Proc. Natl. Acad. Sci. USA* **93**(24), 13902-13907.

Evans, S. K., Sistrunk, M. L., Nugent, C. I. & Lundblad, V. (1998). Telomerase, Ku, and telomeric silencing in Saccharomyces cerevisiae. *Chromosoma* **107**, 352-358.

Eversole, A. & Maizels,N. (2000). In vitro properties of the conserved mammalian protein hnRNP D suggest a role in telomere maintenance. *Mol. Cell. Biol.* **20**(15), 5425-32.

Fang, G. W. & Cech, T. R. (1993). The beta-Subunit of *Oxytricha* Telomere-Binding Protein Promotes G-Quartet Formation by Telomeric DNA. *Cell* **74**(5), 875-885.

Farazi, T., Leichman, J., Harris, T., Cahoon, M. & Hedstrom, L. (1997). Isolation and characterization of mycophenolic acid-resistant mutants of inosine-5'-monophosphate dehydrogenase. *J Biol. Chem.* **272**(2), 961-5.

Fitzgerald-Hayes, M., Clarke, L. & Carbon, J. (1982). Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* **29**(1), 235-44.

Frantz J.D. & Gilbert W., A novel yeast gene product, G4p1, with a specific affinity for quadruplex nucleic acids. *J. Biol. Chem.*, (1995), **270**, 20692-20697.

Fry, M. & Loeb, L. A. (1994). The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. USA* **91**(11), 4950-4.

Gallego, J., Chou, S.-H. & Reid, B.R. (1997). Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad : solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.* **273**(4), 840-856.

Gallego, J., Golden, E. B., Stanley, D. E. & Reid, B. R. (1999). The folding of centromeric DNA strands into intercalated structures: A physicochemical and computational study. *J. Mol. Biol.* **285**(3), 1039-1052.

Gehring, K., Leroy, J. L. & Guéron, M. (1993). A tetrameric structure with protonated cytosine.cytosine base pairs. *Nature* **363**, 561-565.

Gilson, E., Roberge, M., Giraldo, R., Rhodes, D., Gasser, S.M. (1993). Distortion of the DNA double helix by RAP1 at silencers and multiple telomeric binding sites. *J. Mol. Biol.* **231**(2), 293-310.

Gilson, E., Mullers, I., Sogo, J., Laroche, T. & Gasser, S. M. (1994). RAP1 stimulates single- to double-strand association of yeast telomeric DNA: Implications for telomere-telomere interactions. *Nucleic Acids Res.* **22**(24), 5310-5320.

Giraldo, R. & Rhodes, D. (1994). The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA. *EMBO J.* **13**(10), 2411-2420.

Giraldo, R., Suzuki, M., Chapman, L. & Rhodes, D. (1994). Promotion of parallel DNA quadruplexes by a yeast telomere binding protein: A circular dichroism study. *Proc. Natl. Acad. Sci. USA* **91**(16), 7658-7662.

Gottschling D.E., Aparicio O.M., Billington B.L. & Zakian VA. (1990). Position effect *at S. cerevisiae* telomeres : reversible repression of Pol II transcription. *Cell* **63**, 751-62.

Gottschling, D.E. & Stoddard, B. (1999). Telomeres: structure of a chromosome's aglet. *Curr. Biol.* **9**, R164-R167.

Grady, D.L., Ratliff, R.L., Robinson, D.L., McCanlies, E.C., Meyne, J. & Moyzis, R.K. (1992). Highly conserved repetitive DNA sequences are present at human centromeres. *Proc. Natl. Acad. Sci. U S A* **89**(5), 1695-9.

Greider, C.W. (1990). Telomeres, telomerase and senescence. *Bioessays* **12**(8), 363-9.

Greider, C.W. (1991). Telomeres. *Curr. Op. Cell Biol.* **3**, 441-451.

Greider, C. W. (1996). Telomere length regulation. *Annu. Rev. Biochem.* **65**, 337-365.

Greider, C.W. (1998). Telomeres and senescence: the history, the experiment, the future. *Curr Biol*,

8, R178-81.

Greider, C.W. (1999). Telomeres do D-loop-T-loop. *Cell* **97**(4), 419-22.

Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H. & de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell* **97**(4), 503-514.

Guéron, M., Plateau, P. & Decoprs, M. (1991). Solvent signal suppression in NMR. *Progress in NMR Spectroscopy* **23**, 135-209.

Guéron, M. & Leroy, J.L. (2000) The i-motif in nucleic acids. *Curr. Op. Str. Biol.* **10**, 326-331.

Hahn, W. C., Counter, C. M., Lundberg, A. S., Beijersbergen, R. L., Brooks, M. W. & Weinberg, R. A. (1999a). Creation of human tumour cells with defined genetic elements. *Nature* **400**(6743), 464-468.

Hahn, W. C., Stewart, S. A., Brooks, M. W., York, S. G., Eaton, E., Kurachi, A., Beijersbergen, R. L., Knoll, J. H. M., Meyerson, M. & Weinberg, R. A. (1999b). Inhibition of telomerase limits the growth of human cancer cells. *Nature Med.* **5**(10), 1164-1170.

Han, X., Leroy, J. L. & Guéron, M. (1998). An intramolecular i-motif: the solution structure and base pair opeing kinetics of d(5mCCTTTCCTTTACCTTTCC). *J. Mol. Biol.* **278**, 949-965.

Harley, C. B., Vaziri, H., Counter, C. M. & Allsopp, R. C. (1992). The Telomere Hypothesis of Cellular Aging. *Exp. Gerontol.* **27**(4), 375-382.

Holt, S.E., Shay, J.W. & Wright, W.E. (1996). Regulation of telomerase activity in immortal cell Lines. *Mol. Cell. Biol.* **16**, 2932-2939.

Holt, S.E. & Shay, J.W. (1999). Role of telomerase in cellular proliferation and cancer. *J. Cell. Phys.* **180**, 10-18.

Holt, S.E., Aisner, D.L., Baur, J., Tesmer, V.M., Dy, M., Ouellette, M., Trager, J.B., Morin, G.B., Toft, D.O., Shay, J.W., Wright, W.E. & White, M.A. (1999). Functional requirement of p23 and Hsp90 in telomerase complexes, *Genes Dev.* **13**, 817-26.

Huang, S., Elliott, R.C., Liu, P.S., Koduri, R.K., Weickmann, J.L., Lee, J.H., Blair, L.C., Ghosh-Dastidar, P., Bradshaw, R.A., Bryan, K.M., et al. (1987). Specificity of cotranslational amino-terminal processing of proteins in yeast. *Biochemistry* **26**(25), 8242-6.

Imai, S., Armstrong, C.M., Kaeberlein, M. & Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. **403**(6771), 795-800.

Ishikawa, F., Matunis, M. J., Dreyfuss, G. & Cech, T. R. (1993). Nuclear Proteins That Bind the Pre-Messenger RNA 3' Splice Site Sequence r(UUAG/G) and the Human Telomeric DNA Sequence d(TTAGGG)n. *Mol. Cell. Biol.* **13**(7), 4301-4310.

Ivy, J.M., Klar, A.J. & Hicks J.B. (1986). Cloning and characterization of four SIR genes of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **6**(2), 688-702.

Kang, C. H., Berger, I., Lockshin, C., Radliff, R., Moyzis, R. & Rich, A. (1994). Crystal structure of intercalated four stranded d(C3T) at 1.4 Å resolution. *Proc. Natl. Acad. Sci. USA* **91**, 11636-11640.

Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R. & Rich, A. (1995). A stable loop in the crystal structure of the intercalated four-stranded cytosine-rich metazoan telomere. *Proc. Natl. Acad. Sci. USA* **92**(9), 3874-3878.

Kipling, D. & Cooke, H.J. (1990) Hypervariable ultra-long telomeres in mice. *Nature* **347**(6291), 400-2.

Kipling, D. (1995). The Telomere. Oxford University Press.

Klebanow, E. R., Poon, D., Zhou, S. & Weil, P.A. (1996). Isolation and characterization of TAF25, an essential yeast gene that encodes an RNA polymerase II-specific TATA-binding protein-associated factor. *J. Biol. Chem.* **271**, 13706-15.

Kohler, G. A., White, T. C. & Agabian, N. (1997).Overexpression of a cloned IMP dehydrogenase gene of Candida albicans confers resistance to the specific inhibitor mycophenolic acid. *J. Bacteriol.* **179**, 2331-8.

Koll, H., Guiard, B., Rassow, J., Ostermann, J., Horwich, A. L., Neupert, W. & Hartl, F. U. (1992). Antifolding activity of hsp60 couples protein import into the mitochondrial matrix with export to the intermembrane space. *Cell* **68**, 1163-75.

Kota, R. S. & Runge, K. W. (1998). The yeast telomere length regulator TEL2 encodes a protein that binds to telomeric DNA. *Nucleic Acids Res.* **26**, 1528-35.

Kota, R. S. & Runge, K. W. (1999). Tel2p, a regulator of yeast telomeric length in vivo, binds to single-stranded telomeric DNA in vitro. *Chromosoma* **108**, 278-290.

Kyrion, G., Liu, K., Liu, C. & Lustig, A. J. (1993). RAP1 and Telomere Structure Regulate Telomere Position Effects in *Saccharomyces cerevisiae*. *Genes Dev.* **7**(7A), 1146-1159.

La Branche, H., Dupuis, S., Ben-David, Y., Bani, M. R., Wellinger, R. J. & Chabot, B. (1998). Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat. Genet.* **19**(2), 199-202.

Lacroix L., Lienard H., Labourier E., Djavaheri-Mergny M., Lacoste J., Leffers H., Tazi J., Helene C. & Mergny J.L. (2000). Identification of two human nuclear proteins that recognise the cytosine-rich strand of human telomeres in vitro. *Nucleic Acids Res.* **28**(7), 1564-75.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Landry, J., Sutton,A., Tafrov, S.T., Heller, R.C., Stebbins, J., Pillus, L. & Sternglanz, R. (2000). The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc. Natl. Acad. Sci. USA.* **97**(11), 5807-5811.

Lechner, J. (1994). A zinc finger protein, essential for chromosome segregation, constitutes a putative DNA binding subunit of the *Saccharomyces cerevisiae* kinetochore complex, Cbf3. *Embo Journal* **13**, 5203-11.

Leopardi, R. & Roizman, B. (1996). Functional interaction and colocalization of the herpes simplex virus 1 major regulatory protein ICP4 with EAP, a nucleolar-ribosomal protein. *Proc. Natl. Acad. Sci. USA* **93**, 4572-6.

Leroy, J.L., Kochoyan, M., Huynh-Dinh T. & Gueron M. (1988). Characterization of base-pair opening in deoxynucleotide duplexes using catalyzed exchange of the imino proton. *J. Mol. Biol.* **200**(2), 223-38.

Leroy, J. L., Guéron, M., Mergny, J. L. & Hélène, C. (1994). Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res.* **22**(9), 1600-1606.

Lin, J. J. & Zakian, V. A. (1996). The *Saccharomyces* CDC13 protein is a single-strand TG(1-3) telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. *Proc. Natl. Acad. Sci. USA* **93**(24), 13760-13765.

Lingner, J., Cooper, J.P. & Cech, T.R. (1995). Telomerase and DNA end replication: no longer a lagging strand problem? *Science* **269**(5230), 1533-4.

Lingner, J. & Cech, T. R. (1998). Telomerase and chromosome end maintenance. *Curr. Opin. Genet. Dev.* **8**(2), 226-32.

Liu, Y., Bohn, S.A. & Sherley, J.L. (1998). Inosine-5'-monophosphate dehydrogenase is a rate-determining factor for p53-dependent growth regulation. *Mol. Biol. Cell.* **9**(1), 15-28.

Liu, Z. P., Frantz, J. D., Gilbert, W. & Tye, B. K. (1993). Identification and Characterization of a Nuclease Activity Specific for G4 Tetrastranded DNA. *Proc. Natl. Acad. Sci. USA* **90**(8), 3157-3161.

Liu, Z. P. & Gilbert, W. (1994). The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex DNA: Implication of in vivo functions far this novel DNA structure. *Cell* **77**(7), 1083-1092.

Liu, Z. P., Lee A. & Gilbert W. (1995). Gene disruption of a G4-DNA-dependent nuclease in yeast leads to cellular senescence and telomere shortening. *Proc. Natl. Acad. Sci. USA* **92**, 6002-6006.

Lithgow, T., Glick, B. S. & Schatz, G. (1995). The protein import receptor of mitochondria. *Trends Biochem. Sci.* **20**, 98-101.

Loo S. & Rine J. (1995). Silencing and heritable domains of gene expression. *Annu. Rev. Cell. Dev. Biol.* **11**(5), 519-548.

Lundblad, V. (2000). DNA ends : maintenance of chromosome termini versus repair of double strand breaks. *Mutation Research* **451**, 227-240.

Makarov, V. L., Hirose, Y. & Langmore, J. P. (1997). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* **88**(5), 657-666.

Malinoski, F. & Stollar, V. (1981). Inhibitors of IMP dehydrogenase prevent sindbis virus replication and reduce GTP levels in *Aedes albopictus* cells. *Virology* **110**(2), 281-9.

Manzini, G., Yathindra, N. & Xodo, L.E. (1994). Evidence for intramolecularly folded i-DNA structures in biologically relevenat CCC-repeat sequences. *Nucleic Acids Res.* **22**(22), 4634-4640.

Marcand, S., Buck, S. W., Moretti, P., Gilson, E. & Shore, D. (1996). Silencing of genes at nontelomeric sites in yeast is controlled by sequestration of silencing factors at telomeres by Rap1 protein. *Genes Dev.* **10**(11), 1297-1309.

Marcand, S., Brun, C., Ancelin, K. & Gilson, E. (1997). Telomeres, normal and pathological aspects. *M. S. Med. Sci.* **13**(11), 1250-1258.

Marsich, E., Piccini A., Xodo L.E. & Manzini G. (1996). Evidence for a *HeLa* nuclear protein that binds specifically to the single-stranded d(CCCTAA)$_n$ telomeric motif. *Nucleics Acids Res.*, **24**, 4029-4033.

Marsich, E., Xodo L.E. & Manzini G. (1998). Widespread presence in mammals and high binding specificity of a nuclear protein that recognises the single-stranded telomeric motif (CCCTAA)n. *Eur J Biochem*, **258**, 93-99.

Martin, S. G., Laroche, T., Suka, N., Grunstein, M. & Gasser, S. M. (1999). Relocalization of telomeric Ku and SIR proteins in response to DNA strand breaks in yeast. *Cell* **97**, 621-633.

Masutomi, K., Kaneko, S., Hayashi, N., Yamashita, T., Shirota, Y., Kobayashi, K. & Murakami, S. (2000). Telomerase activity reconstituted in vitro with purified human telomerase reverse transcriptase and human telomerase RNA component. *J. Biol. Chem.* **275**(29), 22568-73.

Meyne, J., Ratliff, R.L. & Moyzis, R.K. (1989). Conservation of the human telomere sequence (TTAGGG)n among vertebrates. *Proc. Natl. Acad. Sci. U S A* **86**(18), 7049-53.

Mergny, J.-L., Lacroix, L., Han, X., Leroy, J.L. & Hélène, C. (1995). Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J. Am. Chem. Soc.* **117**, 8887-8898.

Mergny, J.-L. & Hélène, C. (1998). G-quadruplex DNA: A target for drug design. *Nature Med.* **4**(12), 1366-1367.

Mergny, J.-L. (1999). Fluorescence energy transfer as a probe for tetraplex formation: the i-motif. *Biochemistry* **38**, 1573-1581.

Mishra, K. & Shore, D. (1999). Yeast Ku protein plays a direct role in telomeric silencing and counteracts inhibition by Rif proteins. *Curr. Biol.* **9**, 1123-1126.

Miura, T., Benevides, J.M. & Thomas, G.J. (1995). A phase diagram for sodium and potassium ion control of polymorphism in telomeric DNA. *J. Mol. Biol.* **248**(2)**,** 233-238.

Moore H., Greenwell P.W., Liu C.P., Arnheim N. & Petes T.D. (1999). Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc. Natl. Acad. Sci. USA* **96**(4), 1504-9.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L. & Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U S A* **85**(18), 6622-6.

Muraiso, T., Nomoto, S., Yamazaki, H., Mishima, Y. & Kominami, R. (1992). A single-stranded DNA binding protein from mouse tumor cells specifically recognizes the C-rich strand of the (AGG:CCT)n repeats that can alter DNA conformation. *Nucleic Acids Res.* **20**(24), 6631-5.

Murray, A. W., Claus, T. E. & Szostak, J. W. (1988). Characterization of two telomeric DNA processing reactions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8**(11), 4642-50.

Murray P. Deutscher, Methods in Enzymology, Volume 182, Guide to protein purification, Academic Press 1990.

Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B. & Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277**(5328), 955-959.

Nonin, S., Phan, A. T. & Leroy, J.L. (1997). Solution structure and base pair opening kinetics of the i-motif dimer of d(5mCCTTTACC): a noncanonical structure with possible roles in chromosome stability. *Structure* **5**(9), 1231-1246.

Nonin-Lecomte, S. & Leroy J.L. (2001). Structure of a C-rich strand fragment of the human centromeric satellite III : a pH-dependent intercalation topology. *Article accepté : J. Mol. Biol. 2001*

Nugent, C.I. & Lundblad, V. (1998). The telomerase reverse transcriptase: components and regulation. Genes Dev. **12**(8), 1073-85.

Oka, Y., Shiota, S., Nakai, S., Nishida, Y. & Okubo, S. (1980). Inverted terminal repeat sequence in the macronuclear DNA of *Stylonichia pustulata. Gene* **10**, 301-306.

Okazaki, T. & Okazaki, R. (1969). Mechanism of DNA chain growth. IV. Direction of synthesis of T4 short DNA chains as revealed by exonucleolytic degradation. *Proc. Natl. Acad. Sci. U.S.A.* **64**(4), 1242-8.

Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. & Fujiwara, H. (1993). Identification of a pentanucleotide telomeric sequence, (TTAGG)n, in the silkworm *Bombyx mori* and in other insects. Mol Cell Biol. **13**(3), 1424-32.

Olovnikov, A. M. (1973). A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J. Theor Biol.* **41**(1), 181-190.

Olovnikov, A. M. (1996). Telomeres, telomerase, and aging: Origin of the theory. *Exp. Gerontol.* **31**(4), 443-448.

O'Reilly, M., Teichmann, S. A. & Rhodes, D. (1999). Telomerases. *Curr. Opin. Struct. Biol.* **9**(1), 56-65.

Patel, D.J., Bouaziz, S., Kettani, A. & Wang, Y. (1999). In Oxford Handbook of Nucleic Acid Structure, (ed. Neidle, S.), Oxford University Press, Oxford, pp. 389-453.

Perrot, M., Sagliocco, F., Mini, T., Monribot, C., Schneider, U., Shevchenko, A., Mann, M., Jeno, P. & Boucherie, H. (1999). Two-dimensional gel protein database of *Saccharomyces cerevisiae* (Update 1999). *Electrophoresis* **20**, 2280-2298.

Petracek, M.E., Lefebvre, P.A., Silflow, C.D. & Berman, J. (1990). *Chlamydomonas* telomere sequences are A+T-rich but contain three consecutive G-C base pairs. *Proc. Natl. Acad. Sci. U S A* **87**(21), 8222-6.

Petracek M.E., Konkel L.M.C., Kable M.L. & Berman J. (1994). A chlamydomonas protein that binds single-stranded G-strand telomere DNA. *The EMBO journal* **13**, 3648-3658.

Phan A.T. (1999). Structures et mouvements des acides nucléiques en solution: télomères et centromères ; eau et ions associés aux acides nucléiques ; méthodes RMN. Thèse de doctorat de l'Ecole Polytechnique.

Phan A.T., Guéron M. & Leroy J.L. (2000). The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* **299**(1), 123-44.

Poon, D., Bai, Y., Campbell, A. M., Bjorklund, S., Kim, Y. J., Zhou, S., Kornberg, R. D. & Weil, P. A. (1995).Identification and characterization of a TFIID-like multiprotein complex from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U S A* **92**, 8224-8.

Porter S.E., Greenwell P.W., Ritchie K.B. & Petes T.D. (1996). The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **24**, 582-585.

Price C.M. (1990). Telomere structure in Euplotes crassus: Characterization of DNA-Protein Interactions and Isolation of a Telomere-Binding Protein. *Molecular & Cellular Biology* **10**, 3421-3431.

Raghuraman M.K. & Cech T.R. (1989). Assembly and Self-Association of *Oxytricha* Telomeric Nucleoprotein Complexes. *CELL* **59**, 719-728.

Ramakrishnan, S., Sharma, H. W., Farris, A. D., Kaufman, K. M., Harley, J. B., Collins, K., Pruijn, G. J. M., vanVenrooij, W. J., Martin, M. L. & Narayanan, R. (1997). Characterization of human telomerase complex. *Proc. Natl. Acad. Sci. USA* **94**(19), 10075-10079.

Reading, D. S., Hallberg, R. L. & Myers, A. M. (1989). Characterization of the yeast HSP60 gene coding for a mitochondrial assembly factor. *Nature* **337**, 655-9.

Richards, E.J. & Ausubel, F.M. (1988). Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**(1), 127-36.

Robins, R. (1982). Nucleoside and nucleotide inhibitors of inosine monophosphate (IMP) dehydrogenase as potential antitumor inhibitors. *Nucleotides & Nucleotides*, **1**, 35-44.

Rosenblum, J. S., Pemberton, L. F., Bonifaci, N. & Blobel, G. Nuclear import and the evolution of a multifunctional RNA-binding protein. *J. Cell. Biol.* **143**, 887-899 (1998).

Runge, K. W. & Zakian, V. A. (1996). TEL2, an essential gene required for telomere length regulation and telomere position effect in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **16**, 3094-105.

Saenger, W. (1984). Principles of Nucleic Acid Structure. Springer-Verlag, New York.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular Cloning : A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sarig, G., Weisman Shomer, P. & Fry, M. (1997a). Telomeric and tetraplex DNA binding properties of qTBP42: A homologue of the CArG box binding protein CBF-A. *Biochem. Biophys. Res. Commun.* **237**(3), 617-623.

Sarig, G., Weisman Shomer, P., Erlitzki, R. & Fry, M. (1997b). Purification and characterization of qTBP42 a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes. *J. Biol. Chem.* **272**(7), 4474-4482.

Schierer, T. & Henderson, E. (1994). A Protein from *Tetrahymena* Thermophila That Specifically Binds Parallel-Stranded G4-DNA. *Biochemistry* **33**(8), 2240-2246.

Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its applications for meiosis. *Nature* **334**, 364-366.

Sen, D. & Gilbert, W. (1990). A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344**, 410-414.

Shampay, J., Szostak, J.W. & Blackburn, E.H. (1984). DNA sequences of telomeres maintained in yeast. *Nature* **310**(5973), 154-7.

Shay, J.W. (1997). Telomerase in human development and cancer. *J. Cell. Phys.* **173**, 266-270.

Sheng H., Hou Z., Schierer T., Dobbs D.L. & Henderson E. (1995). Identification and Characterization of a Putative Telomere, *Molecular & Cellular Biology* **15**, 1144-1153.

Sherley, J.L. (1991). Guanine nucleotide biosynthesis is regulated by the cellular p53 concentration. *J. Biol. Chem.* **266**(36), 24815-28.

Shore, D. (1998). Cellular senescence: lessons from yeast for human aging?,*Curr. Biol*. **8**, R192-5.

Smith, J.S., Brachmann,C.B., Celic, I., Kenna, M.A., Muhammad, S., Starai, V.J., Avalos, J.L., Escalante-Semerena, J.C., Grubmeyer, C., Wolberger, C. & Boeke, J.D. (2000). A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. USA.* **97**(12), 6658-6663.

Sokoloski, J.A., Blair, O.C. & Sartorelli, A.C. (1986). Alterations in glycoprotein synthesis and guanosine triphosphate levels associated with the differentiation of HL-60 leukemia cells produced by inhibitors of inosine 5'-phosphate dehydrogenase. *Cancer Res.* **46**(5), 2314-9.

Sorger, P. K., Doheny, K. F., Hieter, P., Kopski, K. M., Huffaker, T. C. & Hyman, A. A. (1995) Two genes required for the binding of an essential *Saccharomyces cerevisiae* kinetochore complex to DNA. *Proc. Natl. Acad. Sci. USA.* **92**, 12026-30.

Spelsberg, T.C. (1983). A rapid method for analysis of ligand binding to deoxyribonucleic acid and soluble nucleoproteins using streptomycin: application to steroid receptor ligands. *Biochemistry* **22**(1), 13-21.

Stavenhagen, J.B. & Zakian, V.A. (1998). Yeast telomeres exert a position effect on recombination between internal tracts of yeast telomeric DNA. *Genes & development*, **12**, 3044–3058.

Teng, S.C. & Zakian, V.A. (1999). Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**(12), 8083-93.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., QureshiEmili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M. J., Johnston, M., Fields, S. & Rothberg, J. M. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623-627.

van Steensel, B. & de Lange, T. (1997). Control of telomere length by the human telomeric protein TRF1. *Nature* **385**(6618), 740-743.

van Steensel, B., Smogorzewska, A. & de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**(3), 401-413.

Veselkov, A. G., Malkov, V. A., Frank-Kamenetskii, M. D. & Dobrynin, V. N. (1993). Triplex Model of Chromosome Ends. *Nature* **364**(6437), 496.

Virta-Pearlman, Morris, D.K. & Lundblad V. (1996). EST1 has the properties of a single -stranded telomere end-binding protein. *Genes & development* **10**, 3094-3104.

Wang, Y. & Patel, D. J. (1993). Solution Structure of the Human Telomeric Repeat d[AG3(T2AG3))3] G-Tetraplex. *Structure* **1**(4), 263-282.

Wang, Y. & Patel, D. J. (1994). Solution structure of the *Tetrahymena* telomeric repeat d(T(2)G(4))(4) G-tetraplex. *Structure* **2**(12), 1141-1156.

Watson, J. D. & Crick, F. H. C. (1953a). Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964-967.

Watson, J. D. & Crick, F. H. C. (1953b). Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* **171**, 737-738.

Weber, G. (1983). Biochemical strategy of cancer cells and the design of chemotherapy: G. H. A. Clowes Memorial Lecture. *Cancer Res.* **43**(8), 3466-92.

Weisman-Shomer P. & Fry, M. (1993). QUAD, a protein from Hepatocyte Chromatin That Binds Selectively to Guanine-rich Quadruplex DNA, *The Journal of Biology Chemistry* (1993), **268**, 3306-3312.

Wellinger, R.J., Wolf, A.J. & Zakian, V.A. (1993). *Saccharomyces* telomeres acquire single-strand TG1-3 tails late in S phase. *Cell* **72**(1), 51-60.

Wellinger, R. J., Ethier, K., Labrecque, P. & Zakian, V. A. (1996). Evidence for a new step in telomere maintenance. *Cell* **85**(3), 423-433.

Wellinger, R.J. & Sen D. (1997). The DNA structures at the ends of eukaryotic chromosomes. *Eur. J. Cancer* **33**, 735-749.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., Chu, A.M., Connelly, C., Davis, K., Dietrich, F., Whelen Dow, S., El Bakkoury, M., Foury, F., Friend, S.H., Gentalen, E., Giaever, G., Hegemann, J.H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J.L., Riles, L., Roberts, C.J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R.K., Véronneau, S., Voet, M., Volckaert, G., Ward, T.R., Wysocki, R., Yen, G.S., Yu, K., Zimmermann, K., Phillipsen, P., Johnston, M. & Davis, R.W. (1999). Functional Characterization of the S. cerevisiae Genome by Gene Deletion and Parallel Analysis. *Science* **285**, 901-906.

Wotton, D. & Shore, D. (1997). A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae. *Genes And Development* **11**, 748-60.

Xiang, B., Taylor, J.C. & Markham, G.D. (1996). Monovalent cation activation and kinetic mechanism of inosine 5'-monophosphate dehydrogenase. *J. Biol. Chem.* **271**(3), 1435-40.

Zakian, V.A. (1995). Telomeres: beginning to understand the end. *Science* **270**, 1601-1607.

Zakian, V.A. (1996). Structure, function, and replication of Saccharomyces cerevisiae telomeres. *Annu. Rev. Genet.* **30**, 141-72.

Zhang, R., Evans, G., Rotella, F.J., Westbrook, E.M., Beno, D., Huberman, E., Joachimiak, A., Collart, F.R. (1999). Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry* **38**(15), 4691-700.

Zhong Z., Shiue L., Kaplan S. & de Lange T. (1992). A mammalian Factor That Binds Telomeric TTAGGG Repeats in Vitro. *Molecular & Cellular Biology* **12**, 4834-4843.

Zhu, L., Chou, S.H. & Reid, B.R. (1995). The structure of a novel DNA duplex formed by human centromere d(TGGAA) repeats with possible implications for chromosome attachment during mitosis. *J. Mol. Biol.* **254**(4), 623-37.