HAL
open science

Étude structurale de la RNase Y, une endoribonucléase impliquée dans la dégradation des ARNm chez Bacillus subtilis

Pierre Hardouin

To cite this version:

Pierre Hardouin. Étude structurale de la RNase Y, une endoribonucléase impliquée dans la dégradation des ARNm chez Bacillus subtilis. Interactions entre organismes. Université Pierre et Marie Curie - Paris VI, 2017. Français. NNT : 2017PA066439 . tel-02385674

HAL Id: tel-02385674

https://theses.hal.science/tel-02385674v1

Submitted on 29 Nov 2019

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.

Université Pierre et Marie Curie

Ecole doctorale complexité du vivant (ED 515)

Laboratoire de chimie des processus biologiques (UMR 8229) – Collège de France

Étude structurale de la RNase Y, une endoribonucléase impliquée dans la dégradation des ARNm chez *Bacillus subtilis*

Par Pierre HARDOUIN

Thèse de doctorat spécialité Complexité du vivant

Dirigée par Dr. Béatrice Golinelli-Pimpaneau

Présentée et soutenue publiquement le 27 Novembre 2017

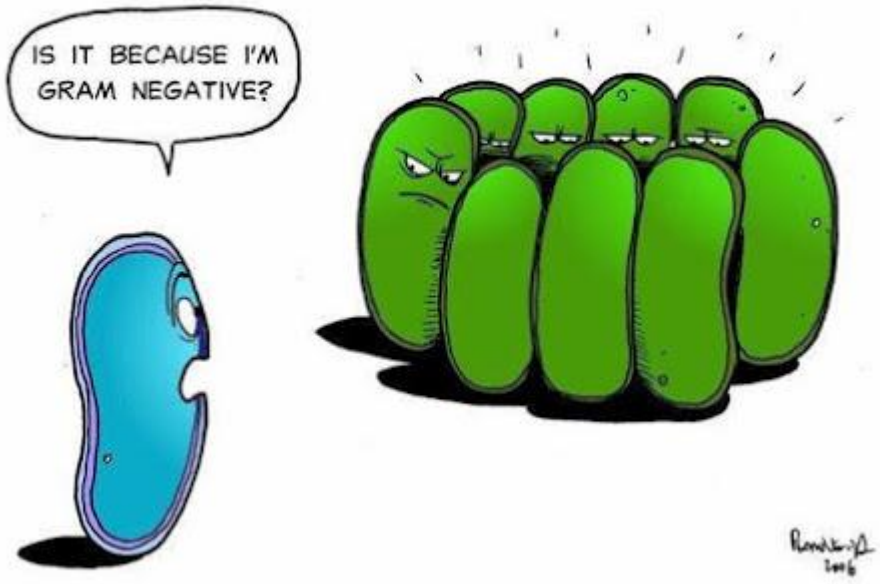
Devant un jury composé de :

Pr. Catherine VENIEN-BRYAN	Professeur (UPMC)	Présidente
Dr. Yves MECHULAM	DR1 - CNRS	Rapporteur
Dr. Philippe GIEGÉ	DR2 - CNRS	Rapporteur
Dr. Harald Putzer	DR1 - CNRS	Examinateur
Dr. Carine VAN-HEIJENOORT	DR2 - CNRS	Examinateur
Dr. Béatrice GOLINELLI-PIMPANEAU	DR2-CNRS	Directrice de thèse

Except where otherwise noted, this work is licensed under
http://creativecommons.org/licenses/by-nc-nd/3.0/

"Pour comprendre la réalité du fonctionnement cellulaire, agrandissons la cellule un milliard de fois jusqu'à ce qu'elle ait un diamètre de 20 km et ressemble à un vaisseau spatial qui aurait la superficie d'une grande ville comme Londres ou New York. Ce que nous verrions ce serait une structure d'une complexité et d'une faculté d'adaptation inégalable. A la surface du vaisseau nous verrions des millions d'ouvertures, comme les hublots d'un vaisseau, s'ouvrant et se fermant afin de permettre à un flot continu de matériel d'entrer et de sortir. Si nous entrions par une de ces ouvertures nous nous trouverions devant un monde d'une technologie et d'une complexité époustouflante."

Michael Denton (Biochimiste), Evolution : Une théorie en crise, *Burnett Books*, 1985



Malgré les différences existantes entre les bactéries à gram négatifs et positifs aux niveaux des fonctions physiologiques et de leurs niches écologiques, elles partagent de nombreuses similarités. C'est ce que nous allons voir pour le processus de dégradation des ARN et en particulier celui des ARNm.

Bonne lecture !

Remerciements

Ce travail de thèse a été réalisé au sein du laboratoire de chimie des processus biologiques du Collège de France dirigé par le Professeur Marc Fontecave et sous la direction de Béatrice Golinelli-Pimpaneau.

Je tiens tout d'abord à remercier Catherine Venien-Bryan d'avoir accepté d'être présidente de mon jury de thèse, ainsi que Philippe Giegé et Yves Mechulam pour avoir accepté d'être les rapporteurs de ce manuscrit. Je remercie également Harald Putzer et Carine Van-Heijnoort d'avoir bien voulu faire partie des membres du jury.

Je remercie également le Labex Dynamo et Francis-André Wollman pour m'avoir financé pendant ces quatres années mais également pour m'avoir permis d'effectuer une quatrième année de thèse qui a été la plus fructueuse en terme de résultats. Cela m'as permis d'obtenbir des publications, ce qui est capital dans la réussite d'une thèse.

Ensuite, je tiens à remercier Béatrice Golineli-Pimpaneau de m'avoir pris comme étudiant en thèse pour travailler sur ce sujet qui fut aussi difficile que passionnant. Je la remercie également pour tous les conseils ainsi que pour tout ce que j'ai apprit en 4 ans sur la biochimie et la structure des protéines. Je la remercie également pour son efficacité et sa patience concernant la correction de ma thèse qui a pris beaucoup de temps. Malgré la difficulté, pour le moement, ces quatre années ont été les plus enrichissantes et intéressantes de ma vie.

Je remercie également le professeur Marc Fontecave pour m'avoir accepté dans son laboratoire pendant 4 ans et m'avoir fait confiance.

Ensuite, je remercie également toute l'équipe du laboratoire pour leur bonne humeur et pour tous les bons moments passés ensemble malgré mon stress permanent à certains moments de ma thèse. Je remercie donc Thibault, Giorgio, Maria, Johnny, Ornella, Charles, Cecilia, Lucie, Myriam, Cameron, Tam, Grégoire, Tanya, Safwan, Xia, David, Huan, Philippe, Daniela. Je remercie également Bruno pour toutes les questions auxquelles il a bien voulu répondre pour la partie technique. Je remercie également les memebres de l'équipe du laboratoire pour la partie biologie, Djemel Hamdane, Murielle Lombard, Ludovic Pecqueur pour l'aide et les conseils qu'ils m'ont apporté.

Ensuite je remercie également le labortaoire d'expression génétique microbienne d'Harald Putzer, pour m'avoir accueillie de mars à juin 2014 afin d'effectuer quelques expériences. Je remercie donc également Harald Putzer et Soumaya Laalami pour leurs soutiens et pour toutes les discussions scientifiques et bons moments passés ensemble

Pour la partie expérimentale, je remercie tout particulièrement Christophe Velours pour toutes les analyses qu'il a effectué au cours de ce projet, pour son aide lors de la rédaction des articles mais également pour sa gentillesse et sa disponibilité.

Je remercie également Catherine Venien-Bryan pour la partie microscopie et d'avoir accepté de se lancer dans ce projet avec moi et Béatrice, sachant que ce n'est pas donné à tout le monde d'avoir accès à une cryo-microscope électronique. Merci beaucoup.

Un grand merci également à Carine Van-Heijenoort et Nadine Assrir pour les analyses RMN et tout ce que j'ai pu apprendre grâce à vous deux. J'ai vraiment beaucoup apprécié passer cette journée du 7 juillet 2017 à faire des analyses RMN dans votre laboratoire. Je commence à me prendre d'affection pour la RMN.

Je remercie également Charles Bou-Nader pour avoir analysé les données SAXS et RMN et d'avoir également prit de son temps pour rédiger les parties SAXS des articles.

Je remercie Johnny Tondeleir pour toute l'aide technique et pour les fous rires que l'on a pu avoir dans le labo.

Merci également à Aude Pouclet, stagiaire de BTS, pour son aide dans les derniers mois de cette thèse.

Également un grand merci à tout mes amis, en particulier à Jean-Baptiste et Laura sans qui je me serai senti bien seul dans cette ville. Merci pour toutes les soirées et les bons moments passés ensemble qui m'ont permis d'évacuer tout le stress de cette thèse. Je remercie également Jean-Yves pour son aide, sa patience et son soutien pendant ces quatres années.

Pour finir, un grand merci à toute ma famille et en particulier mes parents et mes deux frères qui m'ont soutenu pendant ces quatres ans.

Et je remercie également toutes les personnes que j'ai oublié dans ces remerciements mais à qui je pense.

Table des matières

Remerciements **6**

Table des matières **8**

Liste des abréviations **12**

INTRODUCTION **15**

I. Dégradation des ARNm par les Ribonucléases chez les bactéries **17**

I.A. Dégradation des ARNm : deux modèles bactériens différents, *Escherichia coli* et *Bacillus subtilis* 17

i. Les procaryotes au sein du monde vivant 19

ii. *Escherichia coli* et *Bacillus subtilis* : deux modèles d'étude pour la dégradation des ARNm 19

iii. Deux bactéries avec des ribonucléases différentes 23

iv. Identification et caractérisation enzymatique de la RNase E chez E.coli et des RNases J et Y chez B. subtilis 25

v. Mécanisme de clivage de la liaison phosphodiester au sein de l'ARN par les ribonucléases 35

I.B. Dégradation des ARNm : une stratégie conservée 41

i. Protection des ARNm contre la dégradation 41

ii. La dégradation des ARNm chez *E. coli* contrôlée par la RNase E 43

ii. La dégradation des ARNm chez *B. subtilis* : RNase Y vs RNase J 45

iv. Coexistence des endonucléases E, J et Y et évolution convergente 49

v. Dégradation des ARNm et régulation de l'expression des gènes 51

II. Aspects structuraux des ribonucléases d'*Escherichia coli* et de *Bacillus subtilis* **57**

II.A. Les RNase E et J, deux enzymes possédant des similitudes structurales 57

i. Organisation structurale de la RNase E: un domaine catalytique en N-terminal 57

ii. La RNase J, une dualité fonctionnelle expliquée par sa structure 63

iii. La RNase J possède une structure homologue à celle de la RNase E 71

II.B. Un complexe multiproétique macromoléculaire de dégradation des ARNm chez *E. coli*, le dégradosome 73

i. Découverte et caractérisation du complexe 73

ii. Caractérisation du domaine C-terminal intrinsèquement désordonné de la RNase E . 77

iii. Structures cristallographiques de la PNPase et de l'énolase en interaction avec les microdomaines du domaine C-terminal de la RNase E 83

iv. Les hélicases d'ARN de type DEAD-box et leurs rôles dans le dégradosome 87

v. Hypothèse sur l'organisation *in vivo* du degradosome d'*E. coli* 91
III. La RNase Y est l'homologue fonctionnel de la RNase E, mais qu'en est-il de son organisation structurale ? 93
III. A. Les protéines intrinsèquements désordonnées (IDP) : Des protéines pour les interactions moléculaires: 93
i. Qu'est-ce qu'une protéine intrinsèquement désordonnée ? 93
ii. Fonctions et répartition des protéines ID au sein du monde vivant 103
III.B. La RNase Y possède un domaine intrinsèquement désordonné 105
i. La RNase Y possède des caractéristiques fonctionnelles similaires avec la RNase E 105
ii. La RNase Y est une enzyme non essentielle mais importante pour la physiologie cellulaire. 113
III.C. Existence d'un dégradosome chez *B. subtilis* ? 117
i. Un postulat de départ 117
ii. Expériences de double hybride chez *E. coli* 117
iii. Expériences de co-purification in vivo 121
iv. Le dégradosome de Bacillus subtilis éxiste-t-il ? 123
III.D. La RNase Y est un facteur de virulence chez *Staphylococcus aureus* 125
i. Les bactéries font de la résistance 125
ii. De nouvelles cibles thérapeutiques 129
iii. La RNase Y de Staphylococcus aureus 131
RÉSULTATS 137
Introduction des articles 138
Article 1: Characterization of an Fab fragment that promotes ordering of the intrinsically disordered domain of *B. subtilis* RNase Y 140
Article 2: The N-terminal domain of *B. subtilis* RNase Y is a molten-globule 141
Article 3: The elongated dimer of *Bacillus subtilis* RNase Y assembles as a globular oligomer of high molecular weight 142
Annexe de l'introduction des articles 144
ARTICLES 149
Article 1 149
Article 2 171
Article 3 193
DISCUSSION et PERSPECTIVES 211
Bibliographie 219

ANNEXES **237**

Liste des abréviations

5’-UTR : 5’ untranslated region(region non traduite)
5'-triP: 5'-triphosphate
5'-monoP: 5'-monophosphate
a.a : acides aminés
ARN : Acide ribonucléique
ARNt : Acide ribonucléique de transfert
ARNr : Acide ribonucléique ribosomique
ARNm : Acide ribonucléique messager
B. subtilis : *Bacillus subtilis*
Ca: calcium
Co: cobalt
CaN: Calcineurine
CaM : Calmoduline
DNase : désoxyribonucléase
E. coli : *Escherichia coli*
Fab : Fragment antigen-binding
Fe : Fer
ID : intrinsèquement désordonné
IPTG: isopropyl β-D-1-thiogalactopyranoside
kDa: kiloDalton
MM : Masse moléculaire
Mg : Magnésium
Mn : Manganèse
Ni : Nickel
Pi : phosphate inorganique
Protéine ID : protéine intrinsèquement désordonnée
PNPase : Polynucléotide phosphorylase
pb : paire de base
RNase : ribonucléase
RNase : Ribonucléase
RppH : pyrophosphohydrolase
SAM : S-adénosyl-L-méthionine
TM : transmembranaire
Zn : Zinc

INTRODUCTION

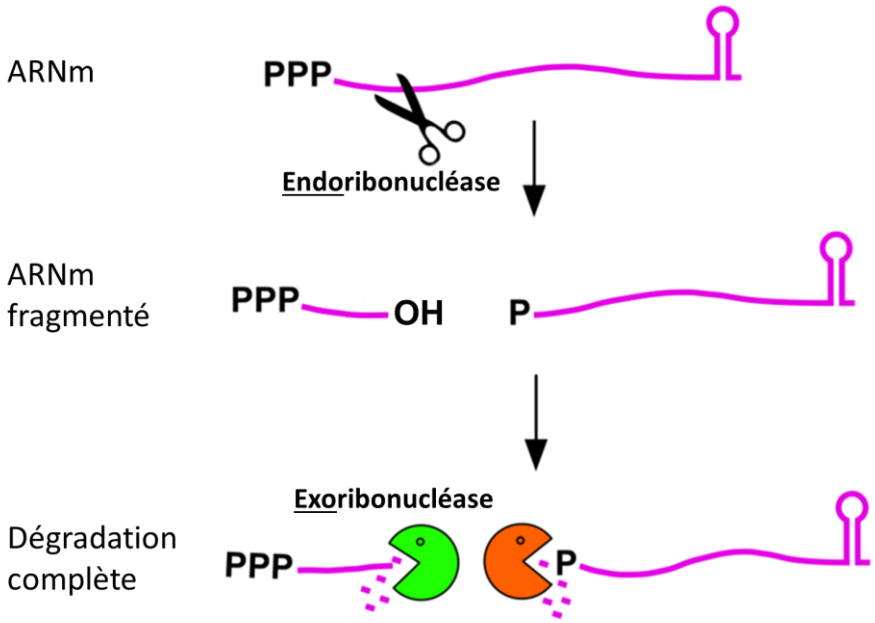


Figure 1 : Comparaison de l'activité catalytique des endoribonucléases et des exoribonucléases. Les endonucléases ont un rôle de fragmentation de l'ARNm via leur activité endonucléolytique. En effectuant un clivage interne au sein de l'ARNm, elles vont créer des fragments d'ARN qui sont substrats pour les exonucléases qui vont dégrader ces fragments en mononucléotides en attaquant les extrémités 3' et 5' (d'après Laalami *et al.*, 2013)

I. Dégradation des ARNm par les Ribonucléases chez les bactéries

I.A. Dégradation des ARNm : deux modèles bactériens différents, *Escherichia coli* et *Bacillus subtilis*

Dans cette introduction, nous allons nous intéresser à la dégradation des ARNm chez les bactéries ainsi qu'aux enzymes impliquées dans cette dégradation en nous focalisant sur les aspects structuraux et mécanistiques. La dégradation des ARNm ayant principalement été étudiée chez deux modèles bactériens, *Escherichia coli* (*E. coli)* et *Bacillus subtilis* (*B. subtilis*), nous verrons d'abord comment se fait cette dégradation chez ces deux bactéries et quelle est son importance dans la régulation de l'expression des gènes lors des divers stress auxquels les bactéries sont soumises. Nous détaillerons ensuite les enzymes responsables de l'initiation de cette dégradation : les endoribonucléases E, J et Y permettent de fragmenter les ARNm afin de générer des fragments d'ARN substrats pour une autre classe de ribonucléases (RNases), les exoribonucléases qui agissent à partir des extrémités 3' ou 5' (ex : PNPase, RNase II, RNase R,…) (Figure 1). Nous verrons comment les structures des RNases E et J ont permis d'obtenir des informations quant aux mécanismes de reconnaissance et de clivage interne de l'ARN. Puis nous rappellerons l'ensemble des études réalisées sur la RNase Y, l'endoribonucléase principale chez *B. subtilis* dont la structure n'est pas connue et nous verrons pourquoi il est primordial d'obtenir des informations structurales sur cette RNase. Des études indiquent que les RNases Y et J font partie d'un complexe macromoléculaire de dégradation des ARNm, le dégradosome. Alors que ce complexe a été largement étudié chez *E. coli*, chez *B. subtilis*, son existence reste encore à démontrer. Pour finir, nous verrons que la RNase Y joue un rôle important dans la virulence de certains pathogènes humains, comme le Staphylocoque doré et que, de manière générale, les RNases sont des cibles thérapeutiques potentielles pour de nouveaux antibiotiques qui pourraient permettre de lutter efficacement contre ces pathogènes qui sont de plus en plus résistants aux traitements actuels.

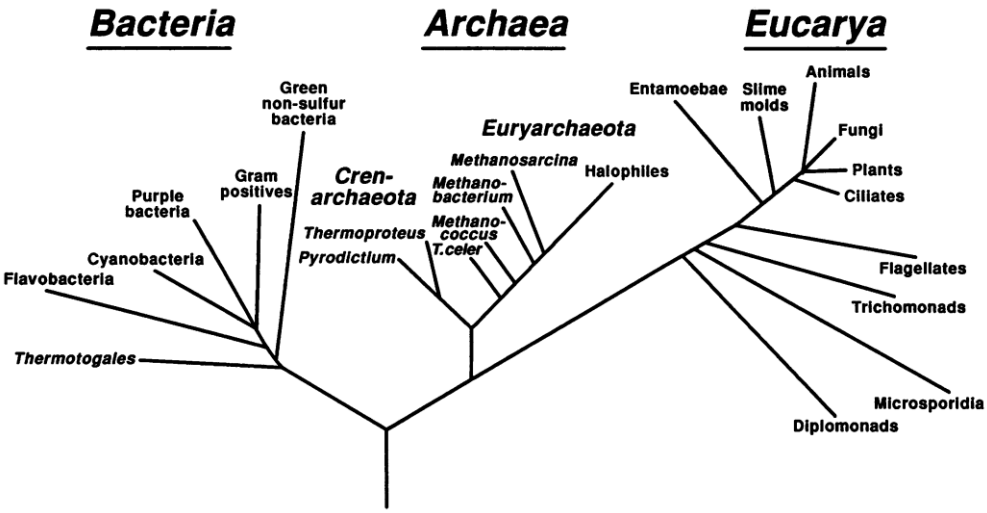


Figure 2: Arbre phylogénétique des trois règnes du vivant. Les archées et bactéries possèdent respectivement leur propre branche au sein de cet arbre phylogénétique. La relation phylogénétique entre les différents organismes a été calculé en comparant les séquences des ARN ribosomiques 16S (ARNr 16S). (d'après Woese, 1994).

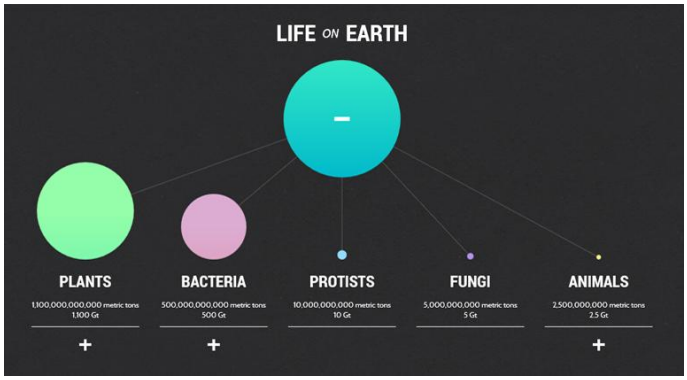


Figure 3 : Répartition de la biomasse en Gigatonnes au sein des êtres vivants. Les archées et bactéries représentent, avec les plantes, la plus grosse partie de la biomasse terrestre (Site internet n°1)

i. Les procaryotes au sein du monde vivant

Les procaryotes (êtres vivants ne possédant pas de noyau), partie intégrante de la biomasse terrestre, rassemblent les archées et les bactéries qui forment deux des trois branches de l'arbre phylogénétique du domaine vivant[1] (figure 2) et représentent, à elles seules, 30,9 % de la biomasse terrestre derrière les plantes (68 %), et devant les protistes (0,62%), les champignons (0,31%) et enfin les animaux et insectes (0,15%) (Figure 3) (Site internet n°1). Les bactéries sont présentes dans tous les environnements : le sol, les eaux (océans, mers, lacs, rivières, sources thermales,...) ainsi qu'au sein des êtres vivants (humains, animaux, ...)[2].

ii. *Escherichia coli* et *Bacillus subtilis* : deux modèles d'étude pour la dégradation des ARNm

Deux modèles principaux bactériens sont utilisés pour les études des processus cellulaires bactériens, comme le métabolisme des ARNm : *Escherichia coli* appartenant à l'embranchement des *Proteobacteria* et à la classe des γ-Protéobactéries et *Bacillus subtilis* appartenant à l'embranchement des *Firmicutes* et à la classe des *Bacilli*. *Escherichia coli* est une bactérie à Gram négatif, présente majoritairement dans le tube digestif des animaux. Cet organisme a d'abord été utilisé pour étudier le processus de conjugaison bactérienne[3] et est à l'heure actuelle un modèle bactérien très largement utilisé en recherche. *Bacillus subtilis* est une bactérie à Gram positif présente dans le sol et capable d'entrer en symbiose avec les racines des plantes. Cette bactérie est donc un excellent modèle pour étudier certains processus cellulaires absents chez *E. coli*, comme la sporulation ou la formation de biofilms[4,5]. Cet organisme est également utilisé pour mieux comprendre le fonctionnement de certains pathogènes humains à Gram positif comme *Staphylococcus aureus*, *Streptococcus pyogenes*, *Bacillus anthracis* (responsable de la maladie du charbon)[6–8] ou encore le genre bactérien *Lysteria*, comme *Lysteria monocytogenes* responsable d'intoxications alimentaires. L'étude de ce modèle devrait permettre de répondre, dans l'avenir, à certaines questions telles que : Quelles pourraient être les nouvelles stratégies thérapeutiques pour combattre ces pathogènes ? Comment ces pathogènes se protègent-t-ils grâce à leur mécanisme de sporulation et de formation de biofilm ? Quel est le mécanisme de résistance à certains antibiotiques, en particulier pour le Staphylocoque doré résistant à la méthicilline (SARM) qui devient multi-résistant et pose de nombreux problèmes dans les hôpitaux[6].

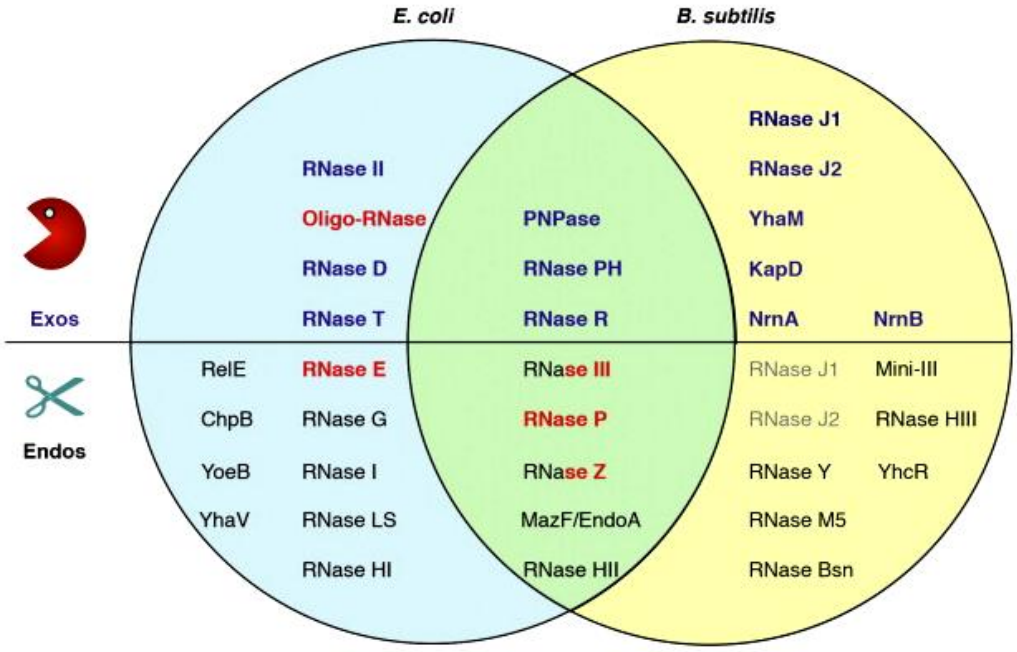


Figure 4 : Endoribonucléases (en bas) et exoribonucléases (en haut) identifiées chez les deux organismes modèles *Escherichia coli* et *Bacillus subtilis*. Certaines de ces enzymes sont présentes chez les deux organismes tandis que certaines sont propres aux bactéries à Gram positif et d'autres aux bactéries à Gram négatif (d'après Condon, 2016)

E. coli et *B. subtilis* vivent dans des environnements complètement différents qui nécessitent des fonctions physiologiques distinctes pour y vive et s'y adapter. De plus, près de trois milliards d'années d'évolution séparent ces deux organismes[9] ce qui permet de penser que des différences existent en ce qui concerne la dégradation des ARNm chez ces deux organismes. Une première différence majeure concernant le métabolisme des nucléotides et de l'ARN a été observée au début des années 70[10].

Une étude a révélé que l'utilisation d'eau marquée ^{18}O conduit à une forte incorporation d' ^{18}O dans les acides nucléiques chez *E. coli* signifiant que le métabolisme des acides nucléiques fait appel à un mode de synthèse et de dégradation hydrolytique[10]. A l'inverse, chez *Bacillus subtilis*, cette incorporation était très faible, indiquant un mécanisme non-hydrolytique[11]. A cette époque, la polynucléotide phosphorylase (PNPase) et la RNase II était les deux enzymes connues fonctionnant respectivement de façon phosphorolytique et hydrolytique. Des tests enzymatiques ont été réalisés à partir d'extraits cellulaires provenant de quatre souches : trois d'*E. coli* (sauvage, PNPase$^-$ et RNase II$^-$) et la souche sauvage de *B. subtilis*, en utilisant un polynucléotide marqué au tritium ([^{3}H]polyA). La comparaison de la quantité de mononucléotides relargués en absence ou présence de phosphate inorganique a montré que, dans la souche sauvage d'*E. coli*, la dégradation hydrolytique était prédominante alors que, chez *B. subtilis*, la dégradation hydrolytique était négligeable par rapport à la dégradation phosphorolytique[12]. De plus, la comparaison de l'activité mesurée pour la souche PNPase – par rapport à la souche RNase II- d'*E. coli* indique que c'est la RNase II qui est responsable de la dégradation de la majeure partie du poly(A) (>90%). Cette étude a suggéré que la différence de mécanisme de dégradation des ARN pour ces deux organismes est due au fait qu'ils vivent dans des milieux différents qui ne dépendent pas des mêmes ressources énergétiques. En effet, la dégradation phosphorolytique, qui permet de relarguer des nucléotides diphosphates et de conserver l'énergie de la liaison phosphodiester, serait présente chez *B. subtilis* qui vit dans le sol où les conditions sont plus difficiles et les ressources énergétiques plus rares. A l'inverse, le mécanisme hydrolytique serait prédominant chez *E. coli*, qui vit dans le tube digestif des animaux ou les sources d'énergie sont abondantes, ce qui ne nécessite pas de conserver l'énergie de la liaison phosphodiester[12]. Cette étude a posé la question de l'influence du milieu environnemental, dans lequel vit la bactérie, sur le mode de dégradation des ARNm.

Tableau 1 : Liste non-exhaustive des RNases impliquées dans la dégradation et la maturation des ARNm, ARNt et ARNr chez *E. coli* et *B. subtilis*. Toutes ces RNases utilisent des cofacteurs métalliques pour leurs activités catalytiques[13–27].

Classe	Enzyme	Fonction	Cofacteurs métalliques	Coordination	Type d'enzyme	Organisme	Références
Endoribonucléases	**RNase E**	Dégradation des ARNm Maturation : ARNr 16S, 5S et ARNt	Un Mg^{2+}	2x Asp	Hydrolytique	*E. coli*	Callaghan *et al.*, 2005 Mackie, 1998 Li et Deutscher, 2002
	RNase J1/J2	Dégradation des ARNm Maturation de l'ARNr 16S	Deux Zn^{2+}	6x His / 2x Asp	Hydrolytique	*B. subtilis*	Britton *et al.*, 2007 Even et al., 2005 Li de la Sierra-Gallay *et al.*, 2008
	RNase Y	Dégradation des ARNm	Un Mg^{2+}, Mn^{2+} ou Zn^{2+} Indéterminé	2x His / 2x Asp	Hydrolytique	*B. subtilis*	Shahbabian *et al.*, 2009
	RNase III	Maturation : ARNr 16S, 23S Dégradation des ARNm (ARNdb)	Deux Mg^{2+}	2x Asp / 2x Glu	Hydrolytique	*E. coli / B. subtilis*	Court *et al.*, 2013
	RNase P	Maturation: extrémité 5' des ARNt	Un Mg^{2+}	Ribonucléotides	Hydrolytique	*E. coli / B. subtilis*	Kazantsev *et al.*, 2005
	RNase Z	Maturation: extrémité 3' des ARNt	Deux Zn^{2+}	6x His / 2x Asp	Hydrolytique	*E. coli / B. subtilis*	Pellegrini *et al.*, 2012

Classe	Enzyme	Fonction	Cofacteurs métalliques	Coordination	Type d'enzyme	Organisme	Références
Exoribonucléases	**PNPase**	Exonucléase **3'=>5'** Dégradation des ARNm Inhibition : structures 2^{daires} des ARN	Mg^{2+} nombre indéterminé	2x Asp	Phosphorolytique	*E. coli / B. subtilis*	Symmons *et al.*, 2000
	RNase PH	Exonucléase **3'=>5'** Maturation des ARNt Dégradation des ARNm (*B. subtilis*)	Mg^{2+} nombre indéterminé	?	Phosphorolytique	*E. coli / B. subtilis*	Harlow *et al.*, 2003
	RNase J1/J2	Exonucléase **5'=>3'** Dégradation des ARNm	Deux Zn^{2+}	6x His / 2x Asp	Hydrolytique	*B. subtilis*	Deikus *et al.*, 2008
	RNase D	Exonucléase **3'=>5'** Maturation des ARNt (3'>5')	Un Zn^{2+} et un Mn^{2+}	3x Asp / 1x Glu / 1x Tyr	Hydrolytique	*E. coli*	Zuo *et al.*, 2006
	RNase II	Exonucléase **3'=>5'** Dégradation des ARNm	Un Mg^{2+}	2x Asp	Hydrolytique	*E. coli*	Frazao *et al.*, 2006
	RNase R	Exonucléase **3'=>5'** Dégradation: structures 2^{daires} (ARNm)	Un Mg^{2+}	?	Hydrolytique	*E. coli / B. subtilis*	Vincent et Deutscher, 2009
	RNase T	Exonucléase **3'=>5'** Maturation : ARNr et ARNt	Deux Mg^{2+}	3x Asp / 1x Glu	Hydrolytique	*E. coli*	Hsiao *et al.*, 2011

iii. Deux bactéries avec des ribonucléases différentes

A l'heure actuelle, 21 ribonucléases ont été identifiées chez *E. coli* et 18 chez *B. subtilis*[28,29]. L'ensemble de ces RNases sont impliquées dans la maturation et la dégradation des ARN cellulaires (ARNm, ARNr, ARNt). Nous n'évoquerons pas ici les petits ARN (ARNp) et ne détaillerons pas toutes les RNases et leurs implications dans le métabolisme des acides nucléiques, mais seulement les principales. Parmi ces ribonucléases identifiées, seulement six sont communes à *E. coli* et *B. subtilis* et, de manière générale, à toutes les eubactéries[29]. Pour les endoribonucléases, il s'agit de la RNase III, la RNase P et la RNase Z. Ces trois enzymes ont la particularité d'être impliquées dans la maturation des ARNr et ARNt, qui sont présents chez toutes les bactéries ce qui explique leurs conservations[29]. Les exoribonucléases communes à *E. coli* et *B. subtilis* sont la PNPase, la RNase R et la RNase PH[29]. La PNPase est une enzyme très conservée[30] chez les bactéries, mais également présente chez certains organismes eucaryotes[31]. Les RNase R et PH sont également impliquées respectivement dans la dégradation des ARNr et la maturation des ARNt[22,23].

En conclusion, les enzymes responsables de la dégradation des ARN sont différentes chez *E. coli* et *B. subtilis* en particulier pour les endonucléases. Au niveau des exonucléases d'ARNm, la RNase II, une enzyme hydrolytique, est absente chez *B. subtilis* et remplacée, lors de la dégradation des ARNm, par une enzyme phosphorolytique, la PNPase. De même, *B. subtilis* utilise la RNase PH, une enzyme phosphorolytique, à la fois pour dégrader les ARNm et maturer les ARNt, tandis que chez *E. coli* elle n'est impliquée que dans la maturation des ARNt. Nous allons maintenant nous intéresser plus particulièrement aux endonucléases principales responsables de l'initiation de la dégradation des ARNm présentes chez ces deux organismes.

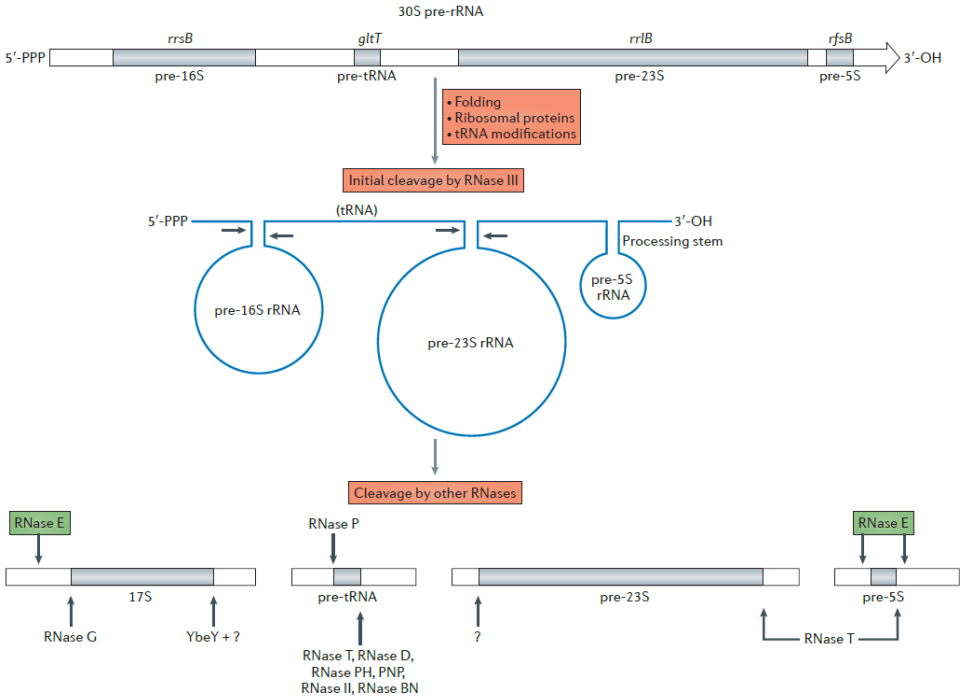


Figure 5: Maturation des ARNr chez *E. coli* par différentes RNases. La RNase E, représentée par un rectangle vert, est responsable de la maturation de l'ARNr 5S mais également de l'ARNr 16S (d'après Mackie, 2013).

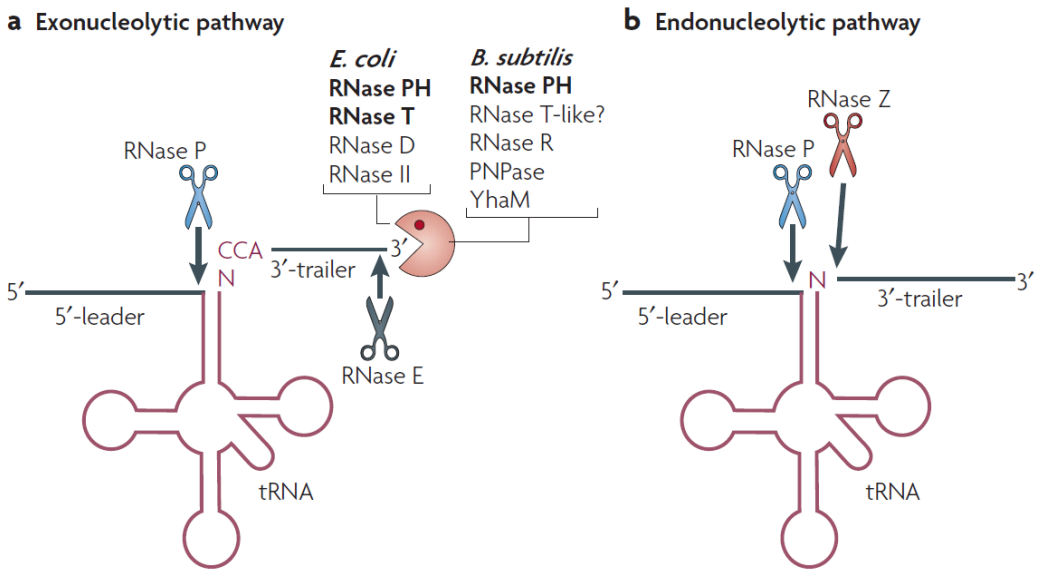


Figure 6: Implication de la RNase E dans la maturation des ARNt chez *E. coli*. La RNase E est désignée par une paire de ciseaux bleu foncé (d'après Redko et al., 2007)

iv. Identification et caractérisation enzymatique de la RNase E chez E.coli et des RNases J et Y chez B. subtilis

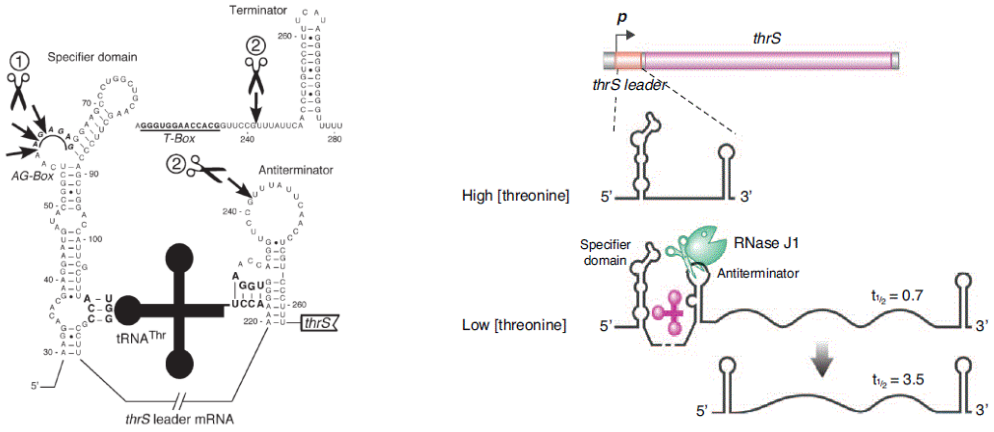
La dégradation des ARNm chez les bactéries a tout d'abord été très étudiée chez *E. coli,* avec l'étude de la RNase E, puis chez *B. subtilis* avec les découvertes des RNases J et Y.

La RNase E d'Escherichia coli

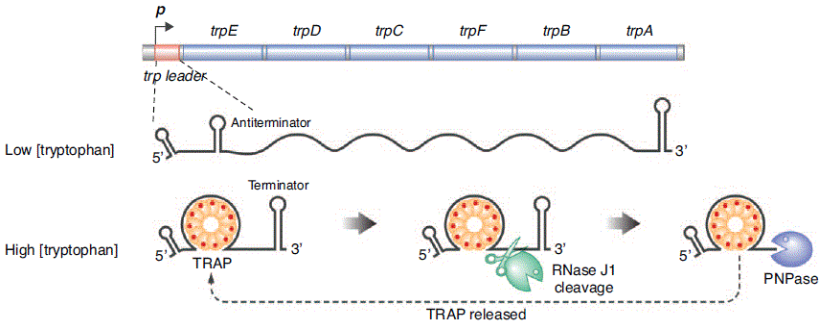
C'est en 1977 que l'endoribonucléase principale d'*E. coli*, la Ribonucléase E (RNase E), responsable de l'initiation de la dégradation des ARNm fut découverte[32]. Son activité fut mise en évidence en étudiant la maturation des ARN ribosomiques (ARNr)[32] en utilisant des souches isogéniques RNase III+ et RNase III-, la RNase III était déjà connue pour être une des endonucléases responsables de la maturation des ARNr [33]. Par Northern blot, l'équipe de David Apirion a montré que la souche RNase III- possédait une autre enzyme capable d'effectuer la maturation du pré-ARNr 5S (Figure 5). Cette enzyme a ensuite été identifiée et renommée RNase E[32]. Il a ensuite été montré qu'une inactivation thermique de la RNase E augmentait la demie-vie des ARNm globaux d'un facteur 5 (passant de 2,5min à plus de 10 min) [34]. Cette expérience indiquait donc que la RNase E était l'enzyme principale responsable de l'initiation de la dégradation des ARNm chez *E. coli*[35]. Après cela, de nombreuses autres études ont permis de caractériser le mécanisme par lequel la RNase E permet la dégradation des ARNm: la RNase E clive dans des régions riches en nucléotides A et U (AU-rich) et simple brin [36]; son activité *in vitro* est 10 fois plus efficace en présence d'un ARNm portant une extrémité 5' monophosphate (5'-monoP) plutôt qu'une extrémité 5' triphosphate (5'-triP)[36]. Pour finir, la RNase E est une RNase essentielle chez *E. coli*. En effet, ceci est due a son activité de maturation des ARN de transfert (ARNt). En absence de RNase E, les ARNt ne sont plus maturés[37,38] (figure 6).

Avec les progrès de la biologie moléculaire dans les années 90 de nombreux génomes ont pu être séquencés dont celui de l'homme, d' *E. coli* et de *B. subtilis* (ces derniers réalisés en 1997)[39,40]. Les résultats de séquençage du génome de *B. subtilis* ont montré que le gène codant pour la RNase E était absent chez *B. subtilis* ainsi que celui codant pour la RNase II, l' exoribonucléase, responsable majoritairement de la dégradation des fragments d'ARNm générés par la RNase E (tableau 1). La dégradation des ARNm était donc catalysée par d'autres enzymes. Les résultats ont conduit les scientifiques à rechercher les enzymes responsables de la dégradation des ARNm chez *B. subtilis*.

A



B



C

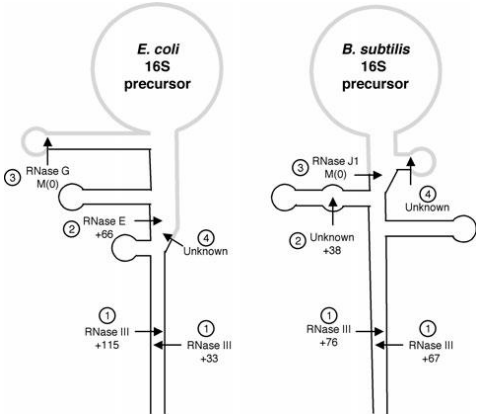


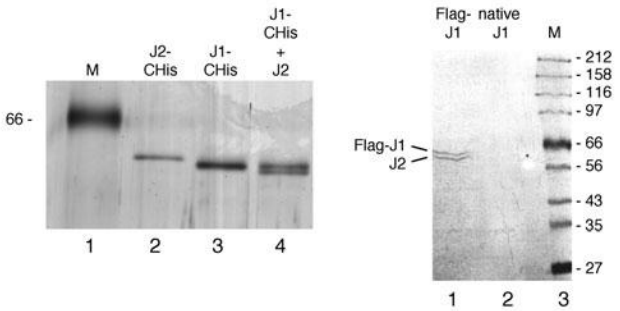
Figure 7 : Implication de la RNase J dans la maturation et la dégradation de certains transcrits chez *B. subtilis* A. Maturation du transcrit *thrS*, le clivage endonucléolytique par la RNase J permet d'augmenter la stabilité du transcrit (d'après Even *et al.*, 2005) **B.** Implication de la RNase J dans la dégradation de l'opéron tryptophane (d'après Condon et Bechhofer, 2011) **C.** Implication de la RNase J dans la maturation de l'ARNr 16S, les sites de coupures sont similaires à ceux de la RNase E chez *E. coli* (d'après Britton *et al.*, 2007)

La RNase J de *Bacillus subtilis*, une enzyme avec une double activité

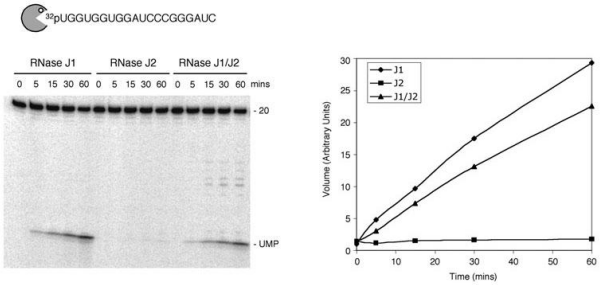
La RNase J fut identifiée en 1997 dans l'équipe de Marianne Grunberg-Manago à l'institut de biologie physico-chimique (IBPC) de Paris. En effet, en étudiant le transcrit codant l'amino acyl-tRNA synthétase (gène *thrS*) de *B. subtilis*, ils ont détecté par Northern blot que ce transcrit était clivé par une activité endonucléolytique chez *B. subtilis*[41]. Ils ont alors cloné le gène *thrS* de *B. subtilis* dans un plasmide, puis ce plasmide a été inséré chez *E. coli*. Un fois transcrit chez *E.coli*, ils observaient un clivage endonucléolytique catalysée par la RNase E. Les sites de clivages par la RNase E et par l'endonucléase de *B. subtilis* étaient identiques. Il a donc été conclu qu'il existait une RNase chez *B. subtilis* portant la même activité que la RNase E[42]. Cette RNase a ensuite été identifiée et fut alors considérée comme l'homologue fonctionnel de la RNase E chez *Bacillus subtilis*. Deux gènes paralogues ont été identifiés, *ykqC* (rebaptisée RNase J1) et *ymfA* (rebaptisée RNase J2). Plus tard, il fut également montré que la maturation de l'opéron *trp* chez *B. subtilis* faisait appel à l'activité endonucléase de la RNase J[43] mais également à une activité 5' exonucléase portée également par la RNase J[43]. Ce fut alors la première enzyme ayant une activité exonucléase en 5' identifiée chez les bactéries. Cette activité étant absente chez *E. coli.* La RNase J est donc la première enzyme possédant une activité double, endonucléase et exonucléase, à avoir été identifiée chez les bactéries. Il existe d'autre enzyme portant deux activités, exo et endonucléase, comme la sous-unité Rrp44p de l'exosome humain, un complexe de dégradation des ARNm[44].

Une autre étude a également montré *in vitro* que cette enzyme était impliquée dans la maturation de l'ARNr 16S et que les sites de clivages endonucléolytiques de la RNase J étaient similaires à ceux de la RNase E[45,46]. Cela conforta l'idée que la RNase J était l'endonucléase principale chez *B. subtilis.* Cependant, une étude ultérieure a montré que la demie-vie des ARNm globaux d'une souche dépleté en RNase J n'était que très faiblement augmentée par rapport à la souche sauvage, passant de 2,6 min à 3,6 min[47]. Il a donc été suggéré que la RNase J n'était sans doute pas la seule RNase responsable de l'initiation de la dégradation des ARNm chez *B. subtilis*.

A



B



C

Enzyme	V_{max} (fmol s^{-1})	K_m (µM)	[Enzyme] (nM)	kcat (s^{-1})
RNase J1	231 ± 10	0.47 ± 0.06	20	0.58 ± 0.03
RNase J2	≤ 74 ± 26	5.96 ± 2.04	800	≤ 0.005 ± 0.002
RNase J1/J2 complex	53 ± 5	0.22 ± 0.03	20	0.13 ± 0.01

Results are the average of three experiments, calculated from double reciprocal plots.
Note that for technical reasons (limited amounts of RNA and high background fluorescence), it was not possible to saturate RNase J2 with substrate. The calculated V_{max} and kcat values for RNase J2 are therefore given as maximal values.

D

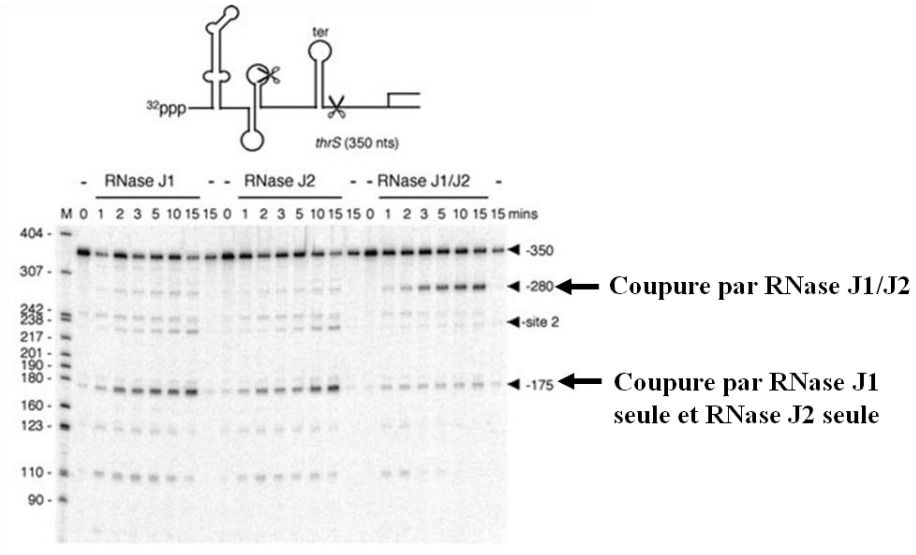


Figure 8: A. Copurification *in vitro* de la RNase J1 et J2. Les puits 2, 3 et 4 correspondent respectivement à la purification de la RNase J1, J2 et du complexe RNase J1/RNase J2 avec la RNase J1 portant une étiquette hexahistidine en C-terminal. **B.** Tests d'activité exonucléase sur un ARN simple brin marqué au phosphore γ 32 (^{32}P). **C.** Paramètres catalytiques calculés pour les RNases J1, J2 et du complexe J1/J2. **D.** Clivage endonucléolytique par la RNase J1, J2 et par le complexe J1/J2 sur le transcrit *thrS* (d'après Mathy *et al.*, 2010)

Les RNases J1 et J2 possèdent respectivement 50% et 70% d'identité et d'homologie de séquence, pour une masse d'environ 62 kDa chacune. De plus, la RNase J1 est une protéine globalement acide avec un pI de 6,2 tandis que la RNase J2 est une protéine plutôt basique avec pI de 9,0. Il a donc été supposé que ces deux enzymes forment un complexe stable. Les deux protéines ont donc été coexprimé *in vitro* chez *E. coli*, la RNase J1 portant une étiquette hexahistidine en C-terminal. Une bande correspondant a la RNase J1 était observée sur gel SDS-PAGE avec une autre bande de masse légèrement plus faible. La même expériences a été réalisé *in vivo* dans laquelle la RNase J1 porte une étiquette FLAG, le même résultat fut observé. Après analyse en spectrométrie de masse, la bande en dessous de la RNase J1 correspondait bien à la RNase J2. Ce qui a prouvé l'existence d'un complexe entre la RNase J1 et J2 *in vitro* et *in vivo*. Une analyse en SEC-MALLS et ultracentrifugation analytique a été réalisé sur le complexe. Les masses déterminées par les deux techniques sont respectivement 219 et 245 kDa, ce qui veut dire que les deux protéines forme un hétérotétramère (2 RNase J1 + 2 RNase J2). Des tests d'activité ont également été réalisé sur ce complexe mais également sur les deux enzymes purifiées séparément. Il s'est avéré que la RNase J2 a une activité exonucléase 200 à 400 fois inférieure à la RNase J1 et au complexe RNaseJ1/J2. De plus, la RNase J2 a une affinité pour le substrat 10 à 30 fois inférieure a la RNase J1 et au complexe RNase J1/J2 respectivement. Il semble donc que ce soit la RNase J1 qui porte l'activité exonucléase. L'activité endonucléase a également été testée sur le transcrit *thrS*, les trois formes de l'enzyme (RNase J1 seule, RNase J2 seule et le complexe RNase J1/J2) clivent le substrat de façon endonucléolytique. Les enzymes seules clivent 175 nt après l'extrémité 5' tandis que le complexe clive à 280 nt de l'extrémité 5'. On a donc une activité endonucléase qui est différente entre les enzymes seules et le complexe[48].

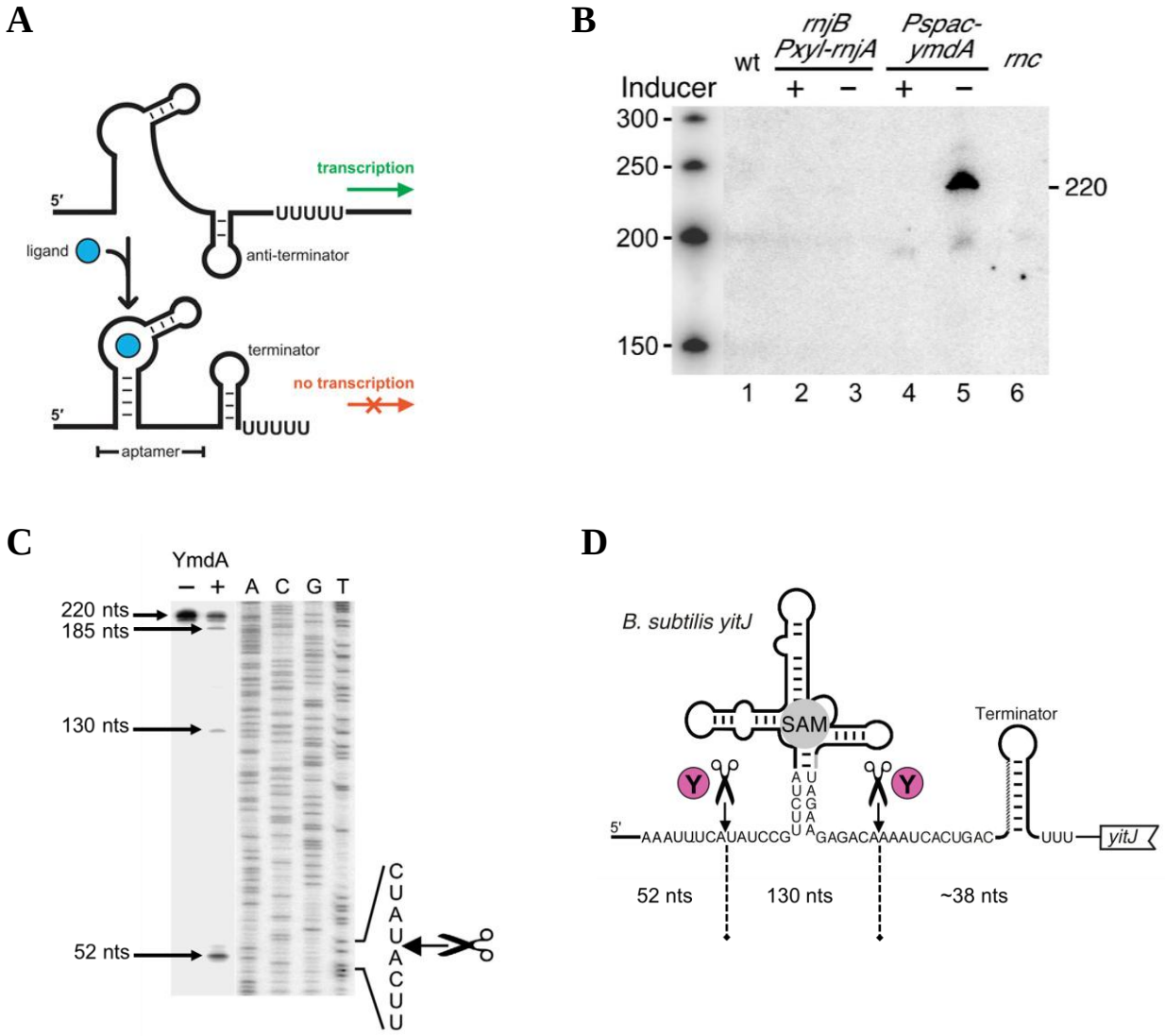


Figure 9: Caractérisation de l'activité endonucléolytique de la RNase Y sur le transcrit *yitJ*. A. Principe de régulation de la transcription par un riboswitch (d'après Kim and Breaker, 2008). **B**. La RNase Y (YmdA) est nécessaire in vivo pour initier la dégradation du transcrit yitJ, en son absence d'inducteur et donc d'expression de la RNase Y, le transcrit s'accumule (puit n°5). **C.** Caractérisation de l'activité endonucléase de la RNase Y (YmdA) sur le transcrit yitJ. **D.** Sites de coupures déterminés pour la RNase Y sur le transcrit yitJ (d'après Laalami et Putzer, 2011).

La RNase Y, l'endoribonucléase principale de *Bacillus subtilis*

En 2009, une autre endonucléase fut identifiée et appelée RNase Y[27]. Son activité a été caractérisé sur le transcrit *yitJ* codant pour une protéine impliquée dans la synthèse de la S-adénosyl-méthionine. (SAM). Ce transcrit est régulé par un riboswitch. Un riboswitch est une structure secondaire présente dans la region 5'UTR capable de changer de conformation via la fixation d'un métabolite, ce qui entraîne prématurément l'arrêt de la transcription (Figure 9A). En général, le métabolite est le produit issus de l'enzyme codée par le gène régulé par le riboswitch. Dans le cas de *yitJ*, la fixation de la SAM dans la région 5'UTR arrête la transcription[49]. Il a été montré que dans une souche dépletée en RNase Y, ce transcrit était accumulé et qu'à l'inverse dans les souches dépletées en RNase J et RNase III, ce transcrit n'était pas présent, donc dégradé (Figure 9B). La RNase Y fut donc identifiée comme étant une endonucléase initiant la dégradation du transcrit prémature *yitJ*, mais également de 10 autres Riboswitch SAM-dépendant. De plus, la stabilité des ARNm globaux est fortement augmentée dans la souche déplétée en RNase Y, la demie-vie croissant de 2,8 min à 6,1 min[27]. Concernant la spécificité de substrat de la RNase Y, il a été montré *in vivo* et *in vitro* que l'enzyme clive dans des sites riches en nucléotides A et U proches de structures secondaires et simple brin, tout comme la RNase E et J[27]. Les séquences des sites de clivage de ces trois endonucléases sont très similaires (annexe 2). Il a également été montré *in vitro* que l'efficacité de clivage est augmentée lorsque le transcrit possède une extrémité 5'monophosphate plutôt qu'une extrémité 5' trisphosphate[27]. Tout comme la RNase E, la RNase Y est donc capable d'effectuer un clivage endonucléotlytqiue sur un ARNm avec une extrémité 5'-triphosphate mais de façon moins efficace[27]. Tout comme la RNase E, il semblerait que la RNase Y empreinte principalement la voie directe pour cliver les ARNm, qui est la dégradation des ARNm en effetcuant un clivage endonucléolytique sans passer par la conversion de l'extrémité 5'-triphosphate en 5'-monophosphate.

Tableau 2 et 3: Tab 2. (Haut) Transcrits dont le clivage endonucléolytique par la RNase Y a été identifié. **Tab 3. (Milieu)** Liste non-exhaustive de transcrits dont la stabilité (demie-vie) est augmentée en absence de la RNase Y (d'après Laalami *et al.*, 2014)

Gène	Protéine codée	Fonction	Rôle du clivage par la RNase Y	Références
yitJ	S-méthyltransférase	Synthèse de la méthionine	dégradation	Shahbabian *et al.*, EMBO, 2009
cwlO	DL-endopeptidase	Synthèse de la paroi bactérienne	dégradation	Noone et al., Journal of Bacteriology, 2013
opéron *gapA*	Glycéraldéhyde 3-phosphate déshydrogénase	Glycolyse	maturation	Commichau *et al.*, Molecular and cellular proteomics, 2009
rpsO	Protéine S15 (sous-unité 30S du ribosome)	Traduction protéique	dégradation	Yao and Bechhofer, Journal of Bacteriology, 2010
opéron *infC-rpmI-tplT*	Facteur d'initiation de la traduction IF3 Protéine ribosomal L35 et L20	Traduction protéique	dégradation	Bruscella et al., Molecular Microbiology, 2011

Gène	Protéine codée	Fonction	Effet de l'absence de la RNase Y	Références
opéron *mreBH*	MreBH	Détermination de la forme de la cellule (Cytosquelette)	Demie-vie du transcrit passe de 9min a plus de 20min	Durand *et al.*, PLOS Genetics, 2012
opéron *spoIISAB*	toxine spoIISA antitoxine spoIISB	Système toxine-antitoxine	Demie-vie du transcrit passe de 2min à 22min	Durand *et al.*, PLOS Genetics, 2012
opéron *trp*	Enzymes de synthèse du tryptophane	Synthèse du tryptophane et des acides aminés aromatiques	Le transcrit est stabilisé (Northern blot)	Lehnik-Habrink *et al.*, Molecular Microbiology, 2011
opéron *gapB-speD*	Glycéraldéhyde 2-phosphate déshydrogénase 2	Néoglucogenèse	Durée de vie du transcit passe de 1,4min à 36min.	Lehnik-Habrink *et al.*, Molecular Microbiology, 2012

Suite à cela de nombreux autres transcrits ciblés par la RNase Y furent identifiés. Ces transcrits codent pour des protéines dont la fonction est primordiale dans la cellule. Par exemple, la RNase Y est responsable de la maturation de l'opéron *gapA* qui code pour la glycéraldéhyde 3-phosphate déshydrogénase (GAPH) qui fait partie de la glycolyse[50]. Le tableau 2 du haut donne une liste non-exhaustive des transcrit ciblés par la RNase Y.

Ce tableau ne cite que des études qui ont été réalisé sur un transcrit en particulier dont le clivage endonucléolytique par la RNase Y a été démontré *in vitro*. D'autres études transcriptomiques à grande échelle ont permis de mettre en évidence de nombreuses autres cibles potentielles de la RNase Y[51–53]. Ces cibles sont également liées à des fonctions biologiques cruciales chez *B. subtilis*. Dans ces études, le gène de la RNase Y est placé sous le contrôle d'un promoteur inductible au xylose, le gène de la RNase Y est donc déplété. On compare alors l'abondance de l'ensemble des transcrits en absence et en présence de xylose, donc en absence ou en présence de RNase Y. Pour chacune des deux conditions, on regarde si la durée de vie des transcrits est augmentée ou diminuée. Des sondes spécifiques de certaines transcrits sont utilisés afin d'étudier la modification de la durée de vie de ces transcrits en absence et en présence de la RNase Y. C'est comme cela que certains opérons dont la fonction est importante dans la cellule ont montré une durée de vie augmentée dans la cellule. Ces opréons sont résumés dans le tableau 3 du bas. La RNase Y a donc été considérée dès lors comme l'endoribonucléase principale chez *B. subtilis* responsable de l'initiation de la dégradation des ARNm et de la maturation de certains transcrits.

Pour finir, ces trois endonucléases ont les mêmes spécificités de substrat et les sites de clivages sur les ARN sont très similaires entre ces trois enzymes (annexe 2).

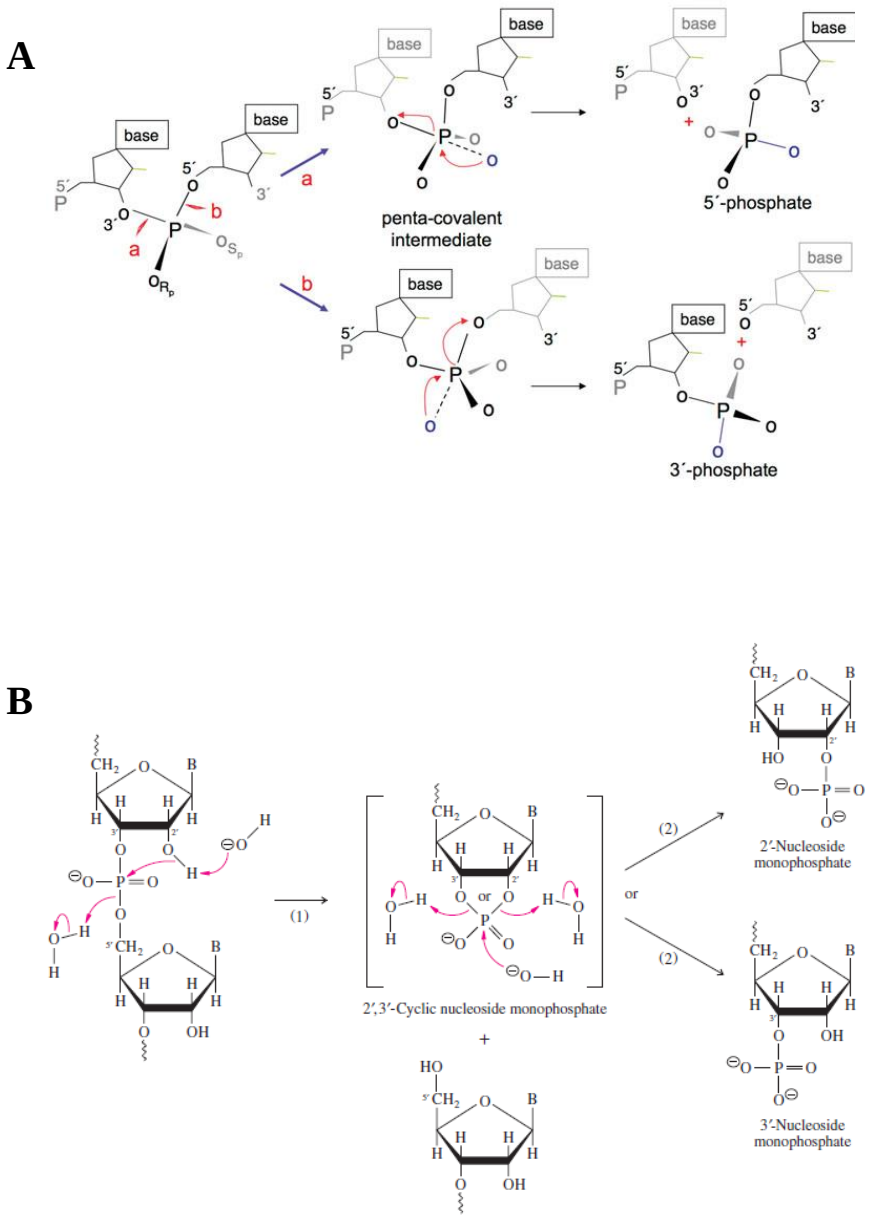


Figure 10: A. Principe du clivage de la liaison phosphodiester par les RNases du côté 5' ou 3' de l'atome de phosphore, donnant respectivement une extrémité 3'-phosphate où 5'-phosphate (d'après Yang, 2010). **B.** Mécanisme de l'hydrolyse alcaline de l'ARN permettant de relarguer des nucléotides 2'- ou 3'-monophosphate (d'après Robert Horton *et al.*, Principes de Biochimie, De Boeck 4ème édition, 2006).

v. Mécanisme de clivage de la liaison phosphodiester au sein de l'ARN par les ribonucléases

Les enzymes qui catalysent l'hydrolyse de la liaison phosphodiester des acides nucléiques sont appelées des nucléases. Ces enzymes sont indispensables au bon fonctionnement de la cellule car impliquées dans des processus biochimiques capitales pour la physiologie cellulaire (annexe 4)[13].

La rupture de la liaison phosphodiester consiste à cliver la liaison covalente entre le phosphate et l'oxygène (P-O), au sein d'un polymère d'acide nucléique. Ce clivage peut se faire du côté 5' ou 3' de la liaison (5'O-P-O3.'), en libérant soit une extrémité 5'-phosphate ou 3'-phosphate (figure 10A). Le mécanisme de clivage est une réaction acide-base dans laquelle la base est responsable de l'activation du nucléophile en le déprotonant. L'acide va faciliter la formation du produit de la réaction en protonant le groupe partant. Les nucléases utilisent divers types de nucléophiles, cela peut être une molécule d'eau, le (désoxy)ribose, un phosphate inorganique, ou les chaines latérales des acides aminés Ser, Tyr ou His[13]. De plus la catalyse peut ou non nécessiter un ion métallique qui est coordonné par des acides aminés au sein de l'enzyme. La réaction peut être divisée en trois étapes :

1) attaque nucléophile du phosphore scissile de la liaison phosphodiester
2) formation d'un intermédiaire penta-covalent chargé négativement
3) rupture de la liaison du phosphate scissile avec l'oxygène du côté 5' ou 3'

Ici, nous allons uniquement nous intéresser au mécanisme des Ribonucléases (RNases) qui clivent l'ARN.

Hydrolyse alcaline de l'ARN

En milieu alcalin, l'ARN va s'hydrolyser sans avoir besoin d'un catalyseur enzymatique. En effet, a pH basique, un ion hydroxyde (HO^-) active la fonction 2'-OH du ribose en la déprotonant. L'oxygène nucléophile va ensuite attaquer la fonction phosphoester pour former un intermédiaire phosphodiester cyclique avec coupure de la chaîne. Un second ion hydroxyde va ensuite venir catalyser la conversion de cet intermédiaire cyclique, ce qui va produire une extrémité 2' ou 3'-monophosphate et 5'-hydroxyle (5'-OH) (Figure 10B). Le mécanisme de clivage par les RNase s'est "inspiré" de l'hydrolyse alcaline des ARN.

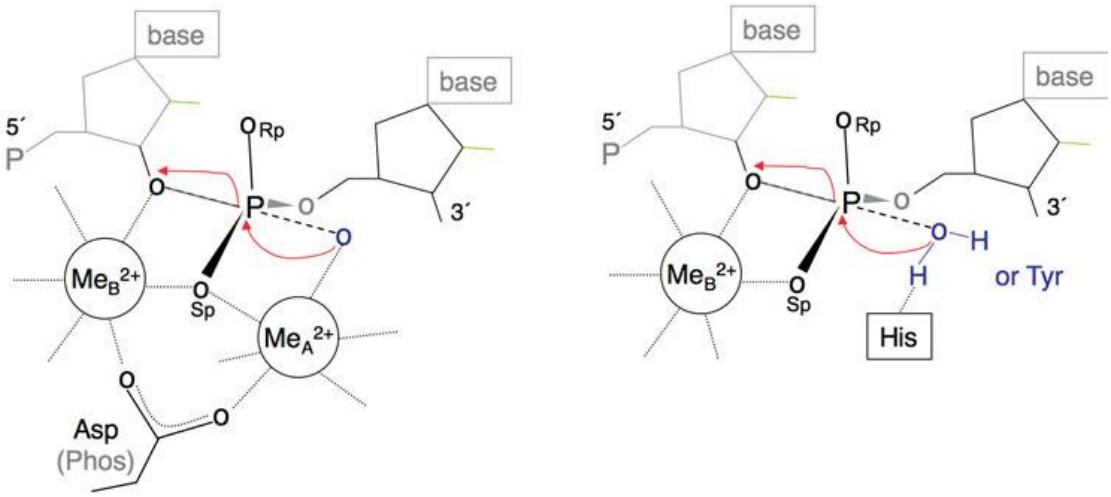


Figure 11: Mécanisme de l'attaque nucléophile par les RNases à deux ions dépendant (gauche) ou à un ion dépendant (droite) (d'après Yang, 2010)

Hydrolyse de l'ARN par les RNases

Dans un premier temps, l'hydrolyse de la liaison phosphodiester par les RNases produit la plupart du temps une extrémité 5'-phosphate et une extrémité 3'-OH. Pourquoi ? Tout d'abord, par ce que la fonction 3'-OH peut directement servir de nucléophile pour d'autres nucléases (ex: ADN et ARN polymérases, ADN ligases, transposases, spliceosomes,...). Ensuite, l'extrémité 5'-phosphate peut également servir directement de susbstrat, lors de la fin de la réplication, pour la ligation de l'ADN, mais également dans les cas de la réparation ou recombinaison de l'ADN. Sur la base du mécanisme catalytique les RNases peuvent être classées en trois groupes[13]:

1) Les RNases métal-indépendant
2) Les RNases métal-dépendant avec deux ions bivalents
3) Les RNases métal-dépendant avec un ion bivalent

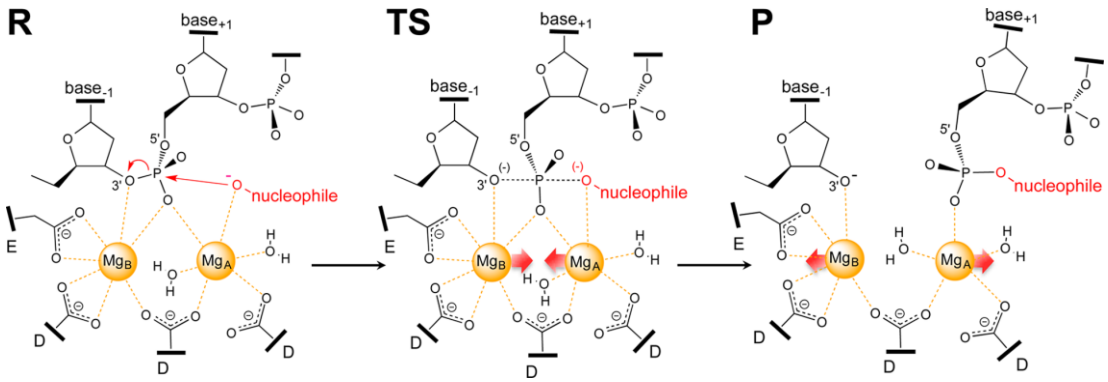
1) RNases métal-indépendant

Ces RNases vont utiliser le groupement 2'-OH du ribose comme nucléophile, son activation va permettre une attaque nucléophile de l'atome de phosphore et la formation d'un groupement 2',3'-phosphate cyclique labile, qui va être hydrolysé et permettre la formation d'une extrémité 3'-phosphate[13]. C'est le cas du mécanisme de clivage par la RNase A[54].

2) Mécanisme dépendant à deux ions métalliques

Dans tous les cas, la réaction d'hydrolyse métal-dépendante de la liaison phosphodiester libère toujours une extrémité 5'-monophosphate et 3'-hydroxyle. Les cations Mg^{2+} et Ca^{2+} sont les deux ions les plus abondants chez les organismes vivants et le Mg^{2+} est celui qui est le plus abondant à l'intérieure des cellules. Le Mg^{2+} est l'ions le plus fréquemment associé aux enzymes impliquées dans le métabolisme des acides nucléiques. Cette utilisation fréquente du Mg^{2+} est probablement due a son abondance dans la cellule, sa stabilité redox et sa solubilité en comparaison du Mn^{2+}, Fe^{2+}, Ca^{2+}, Cu^{2+}, Ni^{2+} et Zn^{2+}. Les avantages d'incorporer deux ions métalliques pour la catalyse sont une plus grande spécificité pour le substrat et une libération du produit plus efficace[13].

A



B

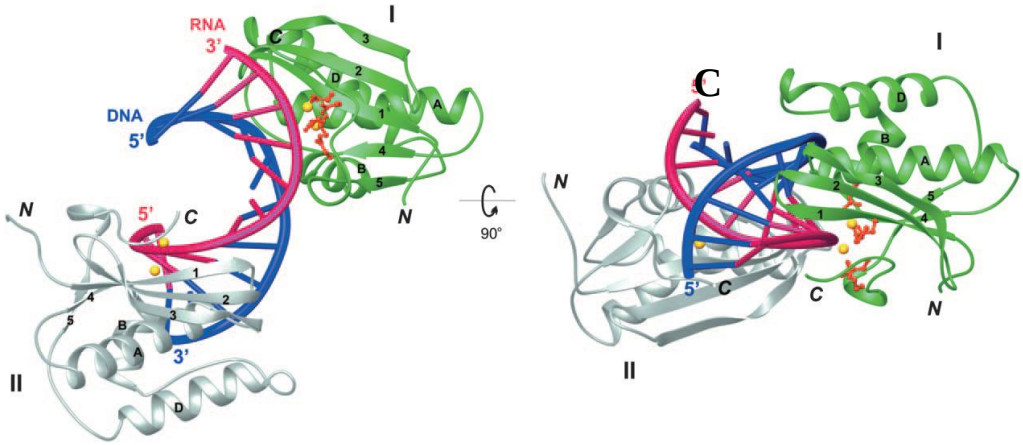
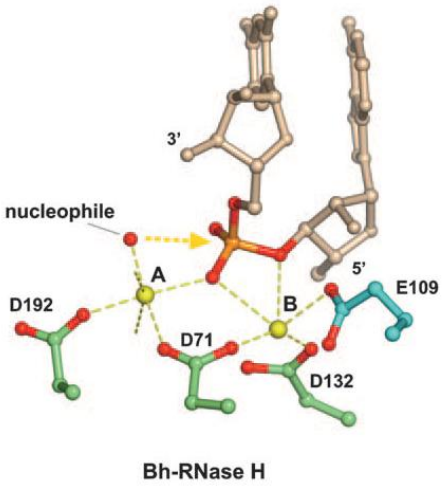


Figure 12: A. Mécanisme enzymatique de la RNase H avec 2 ions métalliques. (R= fixation du substrat; TS = Etat de transition; P = relargage du produit) (d'après Palermo *et al.*, 2015) **B**. Structure cristallographique de la RNase H en interaction avec un duplex ARN/ADN. **C.** Structure atomique du site catalytique de la RNase H avec les deux ions représentés par des sphères jaunes (d'après Nowotny *et al.*, 2005)



Les deux ions métalliques (A et B) sont en général coordonnés par des résidus aspartates. L'ion A est du côté nucléophile tandis que l'ion B est du côté du groupe 3'-O partant. Il est possible que des résisdus autres que l'Asp permettent également la coordination des deux ions et l'exclusion des molécules d'eau de l'ion B. Ces résidus peuvent être des His, Glu ou Tyr[13].

Il a été proposé que l'ion A permette d'initier la réaction par la déprotonation du nucléophile, qui peut être une molécule d'eau ou un phosphate inorganique (Pi), tandis que l'ion B permettrait la stabilisation de l'état de transition penta-covalent ainsi que la déstabilisation du phosphate scissile pour le clivage. Mais des études structurales sur la RNase H ont montré que ces deux ions se rapprochent durant la catalyse (~3,5 Å) pour stabiliser ensemble l'état de transition[13].

La RNase H est une RNase fonctionnant avec deux ions métalliques Mg^{2+}. Cette RNase va cliver les ARN mais également les complexes ARN/ADN. Le mécanisme catalytique de cette enzyme a été résolu via des structures cristallographiques. Dans cette enzyme, l'ion Mg^{2+}(A) va déprotoner le nucléophile lui permettant d'attaquer la liaison phosphodiester scissile, le résultat est la formation d'un état de transition penta-covalent stabilisé par les deux ions. Pour finir, il y a clivage de la liaison et formation des deux extrémités 5'-P et 3'-OH[55].

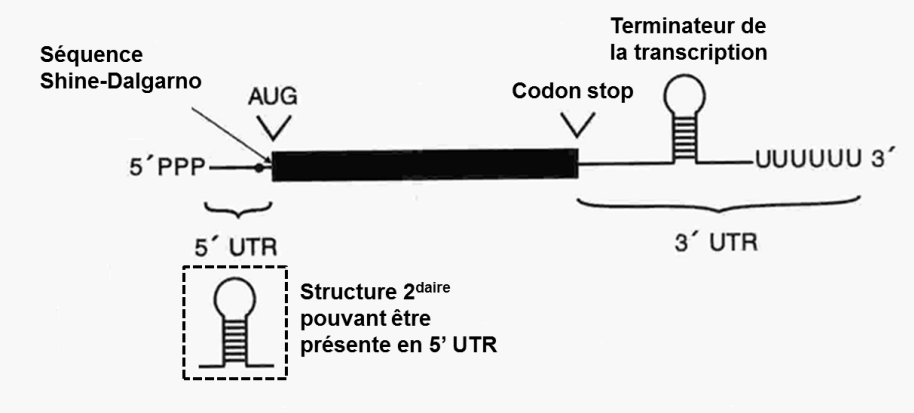
3) Mécanisme dépendant à un ion métallique

Dans le mécanisme dépendant à un ion, l'ion B est présent tandis que l'ion A est absent. En général, c'est une histidine qui coordonne l'ion métallique. L'activation de la molcéule d'eau est en général réalisée par une histidine qui va déprotoner la molécule d'eau, ce qui va créer le nucléophile qui va ensuite aller attaquer le phosphate de la liaison phosphodiester[13].

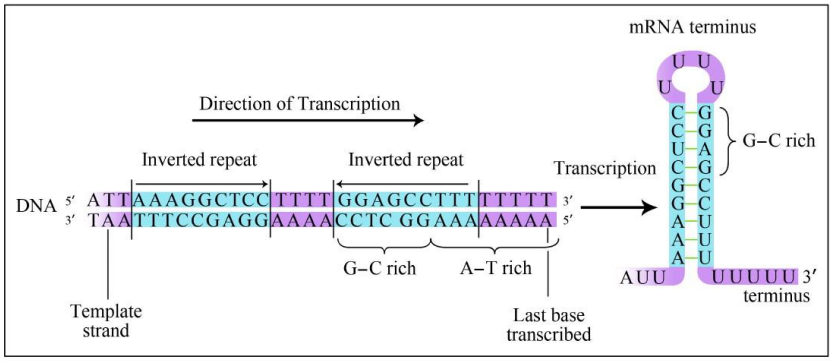
Les RNases bactériennes métal-dépendantes, qu’elles soient des exonucléases ou endonucléases, utilisent un cofacteur métallique pour :

- stabiliser la charge négative du nucléophile
- stabiliser le groupement hydroxyle qui joue le rôle de la base
- stabiliser la géométrie de l’état de transition
- stabiliser la charge négative du produit, le rendant meilleur groupe partant

A



B



C

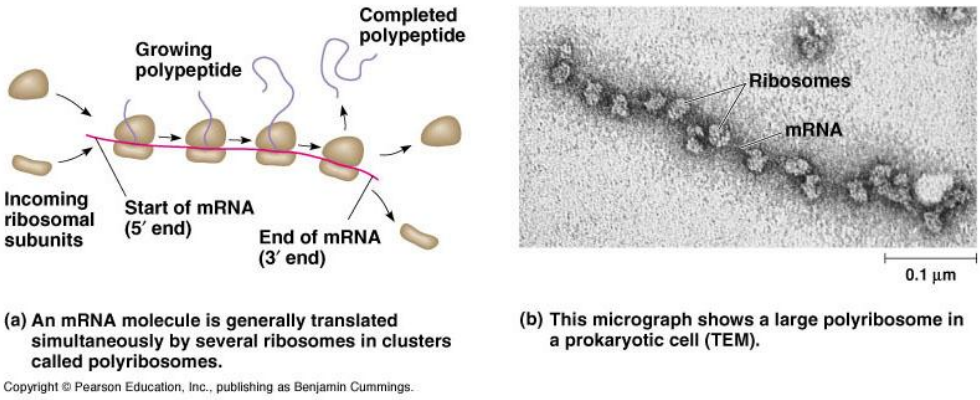


Figure 13: A. Structure générale d'un ARNm procaryote (d'après Garrett et Grisham, Biochimie, 2008). **B.** Structure du terminateur de la transcription chez les procaryotes dans le cas de la terminaison Rho indépendante (d'après Garrett et Grisham, Biochimie, 2008). **C.** Visualisation, au microscope électronique, des ribosomes présents sur un ARNm bactériens (d'après Garrett et Grisham, Biochimie, 2008).

I.B. Dégradation des ARNm : une stratégie conservée

i. Protection des ARNm contre la dégradation

L'instabilité des ARNm a un rôle primordial dans le contrôle de l'expression des gènes. En effet, cette instabilité permet aux bactéries de s'adapter rapidement aux changements de conditions environnementales, mais elle permet également de produire la bonne quantité d'une protéine donnée et le recyclage des ribonucléotides pour les incorporer dans une nouvelle molécule d'ARN[28,56,57]. Les ARNm sont des molécules biologiques dont la durée de vie est relativement courte dans la cellule, de l'ordre de quelques secondes à plusieurs minutes[28]. En général, ce temps de demi-vie est inférieur au temps de division de l'organisme. Par exemple, la demie-vie moyenne des ARNm globaux (~ 80%) chez *Bacillus subtilis* est d'environ 5 min[28]. Cette instabilité est d'abord due à la molécule d'ARN en elle-même, qui possède des liaisons phosphodiester scissiles du fait du groupement 2'-OH sur le ribose. Ensuite, les ARNm possèdent très peu de structures secondaires, contrairement aux ARNt et aux ARNr. De plus, les ARNt et ARNr possèdent de nombreuses modifications post-transcriptionnelles qui les rendent plus résistants à la dégradation et leurs confèrent une durée de vie beaucoup plus longue dans la cellule, pouvant aller jusqu'à plusieurs générations cellulaires[58–60]. Pour pallier à leur instabilité, les ARNm possèdent des systèmes de protection contre la dégradation. Tout d'abord, les ARNm ne sont pas libres dans la cellule mais en grande partie fixés sur les ribosomes (Figure 13C), machinerie de traduction qui empêche les RNases d'accéder au site de clivage[61]. Ensuite, la molécule d'ARNm possède un groupe triphosphate en 5' qui protège cette extrémité contre une attaque par les exonucléases[62]. Des structures secondaires de type tige-boucle présentes en 5', dans la région non traduite (5'UTR) empêchent aussi un clivage interne de l'ARNm[62]. Enfin l'extrémité 3' se termine en général par le terminateur de la transcription qui est une structure tige-boucle protégeant cette extrémité[62](figure 13B). L'initiation de la dégradation des ARNm commence par l'élimination d'une des protections portées par ces ARNm. De façon générale, le clivage débute toujours au niveau de la région 5' UTR, par la conversion du groupement 5' triP en 5' monoP, catalysée par la pyrophosphohydrolase (RppH)[63,64]. L'extrémité 5'monoP est ensuite reconnue par les endonucléases, leurs permettant de fixer l'ARN. De plus, des hélicases d'ARN peuvent éliminer les structures tige-boucle et donc libérer des sites de clivage pour les endonucléases[65].

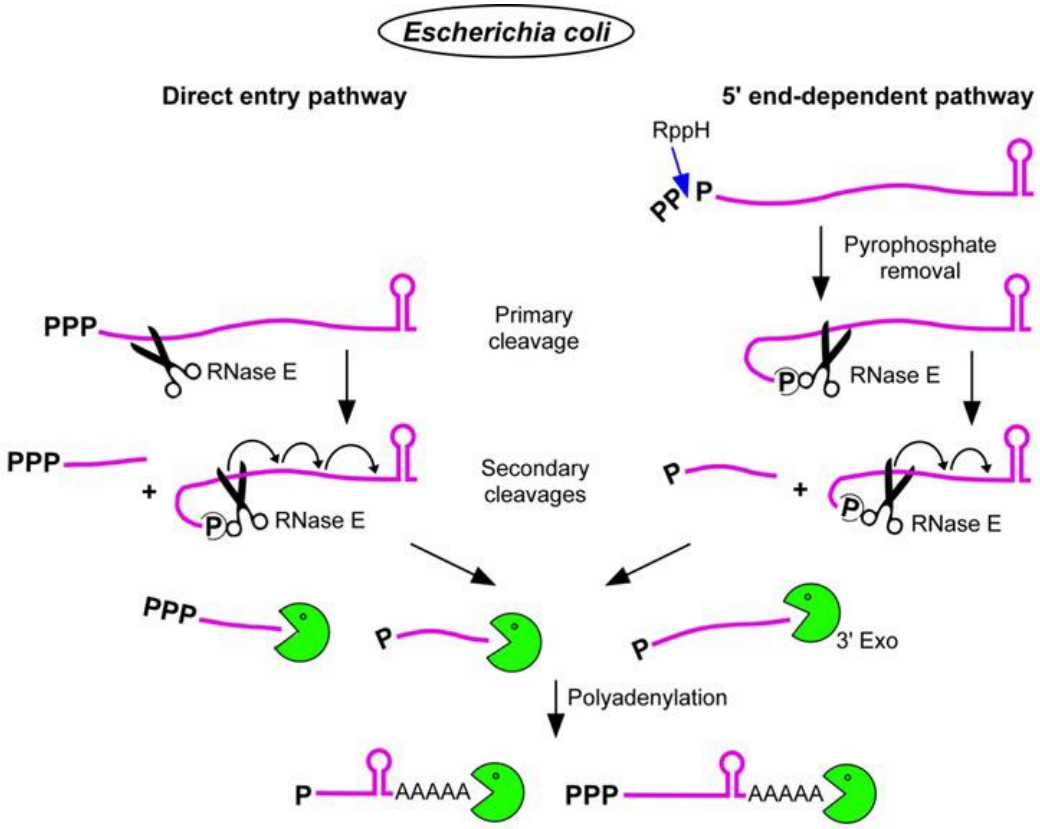


Figure 14 : Dégradation des ARNm chez *E. coli.* La RNase E est responsable de l'initiation de la dégradation par un clivage endonucléolytique. Ce clivage peut être dépendant de l'extrémité 5' (*in vivo/in vitro*) ou peut se faire directement sans nécessité d'avoir une extrémité 5'-monophosphorylée (*in vitro*). Les fragements issus du clivage vont être ensuite dégradés par des exonucléases (RNase II, PNPase). L'extrémité 3' structurée du transcrit est polyadénilé pour permettre à la PNPase de dégrader la structure secondaire ou cette structure secondaire peut être dégradée par la RNase R, une éxonucléase spécifique des structures secondaires (d'après Laalami *et al.*, 2013).

ii. La dégradation des ARNm chez *E. coli* contrôlée par la RNase E

Voie de dégradation RppH-dépendante

Chez *E. coli*, après coupure du groupement 5'-triP par la RppH, l'extrémité 5' monoP est reconnue par la RNase E, via son domaine 5'-senseur, ce qui va induire le clivage endonucléolytique (cf II.A.i). Ce clivage va donc libérer deux fragments, le fragmeent aval portant une extrémité 5' monoP, et le fragment amont portant une extrémité 3' sans structure secondaire (Figure 14). Ce dernier va être rapidement dégradé par des 3' exoribonucléases (RNase II, PNPase, RNase R et oligoribonucléase)[28]. Le fragment portant une extrémité 5' monoP est de nouveau reconnu et clivé par la RNase E afin de répéter le processus de clivage. Les fragments protégés par une structure tige-boucle en 3' sont polyadénylés par la poly(A) polymérase ce qui permet aux exoribonucléases de dégrader cette structure tige-boucle[28,66]. Mais cette structure secondaire peut également être dégradée par la RNase R, une éxonucléase spécifique des structures secondaires.

Voie de dégradation indépendante de la RppH

Il a cependant été montré *in vivo* qu'en absence de la RppH seulement 10% des ARNm chez *E.coli* sont stabilisés, ce qui signifie qu'il existe une voie directe pour initier la dégradation des ARNm indépendante de la RppH. Cette voie est utilisée en général lorsque l'ARN porte des structures secondaires en 5', ou bien lorsque le ribosome est lié à l'ARNm pendant la traduction. Dans ce cas, la RNase E effectue le clivage directement sans conversion de l'extrémité 5'-triP en 5' monoP[28].

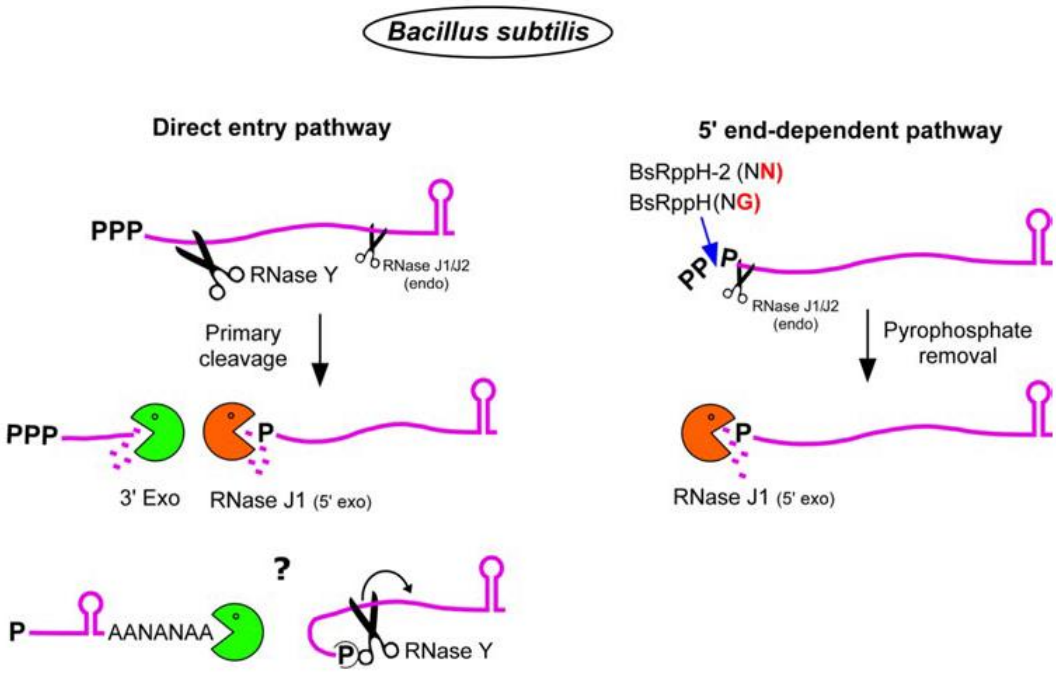


Figure 15 : Dégradation des ARNm chez *B. subtilis*. La RNase Y est responsable de l'initiation de la dégradation par un clivage endonucléolytique. Ce clivage peut être dépendant de l'extrémité 5' (*in vivo/in vitro*) ou peut se faire directement sans nécessité d'avoir une extrémité 5'-monophosphorylée (*in vitro*). Les fragements issus du clivage vont être ensuite dégradés par des exonucléases (RNase J, PNPase) (d'après Laalami *et al.*, 2013).

ii. La dégradation des ARNm chez *B. subtilis* : RNase Y vs RNase J

Les étapes menant à la dégradation complète d'un ARNm chez *B. subtilis* sont les mêmes que pour *E. coli*. Par contre, les acteurs de cette dégradation sont différents : la RNase E est absente chez *B. subtilis* et une nouvelle enzyme, la RNase J est responsable de l'activité 5' exoribonucléase. Comme chez *E. coli*, la dégradation des ARNm chez *B. subtilis* peut se faire via deux voies différentes (une voie directe et une voie dépendante de l'extrémité 5'). Chez *E. coli*, la RNase E est responsable de la dégradation par ces deux voies. Chez *B. subtilis*, ces deux voies sont prisent en charge par la RNase Y, une enzyme découverte en 2009 et responsable de la dégradation de la majorité des ARNm[28]. Mais chez *B. subtilis* la RNase J est également responsable de l'initiation de la dégradation de certains transcrits. Chez *E. coli*, il n'y qu'une seule endonucléase responsable de cette initiation

Implication de la RNase J

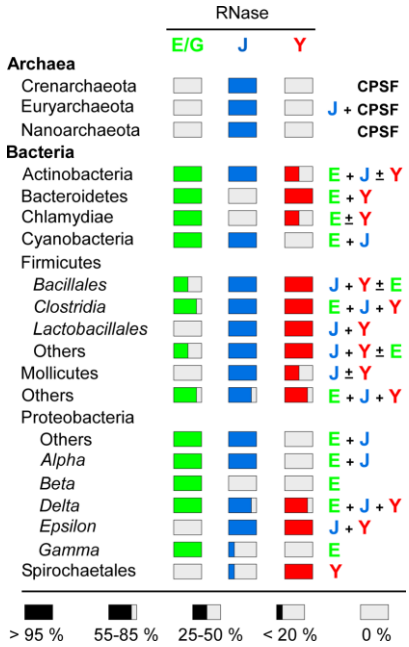
La RNase J est responsable de la dégradation de certains ARNm, en effet, il a été montré que dans certains cas, lorsque l'ARNm possède un groupement triP en 5', la RNase J est capable de reconnaitre ce groupement triphosphate mais ce dernier va bloquer son activité exonucléase. La RNase J va alors effectuer un clivage endonucléolytique au sein de l'ARNm, ce qui va générer une extrémité 5' monoP. La RNase J va alors ensuite dégrader le fragement d'ARN portant une extrémité 5'-monoP grâce à son activité 5' exoribonucléase. L'activité 5'exoribonucléase de la RNase J est également stimulée grâce à l'activité de la RppH de *B. subtilis* (BsRppH)[63,64]. La RppH va éliminer le groupement pyrophosphate (phosphate γ et β) de l'extrémité 5'. Ce qui va permettre à la RNase J de dégrader l'ARN par son activité exonucléase. Pour le moment, le mécanisme de clivage endonucléolytique par la RNase J n'est pas bien connu (cf II.A.ii)[28].

Implication de la RNase Y

La RNase Y est responsable de l'initiation de la dégradation de la majeur partie des ARNm chez *B. subtilis*. Des expériences d'activité *in vitro* sur le transcrit *yitJ* ont montré que la RNase Y était capable d'effectuer un clivage endonucléolytique sans avoir besoin d'une extrémité 5'-monoP. C'est ce qu'on appel la voie d'entrée directe. Mais, dans cette même étude il a également été montré que l'activité de la RNase Y était stimulé par la pla présence d'une extrémité 5'-monophosphate, c'est ce que l'on appel la voie dépendante de l'extrémité 5'[28].

Dans cette voie, tout comme pour le RNase J, c'est la BsRppH qui, en éliminant le pyrophosphate de l'extrémité 5', va stimuler l'activité de la RNase Y[63,64,67]. Lorsque le clivage endonucléolytique de la RNase Y est effectué, c'est la PNPase et la RNase J qui vont dégrader respectivement les fragements en amont et en aval de la coupure endonucléolytique. Ces deux fragements portant respectivement une extrémité 3'-OH et 5'-monoP.

Tableau 4 : Coexistence des RNase E, J et Y au sein des procaryotes. La RNase Y est totalement absente chez les γ-Protéobactéries dont fait partie *E. coli*. Par contre, la RNase E est présente chez 25 à 50% des Bacilles avec les RNase J et Y, mais *B. subtilis* ne possède pas la RNase E. (d'après Laalami *et al*., 2013)



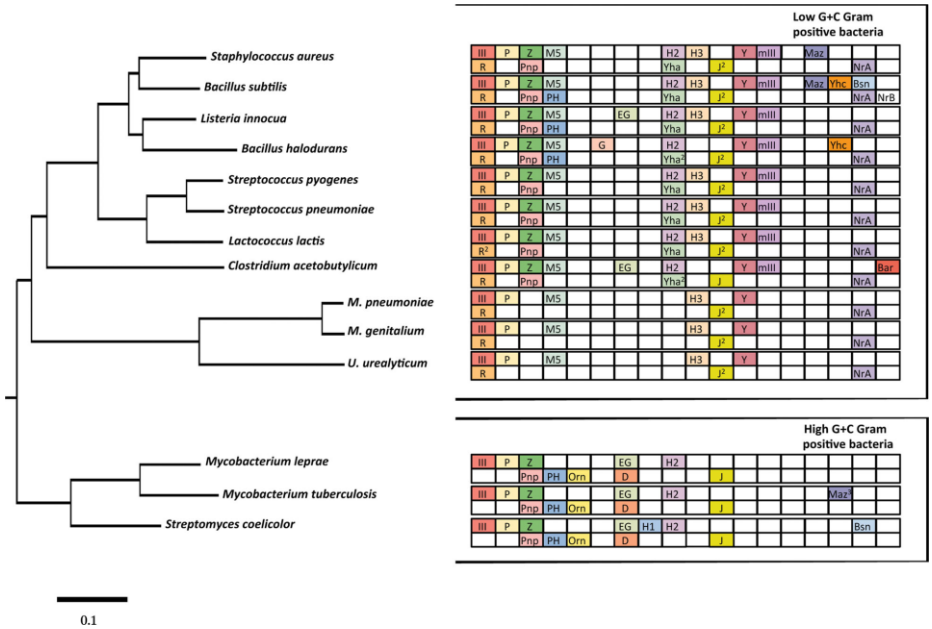
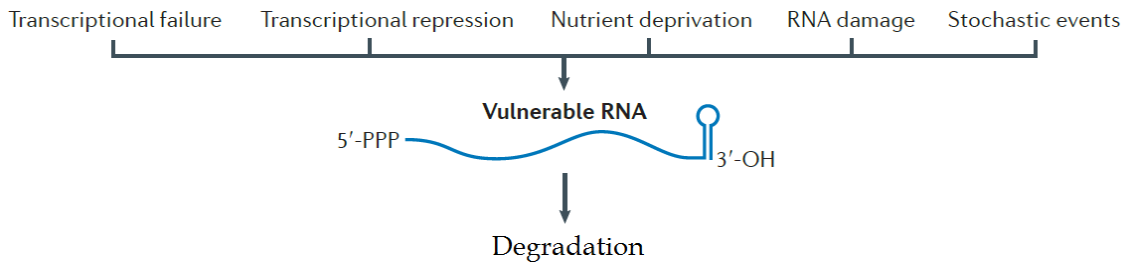


Figure 16: Arbre phylogénétique des RNases chez les bactéries à Gram poitifs. La relation phylogénétique entre les différents organismes a été calculé en comparant les séquences des ARNr 16S en utilisant Clustal X. (III = RNase III; P = RNase P; Z = RNase Z; M5 = RNase M5; G = RNase G; EG = RNase E/G; H1 = RNase HI; H2 = RNase HII; H3 = RNase HIII; Y = RNase Y; mIII = MiniRNase III; Maz = MazF/EndoA; Yhc = YhcR; Bsn = RNase Bsn; Bar = Barnase; R = RNase R; II = RNase II; Pnp = polynucleotide phosphorylase; PH = RNase PH; Orn = oligoribonuclease; T = RNase T; D = RNase D; Yha = RNase YhaM; J = RNase J; NrA = NanoRNase A; NrB = NanoRNase) (d'après Durand *et al.*, 2015)

iv. Coexistence des endonucléases E, J et Y et évolution convergente

Chez les procaryotes, trois endoribonucléases principales permettent d'initier la dégradation des ARNm : les RNases E, Y et J[28]. Chaque bactérie possède au moins l'une de ces trois endoribonucléases. Chez les γ-protéobactéries, c'est la RNase E qui est la plus abondante (*E. coli* ne possède que la RNase E). A l'inverse, chez les firmicutes, ce sont les RNase Y et J qui sont les plus présentes (tableau 4). Chez *B. subtilis*, la RNase E est absente, mais comme nous l'avons vu, les RNases Y et J ont un rôle analogue (Tableau 4).



Fugure 17: Facteurs induisant une vulnérabilité des ARNm face à la dégradation par les RNases (d'après Mackie, 2013)

Tableau 5 : Liste non-exhaustive des stress rencontrés dans le milieu environnemental (d'après Barria *et al.*, 2016)

Stress	Paramètres	Problèmes causés
Thermique	Augmentation : 30 - 42°C	Augmentation trop importante de la croissance cellulaire
	Diminution: ~15°C	Réduction de l'efficacité de la transcription, traduction et dégradation des ARNm
Phase stationnaire	Croissance limitée	Accès aux nutriments limité
Pression osmotique	Choque hyperosmotique	Lyse cellulaire Nécessité d'acquérir des osmoprotectants (glycine, bétaïne et glutamate)
Oxidatif	Accumulation d'intermédiaires réactifs de l'oxygène (ex: H_2O_2)	ARN endommagés
Interactions cellulaires	Invasion de cellules hôtes	Conidtions environnementales changées
Résistance aux métaux	Concentration élevée en métaux (Mg^{2+}, Ni^{2+}, Co^{2+})	Métaux forment des complexes toxiques pour la cellule
Système Toxine - antitoxine	Production d'une toxine	Lyse cellulaire
Biofilm	Formation d'amats cellulaires	Accès à l'oxygène et aux nutriments limité

v. Dégradation des ARNm et régulation de l'expression des gènes

Au sein de leurs niches, les bactéries doivent constamment faire face à de brusques changements de conditions environnementales, et doivent être capables de s'y adapter rapidement. La balance entre la traduction protéique et la dégradation des ARNm va assurer une réponse rapide de la bactérie face à ces nombreux stress[68].

En particulier, la régulation de l'expression des gènes est cruciale pour cette adaptation et fait appel à des mécanismes de régulation complexes, à la fois transcriptionnel et post-transcriptionnel[57,69–71]. Les RNases ont un rôle important dans la régulation post-transcriptionnelle. En effet, ces stress vont être responsables de modification au niveau de la transcription des gènes, en provoquant des arrêts de la transcription pour certains gènes et l'induction de la transcription d'autres gènes. Les transcrits dont la transcription est stoppée; ne pouvant être traduit, doivent être dégradés rapidement afin de permettre le renouvellement des nucléotides pour la synthèse des nouveaux ARN. Les ARN endommagés par des stress, par exemple oxydatif, doivent également être dégradés rapidement (figure 17).

Cette dégradation rapide des ARNm par les RNases permet un ajustement continuel de la population en messager aux besoins de la cellule en protéines spécifiques.

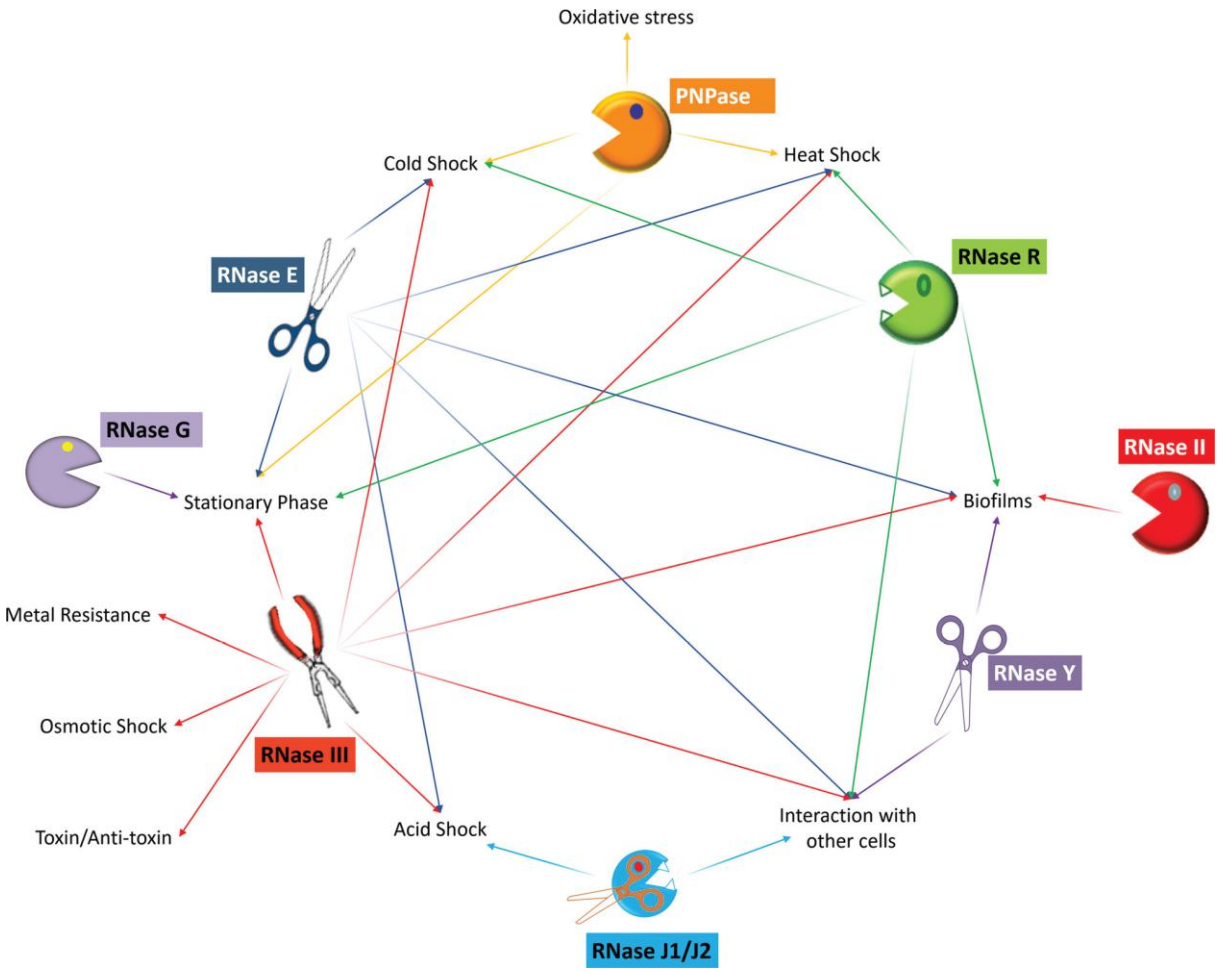


Figure 18 : Réseau d'interactions entre les différents stress et les RNases permettant de fournir une réponse adaptative à ces stress. Les endonucléases sont représentées par des ciseaux tandis que les exonucléases sont représentées par un camembert. Ces RNases permettent à la bactérie de s'adapter face aux différents stress (phase stationnaire / chocs thermique, oxydatif, osmotique, acide / formation de biofilms / interactions bactérie - cellule hôte,...). (d'après Barria *et al.*, 2016)

De nombreux stress sont rencontrés par les bactéries, le tableau 5 présente une liste de certains stress étudiés dans lesquels les RNases sont impliquées dans un but de régulation afin que la bactérie puisse s'adapter à ces conditions de stress.

Par exemple, chez *B. subtilis*, lorsque ce dernier est soumis à un choc thermique avec augmentation de la température, la croissance cellulaire s'en trouve augmentée. Il faut donc que cette croissance soit limitée afin de préserver assez de ressources énergétiques pour les cellules. Dans ce cas-là, il existe un système toxine – antitoxine de type I[72]. Ce système code pour une toxine et une antitoxine qui dans ce cas-ci est un petit ARN, nommé SR4 complémentaire de l'ARNm codant la toxine. Cette toxine provoque la lyse cellulaire. Il a été montré que la RNase Y va, en condition de choc thermique, dégrader l'ARN antisens SR4 afin que l'ARNm codant la toxine puisse être traduit. Ce qui va provoquer la lyse cellulaire et limiter la croissance cellulaire. A l'inverse, lorsque la lyse cellulaire n'est pas nécessaire il faut la stopper. La RNase III est impliquée dans la dégradation des duplex ARNp-ARNm des systèmes toxine – antitoxine chez *B.subtilis*. D'ailleurs cette RNase est essentielle chez *B. subtilis* car elle protège de la lyse cellulaire[68].

Le stress oxydatif est aussi la cause de nombreux dommages au niveau des ARN. Chez *E. coli*, en absence de la PNPase (souche Δ*pnp*), il y a une accumulation des ARN endommagés par le peroxyde d'hydrogène (H_2O_2). La viabilité cellulaire est restaurée lorsqu'un plasmide contenant le gène de la PNPase (*pnp*) est inséré chez la bactérie[68].

Lors d'un choc osmotique, il y a transcription du gène *proU*, qui code pour transporteur membranaire d'osmoprotectants (glycine, bétaïne, proline). Il a été montré que la RNase III mature ce transcrit[68].

En conclusion, les RNases ont un rôle primordial dans la régulation de l'expression des gènes au niveau post-transcriptionnel, qui permet aux bactéries de s'adapter aux milieux dans lesquels elles vivent. La comparaison des deux bactéries, *E. coli* et *B. subtilis*, montre une évolution convergente au niveau du processus de dégradation des ARNm. En effet, l'initiation de la dégradation des ARNm et les étapes suivantes aboutissant à la dégradation complète d'un ARNm sont très similaires. Bien que les enzymes impliquées dans ce processus soient différentes, elles partagent les mêmes propriétés catalytiques (tableau 6). C'est le cas des RNases E, J et Y. Maintenant nous allons voir que ces enzymes sont différentes au niveau des domaines dont elles sont constituées, malgré cela, elles possèdent des similarités structurales

Tableau 6: Caractéristiques des trois endonucléases principales présentes chez les bactéries : les RNases E, J et Y

	RNase E	RNase J	RNase Y
Organisme	*Escherichia coli*	*Bacillus subtilis*	*Bacillus subtilis*
Type d'activité	Endonucléase (*in vivo* /*in vitro*)	Endonucléase (*in vitro*) Exonucléase (*in vivo*/ *in vitro*)	Endonucléase (*in vivo* /*in vitro*)
Spécificité	Séquence AU-rich	Séquence AU-rich	Séquence AU-rich
Type de séquence	Zone simple brin	Zone simple brin	Zone simple brin
Extrémité 5'	Triphosphate (*in vivo*/*in vitro*) Monophosphate (*in vitro*)	Monophosphate (*in vivo*/*in vitro*)	Triphosphate (*in vivo*/*in vitro*) Monophosphate (*in vitro*)

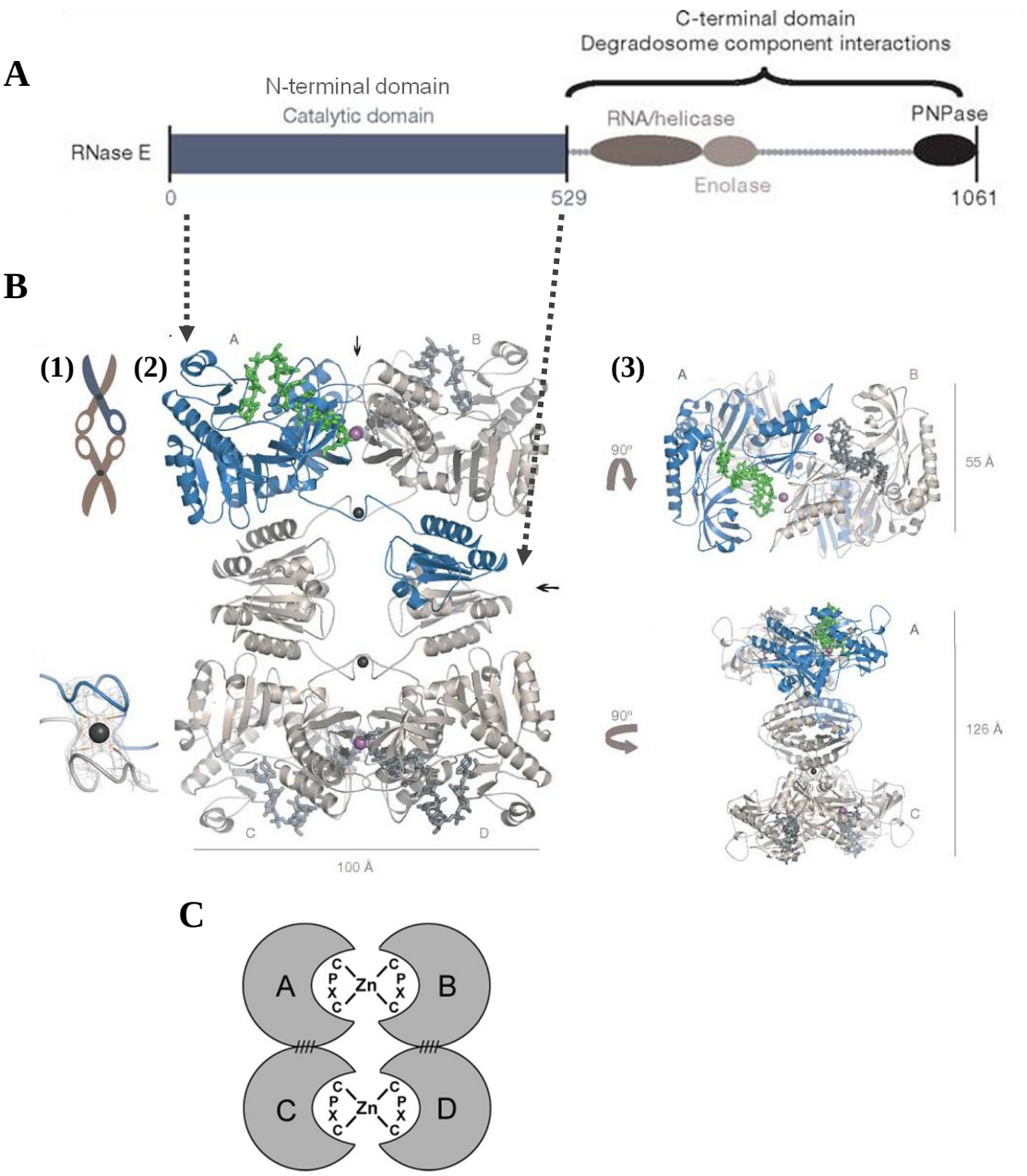


Figure 19 : A. Schéma de l'organisation de la RNase E en deux domaines : le domaine catalytique en N-terminal (1-529) et le domaine d'ancrage des partenaires en C-terminal (529-1061). (D'après Callaghan et *al.*, Nature, 2005). **B.** Structure cristallographique du domaine N-terminal de la RNase E (1-529) en interaction avec un ARN 10-mer. **(1)** Représentation schématique en ciseaux de la RNase E. **(2)** Structure quaternaire de l'homotétramère, avec un protomère (bleu) en interaction avec l'ARN (vert). **(3)** En haut, rotation à 90° à l'horizontale montrant une vue du dessus et en bas rotation de 90° à la verticale montrant une vue de côté de l'homotétramère. **C.** Dimérisation du protomère A avec le protomère B via un ion Zn^{2+} coordonné par deux cystéines du protomère A et deux cystéines du protomère B (d'après Callaghan *et al.*, 2005)

II. Aspects structuraux des ribonucléases d'*Escherichia coli* et de *Bacillus subtilis*

II.A. Les RNase E et J, deux enzymes possédant des similitudes structurales

i. Organisation structurale de la RNase E: un domaine catalytique en N-terminal

La RNase E est une grosse protéine, dont un monomère est constitué de 1061 acides aminés (masse moléculaire de 118 kDa)[26]. Cette enzyme est constituée de deux parties principales (Figure 19A). La partie N-terminale catalytique (1-529) est structurée globulaire et très conservée chez les bactéries[26,73]. La partie C-terminale (529-1061) est, quant à elle, peu structurée et très peu conservée chez les bactéries[74]. Cette partie de la protéine permet l'ancrage des partenaires que sont la PNPase, l'énolase et l'hélicase d'ARN RhlB[75]. Ce complexe macromoléculaire de dégradation des ARNm est appelé dégradosome. Des structures cristallographiques de la partie N-terminale (1-520) ont été résolues sous forme libre (Apo)[76] et complexée (Holo)[26]. La structure du domaine catalytique (1-529) de la RNase E, en interaction avec un ARN 10-mer, résolue en 2005 à 2,9 Å de résolution, a révélée que la RNase E est organisée sous forme d'un homotétramère (dimère de dimère) (Figure 19B). L'organisation de cette structure quaternaire peut être représentéee par deux paires de ciseaux dos à dos organisés en tandem. (figure 19B). La partie catalytique (1-529) est constituée de deux domaines : le grand domaine (1-400) et le petit domaine (415-520). Entre ces deux domaines, un motif conservé CPxCxGxG, contenant deux cystéines, permet de lier un cation Zn^{2+} (Figure 19C). Ce Zn^{2+} est coordonné par les deux cystéines d'un protomère (A) et les deux cystéines d'un autre protomère (B) permettant leurs dimérisation (figure). Le grand domaine est subdivisé en quatre sous domaines : le sous-domaine RNase H (1-35 et 215-279) ayant un rôle structural uniquement ; le sous-domaine DNase I (280-400) contenant un cation bivalent Mg^{2+} et portant l'activité catalytique ; les sous-domaines 5' senseur (119-224) et S1 (36-118) permettant, respectivement, la liaison et l'orientation du substrat pour le clivage (figure 20A).

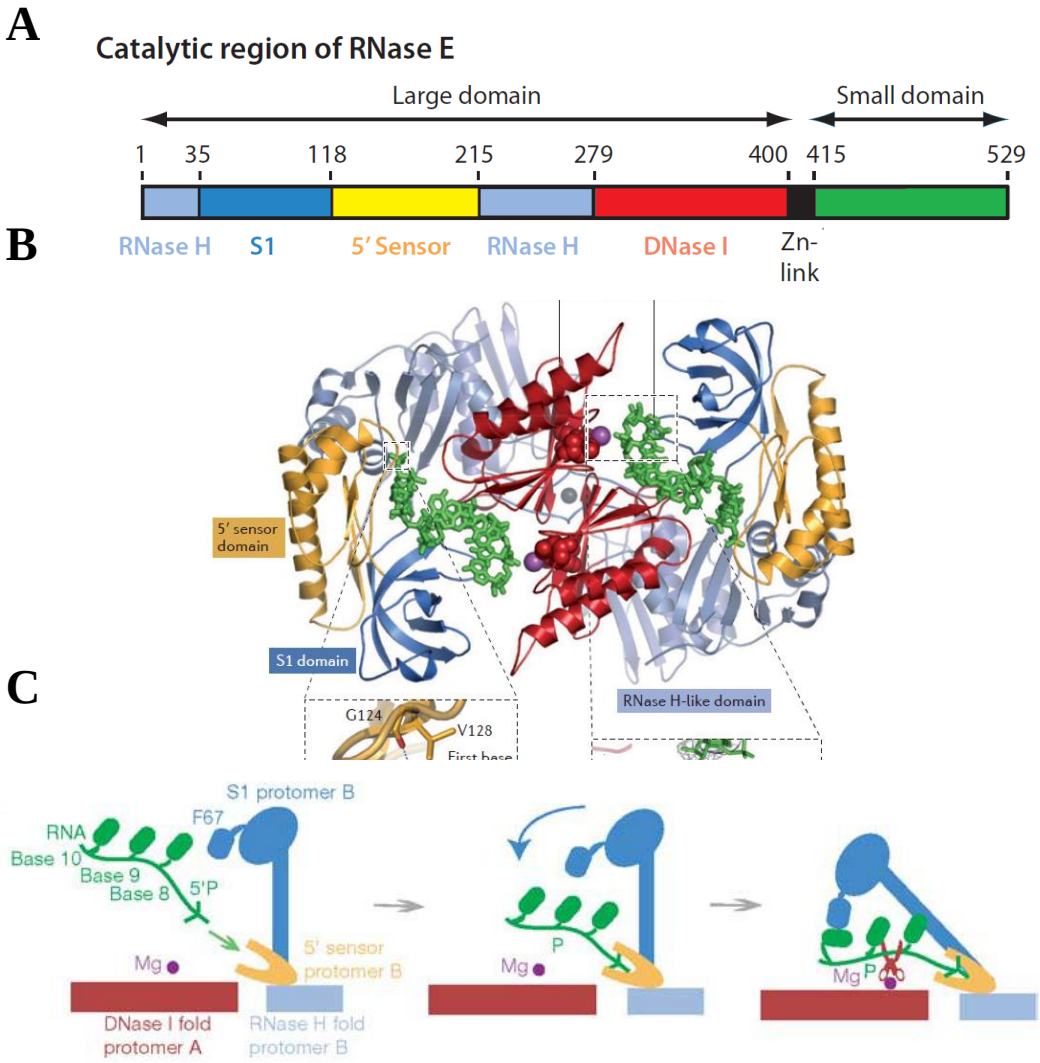


Figure 20: A. Organisation schématique de la partie catalytique de la RNase E. Le grand domaine et le petit domaine sont liés par le domaine à Zn. **B.** Vue du dessus d'un des deux dimères de la RNase E en interaction avec l'ARN 10-mer. Cet ARN est en contact avec les sous-domaines 5' senseur et S1, mais également avec les sous-domaines RNase H et DNase I. Les cations bivalents Zn^{2+} pour la dimérisation et Mg^{2+} pour la catalyse sont représentés par une sphère grise ou violette, respectivement. **C.** Modèle du mécanisme catalytique de la RNase E dit de "tapette à souris" montrant la communication entre le sous domaine 5' senseur et le site catalytique. (d'après Callaghan *et al.*, 2005)

La structure révèle que chaque molécule d'ARN est liée par les sous-domaines 5' senseur et S1 d'un protomère (B) et par le sous-domaine DNase I du protomère (A). Ainsi, au sein du dimère, un protomère va effectuer la liaison et l'orientation du substrat tandis que l'autre protomère va permettre le clivage grâce au domaine DNase I (figure 20C). Un modèle du mécanisme catalytique par un dimère de RNase E a donc été proposé. Dans ce mécanisme, l'extrémité 5' monophosphate de l'ARN va se lier au domaine 5' senseur, induisant un changement conformationnel du sous-domaine S1 lui permettant de fixer l'ARN. Cette fixation oriente le substrat pour le clivage au niveau de l'ion Mg^{2+} du domaine DNase I (figure 20C). Au niveau moléculaire, le mécanisme consisterait en une attaque nucléophile du phosphate scissible par un groupe hydroxyle (HO^-) fourni par une molécule d'eau coordonnée par l'ion Mg^{2+}. Les groupes carboxylates des aspartates 303 et 346 pourraient servir de base générale pour activer la molécule afin de réaliser cette attaque nucléophile. Cette première structure cristallographique a donc permis de formuler une première hypothèse quant au mécanisme catalytique mais également de confirmer la préférence de la RNase E pour les ARNm possédant une extrémité 5' monophosphate.

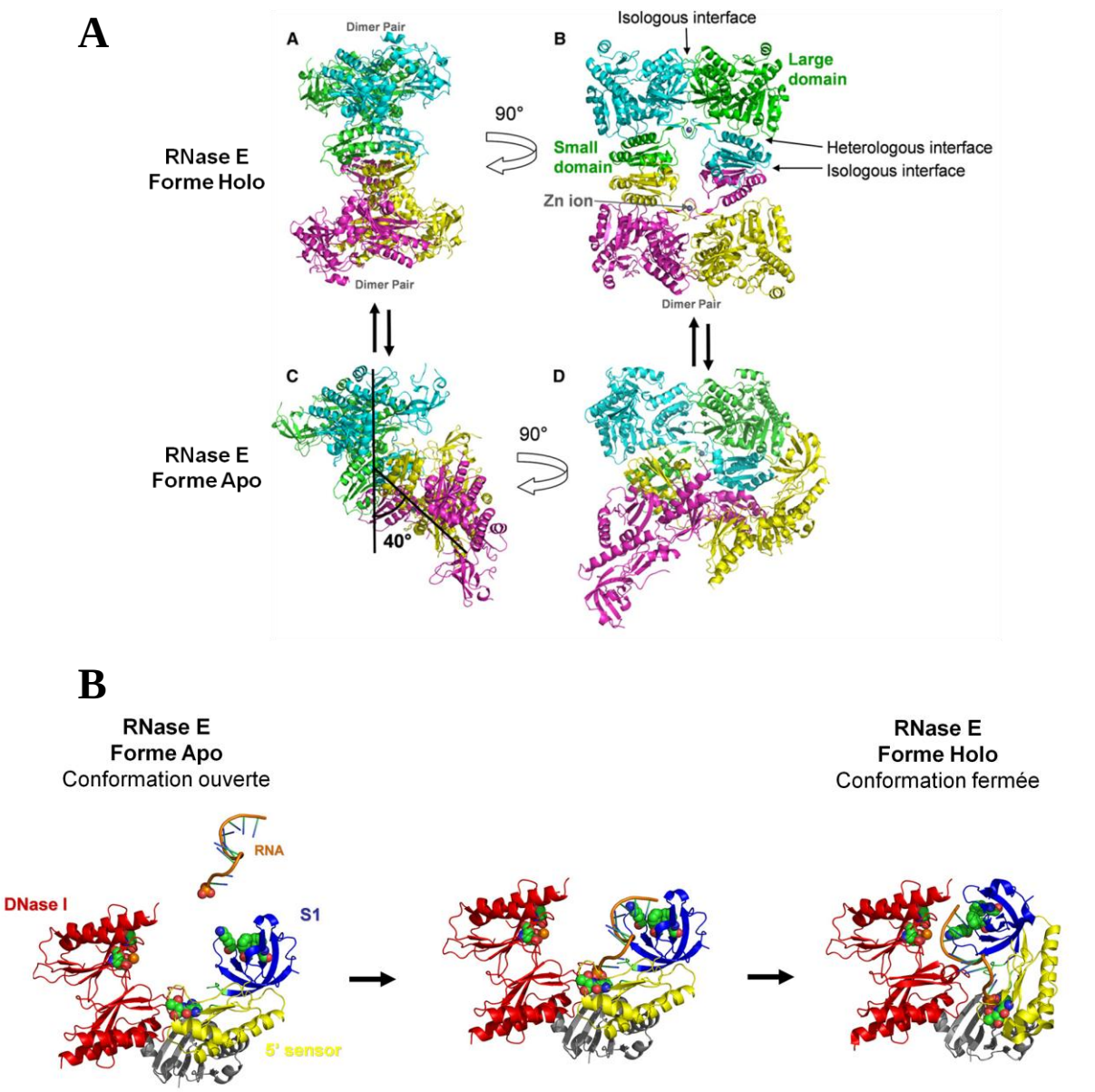


Figure 21: A. Comparaison de l'organisation quaternaire de la RNase E dans sa forme complexée (Holo) et libre (Apo). Les protomères sont représentés en bleu, vert, jaune et rose. Deux protomères (bleu et vert) forment un dimère et deux autres protomères (rose et jaune) forment l'autre dimère. Le tétramère de RNase E a une symétrie de type D2 dans la structure complexée, mais dans la structure libre une inclinaison du tétramère de 40° casse cette symétrie. **B.** Modèle du mécanisme de reconnaissance du substrat par un protomère de RNase E . A. En absence d'ARN, le protomère est en conformation ouverte. Les sous-domaines S1 et 5' senseur sont fortement exposés au solvant, permettant la fixation du substrat. Après liaison de l'ARN, le 5' senseur va lier le 5' monophosphate et le sous-domaine S1 se lie à l'ARN. Puis, les deux sous-domaines liant l'ARN se déplacent d'un bloc afin de rapprocher le substrat du domaine DNase I pour le clivage (d'après Koslover *et al.*, 2008)

En 2008, une deuxième structure catalytique de la RNase E, sous forme libre, à 3,3 Å de résolution [76], a permis de montrer que des changements conformationnels ont lieu lors de la liaison du substrat. En effet, la comparaison des structures libre et complexée montre une torsion de 45° de l'interface entre les deux dimères (interface isologue), ce qui engendre un mouvement de rotation d'environ 40° pour le tétramère au niveau de l'interface hétérologue. Les changements conformationnels observés dans la structure quaternaire sont donc dus à un changement d'orientation du petit domaine par rapport au grand domaine et la région de coordination du Zinc n'est pas altérée (figure 21A). Il y a également des changements conformationnels au niveau des sous domaines du grand domaine: S1 et 5' senseur (figure 21B). Lors de la liaison du substrat, les sous domaines 5' senseur et S1 sont déplacés d'un angle de 60° permettant de passer à la conformation fermée. (figure 21B). Ce mouvement permet de rapprocher le domaine 5' senseur du domaine catalytique DNase I. Ainsi, cette structure de la RNase E libre a permis d'affiner le premier modèle proposé du mécanisme catalytique dans lequel la fixation du substrat au 5' senseur induisait un changement conformationnel du sous-domaine S1 lui permettant d'interagir avec l'ARN. Pour finir, il a été supposé que la grande flexibilité de l'homotétramère au niveau de l'interface hétérologue permettrait le clivage d'ARNm hautement structuré dans la région 5'UTR, même en présence d'une extrémité 5'-trisphoshate. Mais le détail de la catalyse reste encore à étudier.

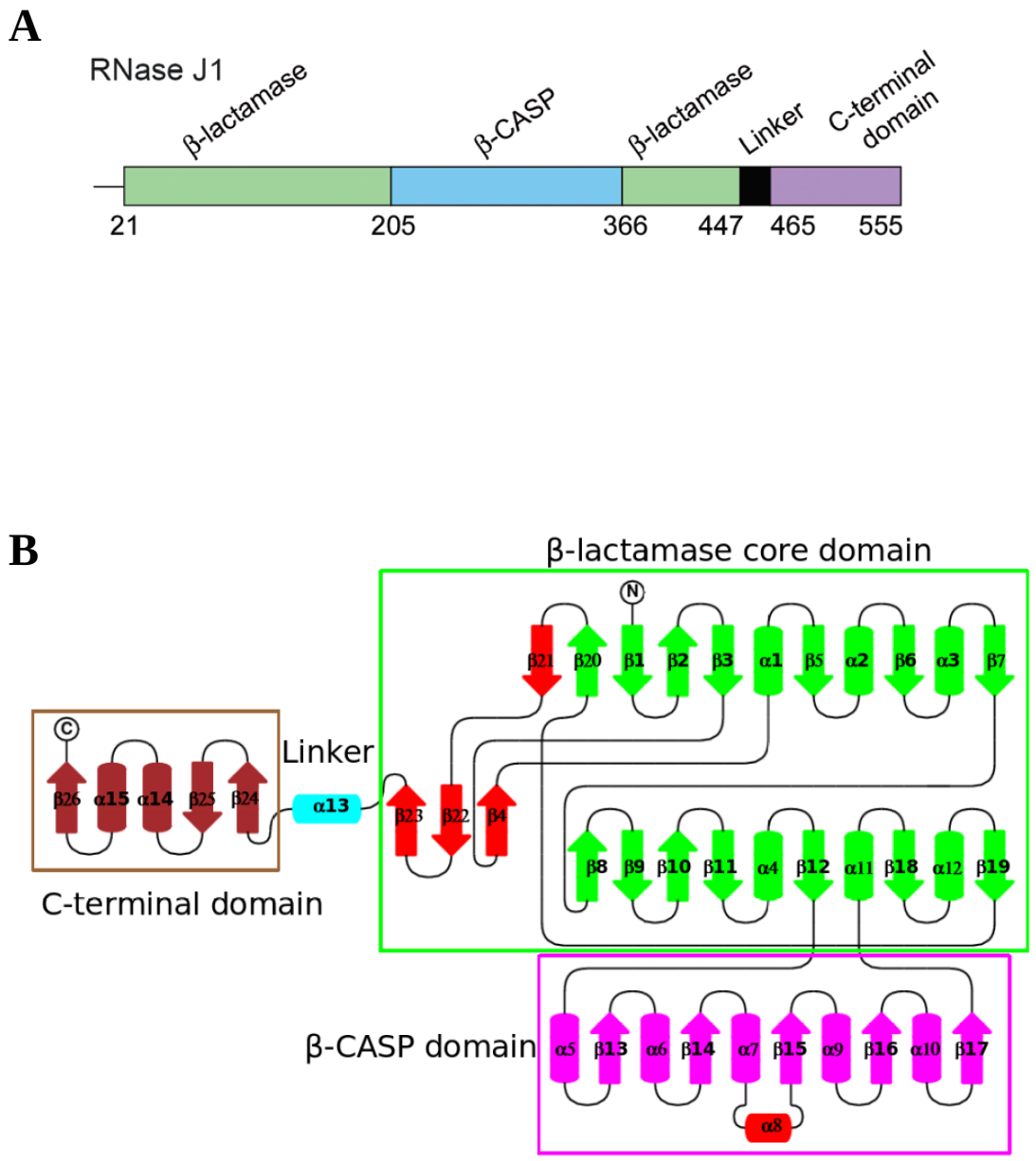


Figure 22: A. Organisation schématique de la RNase J de *Thermus thermophilis* **B**. Organisation en deux domaines structuraux de la RNase J et topologie des structures secondaires. Les brins β sont représentés par des flèches et les hélices α par des cylindres (d'après Li de la Sierra-Gallay *et al.*, 2008)

ii. La RNase J, une dualité fonctionnelle expliquée par sa structure

Comme nous l'avons mentionné précédemment, la RNase J est la première RNase possèdant une double activité (5'exo/endo) à avoir été découverte[77]. De plus, c'est la première RNase portant une activité 5'-exoribonucléase découverte chez les bactéries. La RNase J possède un homologue fonctionnel à activité 5'-exoribonucléase chez les procaryotes, Xrn-1[78]. Deux structures de la RNase J de *Thermus thermophilus (T. thermophilus)*, une bactérie à Gram-négatif, ont été résolues par cristallographie : Une en complexe avec de l'uridine monophosphate (UMP)[19] et une autre en présence d'un ARN 4-mer[79].

La première structure a permis de mettre en évidence l'organisation oligomérique de la RNase J ainsi que l'organisation de ses domaines. La RNase J fait partie de la famille des metallo-β-lactamases. Cette famille est divisée en plusieurs sous-familles comme par exemple les β-CASP, dont fait partie la RNase J. Le nom β-CASP provient de membres de cette classe de protéines impliquées dans la réparation de l'ADN, les protéines CPSF-73, Artemis et Snm1/Pso2. En effet, cette famille d'enzyme regroupe à la fois des RNases mais également des DNases[80].

La RNase J est constituée de trois domaines: le domaine β-CASP, le domaine β-lactamase possèdant l'activité catalytique et le domaine C-terminal dont la fonction reste encore inconnue. On remarque que le domaine β-CASP est ancré dans le domaine β-lactamase (figure 22A). Un connecteur relie le domaine β-CASP/β-lactamase et le domaine C-terminal. La RNase J de *T. thermophilus* possède 61% d'identité de séquence avec celle de *B. subtilis.* La RNase J de *B. subtilis* a une structure très proche de celle de *T. thermophilus.* La structure de la RNase J de *B. subtilis* a été résolue en 2011[77].

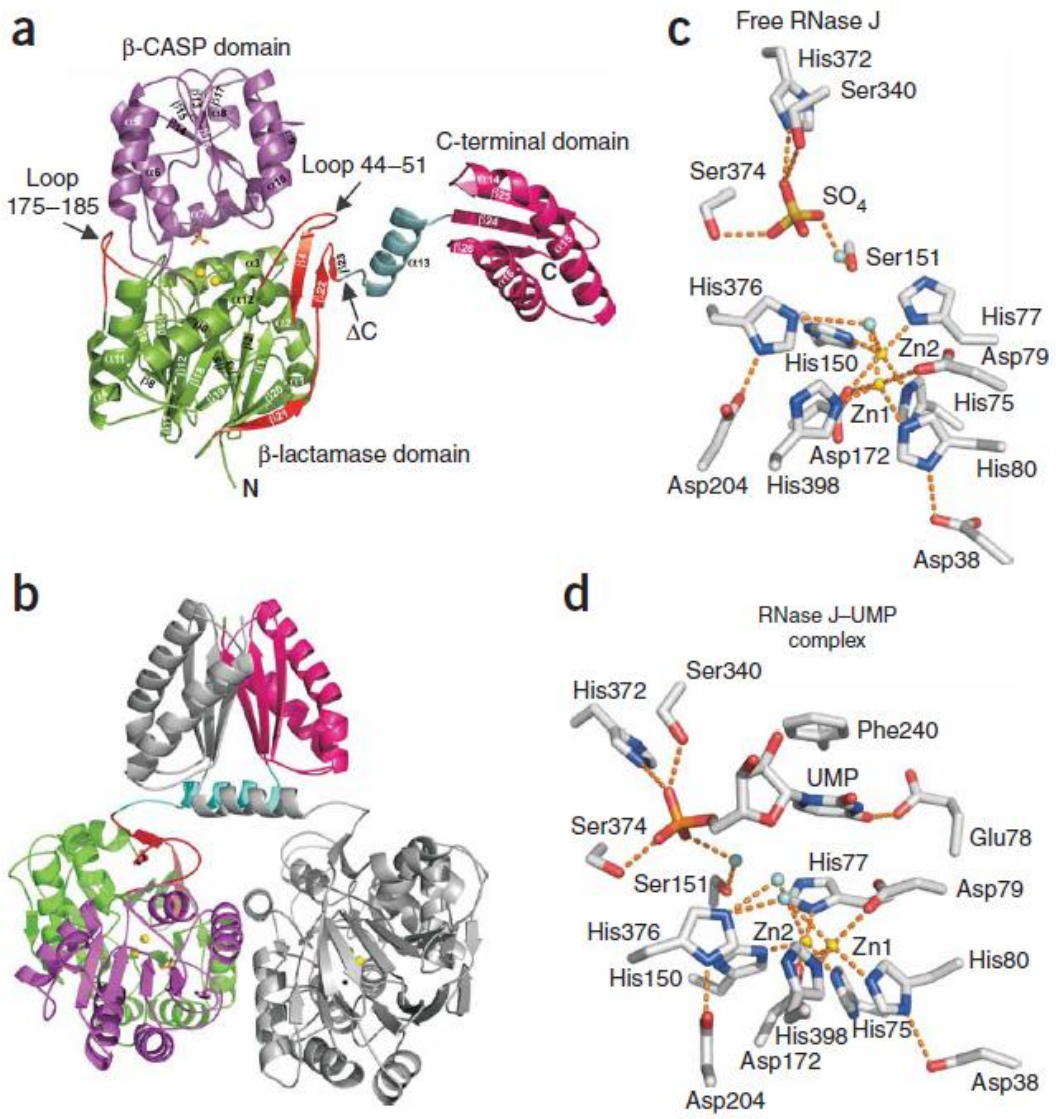


Figure 23: Structure de la RNase J dans sa conformation fermée. **A.** structure d'un monomère de la RNase J. Les deux ions Zinc sont représentés par des sphères jaunes, l'ion sulfate est représenté sous forme bâton. **B.** Structure du dimère de RNase J, un monomère est représenté avec les mêmes couleurs qu'en A, l'autre monomère est représenté en gris. **C.** Liaison du zinc à un ion sulfate présent dans la condition de cristallisation en absence de ligand. Les ions zinc sont représentés par des sphères jaunes. Les liaisons hydrogènes permettant l'interaction avec l'ion sulfate sont représentées en traits pointillés oranges. **D.** Centre catalytique de la RNase J lié à de l'UMP. Le phosphate en 5' de l'UMP est lié par les mêmes résidus que ceux impliqués dans la liaison du sulfate en C. La base uridine est coordonnée par le glutamate 78 et la phénylalanine 240, tandis que le phosphate est tenu par la Serine 151, 340 et 374 et par l'histidine 372 et la glycine 375 (d'après Li de la Sierra-Gallay *et al.*, 2008).

La RNase J est organisée sous forme d'un dimère en solution via son domaine C-terminal. Le site catalytique est constitué de deux ions Zn^{2+} coordonnés par 5 histidines et 2 aspartates. Ces deux ions sont enfouis entre les domaine β-CASP et β-lactamase[19].

Dans cette première structure en complexe avec l'UMP, le nucléotide présent dans la poche de liaison du 5'-monophosphate est localisé à un nucléotide de distance du site catalytique. Cette position de l'UMP est concordante avec la translocation de l'ARN d'un nucléotide nécessaire pour l'activité éxonucléase (figure 23D)

L'arrangement spatial de la poche de liaison du nucléotide en position 5' exclut la possibilité de liaison d'un nucléotide triP, ce qui explique la préférence de la RNase J pour les substrats ayant une extrémité 5'-monoP (figure 23D). La structure liée à l'UMP présente une conformation fermée qui ne peut pas lier une molécule d'ARN. Une structure de la RNase J en complexe avec l'ARN est nécessaire pour comprendre comment se fait la liaison de l'ARN et expliquer le mécanisme catalytique. Dans la structure de l'enzyme libre un ion sulfate occupe la même position que le monoP de l'UMP au niveau de la poche de liaison (figure 23C).

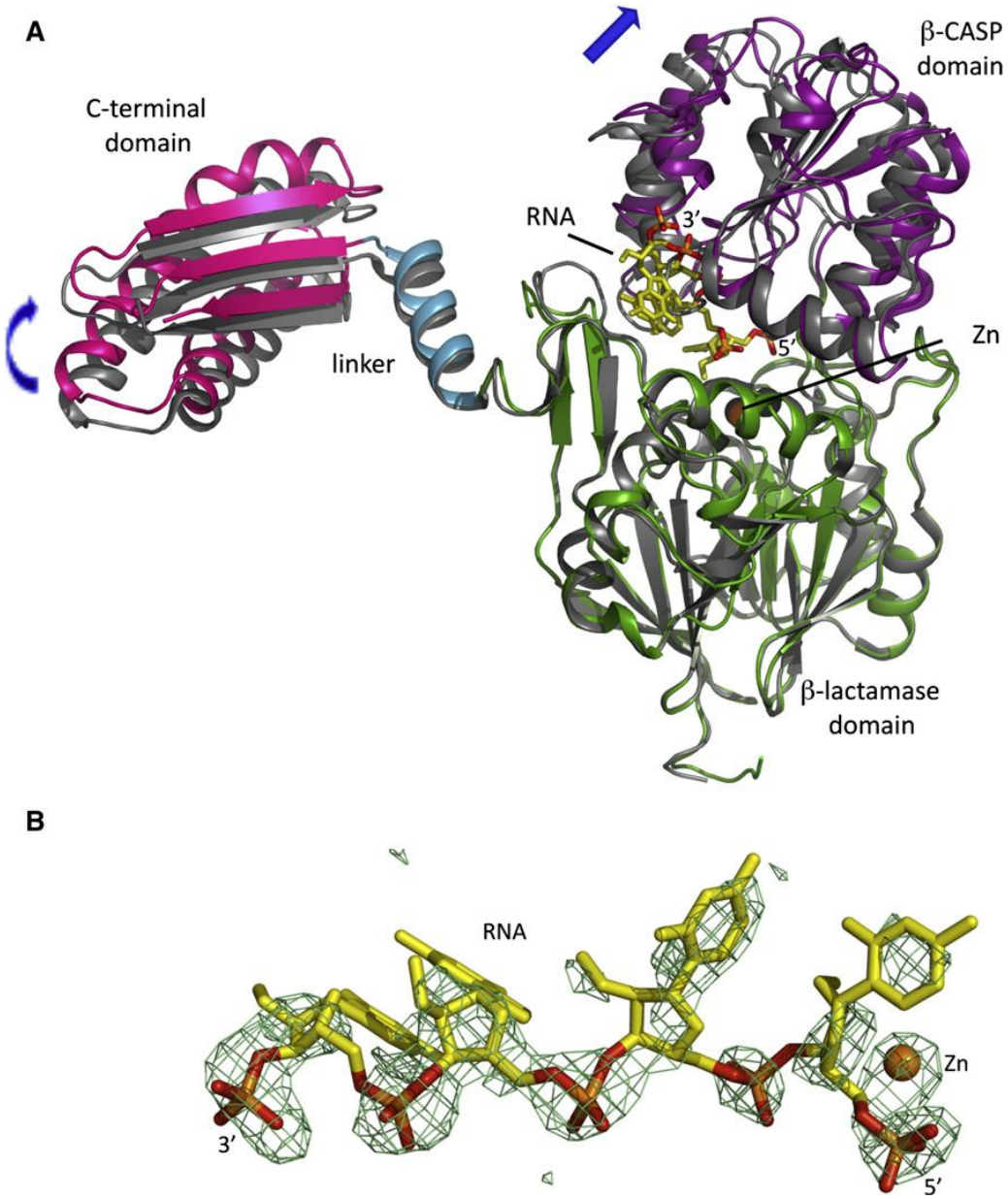


Figure 24: A**.** Comparaison des formes fermée (gris) et ouverte (colorée) d'un monomère de la RNase J avec le domaine β-CASP en violet, le domaine β-lactamase en vert, le linker en bleu et le domaine C-terminal en violet. L'ARN 4-mer est représenté en jaune sous forme bâton et les ions Zinc par des sphères oranges. Les flèches bleues indiquent les mouvements des domaines β-CASP et C-terminal par rapport au domaine β-lactamase lors de la liaison du substrat ARN. **B.** Visualisation de la densité électronique de l'ARN avec le squelette phosphate en orange et rouge et les bases en jaune (d'après Dorléans *et al.*, 2011)

Une deuxième structure de la RNase J de *T. thermophilus* a été résolue en présence d'un ARN 4-mer. La superposition de cette structure avec celle en complexe avec l'UMP montre certains changements conformationnels (rmsd 1,9 Å)[79].

Les rmsd sont de 2,3; 1,1 et 1,3 Å pour les domaines β-CASP, β-lactamase et C-terminal. Le changement de conformation principal a lieu au niveau du domaine β-CASP impliqué dans la liaison du substrat (figure 24A). En effet, les domaines β-CASP et β-lactamase s'éloignent d'un angle de 4° permettant de créer un tunnel pour la fixation d'un ARN simple brin. La taille de cette poche dans la conformation ouverte reflète bien la préférence de la RNase J pour les ARN simple brin. La densité électronique est de moins bonne qualité pour les bases que pour le squelette phosphate de l'ARN, ce qui suggère une plus grande flexibilité des bases. La RNase J se lie au squelette phosphate de l'ARN puisque c'est une exonucléase d'ARN simple brin, non spécifique de la séquence de son substrat.

Cette structure en interaction avec un ARN 4-mer dans une conformation ouverte montre que le tunnel de liaison à l'ARN se prolonge au-delà du site catalytique représenté par le Zn et la poche de liaison du 5' monophosphate (figure 24A et B). Plusieurs résidus chargés positivement sont présents dans cette extension du tunnel et pourraient lier un ARN dont l'extrémité 5' aurait dépassée la poche de liaison du phosphate et le site catalytique.

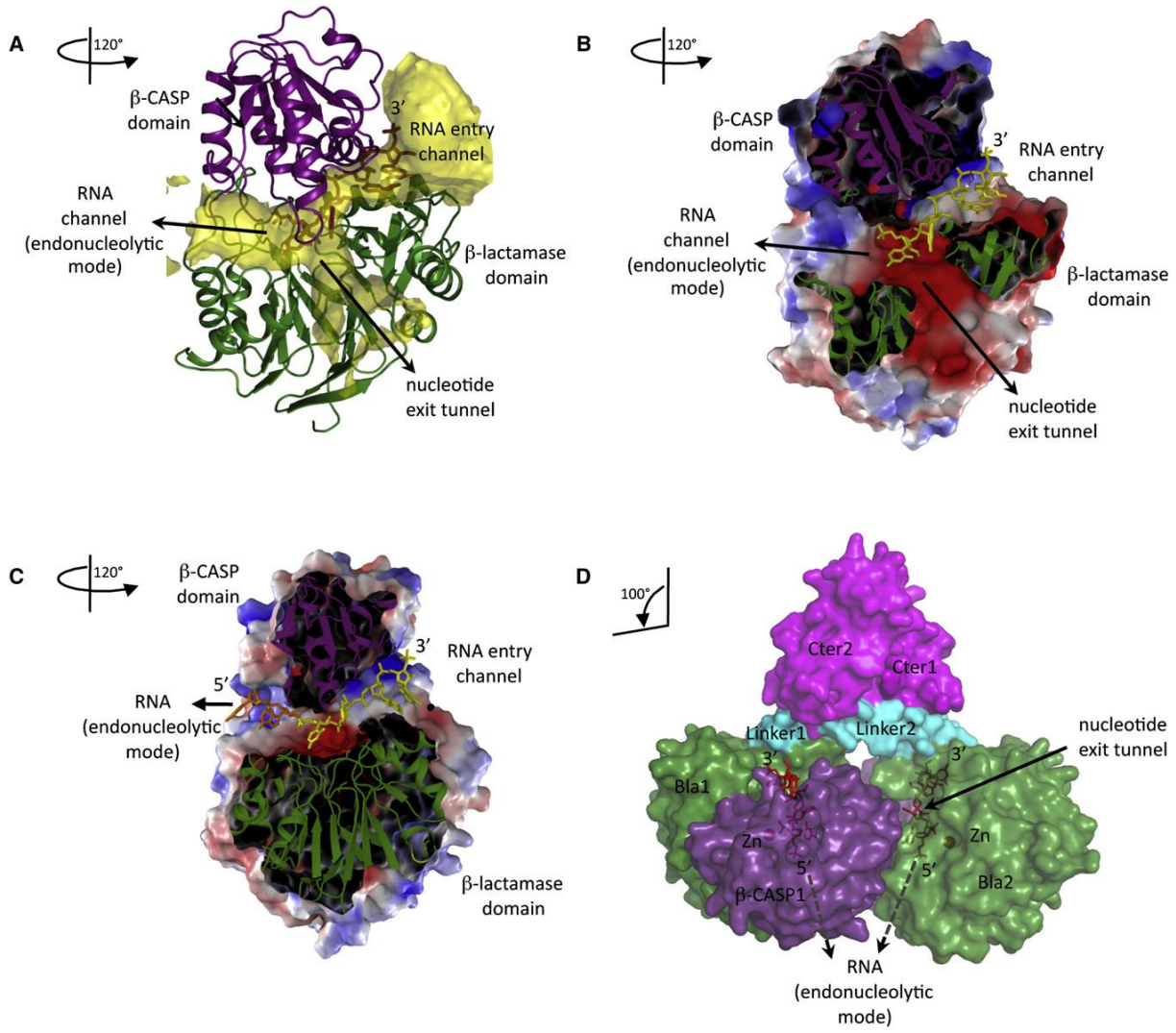


Figure 25: A. Représentation des trois tunnels présents dans un monomère de RNase J. Le domaine β-lactamase est représenté en vert, le domaine β-CASP en violet, l'ARN en rouge et les tunnels en jaune. **B.** Représentation de la surface électrostatique dans la même orientation qu'en (A). Les charges positives sont en bleu et les négatives sont en rouge. L'ARN est représenté en jaune. Les tunnels de liaison de l'ARN et le tunnel proposé pour la sortie des nucléotides sont indiqués par des flèches. **C.** Même représentation avec un ARN possédant deux nucléotides supplémentaires modélisés en orange au sein du tunnel proposé pour le clivage endonucléolytique. **D.** Vue de la surface du dimère de la RNase J. Les domaines β-lactamase de chaque monomère sont en vert, les domaines β-CASP en violet, les domaines C-terminal en rose et les linker en bleu, l'ARN est en rouge (d'après Dorléans *et al.*, 2011).

Les auteurs ont modélisé deux nucléotides supplémentaires en 5' de l'ARN 4-mer afin de voir comment un ARN plus long pourrait se loger dans ce tunnel (figure 25C)[79]. Pour le moment, il n'est pas encore compris comment l'ARN acquiert l'accés à ce tunnel lorsque l'enzyme est en mode endonucléase.

Pour finir, l'enzyme ne possède qu'un seul site catalytique, ce qui suggère que les clivages exonucléolytique et endonucléoytique sont catalysés par ce même site. L'hypothèse sur le fonctionnement endonucléase de cette enzyme est que la RNase J va fixer un ARN 5' triphosphorylé au dela du site catalytique en passant par le tunnel endo (figure 25A). L'enzyme va effectuer le clivage endonucléolytique sur cet ARN et une fois le clivage effectué l'enzyme va passer en mode exonucléolytique afin de commencer la dégradation processive du substrat[79].

Cette hypothèse est valable pour l'opéron *trp* qui est clivé de façon endonucléolytique par la RNase J puis ensuite dégradé par son activité exonucléase (figure partie I). De plus, il est possible que la RppH de *B. subtilis* (BsRppH) découverte récemment, puisse également déphosphoryler l'ARN afin de permettre directement une dégradation par l'activité exonucléase de la RNase J[63,64,67].

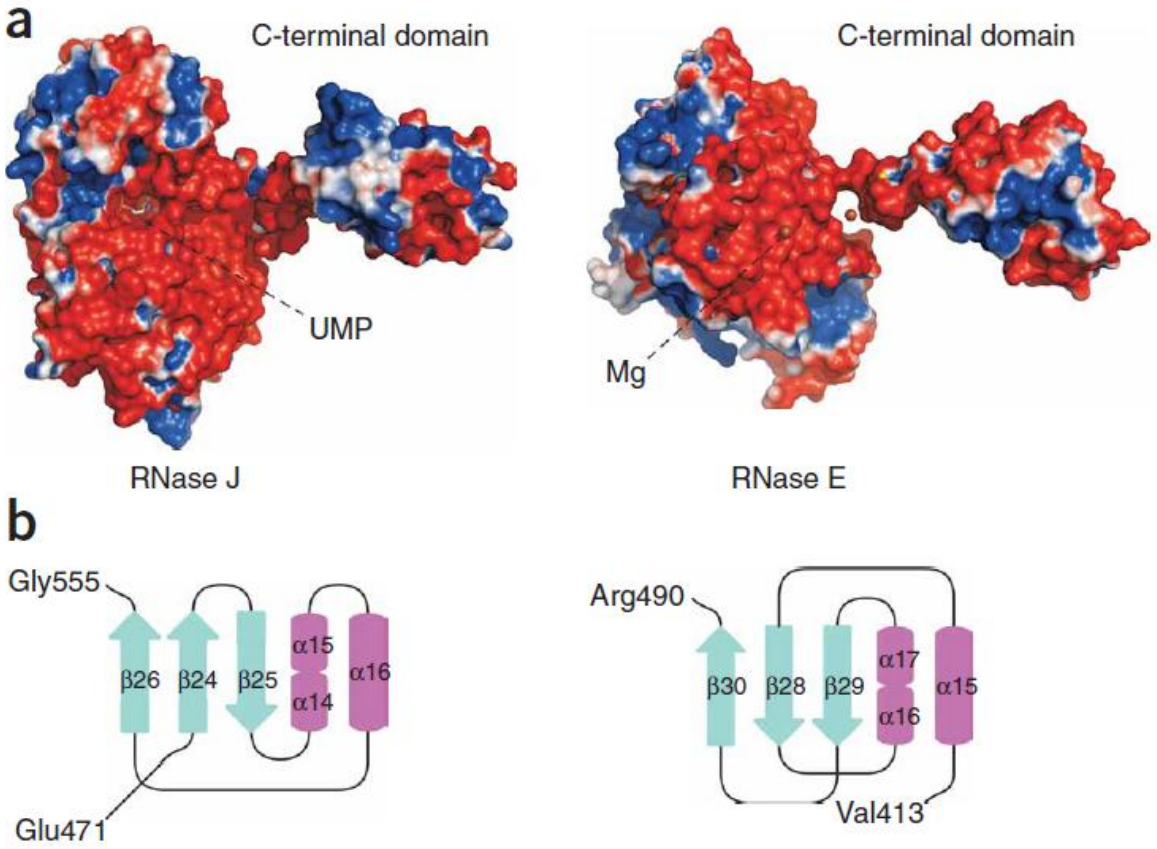


Figure 26: Comparaison des structures d'un monomère de la RNase J de *T. thermophilus* (à gauche) et de la RNase E d'*E. coli* (à droite). **a.** Comparaison des surfaces électrostatiques des deux enzymes. L'UMP et l'ion Mg^{2+}, présents respectivement dans les sites catalytiques de la RNase J et E, sont indiqués. **b.** Topologie du domaine C-terminal de la RNase J et du petit domaine de la RNase E. Ces deux domaines présentent la même architecture (trois brins β avec deux hélices α) (d'après Li de la Sierra-Gallay *et al.*, 2008).

iii. La RNase J possède une structure homologue à celle de la RNase E

Malgré leur l'absence d'homologie de séquence, les RNases J et E présentent des similitudes structurales. Nous avons déjà dit qu'elles possèdent la même spécificité de clivage endonucléolytique, au niveau de séquences riches en nucléotides AU et qu'elles préfèrent les ARN simple brin[36,47]. Par exemple, ces deux enzymes clivent le transcrit *thrS* au niveau du même site[42].

Premièrement, les deux enzymes ont la même forme globale, avec une distribution des charges électrostatiques en surface très similaires entre les deux enzymes (figure 26a). Enfin, le domaine C-terminal de la RNase J est constitué de trois brins β et de deux hélices α, architecture que l'on retrouve dans le petit domaine de la RNase E (figure 26b). Ces domaines semblent être impliqués dans la même fonction: le petit domaine de la RNase E permet l'association des deux dimères pour former l'homotétramère et le domaine C-terminal de la RNase J est également impliqué dans la dimérisation.

La similitude structurale des RNase J et E dont les séquences ne sont pas homologues indique que le mécanisme de dégradation des ARNm a subi une évolution convergente. Les mécanismes de catalyse sont les mêmes malgré que les domaines de liaison du substrat et catalytique soient complètement différents: la RNase J appartient à la famille des metallo-β-lactamase avec un domaine β-CASP, tandis que la RNase E est une hydrolase avec un domaine DNAse I.

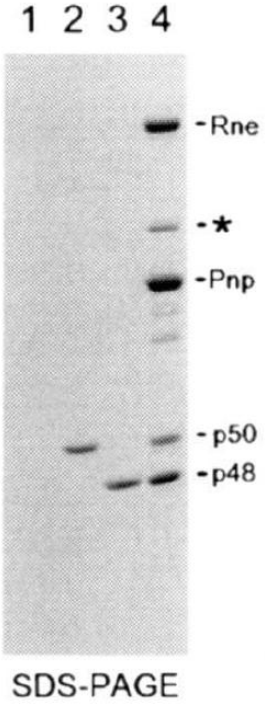


Figure 27: Purification du dégradosome et analyse sur gel SDS-PAGE. Les colonnes 1, 2 et 3 correspondent respectivement à l'expression dans des BL21(DE3) du plasmide pET11a seul (contrôle), du plasmide contenant l'hélicase RhlB (p50) et du plasmide contenant l'énolase (p48). La colonne 4 correspond à la reconstitution du dégradosome après passage sur colonne d'exclusion de taille (SP-Sepharose) (d'après *Py et al.*, 1996)

Tableau 7 : Résumé des zones d'interactions de la RNase E avec ses partenaires déterminés par Vanzo *et al.*, 1998.

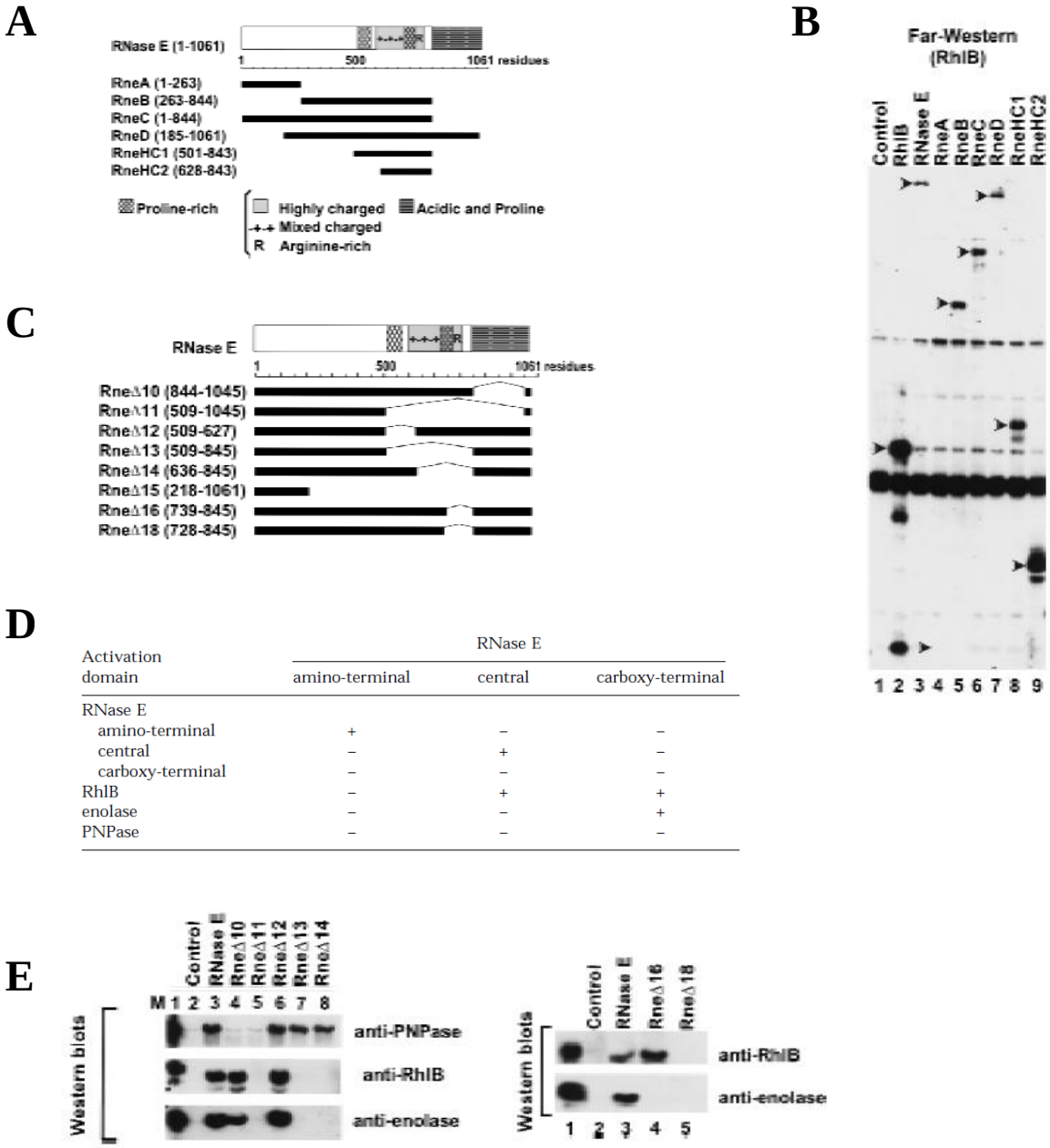
	Zones d'interactions sur la RNase E (acides aminés)
PNPase	846 à 1061
Enolase	739 à 845
RhlB	734 à 738

II.B. Un complexe multiproétique macromoléculaire de dégradation des ARNm chez *E. coli*, le dégradosome

i. Découverte et caractérisation du complexe

Le dégradosome est le nom donné au complexe de dégradation des ARNm chez les bactéries. Il a été découvert en 1994 dans deux laboratoires différents. Les équipes d'Henry Krisch et Christopher Higgis ont toutes deux montré une interaction entre la RNase E et la PNPase par copurification et immunoprécipitation[81,82]. Deux ans plus tard, il a été montré par coimmunopurification qu'une hélicase à ARN appelée RhlB faisait aussi partie de ce complexe[65]. Cette hélicase s'est révélée être indispensable à l'activité exoribonucléolytique de la PNPase pour la dégradation des structures secondaires au sein des ARN[65]. C'est également dans cette même étude, par la même méthode, que l'énolase, une enzyme impliquée dans le métabolisme des sucres (glycolyse), a été identifiée comme interagissant également avec la RNase E . Le rôle de l'énolase au sein du dégradosome reste encore inconnu.

Après avoir identifié les partenaires de la RNase E (PNPase, l'hélicase d'ARN RhlB et l'énolase), il était important de déterminer les zones d'interactions avec ces partenaires au sein de la séquence de la RNase E. C'est l'équipe d'Agamemnon Carpousis, en 1998, qui a déterminé que la partie C-terminale de la RNase E était la région impliquée dans les interactions avec les partenaires[75]. Pour cela, ils ont combiné diverses méthodes. Ils ont utilisé la méthode de Far-Western blot permettant de détecter des interactions protéine-protéine par immunoblot, mais ils ont également utilisé la technique de double-hybride (annexe 1) chez la levure ainsi que la coimmunopurification. Pour cela ils ont réalisé diverses constructions de la RNase E en déletant certaines parties de la protéine, afin de déterminer grossièrement les zones impliquées dans les interactions avec chacuns des partenaires (Figure 28A).



Activation domain	RNase E amino-terminal	RNase E central	RNase E carboxy-terminal
RNase E			
amino-terminal	+	–	–
central	–	+	–
carboxy-terminal	–	–	–
RhlB	–	+	+
enolase	–	–	+
PNPase	–	–	–

Figure 28: A. Constructions réalisées pour le Far-Western blot et exprimée avec le plasmide pET11a chez des BL21 (DE3). **B.** Far-Western blot des diverses constructions réalisées. Après transfert sur membrane, cette dernière est incubée avec une solution de RhlB puis l'interaction avec les diverses constructions de la RNase E est détectée à l'aide d'un anticorps anti-RhlB. La détection de RhlB est montrée par une flêche noire. **C.** Constructions réalisées pour la coimmunopurification et exprimée avec le plasmide pAM chez la souche PBRN1. **D.** Résultats d'interactions de la RNase E avec ses partenaires par double-hybride chez la levures. Pour cela trois constructions de la RNase E ont été créées : RNase E amino-terminale (1-528), centrale (500-752) et amino-terminale (734-1061). **E.** Résultats de coimmunopurification vérifiées par Western-blot par utilisation d'anticorps anti-PNPase, anti-énolase et anti-RhlB. (d'après Vanzo *et al.*, 1998)

Dans le cas de la méthode de Far-western blot, plusieurs constructions ont été utilisées afin de déterminer la zone d'interactions avec l'hélicase RhlB (Figure 28A). Dans cette technique, les diverses constructions de la RNase E sont déposées sur gel d'acrylamide dénaturant et transférées sur membrane de nitrocellulose, cette membrane est ensuite incubée avec l'hélicase sous forme native puis l'hélicase est détectée avec de anticorps anti-hélicase RhlB. Ils ont constaté que l'hélicase interagit avec la région centrale de la RNase E, en limitant cette zone d'interaction entre les acides aminés 628 à 843 (Figure 28B).

Ensuite, d'autres constructions ont été réalisése afin de déterminer les interactions avec les autres partenaires, enolase et PNPase, par coimmunopurification (Figure 28C). Dans cette technique, les auteurs ont utilisé des anticorps anti-RNase E afin de purifier les complexes pour chaque construction. Ensuite les interactions sont détectées par Western blot (figure 28E) en utilisant des anticorps contre chaque partenaire (anti-PNPase, anti-RhlB et anti-enolase). Pour la PNPase, l'interaction est détectée sauf avec les constructions RneΔ10 et RneΔ11, ce qui a permis de montrer que la zone d'interaction pour la PNPase est présente dans la partie C-terminale de la RNase E, dans les 217 derniers acides aminés (Tableau 7).

Ensuite, pour l'énolase l'interaction avec la RNase E a d'abord été détectée par double hybride avec la partie C-terminale de la RNase E uniquement (acides aminés 734 à 1061), mais cette interaction est également détectée en coimmunopurification avec les constructions RneΔ10 et RneΔ12, et l'on perd l'interaction avec les délétions courtes RneΔ16 et RneΔ18, ce qui a permis de montrer que la partie 734 à 845 contient le site d'interaction avec l'énolase (figure 28E). La méthode par double hybride a également permis, pour l'hélicase RhlB, de restreindre le site d'interaction pour RhlB aux acides aminés 734 à 752 (Figure 28D). Puis par coimmunopurification, on constate que pour la construction RneΔ16 que l'interaction est conservée, tandis qu'elle est perdue pour la construction RneΔ18, ce qui a permis de montrer que les acides aminés importants pour l'interaction avec RhlB est un segment allant de 734 à 738 (figure 28E). Les interactions avec la PNPase et l'énolase n'ont pu être détectées par la méthode de Far-Western car, comme nous allons le voir par la suite, ces zones d'interactions sont de courts segments structurés, (hélice α et brin β), dont la structure est perdue en gel d'acrylamide dénaturant. Pour conclure, après cette étude, il fut claire que la partie C-terminale de la RNase E est la zone impliquée dans les interactions avec les partenaires. Les sites d'interactions précis restant encore un peu flous.

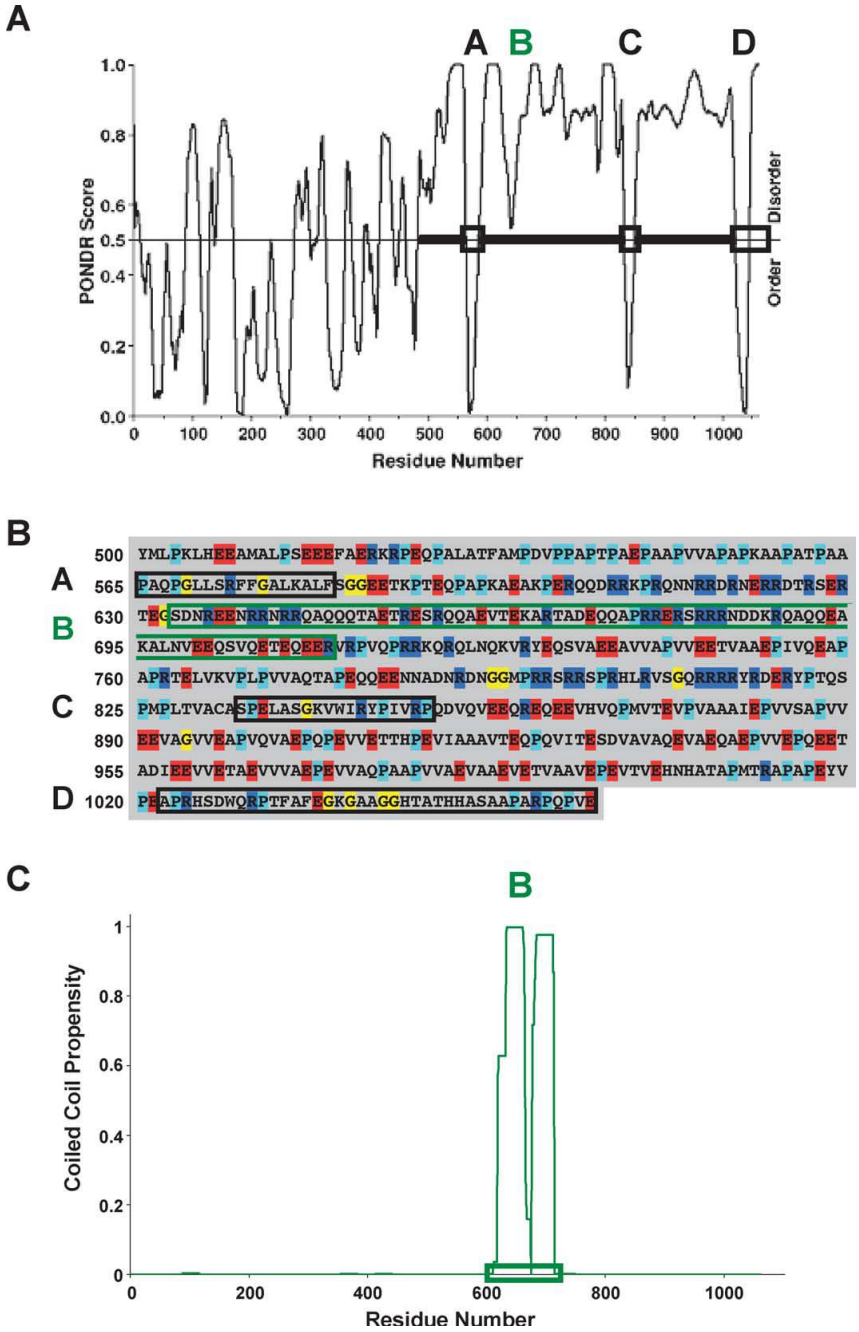
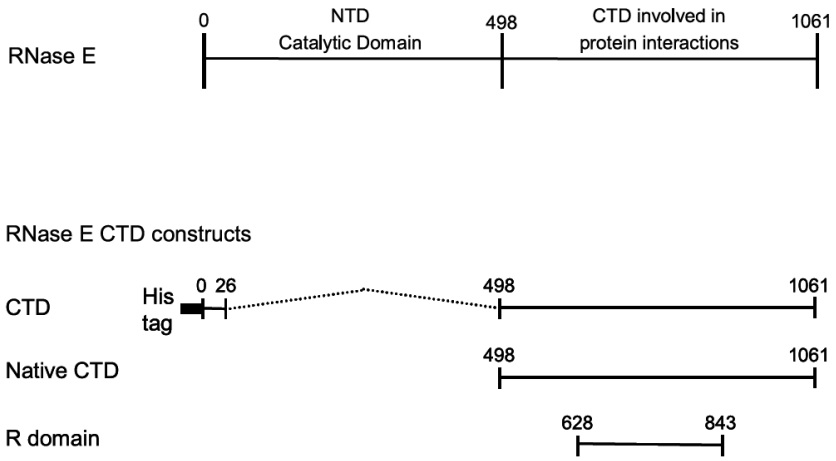


Figure 29: A. Prédiction du désordre au sein de la séquence de la RNase E (1-1061) par le prédicteur de désordre PONDR. Les zones déstructurées sont au-desus de la ligne horizontale. **B.** Séquence primaire du domaine CTD de la RNase E (500-1061). Les résidus prolines (Cyan), arginines (Bleu), glycines (Jaune) et glutamates (rouge) sont surlignés afin de mettre en évidence leurs fortes présences au sein du domaine CTD. **C.** Prédiction des structures coiled-coil au sein de la séquence de la RNase E. Prédiction faite à partir du serveur COILS. (d'après Callaghan *et al.*, 2004)

ii. Caractérisation du domaine C-terminal intrinsèquement désordonné de la RNase E

En 2004, l'équipe de Ben Luisi a étudié les domaines N-terminal (NTD) et C-terminal (CTD) de la RNase E afin de déterminer précisément les acides aminés impliqués dans les interactions avec les partenaires du dégradosome[83]. Tout d'abord, une analyse *in silico* de la séquence de la RNase E par le prédicteur de désordre PONDR montre clairement que le domaine C-terminal est déstructuré par rapport au domaine N-terminal (figure 29A) qui lui est prédit pour être structuré. De plus, une analyse par le serveur COILS montre qu'une partie du domaine CTD, acides aminés 633-662 et 685-712, est prédit pour se structurer en domaine coiled-coil, cette région est notée segment B (figure 29C). Cette région se superpose avec le domaine d'interaction avec l'ARN (RBD, RNA binding domain). Il a donc été supposé que ce domaine coiled-coil serait impliqué dans l'interaction avec l'ARN. De plus, cette analyse a également révélé que ce domaine CTD est très riche en acides aminés Arg, Glu, Pro et Gly, des acides aminés que l'on trouvent en abondance dans les protéines intrinsèquements désordonnées (figure 29B), ce qui est en accord avec la prédiction du désordre de ce domaine (figure 29A). Ensuite, l'analyse par PONDR a montré qu'au sein de ce domaine CTD il y aurait de courtes régions structurées, notées A, C et D (figure 29A), qui correspondent respectivement aux acides aminés 565-585, 839-850 et 1021-1061.

A



B

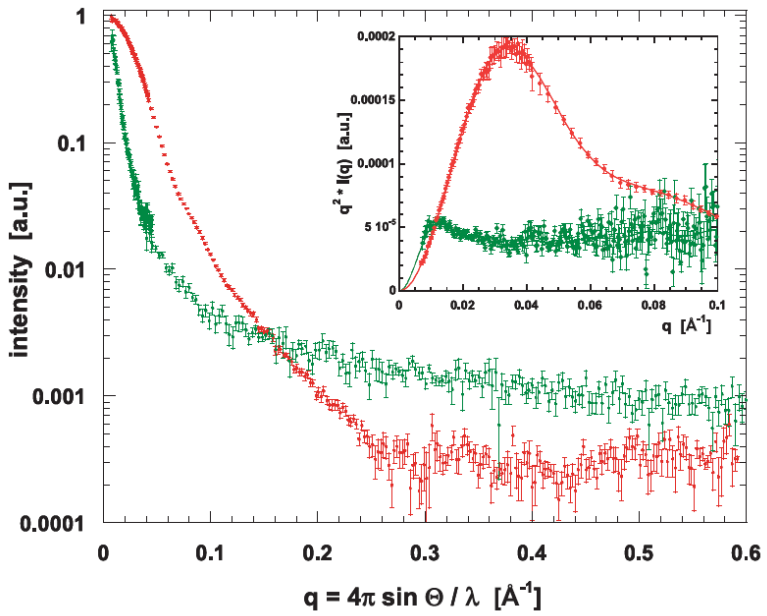


Figure 30: A. Constructions réalisées pour l'étude en SAXS du domaine C-terminal (CTD) de la RNase E. **B.** Profil de diffusion des rayons X du domaine N-terminal (1-498) en rouge et du domaine C-terminal en vert (6His-498-1061). La représentation de Kratky est insérée en haut à droite de la courbe de diffusion. **C.** Sites d'interactions des partenaires identifiés par spectrométrie de masse. Les segments A, B, C et D correspondent respectivement aux zones d'interaction avec l'ARN, avec l'hélicase RhlB (633-662) et l'ARN (685-712), avec l'énolase et la PNPase. (d'après Callaghan *et al.*, 2004)

Suite à cela des constructions ont été réalisées, une du domaine NTD (1-498) et trois pour le domaine CTD (498-1061) (figure 30A). Les trois constructions du domaine CTD avaient chacunes un objectif :

- CTD (498-1061) avec étiquette histidine en N-terminal pour l'analyse SAXS
- CTD (498-1061) pour la détermination précise des zones interagissants avec l'énolase et la PNPase
- Domaine R (628-843) pour l'analyse des interactions avec RhlB et ARN (Vanzo *et al.*, 1998)

L'analyse par diffusion des rayons X aux petits angles (SAXS) du domaine CTD montre un profil de diffusion caractéristique d'une protéine non-structurée à en juger par le comportement asymptotique de la courbe issue du graphique de Krathy. A l'inverse le domaine NTD est très structuré et ne possède pas cette caractéristique asymptotique sur la représentation de Krathy (figure 30B). Cette analyse en SAXS a donc confirmé la prédiction de désordre faite avec le serveur PONDR, et il a été supposé que les segments identifiés pour être structurés au sein de ce complexe seraient potentiellement les zones d'interactions avec les partenaires protéiques et l'ARN. Ceci à donc été étudié pour l'énolase et la PNPase en utilisant la forme native du CTD (498-1061 sans étiquette). Ce domaine c'est avéré être digéré lors de la purification malgré la présence d'anti-protéase. Après analyse par séquençage amino-terminal, le domaine CTD commence à l'acide aminé 580 et non 498. Mais une analyse en protéolyse ménagée à également montré qu'il finissait à l'acide aminé 1038.

Le complexe CTD-énolase fut observé en gel filtration et il c'est avéré sensible à la protéolyse par la protéase K et la chymotrypsine. Après digestion, puis migration sur gel natif du complexe et séquençage amino-terminal, le fragment du domaine CTD lié à l'énolase commence à l'acide aminé 816 jusqu'à l'acide aminé 1038. Cette zone recouvre le segment C déterminé par le serveurs PONDR. Les auteurs ont donc créée un peptide synthétique correspondant au segment C (833-850). L'interaction fut analysée par spectrométrie de masse (Nanoflow-ESI). Le pic prédominant ayant une masse de 93 158 (± 14) Da et correspondant au dimère d'énolase lié à un peptide (Figure 31A). Le domaine CTD protéolysée naturellement (580-1038) a également été utilisé pour tester l'interaction avec la PNPase, mais aucun complexe n'a été observé en gel filtration. Il est donc possible que la PNPase interagisse avec la partie digérée se situant après l'acide aminé 1038 et correspondant au segment D prédit pour être structuré.

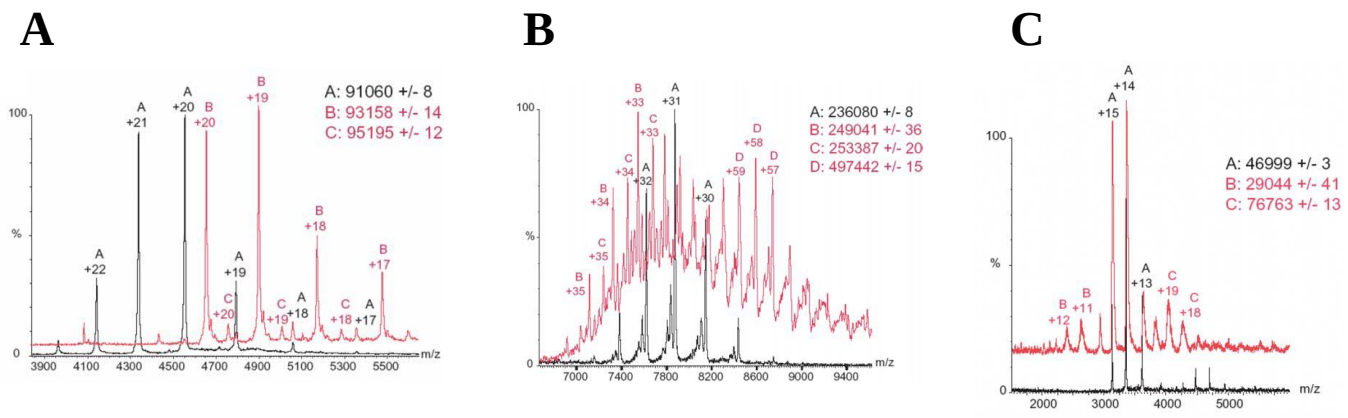


Figure 31: Étude par spectrométrie de masse (Nanoflow-ESI) des interactions. **A.** Spectre de masse ESI de l'énolase seule (noire - série A) et en complexe avec le peptide (833 - 850) (rouge), la série B (majoritaire) correspond au dimère d'énolase en interaction avec le peptide (un peptide $MM_{théo}$ = 2042 Da) et la série C (minoritaire) correspond au dimère d'énolase en interaction avec deux peptides. **B.** Spectre de masse ESI de la PNPase seule (noire - série A) et en complexe avec le peptide (1021 - 1061) (rouge). Les série B et C correspondent respectivement à l'homotrimère de PNPase en interaction avec trois et quatre peptides (un peptide $MM_{théo}$ = 4333 Da. La série D correspond à un dimère de trimère de la PNPase en interaction avec six peptides. **C.** Spectre de masse ESI de RhlB seule (noire - série A) et en complexe avec le domaine R (rouge). La série B correspond au domaine R seul ($MM_{théo}$ = 28 822 Da) et la série C correspond au complexe RhlB - domaine R (d'après Callaghan et al., 2004)

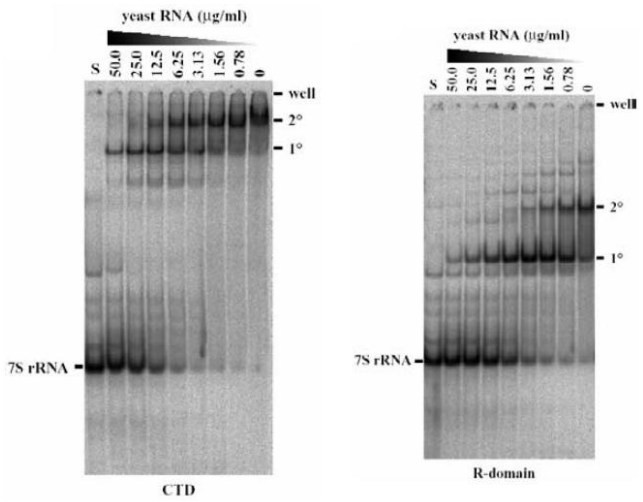


Figure 32: Gel retard (EMSA) du domaine CTD et R en présence de l'ARNr 7S. Le complexe est formé, pour chaque domaine, en présence d'un compétiteur, un ARN de levure. Les complexes sont annotés sur le gel par les numéros 1° et 2°, correspondant respectivement à deux complexes formés avec l'ARNr 7S par chaque domaine. Le complexe 1° se forme même en présence de 5 à 10 excès de l'ARN de levure.

Un peptide synthétique (1021-1061) fut donc créé afin de mettre en évidence l'interaction. Cette interaction, analysée par spectrométrie de masse (Nanoflow-ESI), montre la présence d'un pic prédominant (Figure 31B, pic B) avec une masse exacte de 249 041 (± 36) Da. Cette masse correspond donc a l'homotrimère de la PNPase lié à trois peptides. Chaque monomère de PNPase lie donc un peptide (figure 31B). Enfin, le domaine R (628-843) a été utilisé pour étudier l'interaction avec l'hélicase RhlB. Cette interaction a également été analysé par spectrométrie de masse (Nanoflow-ESI) et un pic majoritaire a été détecté avec une masse de 76 763 (± 13) correspondant au monomère de RlB lié au domaine R (figure 31C). Des expériences d'activités ont également montré que l'hélicase RhlB était 25 fois plus active en complexe avec le domaine R. Cette interaction protéine - protéine a donc un effet stimulateur sur l'activité d'RhlB. Ce qui a confirmé une étude précédemment faite sur ce même domaine en interaction avec RhlB par Vanzo *et al*., 1998. De plus, une étude par CD a montré qu'aucun changement structural n'avait lieu au niveau du domaine R après interaction de l'hélicase. Enfin, il a également été montré que les interactions de l'énolase et de la PNPase avec la RNase E n'avaient aucuns effets sur leurs activités.

Cette étude à permis de restreindre les zones d'interactions, pour l'énolase et la PNPase, au sein du domaine CTD de la RNase E. En effet, ces zones correspondent respectivement à deux peptides de 18 et 40 acides aminés. De plus, cette étude a permis de déterminer la stœchiométrie des complexes et montre que l'activité de l'hélicase d'RhlB est stimulée lorsqu'elle est en interaction avec le domaine R. De plus, cette étude montre également que le domaine CTD lie l'ARN, ainsi que le domaine R qui contient le domaine d'interaction avec l'ARN riche en arginine (AR2) ainsi que le site d'interaction à l'hélicase. On peut donc supposer que la présence de l'hélicase au sein du dégradosome permet d'éliminer les structures secondaires afin de faciliter la dégradation de l'ARN. Pour finir, au sein du domaine CTD de la RNase E, il y a deux domaines d'interactions avec l'ARN, le RBD (RNA binding domaine) et le domaine riche en arginine (AR2).

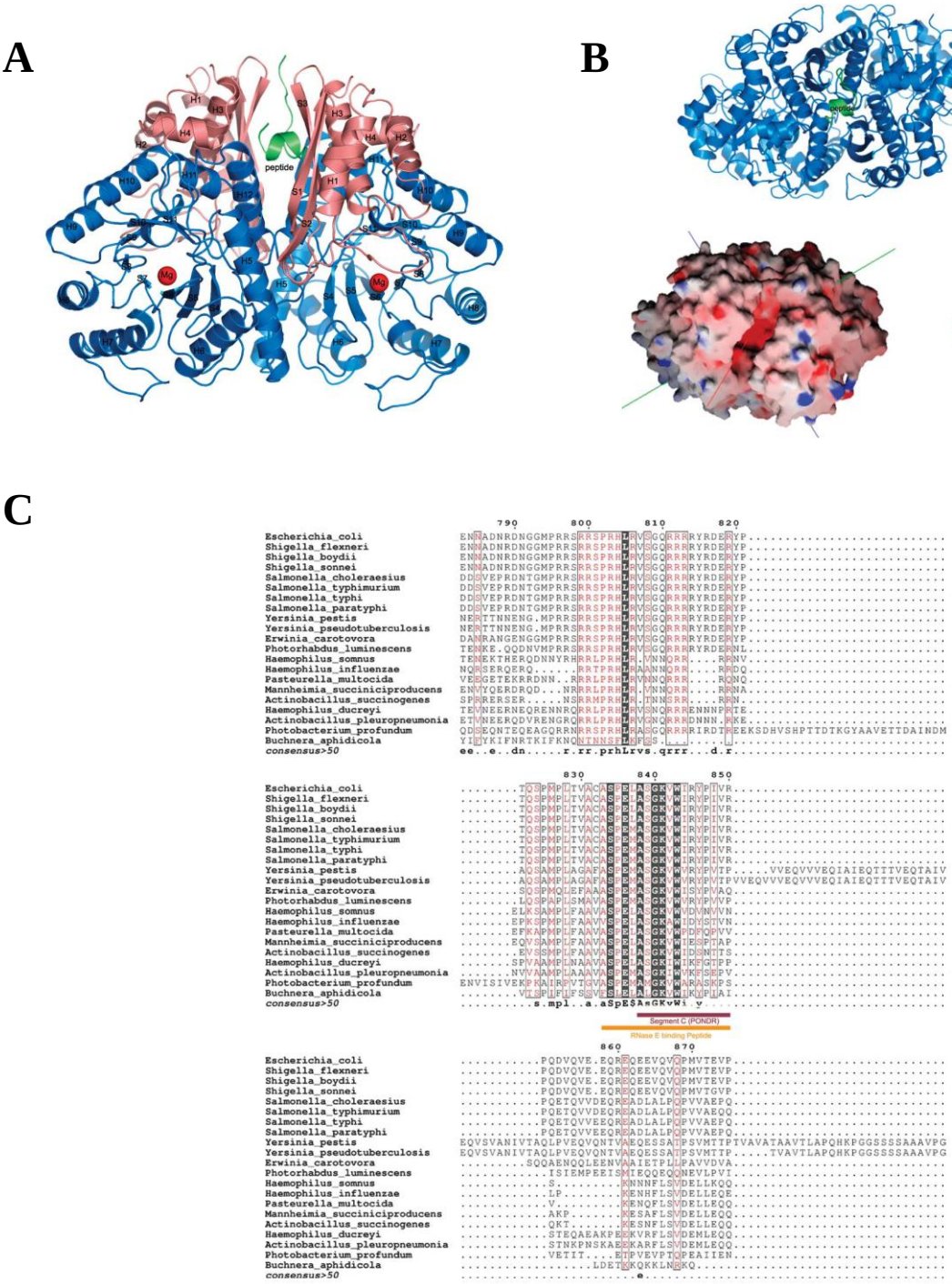


Figure 33: A. Structure cristallographique de l'énolase en interaction avec le peptide de la RNase E (833-850) représenté en vert. Les sphères rouges représentent les ions Mg^{2+} présent au niveau des sites actifs. **B.** Vue du dessus du dimère d'énolase en interaction avec le peptide. En bas à gauche, la surface électrostatique de l'énolase avec le sillon d'interaction chargé positivement, et en bas à droite le peptide dont la surface électrostatique est négative. **C.** Alignement de séquence du domaine CTD de la RNase E chez différentes γ-Protéobactéries. Le segment surligné en rouge correspond à la zone prédite pour former une structure secondaire et le segment surligné en jaune est celui utilisé pour l'alignement avec BLAST (d'après Chandran et Luisi, 2006).

iii. Structures cristallographiques de la PNPase et de l'énolase en interaction avec les microdomaines du domaine C-terminal de la RNase E

Structure de l'énolase en interaction avec son microdomaine (833-850)

L'énolase est un enzyme présente au sein des organismes des trois domaines du vivants. Elle intervient dans la voie métabolique appelée la glycolyse, en effet, elle catalyse la réaction d'interconversion du phosphoénolpyruvate en 2-phospho-D-glycérate[84]. La glycolyse est une voie métabolique importante car elle permet, par le catabolisme du glucose, la production d'ATP qui fournit l'énergie nécessaire à de nombreux processus biochimiques dans la cellule.

Cette protéine de 48 kDa forme un dimère et 1/10ème de l'énolase présente chez *E. coli* est séquestrée au sein du complexe du dégradosome. La structure de l'énolase en interaction avec le peptide de la RNase E (833-850) a été résolue par cristallographie aux rayons X à 1,6 Å de résolution[85]. Cette structure à permis de montrer, dans un premier temps, comment se faisait cette interaction et de valider la prédiction de désordre faite par PONDR sur le domaine CTD de la RNase E[83]. En effet, le peptide de la RNase E cristallisé avec l'énolase était prédit pour être structuré et on constate que ce peptide forme une petite hélice α (figure 32A). Cette interaction a lieu à l'interface de dimérisation de l'énolase, très éloignée du site actif, ce qui explique pourquoi l'interaction de l'énolase avec la RNase E n'impacte pas l'activité de l'énolase[83].

Pour le moment le rôle précis de l'énolase au sein d'un complexe de dégradation des ARNm reste encore imprécis, mais il a été montré que les acides aminés du site de fixation de l'énolase chez la RNase E sont très conservés chez la plupart des γ-Protéobactéries[85]. Ce qui laisse supposer que cette enzyme est également présente au sein du dégradosome d'autres bactéries. Un alignement a été réalisé avec différentes RNase E appartenant à diverses γ-Protéobactéries et on constate que le site d'interaction est conservé chez certains pathogènes comme *Yersinia pestis*, *Salmonella typhimurium*, *Shigella flexneri*,... De plus lorsque l'interaction de l'énolase avec la RNase E est perturbée, cela a des conséquences sur la dégradation des ARNm codant pour des enzymes impliquées dans le métabolisme énergétique. Il est supposé que l'énolase aurait un rôle de régulation et ferait le lien entre le métabolisme des sucres et celui des acides nucléiques. Il serait donc intéressant de regarder qu'elle effet aurait cette perturabtion de l'interaction chez certaines bactéries pathogènes.

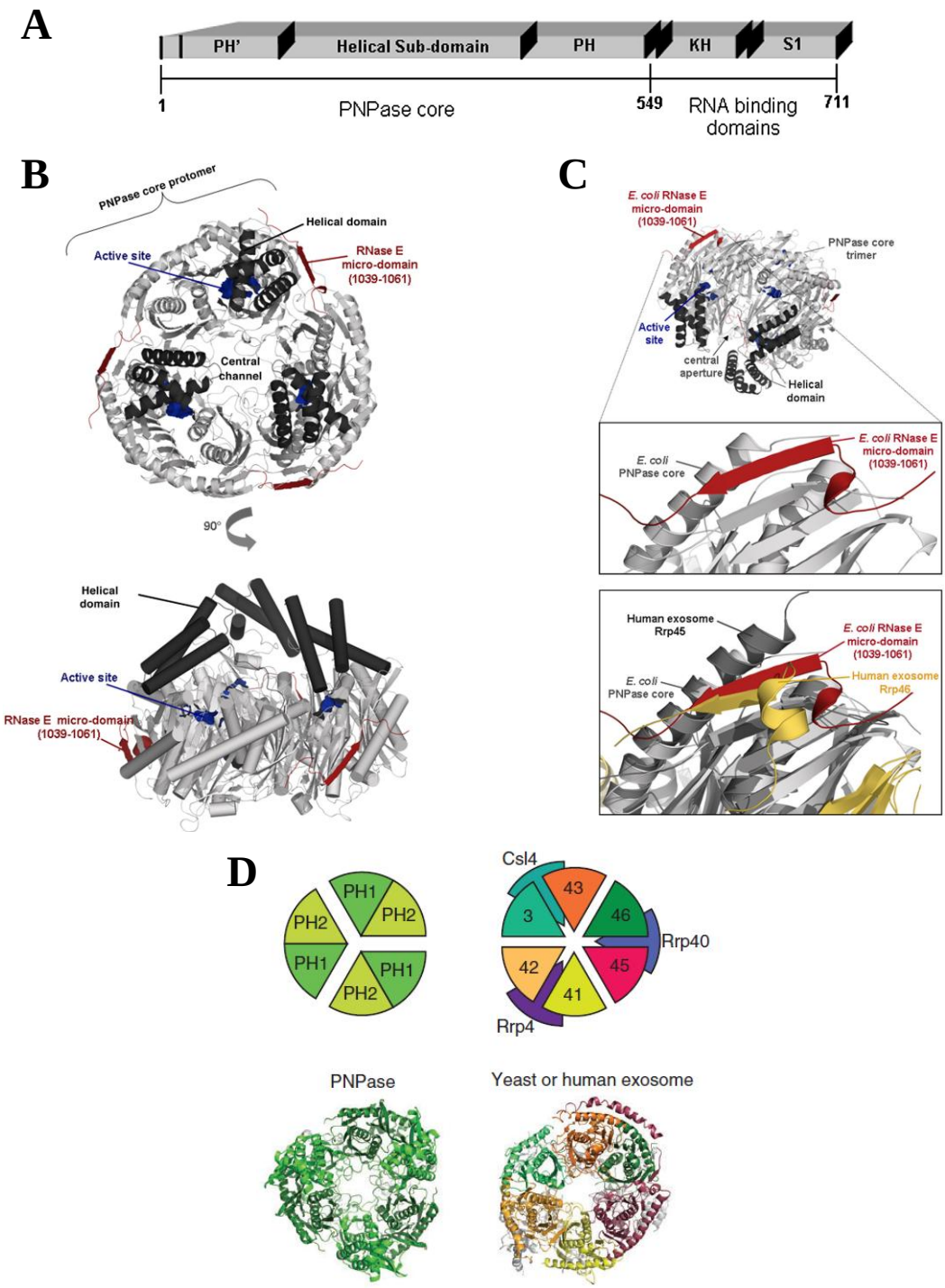


Figure 34: A. Organisation des domaines de la PNPase (PH' = RNase PH1, PH = RNase PH2). **B.** Structure cristallographique de l'homotrimère de la PNPase en interaction avec le microdomaine représenté en rouge. **C.** En haut, visualisation de l'interaction du microdomaine avec le feuillet β du domaine RNase PH1. En bas, comparaison de l'interaction du microdomaine de la RNase E avec la PNPase et de l'interaction Rrp45-Rrp46 au niveau de l'éxosome humain (d'après Nurmohamed *et al.*, 2009)) **D.** Comparaison de la structure de la PNPase avec celle de l'exosome humain (d'après Lykke-Andersen *et al.*, 2009).

Structure de la PNPase en interaction avec son microdomaine (1021-1061)

La PNPase (3'-exonucléase) est une protéine de 77kDa formant un homotrimère. Plusieurs structures cristallographiques de cette enzyme ont été résolues chez différents organismes (*Streptomyces antibioticus*, *E. coli*, *Coleobacter crescentus* et *Homo sapiens*)[31,86–88]. La PNPase est présente chez toutes les bactéries mais égalment au niveau des organelles des eucaryotes (mitochondrie et chloroplaste)[30]. Ces structures révèlent que l'organisation générale de l'enzyme est conservée. Une de ces structures a également été résolue avec un ARN, ce qui a permis d'obtenir des informations sur le mécanisme de cette enzyme. Le trimère forme un anneau compacte avec au centre un tunnel par ou passe l'ARN et dans lequel est présent le site catalytique. La PNPase est constituée de deux domaines RNase PH (PH1 et PH2), seul le domaine PH2 est actif. Ensuite, il y a également un domaine KH et un domaine S1 qui permettent de lier l'ARN (figure 34A). La structure cristallographique de la PNPase en interaction avec le microdomaine de la RNase E (1021-1061) a été résolue à 2,4 Å de résolution[89]. Dans cette structure, seul les acides 1039 à 1061 du microdomaine montraient une densité interprétable. Cette structure montre bien que chaque monomère de PNPase lie un microdomaine, ce qui correspond à la stoechiométrie déterminée par spectrométrie de masse, trois peptides pour un trimère de PNPase[83]. L'affinité du peptide a été mesurée par titration calorimétrique isotherme (ITC), cette affinité est 0,9 µM. Cette affinité n'a pour le moment pas été mesurée avec la RNase E entière. De plus cette interaction se fait au niveau du domaine RNase PH1 via des liaisons hydrogènes. L'interaction est très éloignée du site actif, ce qui permet de comprendre pourquoi cette interaction n'influence pas l'activité de la PNPase. On peut supposer que ce domaine PH1, dont l'activité n'a pas été démontrée, n'est peut être présent que pour effectuer cette interaction avec la RNase E.

Le point important de cette structure est que cette interaction est très similaire à celle rencontrée entre les sous-unités Rrp45 et Rrp46 de l'exosome, un complexe de dégradation des ARNm à activité 3'-exonucléase et présent chez les cellules eucaryotes et les archées[90]. En effet, l'interaction du micrdomaine de la RNase E a lieu au niveau d'un feuillet β du domaine RNase PH1, le microdomaine a d'ailleurs une structure secondaire en brin β permettant l'interaction. Au niveau de l'exosome, c'est la même chose, l'interaction a lieu au niveau d'un feuillet β de la sous-unité Rrp45, et le domaine d'interaction de la sous-unité Rrp46 a la même structure secondaire que le microdomaine de la RNase E.

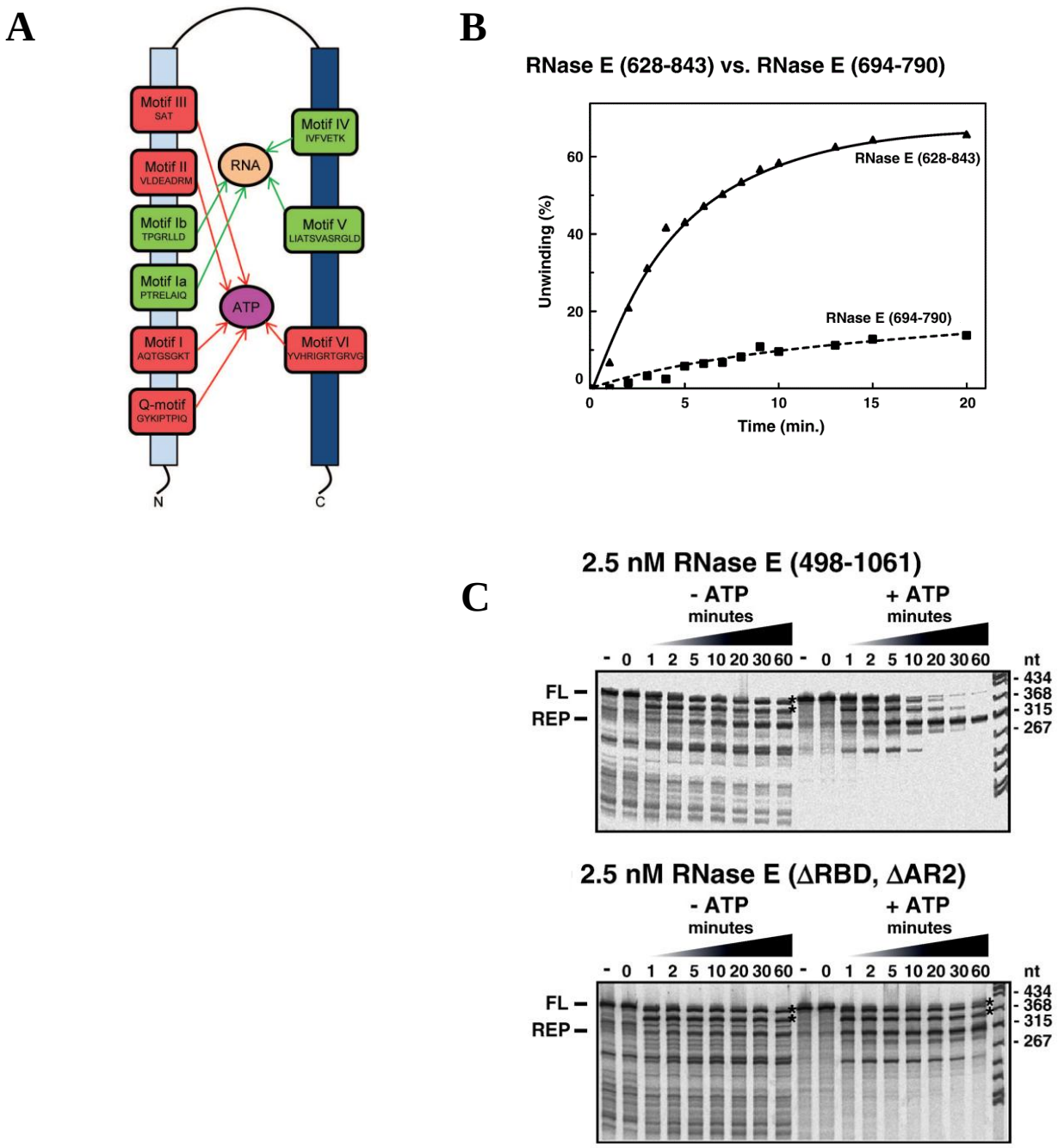


Figure 35: A. Organisation des domaines des hélicases de type DEAD et des motifs impliqués dans la liaison de l'ATP et de l'ARN. Ces hélicases sont constituées de deux domaines, un domaine RecA n°1 (N-ter) et RecA n°2 (C-ter). Le motif DEAD est impliqué dans la liaison de l'ATP. **B.** Dépliement du duplex d'ARN par l'helicase RhlB en interaction avec le domaine CTD 628-843 (triangles) ou 694-790 (carrés). Les mesures en pourcentage ont été faites à partir d'un gel retard en condition native après coloration de l'ARN. **C.** Activité 3'exonucléase de la PNPase après reconstitution du mini-dégradosome contenant la PNPase, l'hélicase RhlB et le domaine CTD 498-1061 avec (en haut) ou sans (en bas) les domaines d'interactions de l'ARN.

iv. Les hélicases d'ARN de type DEAD-box et leurs rôles dans le dégradosome

Les hélicases d'ADN ou d'ARN sont connues depuis les années 80[91]. Ces enzymes NTP-dépendantes permettent d'éliminer les structures secondaires présentent au sein des ARN. Elles sont impliquées dans de nombreux processus cellulaire important comme la réplication et la transcrption de l'ADN, le repliement de l'ARN ainsi que la biogenèse du ribosome[92]. Ici, nous allons nous intéresser aux hélicases d'ARN et plus spécifiquement à celle impliquée dans la dégradation des ARNm. En effet, chez *E. coli* il a été montré qu'une hélicase d'ARN, RhlB, s'associe avec la RNase E dans le cadre de la formation du dégradosome[65]. Cette hélicase appartient à la superfamille des hélicases DEAD. DEAD faisant référence à un motif conservé au sein du domaine ATPase de l'enzyme (figure 35A). Ces hélicases sont organisées en deux domaines principaux, le domaine RecA n°1 en N-terminale et le domaine RecA n°2 en C-terminale. RecA est une ATPase ADN-dépendant impliquée dans la recombinaison de l'ADN[91].

Tout comme la PNPase et l'énolase, l'hélicase RhlB fait partie du dégradosome d'*E. coli*. Il a été montré rapidement après la découverte du dégradosome, que cette hélicase facilite la dégradation des fragments d'ARN par la PNPase via l'élimination des structures secondaires[65]. En effet, en absence de l'hélicase, la PNPase est incapable dégrader un fragment d'ARN portant une structure secondaire[65]. La zone d'interaction de l'hélicase au niveau du domaine CTD de la RNase E est présente dans la région 628 à 843. Cette région comprend également deux domaines d'interactions avec l'ARN, le RNA binding domain (RBD) ainsi qu'un domaine riche en arginine (AR2). Des expériences de protéolyse ménagée à la trypsine, chymotrypsine et protéase V8 sur le complexe RhlB-RNase E(628-843) couplées à des séquençages amino-terminal, ont montré que le site d'interaction serait resserré aux acides aminés 698-762. Ce résultats a été conforté par un aligenement de séquence qui montre que cette région est très conservée au sein du domaine CTD de la RNase E[93].

De précédente étude ont montré que l'interaction de RhlB avec la RNase E stimulait son activité par 15[75]. D'autres études sont venues compléter celle-ci en montrant que cette interaction stimulait bien le dépliement de l'ARN mais qu'également les domaines d'interactions à l'ARN portés par le domaine CTD de la RNase E contribuaient à ce dépliement des structures secondaires de l'ARN. Pour cela, deux constructions du domaine CTD de la RNase E ont été utilisées, celle précédemment vue 628 à 843 et une autre allant de 694 à 790, ne contenant donc pas les domaines d'interaction à l'ARN (RBD et AR2) mais seulement le site d'interaction à RhlB. Un duplex d'ARN a été utilisé pour cette étude.

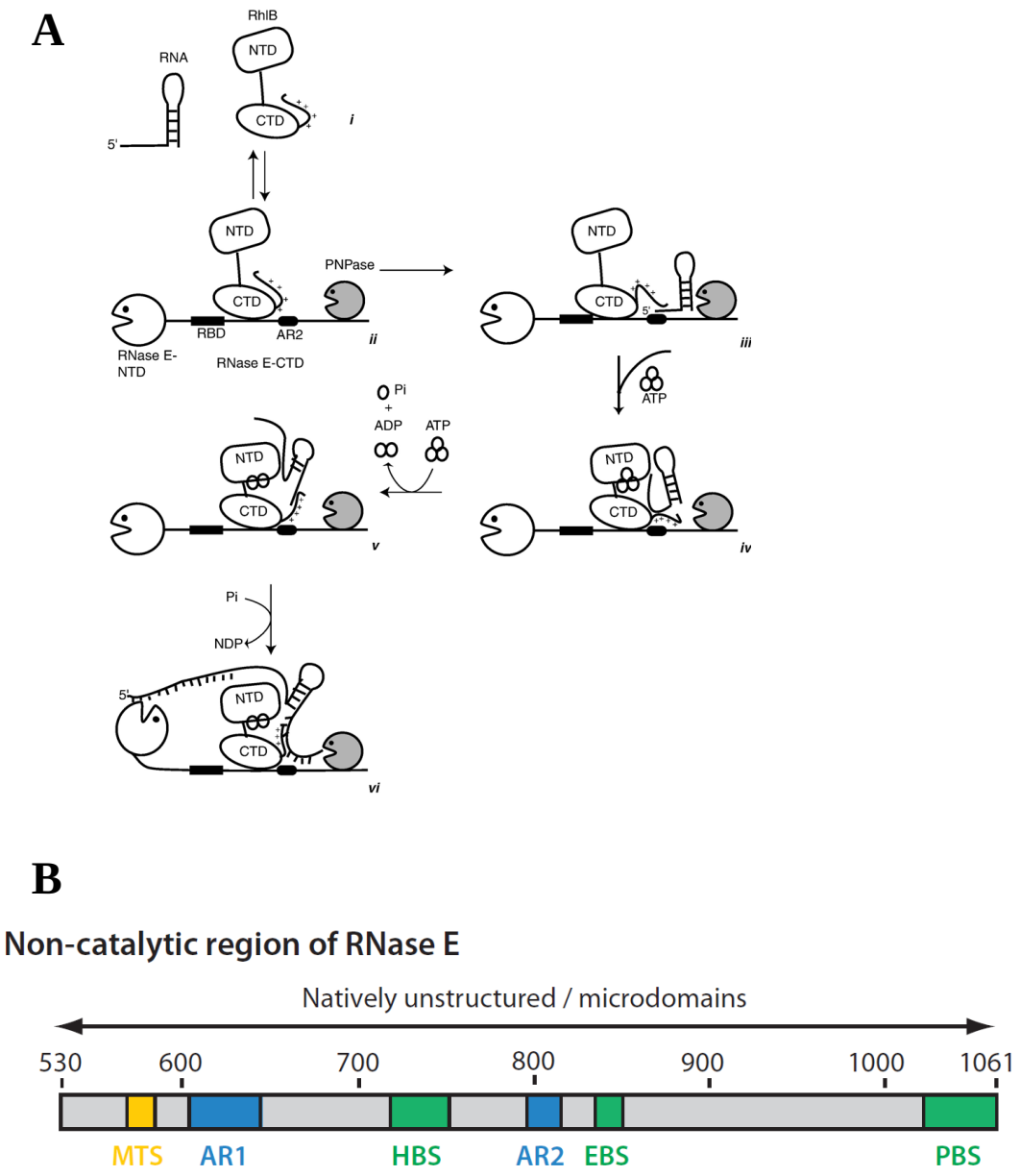


Figure 36: A. Modèle de dégradation des ARNm portants des structures secondaires. (i) L'hélicase RhlB interagit faiblement avec l'ARN donc elle va plutôt avoir tendance à s'associer avec la RNase E (ii). L'ARN va ensuite se lier à la RNase E (interaction avec le domaine RBD ou AR2) (iii). Plusieurs cycles d'hydrolyse de l'ATP permettent aux domaines de RhlB d'interagir avec l'ARN et de stabiliser son interaction pendant le dépliement de l'ARN (iv et v). Enfin, l'ARN linéarisé va être dégradé par l'activité 3'-exonucléase de la PNPase ou par l'activité endonucléase de la RNase E (vi) (d'après Chandran *et al.*, 2007). **B.** Schéma de l'organisation du domaine intrinsèquement désordonné de la RNase E (CTD) avec, le domaine d'interaction à la membrane (MTS, 565-585), les deux domaines d'interaction à l'ARN (RBD ou AR1, 604-644 et AR2, 796-814) et les domaines d'interaction avec l'hélicase, l'énolase et la PNPase (HBS, 719-731; EBS, 834-850; PBS, 1021-1061) (d'après Bouvier et Carpousis, 2011) .

Il a été montré que lorsque l'hélicase est en interaction avec le domaine 628-843, et en présence d'ATP, que 66% de l'ARN est déplié. Contre, 4 et 8% respectivement lorsque le domaine 628-843 est sans hélicase et lorsque l'hélicase est seule. Ce qui montre que l'activité de l'hélicase est stimulée par l'interaction avec le domaine CTD (figure 35B). Ce qui avait déjà été prouvé[75].

Ensuite, lorsque l'hélicase est en interaction avec le domaine 694-790, et en présence d'ATP, il n'y a que 14% de l'ARN qui est déplié (figure 35B). Cette étude montre bien que les domaines d'interactions de l'ARN du domaine CTD de la RNase E sont importants pour permettre le depliement de l'ARN par l'hélicase RhlB. Ces expériences suggèrent qu'il y a une coopération des domaines RBD et AR2 de la RNase E et de l'hélicase pendant le processus de dépliement des structures secondaires portés par l'ARN.

Pour finir, les auteurs ont également effectué des tests d'activités sur un dégradosome reconstitué *in vitro*, en mélangeant la PNPase, l'hélicase RhlB, le domaine CTD (498-1061) ou ce même domaine déleté des zones d'interactions avec l'ARN (ΔRBD, ΔAR2). la deletion des domaines d'interactions avec l'ARN n'affecte pas la formation du dégradosome (figure 35C)[93]. En comparant les deux mini-dégradosome, les auteurs ont constaté quand absence des zones d'interactions avec l'ARN, la PNPase n'est plus capable de dégrader les fragments d'ARN. Ces expériences montre que la délétion des domaines RBD et AR2 empêche le depliement des structures secondaires par RhlB et donc de faciliter la dégradation de ces fragments d'ARN par la PNPase. Un modèle du mécanisme de dégradation des ARNm par le dégradosome a donc été déterminé dans lequel RhlB déplie les structures secondaires de l'ARN afin de faciliter la dégradation de l'ARN(figure 36A).

RhlB est donc une hélicase présente au sein du dégradosome, son site de fixation est localisé entre les deux domaines d'interactions à l'ARN (RBD et AR2). Cette enzyme est cruciale pour permettre la digestion des fragments d'ARN, portants des structures secondaires, par la PNPase. A ce jour, à l'inverse de la PNPase et de l'énolase, aucune donnée structurale n'a été obtenue sur RhlB seule ou en interaction avec son site de fixation sur le domaine CTD de la RNase E, mais son implication dans la dégradation des ARN au sein du dégradosome chez *E. coli* a été largement prouvé. Nous allons voir dans la partie suivante qu'il existe égalemnt un homologue de RhlB chez *B. subtilis*, CshA, une hélicase d'ARN appartenant à la même famille qu' RhlB.

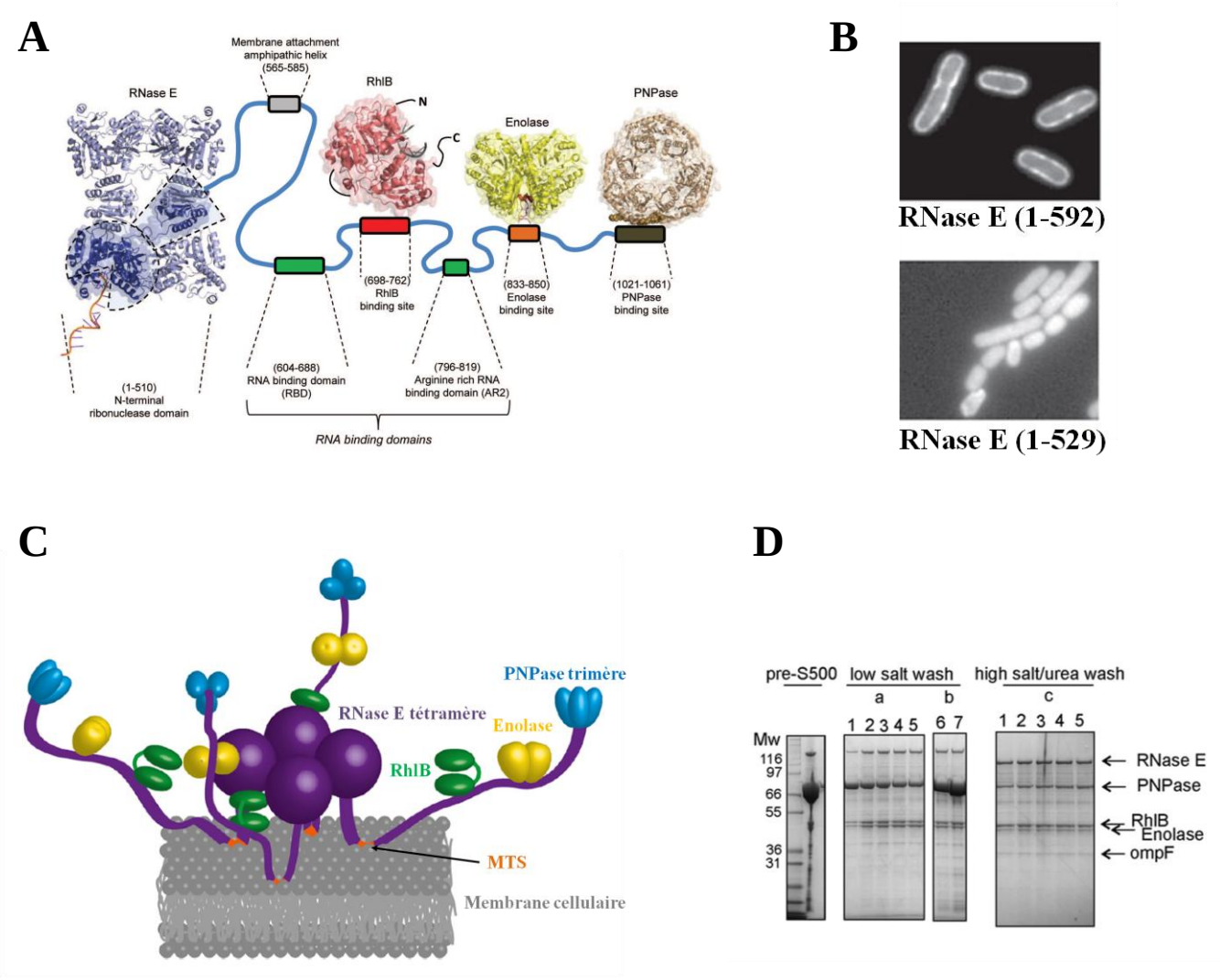


Figure 37: A. Représentation de la RNase E en interaction avec ses partenaires (PNPase, hélicase et énolase) (d'après Hardwick et Luisi, 2013). **B.** Microscopie à fluorescence montrant la localisation à la membrane de la RNase E via son domaine MTS (1-592) et en absence de ce domaine (1-529) la RNase E est localisée dans le cytoplasme cellulaire (d'après Khemici at al., 2008). **C.** Modèle du dégradosome dans lequel la RNase E est localisée à la membrane via son domaine d'interaction avec les phospholipides. Dans ce modèle, la RNase E est représentée avec chaque protomère interagissant avec l'ensemble de ses partenaires (PNPase, hélicase et énolase) (d'après Bandrya *et al.*, 2013). **D.** Gel SDS-PAGE montrant la copurification des quatres protéines à faible et haute concentration en sel, les fractions 1 à 7 correspondent à l'élution de la colonne de chromatographie d'exclusion de taille (Sephacryl S500). La RNase E et l'hélicase RhlB sont clonées dans un même plasmide, pRSFDuet-1, sans aucune étiquette. La PNPase et l'énolase sont clonées dans un même plasmide, pETDuet-1, avec la PNPase portant l'unique étiquette hexahistidine en C-terminal (d'après Worrall et al., 2008).

v. Hypothèse sur l'organisation *in vivo* du degradosome d'*E. coli*

A l'heure actuelle, il n'exsite aucune structure cristallographique de la RNase E entière ou de son domaine CTD en interaction avec l'ensemble des partenaires en interactions (PNPase, énolase et RhlB). L'ensemble des données structurales actuelles sont représentées dans la figure 37A. Cependant, il a été montré que l'énolase, liée au domaine CTD (498-1061), empêche la liaison de l'hélicase RhlB[83]. Pour autant, les partenaires sont tous copurifiés avec la RNase E. Cela indique que chaque protomère de RNase E ne va pas lier l'ensemble des partenaires. Il est possible, en considérant le tétramère de RNase E, qu'un protomère va lier un dimère d'énolase et qu'un autre protomère va lier la PNPase et l'hélicase.

De même, pour la PNPase, nous savons qu'un trimère de PNPase va lier trois peptides (1021-1061) de la RNase E[83], mais dans le contexte du dégradosome aucune étude n'a encore été faite pour savoir si ces trois sites d'interaction au sein de la PNPase sont occupés par trois protomères d'un seul tétramère de RNase E ou si chaque site de liaison va lier trois tétramères différents de RNase E. Nous n'avons donc pour le moment aucune information quant à l'organisation structurale globale du dégradosome *in vivo.* Cependant, il a été montré en microscopie à fluorescence, en utilisant des formes tronquées de la RNase E (1-529 et 1-592), que cette dernière est localisée à la membrane via un domaine d'interaction avec les phospholipides présent lui aussi au sein du domaine CTD, entre les acides aminés 565 et 582 (figure 37B)[94]. Des expériences de microscopie à fluorescence ont également montré que le dégradosome et en particulier la RNase E et RhlB joueraient un rôle dans l'organisation du cytosquelette de la bactérie[95,96]. Pour le moment, seul des modèles du dégradosome sont proposés, toujours en interaction avec la membrane cellulaire (figure 37A)[97]. Maintenant, nous manquons de preuves scientifiques pour valider ce modèle. Ce dégradosome n'a pour le moment été que reconstitué *in vitro* en co-exprimant la RNase E et ses partenaires à l'aide des plasmides pRSFDuet-1 et pETDuet-1 permettant chacun de co-exprimer deux protéines (figure 37D)[98].

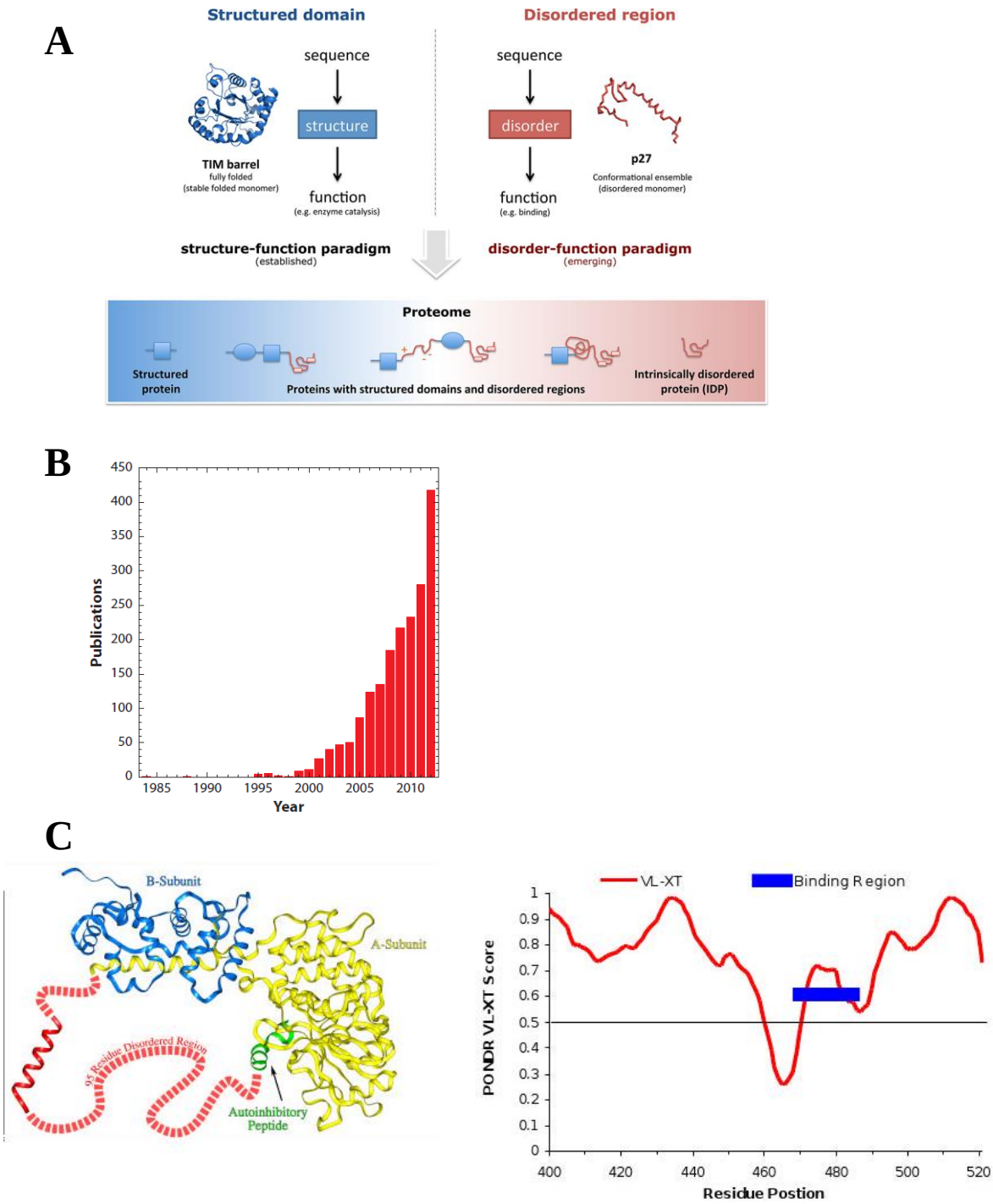


Figure 38 : A. Relation structure fonction pour les protéines structurées et ID (d'après Van der lee *et al.*, 2014) **B.** Augmentation du nombre de publications traitant de protéines ID depuis 1985 (d'après Oldfield et Dunker., 2014). **C.** A gauche, structure cristallographique de la calcineurine dans laquelle une région de 95 acides aminés (boucle rouge) impliqués dans la liaison de la calmoduline n'a pu être résolue car la densité électronique étant absente (d'après Dunker *et al.*, 2001). A droite, prédiction du désordre par PONDR de la région des 95 acides aminés de la calcineurine. Une zone semble plus structurée que le reste (rectangle bleu), il s'agit du domaine d'interaction avec la calmoduline (hélice α rouge à gauche) (d'après Uversky et Dunker, 2010)

III. La RNase Y est l'homologue fonctionnel de la RNase E, mais qu'en est-il de son organisation structurale ?

III. A. Les protéines intrinsèquements désordonnées (IDP) : Des protéines pour les interactions moléculaires:

Cela fait plus d'une vingtaine d'années que le dogme central de la biologie « une séquence – une structure – une fonction » a été remis en cause (figure 38A)[99,100]. En effet, c'est en 1995 qu'une étude pionnière dans le domaine des protéines ID a été publiée : Kissinger et son groupe ont résolu la structure de la calcineurine humaine (CaN) dont une partie importante de la protéine n'était pas définie dans la densité électronique[101]. Depuis 1995, de nombreuses protéines intrinsèquement désordonnées ont été étudiées et caractérisées (figure 38B). Elles constituent, en fait, une partie importante du génome des organismes vivants.

i. Qu'est-ce qu'une protéine intrinsèquement désordonnée ?

Définition et caractéristiques : les différents niveaux de désordre

En 1995, les protéines ayant une absence de structuration secondaire étaient appelées des protéines non-structurées ou déstructurées[102]. Puis, lorsque la communauté scientifique s'est aperçue que, malgré ce désordre apparent, ces protéines étaient capables d'assurer certaines fonctions particulières et de subir des transitions désordre/ordre, on a commencé à parler de protéines intrinsèquements désordonnées[103]. Le terme "intrinsèquement" est lié au fait que ces protéines ont la capacité de se replier ou de former des structures secondaires lorsqu'elles interagissent avec leurs cibles biologiques, qui peut être une petite molécule (ligand) ou un partenaire macromoléculaire (protéine, ADN, ARN) comme la calcineurine qui interagit avec la calmoduline[101]. Cette structuration peut également avoir lieu via des modifications post-traductionnelles (ex : phosphorylation/déphosphorylation via les kinases et phosphatases). Il est maintenant possible d'affiner la notion de désordre et de définir différents niveaux de désordres. Premièrement, les protéines peuvent être complètement déstructurées, on parle alors de protéines intrinsèquement désordonnée (IDP) ou contenir seulement un domaine déstructuré, on parle de domaine ou région intrinsèquement désordonnée (IDR).

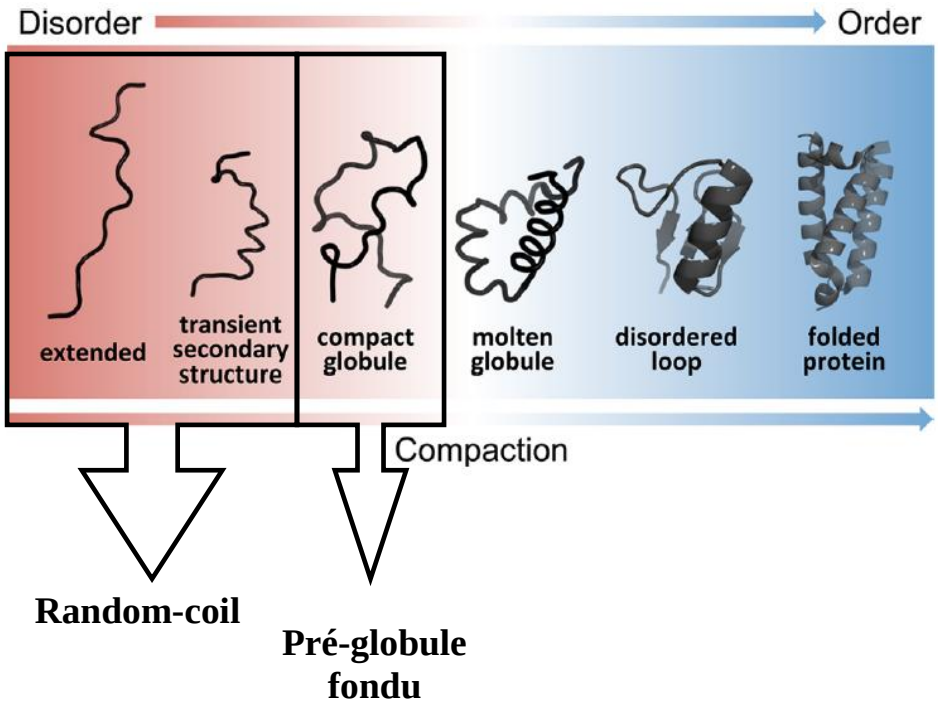


Figure 39: Les différents niveaux de désordre des protéines (d'après Van der lee et al., 2014)

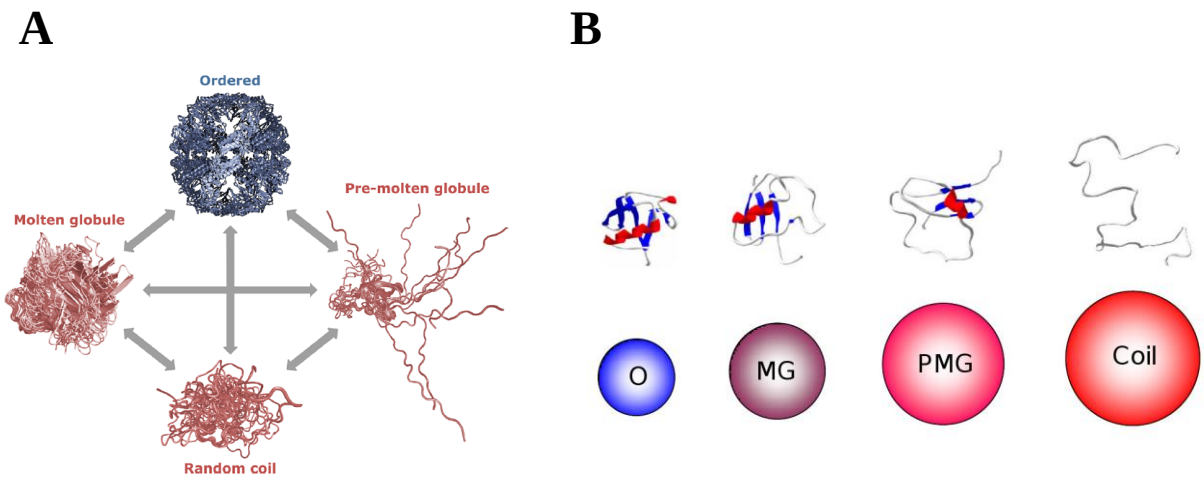


Figure 40: A. Echange entre les trois niveaux principaux de désordre des protéines (état random-coil, pré-globule fondu et globule fondu) et l'état structuré (d'après Van der lee et al., 2014). **B.** Visualtion des différents sous-états structuraux d'une protéine de 100 acides aminés (O = Structurée; MG = globul-fonud; PMG = pré-globul fonu; Coil = radom-coil) avec en dessous le volume hydrodynamique occupé par ces 100 résidus dans ces quatre conformations. (d'après Uversky et Dunker, 2010)

Ensuite, ce désordre peut se classer en différents sous-types conformationnels. Il a été proposé que les protéines, adoptent un continuum conformationnel, c'est-à-dire que les protéines passent par divers sous-types de désordre conformationnel entre l'état structuré et l'état déstructuré (Figure 39)[100]. L'état" random-coil" correspond à l'état dénaturé d'une protéine globulaire. Il ne contient pas ou très peu de structure secondaire[100]. Ensuite, dans l'état "pré-globule fondu" (Pre-molten globule), des structures secondaires apparaissent. Cet état possède une forte tendance à se structurer s'il y a une interaction avec un partenaire. Dans l'état suivant, "globule fondu" (Molten globule), il y a des structures secondaires que l'on retrouve dans l'état natif (structuré) de la protéine mais la structure tertiaire est extrêmement dynamique par rapport à l'état natif. Cet état est un état transitoire lors du repliement de la protéine. Ces différentes transitions (random-coil, pré-globule fondu, globule fondu) s'accompagnent d'une augmentation de la compaction de la chaîne polypeptidique. Ceci se reflète sur les paramètres biophysiques comme le rayon de gyration (Rg) ou le rayon hydrodynamique (Rh) qui vont diminuer au fur et à mesure que la compaction augmente[100] (Figures 39 et 40A et B)[104]. Ainsi des méthodes comme le SAXS (diffusion des rayons X aux petits angles)[105,106], la RMN (Résonance magnétique nucléaire)[107,108] ou encore la chromatographie par exclusion stérique peuvent permettre de mettre en évidence la présence d'états partiellement ou non structurés des protéines. Grâce à ces techniques biophysiques, la communauté scientifique a compris que certaines protéines pouvaient assurer leurs fonctions sans avoir une structure native pleinement structurée.

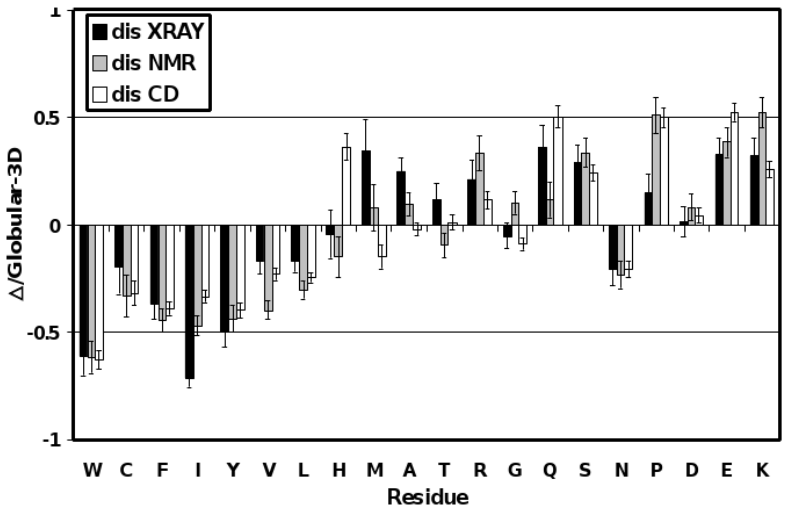
Tableau 8 : A. Proportion des acides aminés présents dans une protéine ID ou structurée (d'après Tompa, 2002) **B.** Particularités de la composition en acide aminés des protéines ID. Comparaison de la composotion en acides aminés de protéines structurées avec trois bases de données de protéines déstructurées (dis XRAY, dis NMR, dis CD). Un acide aminés est dit promoteur d'ordre ou promoteur de désordre par le calcul suivant : (% d'a.a désordonnés dans la base de donnée - % d'a.a structurés dans la base de donnée) / (% d'a.a structurés dans la base de donnée) = Δ/globuarl_3D (d'après Uversky et Dunker, 2010).

A

Amino acid	Intrinsically unstructured proteins	Swiss-Prot[a]	Globular[b]
Ala (A)	7.15	7.62	8.15
Arg (R)	4.21	5.19	4.61
Asn (N)	2.06*	4.36	4.66*
Asp (D)	5.05	5.25	5.78
Cys (C)	0.61*	1.62	1.64*
Gln (Q)	4.46	3.93	3.69
Glu (E)	14.26*	6.47	5.98*
Gly (G)	4.31	6.85	7.99
His (H)	1.51	2.25	2.33
Ile (I)	3.67	5.85	5.43
Leu (L)	5.44	9.54	8.37
Lys (K)	10.43	5.97	6.05
Met (M)	1.30	2.37	2.03
Phe (F)	1.66*	4.10	3.95*
Pro (P)	12.07*	4.89	4.61*
Ser (S)	6.91	7.08	6.31
Thr (T)	5.14	5.57	6.15
Trp (W)	0.32*	1.21	1.55*
Tyr (Y)	1.42*	3.16	3.64*
Val (V)	8.02	6.61	7.00

Amino acid frequencies in % are given for globular proteins, Swiss-Prot proteins and for the intrinsically unstructured proteins and regions listed in Table 1. Asterisks identify amino acids that are at least two times more or less frequent than in an average globular protein in the Protein Data Bank.
[a]Data from http://us.expasy.org/sprot.
[b]Data 'globular-3D' from http://disorder.chem.wsu.edu.

B



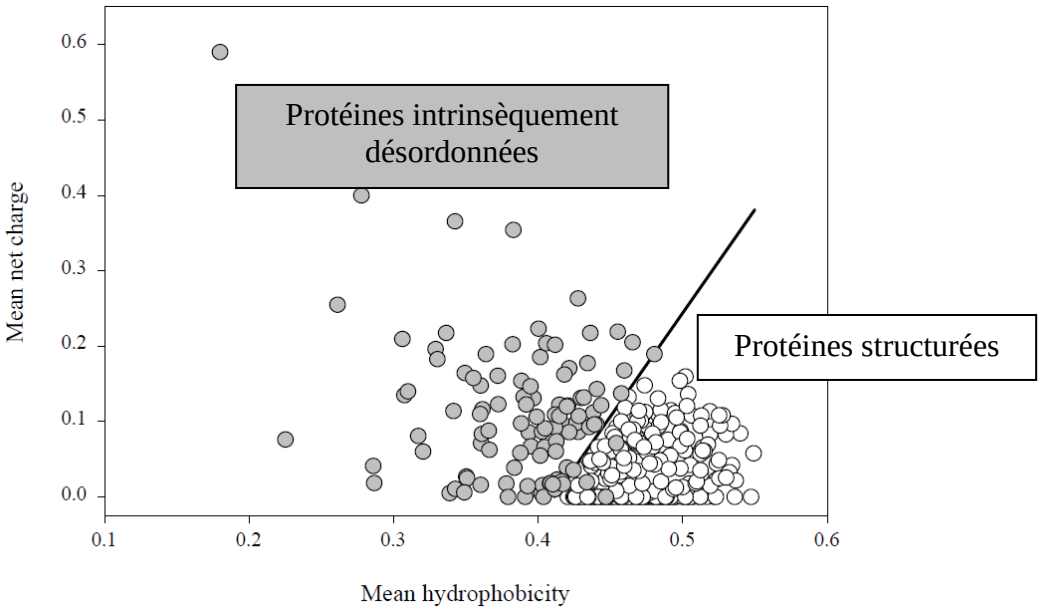


Figure 41: Répartition des protéines structurées et des protéines ID par un graphe montrant la charge nette des protéines en fonction de leur hydrophobicité (d'après Uversky, 2002)

Identification et prédiction du désordre des protéines

L'identification de ces protéines ID au sein du protéome des organismes vivants a nécessité de comparer les séquences de plusieurs protéines ID répertoriées sur la base de données DisProt (http://www.DisProt.org) avec celles de protéines structurées. Il est ressorti de ces études que certains acides aminés sont promoteurs d'ordre alors que d'autres sont promoteurs de désordre et que plus une protéine contient d'acides aminés promoteurs de désordre, plus elle a une forte probabilité d'être intrinsèquement désordonnée (tableau 8). Les acides aminés promoteurs de désordre, classés dans l'ordre croissant de désordre sont : K, Q, S, E et P. Les acides aminés promoteurs d'ordre, classés par ordre croissant par rapport à l'ordre sont : N, H, M, L, F, Y, I, W et C (Tableau 8A et B). Les acides aminés A, G, T, R et D sont présents en même quantité dans les protéines ID ou structurées[109–112].

Les trois acides aminés E, P et K sont deux fois plus présents chez les protéines ID par rapport aux protéines structurées. Les acides aminés Tyr (Y), la Cys (C) et Asn (N) sont deux fois moins présents chez les protéines ID par rapport aux protéines structurées et le Trp (W) est quatre fois moins présent[109,110,113].

La charge nette d'une protéine est le paramètre le plus discriminant pour déterminer si une protéine est intrinsèquement désordonnée ou non. Les protéines ID sont riches en résidus polaires chargés et contiennent peu de résidus hydrophobes. Elles ont donc une charge nette plus forte que les protéines structurées. Plus la charge nette d'une protéine est forte, plus les forces de répulsions électrostatiques sont importantes et donc plus la chaîne polypeptidique aura tendance à se déplier et à être désordonnée. Il est donc possible de séparer les protéines ID et structurées sur un graphe représentant la charge nette moyenne d'une protéine en fonction de son hydrophobicité moyenne : **GRAVY** pour **GR**and **AV**erage of h**Y**dropathy[114] (figure 41)

Tableau 9: Liste non-exhaustive des différents serveurs de prédiction du désordre disponibles sur le web (d'après Meng *et al.*, 2017)

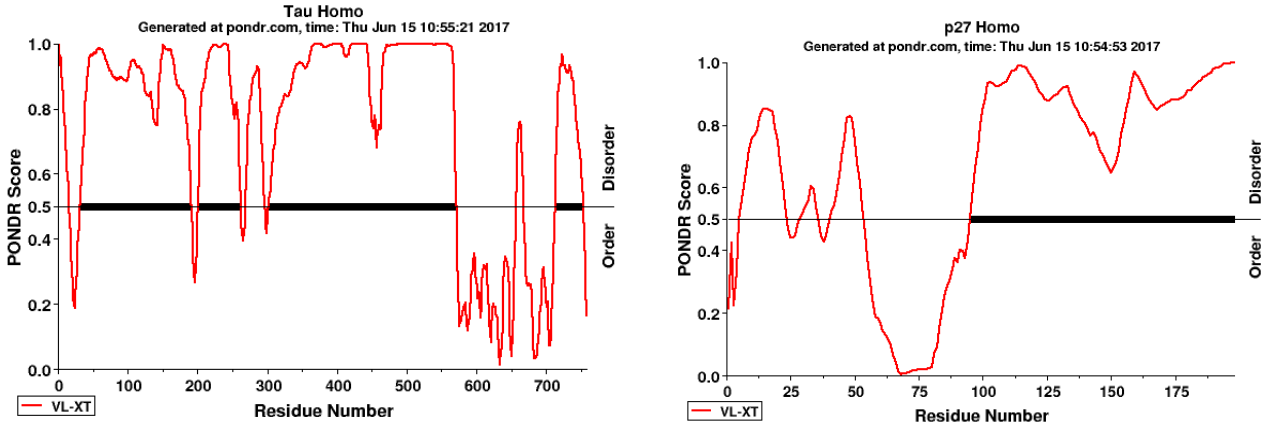
Name	Year last published	Availability[a]	Batch submission (max # proteins)	High-throughput	URL
DISOclust	2015	WS + SP	No		http://www.reading.ac.uk/bioinf/IntFOLD/IntFOLD2_form.html
DISOPRED	2015	WS + SP	No		http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1
DeepCNF-D	2015	SP	No	Yes	http://ttic.uchicago.edu/~wangsheng/DeepCNF_D_package_v1.00.tar.gz
DisMeta	2014	WS	No		http://www-nmr.cabm.rutgers.edu/bioinformatics/disorder/
disCoP	2014	WS	Yes (up to 5)		http://biomine.cs.vcu.edu/servers/disCoP/
DNDisorder	2013	WS	No		http://iris.rnet.missouri.edu/dndisorder/
MFDp2	2013	WS	Yes (up to 100)		http://biomine.cs.vcu.edu/servers/MFDp2/
ESpritz	2012	WS + SP	Yes (no limit)	Yes	http://protein.bio.unipd.it/espritz/
MetaDisorder	2012	WS	No		http://iimcb.genesilico.pl/metadisorder/
SPINE-D	2012	WS + SP	No		http://sparks-lab.org/SPINE-D/
CSpritz	2011	WS	Yes (no limit)		http://protein.bio.unipd.it/cspritz/
IsUnstruct	2011	WS	No	Yes	http://bioinfo.protres.ru/IsUnstruct/
MFDp	2010	WS	Yes (up to 5)		http://biomine.cs.vcu.edu/servers/MFDp2/
PONDR-FIT	2010	WS	No	Yes	http://disorder.compbio.iupui.edu/metapredictor.php
MD	2009	WS + SP	No		https://ppopen.rostlab.org/
PreDisorder	2009	WS + SP	No		http://sysbio.rnet.missouri.edu/predisorder.html
metaPrDOS	2008	WS	No		http://prdos.hgc.jp/cgi-bin/meta/top.cgi
OnD-CRF	2008	WS	No		http://babel.ucmp.umu.se/ond-crf/
Norsnet	2007	WS + SP	No		https://ppopen.rostlab.org/
Ucon	2007	WS + SP	No		https://ppopen.rostlab.org/
PrDOS	2007	WS	No		http://prdos.hgc.jp/cgi-bin/top.cgi
PROFbval	2006	WS + SP	No		https://ppopen.rostlab.org/
PONDR-VSL2B	2006	WS + SP	No	Yes	http://www.dabi.temple.edu/disprot/predictor.php
FoldUnfold	2006	WS	No	Yes	http://bioinfo.protres.ru/ogu/
DISpro	2005	WS + SP	No		http://scratch.proteomics.ics.uci.edu/
FoldIndex	2005	WS	No	Yes	http://bioportal.weizmann.ac.il/fldbin/findex
IUPred	2005	WS + SP	No	Yes	http://iupred.enzim.hu/
RONN	2005	WS + SP	No	Yes	https://www.strubi.ox.ac.uk/RONN
PONDR-VL3	2005	WS	No	Yes	http://www.dabi.temple.edu/disprot/predictor.php
DisEMBL	2003	WS + SP	No	Yes	http://dis.embl.de/
GlobPlot	2003	WS	No	Yes	http://globplot.embl.de/
NORSp	2003	WS + SP	No		https://ppopen.rostlab.org/

[a] Availability: SP (standalone package); WS (webserver)

Basés sur les caractéristiques de séquences des protéines ID et structurées, de nombreux algorithmes ont été créés afin de prévoir, à partir d'une séquence, le niveau de désordre d'une protéine donnée[115]. Ceux-ci sont disponibles gratuitement sur le web (Tableau 9), le plus utilisé étant appelé PONDR[116]. La précision de ces algorithmes varie selon l'algorithme utilisé: par exemple pour PONDR, la précision est supérieure à 86 %. Il existe aussi des logiciels qui combinent plusieurs algorithmes de prédiction, le plus utilisé étant metaPrDOS[115].

La figure 42A présente la prédiction de désodre par PONDR pour deux protéines ID, la protéine Tau et la protéine p27. La protéine Tau est une protéine qui va stabiliser les microtubules en interagissant avec. La prédiction de désordre montre bien la présence de zones déstructurées dans la majeure partie de sa séquence. Une étude en RMN a montré que cette protéine est bien déstructurée et quelle possède une grande flexibilité due à divers sous états conformationnels adoptés par la protéine (figure 42B)[117]. La prédiction du désordre pour p27 montre également que cette protéine possède une zone complètement déstructurée. La protéine p27 est une protéine qui va inhiber l'activité kinase des complexes cyclines-CDK du cycle cellulaire. La structure de cette protéine a été obtenue en interaction avec le complexe CyclineA-CDK2 (figure 42B)[118]. Les prédicteurs de désordre sont donc très précis et permettent de mettre en évidence la présence d'une zone déstructurée au sein d'une protéine. Cela se voit également pour la CaN qui possède un domaine d'interaction avec la CaM dans sa région déstructurée. En effet, au sein de cette région de 95 a.a une courte zone d'interaction se structure en hélice α, et la prédiction de désordre sur cette région de 95 acides aminés déstructurés montre une zone plus structurée que le reste (figure 38C, à droite) correspondant à la zone indiquée par le rectangle bleu. On peut donc déterminer via les prédicteurs de désordre, les zones d'interactions potentielles avec les partenaires au sein d'un domaine prédit pour être déstructuré.

A



B

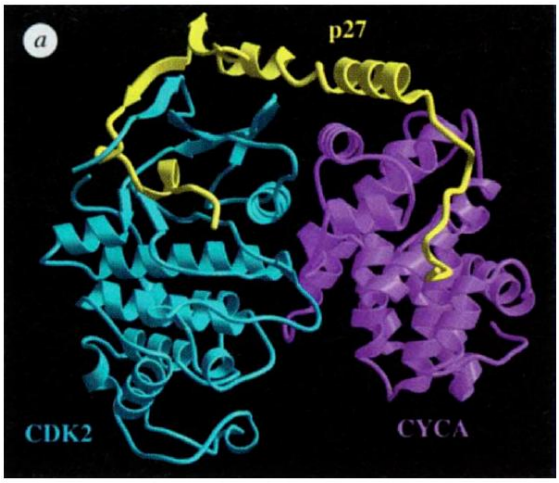


Figure 42 : A. Prédiction du désordre par le prédicteur PONDR pour la protéine Tau (gauche) et p27 (droite). **B.** Structure de la protéine Tau obtenu par RMN (gauche) et structure cristallographique de la protéine p27 obtenu en interaction avec le complexe cycline A/CDK2 (d'après Russo et al., 1996 et Wintjens et al., 2001)

Rôle du désordre au sein des protéines

En s'associant avec leurs partenaires, qui peuvent être multiples et de conformation différentes[119], les IDP vont se structurer. Cette association est en générale de faible affinité mais de forte spécificité, grâce à la très grande flexibilité des IDP, permet des interactions transitoires[120,121]. De nombreuses études bio-informatiques fonctionnelles ont montré que les IDP/IDR étaient très largement impliquées dans des interactions protéine – protéine et protéine – acide-nucléique. Très peu d'IDP/IDR possèdent une activité catalytique, probablement à cause de la faible affinité d'une IDP pour son partenaire (petites molécules ou macromolécules), alors que la catalyse enzymatique nécessite en général une forte affinité pour le substrat. La calcineurine (CaN) est une phosphatase calcium- et calmoduline-dépendante, critique pour plusieurs processus cellulaires dont l'activation des lymphocytes T. Elle possède une région non structurée où sont présents deux sites de liaison: un pour la calmoduline (CaM) et un autre impliqué dans l'auto-inhibition de la protéine. La liaison de la CaM à la CaN structure la partie désordonnée du domaine auto-inhibiteur en hélice α, permettant l'activation de la protéine [122,123] (Figure 38C). Il a été montré, par exemple, que de nombreuses IDP font partie de complexes multi-protéiques comme le spliceosome permettant la maturation des ARNm , les complexe Cycline/CDK de régulation du cycle cellulaire[119], les complexes de certaines protéines ribosomiques permettent une régulation de la traduction. Les IDP comprennent de nombreuses protéines de virus, comme, par exemple, la protéine Tat (Trans Activator of Transcription) qui permet de réguler la transcription du virus HIV[124], la calcineurine, décrite plus haut, interagit avec la calmoduline dans un processus de régulation des défenses immunitaires. De plus, de nombreuses IDP se structurent après phosphorylation par une kinase, ce qui leur permet d'assurer leurs fonctions. En effet, les kinases sont des enzymes connues pour être impliquées dans des mécanismes de régulation de la signalisation intracellulaire chez les eucaryotes. De manière générale, après interaction avec leurs partenaires, les IDP vont se structurer, en hélice alpha. La structure des hélices α est beaucoup moins stable que les feuillets β, et les acides aminés polaires sont généralement en surface, permettant d'établir des liaisons électrostatiques avec un partenaire potentiel. Ces deux caractéristiques permettent des interactions transitoires entre l'IDP et son partenaire avec un dissociation rapide du complexe.

Tableau 10: (Gauche) Processus biologiques dans lesquels les protéines ID sont impliquées (gauche) et fonctions portées par les protéines ID (droite) (d'après Xie et al., 2007)

keywords	number of proteins	number of families	average sequence length	*Z*-score	*P*-value
Differentiation	1406	422	439.25	18.81	1
Transcription	11223	1653	442.64	14.62	1
Transcription regulation	9758	1554	413.31	14.33	1
Spermatogenesis	332	189	280.49	13.9	1
DNA condensation	317	130	300.06	13.34	1
Cell cycle	4278	612	494.17	12.17	1
mRNA processing	1575	249	515.55	10.92	1
mRNA splicing	716	180	459.06	10.13	1
Mitosis	718	215	620.43	9.42	1
Apoptosis	810	211	465.48	9.35	1
Protein transport	3081	579	421.73	8.77	1
Meiosis	284	170	639.16	8.7	1
Cell division	3466	385	451.63	8.51	1
Ubl conjugation pathway	1254	244	525.99	8.13	1
Wnt signaling pathway	417	41	476.84	6.58	1
Neurogenesis	322	74	667.4	6.56	1
Chromosome partition	556	67	495.39	6.39	1
Ribosome biogenesis	319	71	391.79	5.9	1
Chondrogenesis	64	6	332.85	5.58	1
Growth regulation	155	45	354.9	5.14	1

keywords	number of proteins	number of families	average sequence length	*Z*-score	*P*-value
Ribonucleoprotein	12236	412	150.55	22.13	1
Ribosomal protein	11692	330	140.58	20.63	1
Developmental protein	3260	721	477.93	19.28	1
Hormone	1187	161	141.13	15.58	1
Growth factor	785	84	255.7	11.16	1
Cytokine	899	110	213.28	10.21	1
Neuropeptide	268	209	95.08	9.65	1
Activator	3086	573	428.47	9.04	1
GAP protein	47	2	232.96	7.42	1
Antigen	1113	455	437.48	6.99	1
Repressor	2309	449	374.46	6.92	1
Chromatin regulator	334	100	801.24	6.7	1
Pyrogen	37	2	262.59	6.44	1
Vasoactive	125	39	160.39	5.56	1
Amphibian defense peptide	123	148	50.64	5.44	1
GTPase activation	311	70	831.03	5.36	1
Endorphin	42	4	226.68	5.35	1
Opioid peptide	24	4	216.96	5.14	1
Protein phosphatase inhibitor	47	8	366.51	5.07	1
Cyclin	182	25	430.58	4.88	1

Tableau 11: Proportion des protéines ID au sein des différents organismes procaryotes et eucaryotes (d'après Dunker *et al.*, 2000)

Table 10. Genomic Disorder

Kingdom	Species	# seqs	Length ≥ 50		CDF*	
Archaea	*Methanococcus jannaschii*	1 714	71	4%	26	2%
Archaea	*Pyrococcus horikoshii*	2 062	164	8%	70	3%
Archaea	*Pyrococcus abyssi*	1 764	157	9%	62	4%
Archaea	*Archaeoglobus fulgidus*	2 402	244	10%	93	4%
Archaea	*Methanobacterium thermoautotrophicum*	1 869	365	20%	140	7%
Archaea	*Aeropyrum pernix K1*	2 694	637	24%	490	18%
Eubacteria	*Ureaplasma urealyticum*	611	14	2%	9	1%
Eubacteria	*Rickettsia prowazekii*	834	23	3%	5	1%
Eubacteria	*Borrelia burgdorferi*	845	26	3%	14	2%
Eubacteria	*Campylobacter jejuni*	2 309	80	3%	21	1%
Eubacteria	*Mycoplasma genitalium*	480	20	4%	10	2%
Eubacteria	*Helicobacter pylori*	1 532	69	5%	24	2%
Eubacteria	*Aquifex aeolicus*	1 522	94	6%	29	2%
Eubacteria	*Haemophilus influenzae*	1 708	126	7%	27	2%
Eubacteria	*Bacillus subtilis*	4 093	323	8%	87	2%
Eubacteria	*Escherichia coli*	4 281	363	8%	107	2%
Eubacteria	*Vibrio cholerae*	3 815	333	9%	93	2%
Eubacteria	*Mycoplasma pneumoniae*	675	60	9%	14	2%
Eubacteria	*Xylella fastidiosa*	2 761	246	9%	103	4%
Eubacteria	*Thermotoga maritima*	1 842	165	9%	53	3%
Eubacteria	*Neisseria meningitidis MC58*	2 105	190	9%	64	3%
Eubacteria	*Chlamydia pneumoniae*	1 052	100	10%	40	4%
Eubacteria	*Synechocystis sp*	3 167	338	11%	104	3%
Eubacteria	*Chlamydia trachomatis*	894	99	11%	42	5%
Eubacteria	*Treponema pallidum*	1 028	115	11%	37	4%
Eubacteria	*Mycobacterium tuberculosis*	3 916	747	19%	293	7%
Eubacteria	*Deinococcus radiodurans chr I*	2 580	534	21%	212	8%
Eucaryote	*Caenorhabditis elegans*	17 049	4 636	27%	1 322	8%
Eucaryote	*Arabodiopsis thaliana chrII, IV*	7 849	2 248	29%	653	8%
Eucaryote	*Saccharomyces cerevisiae*	6 264	1 858	30%	356	6%
Eucaryote	*Drosphila melanogaster*	13 885	5 651	41%	2 403	17%

Procaryotes ~ 6,3 %

Eucaryotes ~32 %

* Sequences in each genome judged to be wholly disordered, using a method based on cumulative distribution functions.[261]

ii. Fonctions et répartition des protéines ID au sein du monde vivant

Fonctions des protéines ID

A partir des outils mentionnés plus haut, des études ont répertorié les protéines ID présentes dans les génomes des organismes vivants. Une étude a été réalisée en 2007 afin de déterminer dans quels processus biologiques et dans quelles fonctions étaient impliquées les protéines ID[125]. 200 000 séquences protéiques de la base de données SwissProt ont été analysées avec le logiciel PONDR_VL3E (The Predictor Of Natural Protein Disorder)[125]. Sur les 710 fonctions répertoriées dans SwissProt, 238 sont associées à des protéines intrinsèquement désordonnées contre 302 pour les protéines structurées. Deux fonctions sont principalement représentées dans les protéines structurées : « Ligand » et « Biological processes ». A l'inverse, les protéines ID sont présentes dans toutes les catégories. Il s'avère donc que les protéines ID ont un répertoire fonctionnel beaucoup plus large que les protéines structurées[125,126]. On remarque que les protéines ID sont fortement impliquées dans des processus biologiques importants pour la cellule comme la différenciation cellulaire, la régulation de la transcription, le cycle cellulaire, la signalisation cellulaire (ex : p53, p27, Tau,…) ainsi que dans la maturation et l'épissage des ARN (Tableau 10, gauche). Ainsi 70% des protéines ID sont impliquées dans la signalisation cellulaire. Les ribonucléoprotéines sont les protéines les plus représentées parmi les protéines ID (Tableau 10, droite). Ainsi, les interactions avec les acides nucléiques nécessitent souvent l'intervention de protéines intrinsèquement désordonnées. Par exemple, la formation du complexe de maturation des ARNm chez les eucaryotes appelé spliceosome nécessite l'intervention de protéines ID.

Répartition des protéines ID au sein du monde vivant

Une protéine est dite ID si elle possède un domaine ID d'au moins 50 résidus. Une analyse de 31 génomes par PONDR a montré que les organismes eucaryotes possèdent un génome contenant 27 à 41 % de protéines ID, et que les organismes procaryotes en possèdent moins de 10%[127](tableau 11). En conclusion, la proportion de protéines ID dans un organisme vivant augmente avec sa complexité car il y a plus de fonctions biologiques complexes à réguler[128].

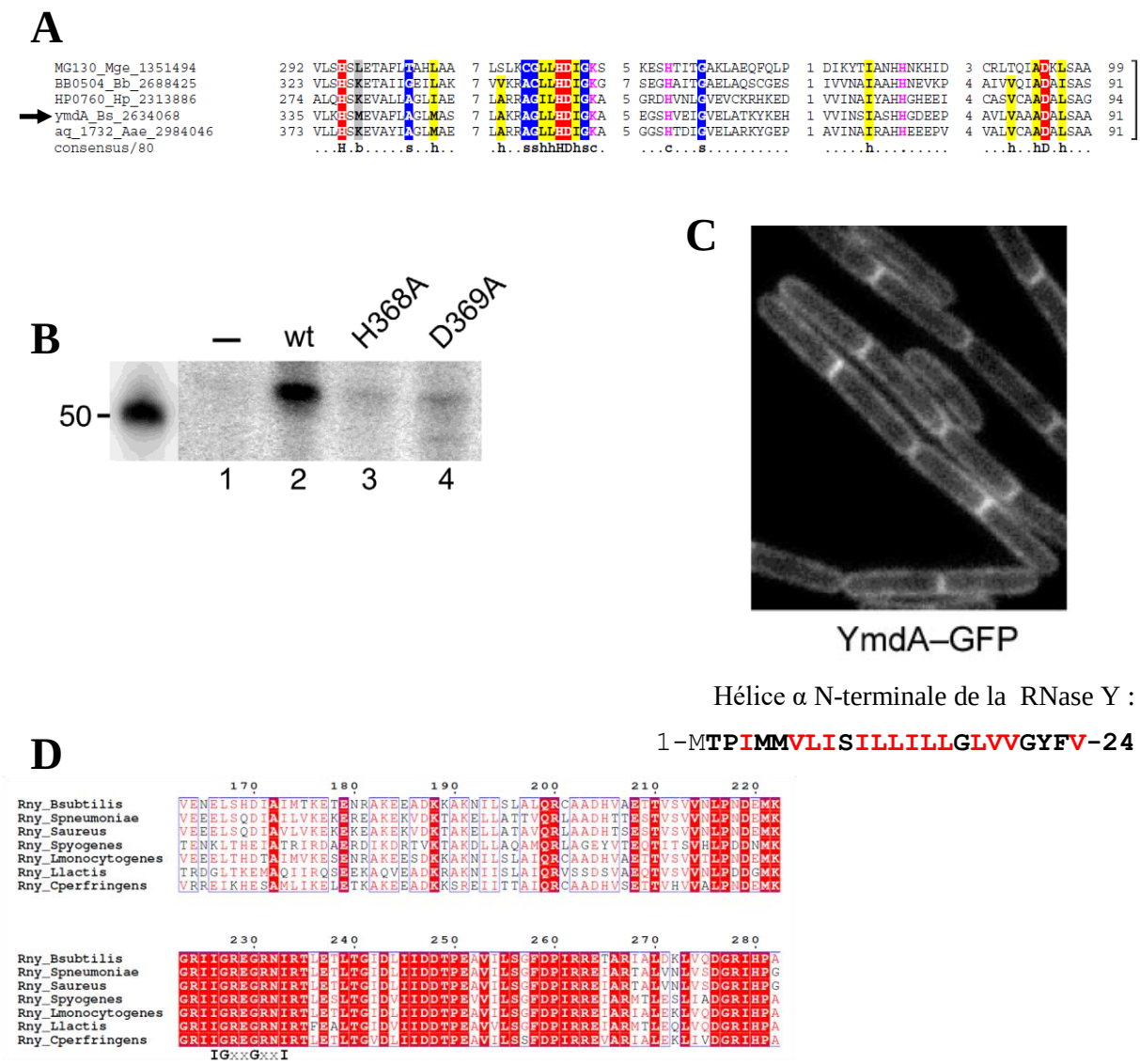


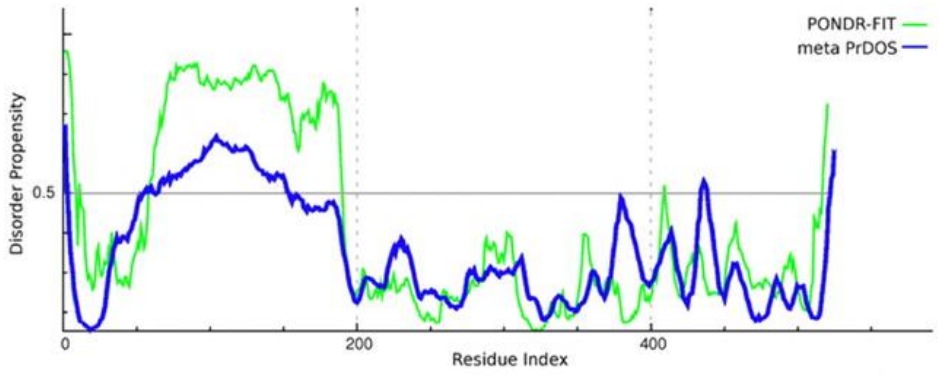
Figure 43: A. Identification par alignement multiple d'un motif conservé (HxHDxD) chez des protéines de fonctions inconnues. Le domaine portant ce motif a été appelé domaine HD et permet l'hydrolyse de l'ADN, l'ARN ou de monomucléotide cyclique (d'après Aravin et Koonin, 1998). **B.** Clivage du transcrit *yitJ* par la RNase Y sauvage et par deux mutants du domaine HD de la RNase Y (H368A et D369A) (d'après Shahbabian *et al.*, 2009). **C.** Image de microscopie à fluorescence montrant la localisation de la RNase Y – GFP dans les cellules de *B. subtilis* (d'après Hunt et al., 2006). **D.** Alignement de plusieurs séquences de RNase Y provenant de bactéries à Gram positifs et montrant la conservation du motif IGxxGxxI signifiant la présence d'un domaine KH (Bsubtilis = *Bacillus subtilis*; Spneumoniae = *Streptococcus poneumoniae*; Saureus = *Staphylococcus aureus*; Spyogenes = *Streptococcus pyogenes*; Lmonocytogenes = *Listeria monocytogenes*; Llactis = *Lactococcus lactis*; Cperfringens = *Clostridium perfringens*).

III.B. La RNase Y possède un domaine intrinsèquement désordonné

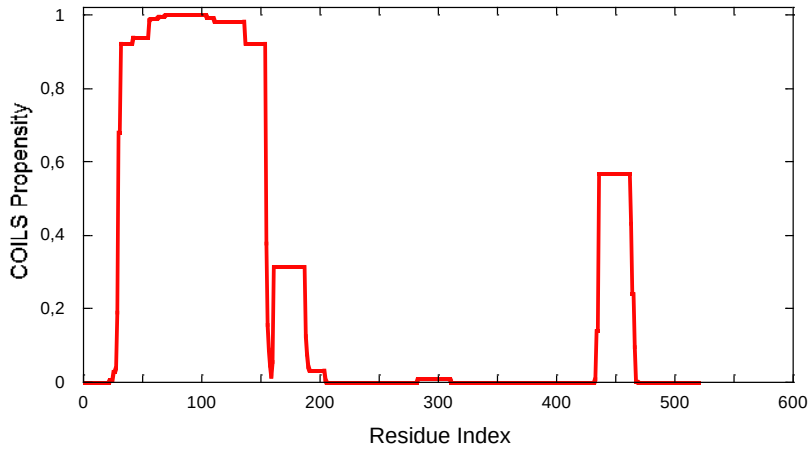
i. La RNase Y possède des caractéristiques fonctionnelles similaires avec la RNase E

Avant 2009, la RNase Y faisait partie des protéines dont la fonction n'avait pas été caractérisée[62,129–132]. Mais le domaine catalytique de cette protéine de fonction inconnue avait été identifié comme appartenant à la famille des phosphohydrolases métal-dépendantes à domaine HD. Un alignement de plusieurs dizaines de protéines appartenant à des organismes différents avait été réalisé. Toutes ces protéines portaient un motif conservé, qui est le suivant **H**x_1**HD**x_2**D** (pour la RNase Y x_1 = 28 acides ami,és; x_2= 54 acides aminés, acides aminés qui séparent les motifs du domaine HD). Ce domaine fut appelé domaine HD et a permis de définir une nouvelle famille de phosphohydrolase métal-dépendante. Une partie de cette alignement est présent dans la figure 43A, la RNase Y, appelée YmdA (taxonomie génomique), était placée dans le groupe 5, un groupe de protéine à domaine HD de fonction inconnue[129]. Ce motif, composé de deux histidines et deux aspartates, permet de coordonné un cation bivalent qui selon les protéines peut être un Mg^{2+}, Mn^{2+} ou encore un Zn^{2+}, et parfois même un Ni^{2+}[133–136]. Cette famille de protéines à domaine HD rassemble des enzymes impliquées dans le métabolisme des acides nucléiques (ADN, ARN ou de mononucléotides)[129], ainsi que dans la transduction du signal[137]. La présence de ce domaine HD au sein de la RNase Y a été confirmée par des études d'activité faites sur deux mutants de la RNase Y. Ces mutations ont été faites au niveau du motif Hx**HD**xD, l'histidine 368 et l'aspartate 369 on été mutés en alanine. Suite à cela des tests d'activités *in vitro* sur le transcrit *yitJ* ont été réalisés, dans le cas des mutants aucun produit de clivage n'est observé, ce qui prouve que ce domaine HD est important pour l'activité catalytique (figure 43B)[27]. En 2006, une étude fonctionnelle sur 11 gènes de fonction inconnue dont la RNase Y a été réalisée chez *B. subtilis*[131]. Les auteurs ont montré, par des expériences de microscopie à fluorescence en utilisant la "Green fluorescent protein" (GFP) que la RNase Y (YmdA) était localisée à la membrane (figure 43C). Dans cette expérience la GFP avait été fusionnée à l'extrémité N-terminale de la RNase Y (figure 43C). En effet, cette protéine possède en C-terminale une séquence de 21 acides aminés majoritairement hydrophobes, riche en résidus valine (4), leucine (6) et isoleucine (4) (figure 43C).

A



B



C

Désordre

Acides aminés (a.a)	RNase Y entière (520 a.a)	Domaine ID (169 a.a)	a.a au sein du domaine ID par rapport à la totalité
E	69 (13 %)	32 (19 %)	32 sur 69 (**46%**)
S	24 (5 %)	6 (4 %)	6 sur 24 (25%)
Q	15 (3 %)	10 (6 %)	10 sur 15 (**67%**)
Y	8 (1,5 %)	0	0 sur 8 (**0%**)
I	48 (9,2 %)	13 (7,2 %)	13 sur 48 (27%)
W	0	0	0

Ordre

Figure 44: Analyse in silico de la séquence de la RNase Y. A. Prédiction de région non-structurée au sein de la séquence de la RNase Y par les prédicteur PONDR et metaPrDOS. Ces deux prédicteurs ont permis de définir un consensus de 100 acides aminés déstructurés. **B.** Prédiction de la présence de structures secondaires de type coiled-coil par le prédicteur COILS. Il y aurait une forte probabilité pour qu'il y ai un tel domaine entre les acides aminés 30 et 150 chez la RNase Y. **C.** Proportion en acides aminés promoteur d'ordre (Y, I et W) et promoteur de désordre (Q, S et E) au sein de la RNase Y entière et du domaine ID. Le domaine ID contient 67 % et 46 % des Gln et Glu de la RNase Y, et il ne contient aucune Tyr.

Cette protéine a également été prédite pour contenir un domaine KH (ribonucléoprotein K homology domain) (figure 43D). Ces domaines sont connus pour interagir avec des acides nucléiques, en particulier avec les ARN ou les ADN simple brin, de plus un motif est présent au sein de ces domaines permettant d'identifier qu'il s'agit d'un domaine KH, ce motif est le suivant **VIG**xx**G**xx**I**[138,139]. Lorsque plusieurs séquence de RNase Y sont alignées, ce motif est présent et très conservé, ce qui est la signature de la présence d'un domaine KH (figure 43D). Ce domaine KH est suivi du domaine HD. Après cela, YmdA a donc été considérée comme une potentielle RNase.

Après la caractérisation de son activité en 2009, une étude publiée en 2011 a été plus loin dans la caractérisation *in silico* de cette protéine[140]. En effet, la séquence de la RNase Y a été soumise à des prédicteur de désordre (PONDR et metaPrDOS) et il c'est avéré que la RNase Y possède une région non-structurée enN-terminale allant du résidus 30 à 192 (figure 44A). De plus, les auteurs ont également soumis cette séquence au prédicteur de domaine coiled-coil. Ces domaines sont connus pour interagir avec des partenaires protéiques ou des acides nucléiques. Les leucine-zipper, présent notamment au niveau des facteur de transcription leurs permettant d'interagir avec l'ADN font partie des domaines coiled-coil[141,142]. Cette prédiction a donc montré la présence d'un potentiel domaine coiled-coil dans la partie N-terminale de la RNase Y, entre les acides aminés 30 et 150 (figure 44B). La RNase Y possède donc un domaine prédit pour être intrinsèquement désordonné, dont la fonction serait d'interagir avec des partenaires cellulaires. L'hypothèse est que ce domaine va acquérir sa structure en hélice α de type coiled-coil après interaction avec des partenaires cellulaires. Il a donc été supposé que ce domaine serait impliqué dans la formation du dégradosome chez *B. subtilis*[140]. De plus, lorsque l'on analyse la proportion en acide aminés promoteurs d'ordre (Y, I et W) et de désordre (Q, S et E) au sein de la séquence de la RNase Y (1 à 520) et de la séquence du domaine ID (24 à 192) (figure 44C), on constate que le domaine ID contient 67 % et 46 % de la totalité des résidus Gln (Q) et Glu (E), sachant que ce domaine ne réprésente que 30% de la séquence totale. La RNase Y possède donc bien une proportion importante d'acides aminés promoteurs de désordre. De plus, sur les huit Tyr (Y) présentes au sein de la séquence de la RNase Y aucune n'est présente dans le domaine ID et seulement 27% des Ile (I) sont présentes dans le domaine ID. Le domaine ID de la RNase Y possède donc les proportions typiques en acides aminés d'une protéine ID (tableau 9A et B, cf III.A.i).

.

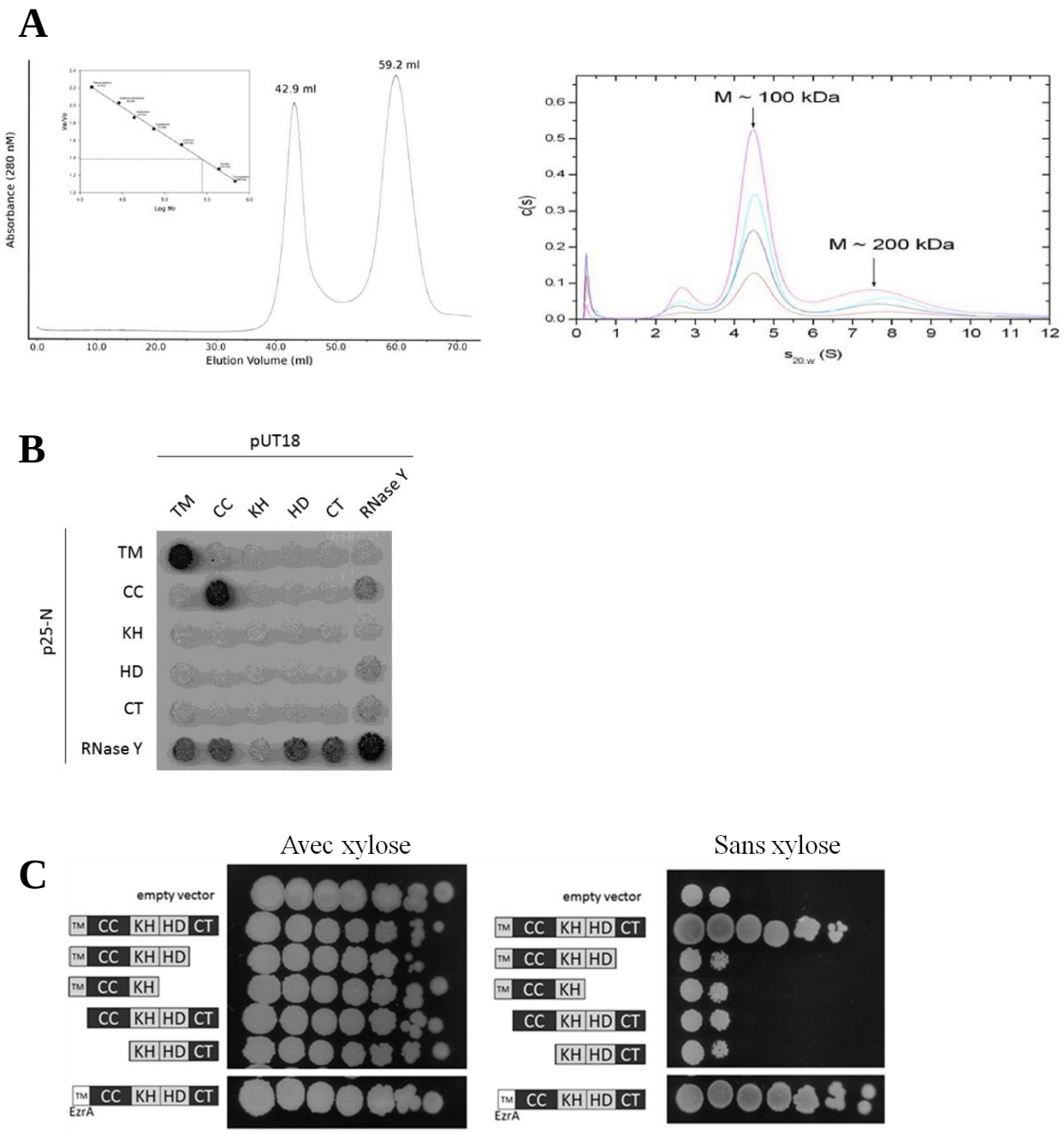


Figure 45: A. Purification sur colonne de chromatographie d'exclusion stérique (à gauche) de la RNase Y (Δ1-25). Deux pics principaux sont observés dont un correspond à une espèce de 280 kDa. Détermination de la masse réelle du tétramère apparent (280 kDa) par ultracentrifugation analytique (à droite). Deux espèces ont été observé dont une majoritaire à 100 kDa. **B.** Expérience de double hybride chez E. coli montrant l'auto-interaction de l'hélice α TM et du domaine ID de la RNase Y. **C.** Expérience de complémentation in vivo. Le gène rny du chromosome bactérien est placé sous le contrôle d'un promoteur inductible au xylose. Cette souche est transformée avec diverses constructions de la RNase Y clonés dans un plasmide (pGP1370).

Dans cette même étude les auteurs ont cloné et purifiée la RNase Y sans son hélice α transmembranaire (TM) (résidus 25 à 520 ou Δ1-24) et avec une étiquette hexahistine en C-terminale de la protéine. La purification sur chromatographie d'exclsion stérique a montré la présence de deux espèces principales, une espèce de très hauts poids moléculaire éluée dans le volume mort de la colonne et une deuxième espèce éluée à une masse moléculaire calculée, à partir de la calibration de la colonne, de 280 kDa, ce qui laissait penser que la RNase Y était organisée sous forme d'un tétramère (figure 45A, gauche). L'analyse du deuxième pic par ultracentrifugation analytique (AUC) montre un pic majeur avec un coefficient de sédimentation de S20,w = 4,5 correspondant à une masse de 100 kDa et un pic minoritaire avec un coefficient de sédimentation de S20,w = 7,6 correspondant à une masse de 200 kDa (figure 45A, droite). Sachant qu'un monomère de RNase Y a une masse théorique de 57 kDa, la protéine serait donc exprimée principalement sous forme dimérique, avec un peu de forme tétramérique

Pour déterminer quel domaine est impliqué dans la dimérisation de la RNase Y *in vivo*, des expériences de double hybride chez *E. coli* utilisant l'adénylate cyclase ont été effectuées (annexe1, figure 45B)[143]. Les résultats indiquent que le domaine prédit comme étant intrinsèquement désordonné (ID) ainsi que la région N-terminale prédite pour être une hélice α TM auto-interagissent. Le domaine N-terminal serait donc responsable de la dimérisation de la RNase Y. On constate également que la RNase Y entière serait capable d'interagir avec les domaines de façon séparée à l'exception du domain KH. Des expériences de complémentation d'une souche délétée en *rny* par des plasmides portant le gène codant pour différentes constructions de la RNase Y chez *B. subtilis* ont permis aussi de montrer, qu'en absence d'un des domaines de la RNase Y, le phénotype sauvage n'est pas rétabli et que seule la RNase Y entière permet une complémentation fonctionnelle de la souche délétée (Figure 45C).

Tableau 12 : Similarités et différences existantes entre la RNase Y et la RNase E.

RNase Y vs RNase E	
Similarités	**Différences**
1) Localisation membranaire. 2) Même organisation : un domaine structurée et un domaine ID. 3) Même activité (endonucléase principale).	1) Aucune homologie de séquence. 2) Domaines structuraux différents. ex: RNase E (domaine DNase I) / RNase Y (domaine HD) 3) RNase E = tétramère / RNase Y = dimère. 4) IDD RNase E en C-ter / IDD RNase Y en N-ter. 5) RNase E = 1061 a.a / RNase Y = 520 a.a.

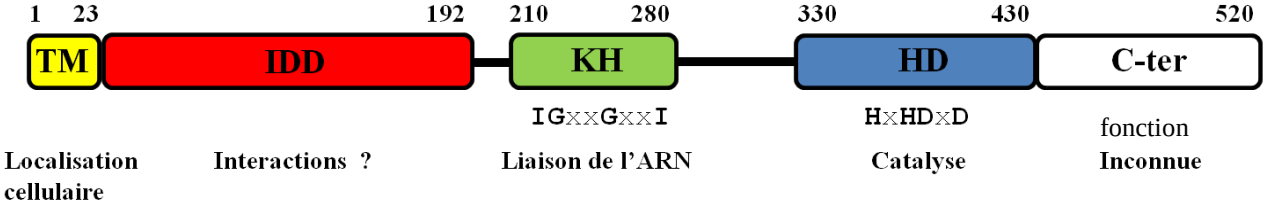


Figure 46: Schéma à l'échelle de l'organisation des domaines de la RNase Y obtenu suite à l'étude réalisée par Lehnink-Habrink *et al.*, 2011. Les numéros des acides aminés au-dessus du schéma indique la délimitation de chaque domaine (IDD = Domaine intrinsèquement désordonné; TM = hélice α transmembranaire; KH = domaine KH; HD = domaine HD; C-ter = domaine C-terminal).

Cette première étude biochimique de caractérisation structurale et fonctionnelle de la RNase Y a permis de mettre en évidence l'organisation de ses domaines. Cette enzyme est donc constituée d'une hélice α TM en N-terminale (1-23), d'un domaine ID (30 -192), d'un domaine KH (210-280), d'un domaine HD (330-430) et d'un domaine C-terminal dont la fonction est inconnue (430-520) (figure 46). Des points de similarités entre la RNase Y et la RNase E ont également été remarqués grâce à cette étude. En effet, ces deux enzymes ne partagent aucune homologie de séquence mais partagent des fonctionnalités similaires. Les deux enzymes sont localisées à la membrane, elles possèdent chacune un domaine catalytique structuré et un domaine ID et elles ont toutes les deux les mêmes caractéristiques au niveau de leurs activités. Ce sont toutes les deux des endonucléases, qui clivent dans les zones riches en nucléotides AU, et qui initient la dégradation des ARNm (tableau 12).

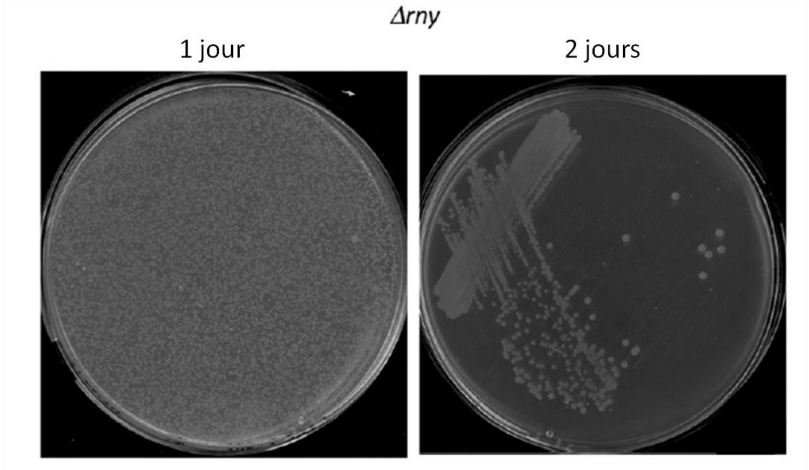


Figure 47: Observation de la souche Δ*rny* sur milieu gélosé après 1 jour d'incubation (gauche). Ces cellules ont ensuite été étalées de nouveau sur milieu gélosé à partir de la première boîte (droite) (d'après Figaro *et al.*, 2013)

ii. La RNase Y est une enzyme non essentielle mais importante pour la physiologie cellulaire.

Durant la période ou la fonction de la RNase Y était inconnue, cette protéine était considérée comme éssentielle. Une étude en 2003 a porté sur la détermination de l'essentialité de 150 gènes de *B. subtilis*[130]. La RNase Y en faisait partie. Pour cela, les auteurs ont inséré, pour chaque gène, un plasmide en amont de la séquence du gène, le plasmide pMUTIN[144]. Ce qui engendre trois conséquences pour la souche :

1) le gène ciblé est inactivé.

2) le gène *lacZ* est fusionné au gène inactivé.

3) le promoteur *Pspac* contrôle la transcription du gène inactivé qui devient inductible uniquement à l'IPTG.

Cette inactivation est appelée une déplétion. Si cette inactivation du gène conduit la souche à devenir dépendante à l'IPTG, cela signifie que le gène est essentiel. Ce fut le cas pour la RNase Y. Mais 10 ans plus tard, la communauté scientifique est revenue sur sa décision. En effet, une étude parue en 2013 prouve que la RNase Y n'est pas essentielle pour *B. subtilis*[145]. Dans cette étude le gène codant la RNase Y, *rny*, a été délété (Δ*rny*) chez une souche de *B. subtilis* (souche 168) et remplacé par une cassette de résistance à la spectinomycine. Puis un plasmide contenant le gène de la RNase Y, sous contrôle d'un promoteur inductible au xylose, a été inséré chez cette souche. Les cellules se sont avérées viables sur milieu gélosé en absence d'induction du gène au xylose (figure 47). Par contre, en absence de RNase Y, le phénotype cellulaire est modifié. En effet, le temps de division cellulaire est augmenté, passant de 26 min pour la souche sauvage à 56 min pour la souche Δ*rny*. La croissance cellulaire est donc fortement altérée en absence de la RNase Y. Cette délétion a également des conséquences sur la compétence et la capacité de sporulation de *B. subtilis*. En effet, il y a une perte totale de la compétence bactérienne, et la capacité de sporulation est diminué de quatre ordre de grandeur.

Tableau 13: Sensibilité aux antibiotiques de la souche délétée en RNase Y (d'après Figaro *et al.*, 2013)

Antibiotique	Sensibilité (augmentation en %)
Tétracycline	150
Érythromycine	45
Ampicilline	40
Chloramphénicol	35

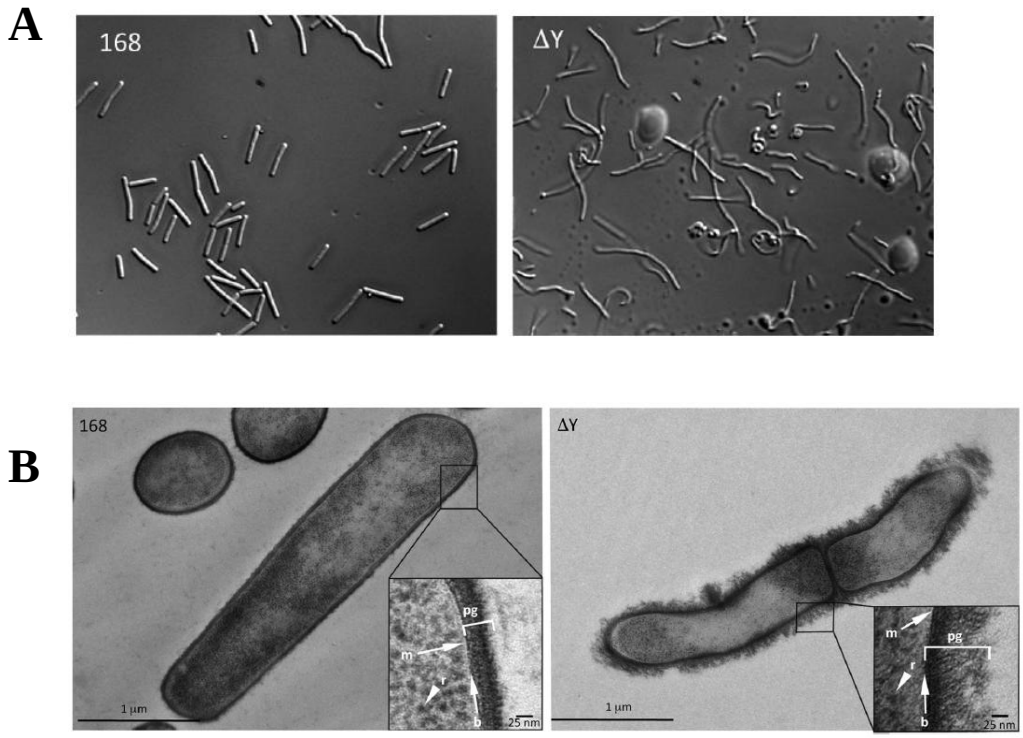


Figure 48: A. Comparaison de la morphologie cellulaire de la souche sauvage (gauche) et de la souche Δ*rny* (droite) au microscope optique. **B.** Comparaison de la morphologie de l'enveloppe cellulaire chez la souche sauvage (gauche) et chez la souche Δ*rny* (droite) au microscope électronique à transmission (d'après Figaro *et al.*, 2013)

Une autre conséquence, très importante, causée par cette délétion est l'hypersensibilité de la souche Δ*rny* à certains antibiotiques. La souche Δ*rny* a une sensibilité accrue aux antibiotiques suivant : tétracycline, érythromycine, chloramphénicol et ampicilline (tableau 13). Cette augmentation de la sensibilité a été calculée en déterminant le pourcentage d'augmentation du diamètre de la zone d'inhibition de l'antibiotique, comparé à celui de la souche sauvage. Ces données suggèrent que l'enveloppe cellulaire est altérée et que la perméabilité à certains antibiotiques soit modifiée. Ceci qui a été conforté par des expériences de microscopie optique et électronique. La souche Δ*rny* a été observée en microscopie optique et en microscopie électronique à transmission afin de comparer la morphologie cellulaire avec celle de la souche sauvage.

En microscopie optique, on constate que les cellules Δ*rny* ont une morphologie cellulaire altéré, les cellules n'ont plus la capacité de former des bâtonnets de taille régulière caractéristique des bacilles. On observe des cellules avec une grande diversité de forme et de longueur (figure 48A). De plus, en microscopie électronique à transmission on constate que le peptidoglycane des cellules Δ*rny* est désorganisé, il ne forme pas une monocouche d'épaisseur régulière (figure 48B). Cette différence au niveau de la morphologie de l'enveloppe cellulaire peut causer une altération au niveau de la perméabilité cellulaire aux antibiotiques.

En conclusion, la RNase Y n'est donc pas une enzyme essentielle mais, en son absence, la physiologie cellulaire est altérée. La RNase Y est également présente chez certains pathogènes à Gram positifs, comme *Staphylococcus aureus* ou *Streptococcus spyogenes*. Il serait donc intéressant de déterminer l'effet de la délétion du gène *rny* chez ces bactéries pathogène pour l'homme.

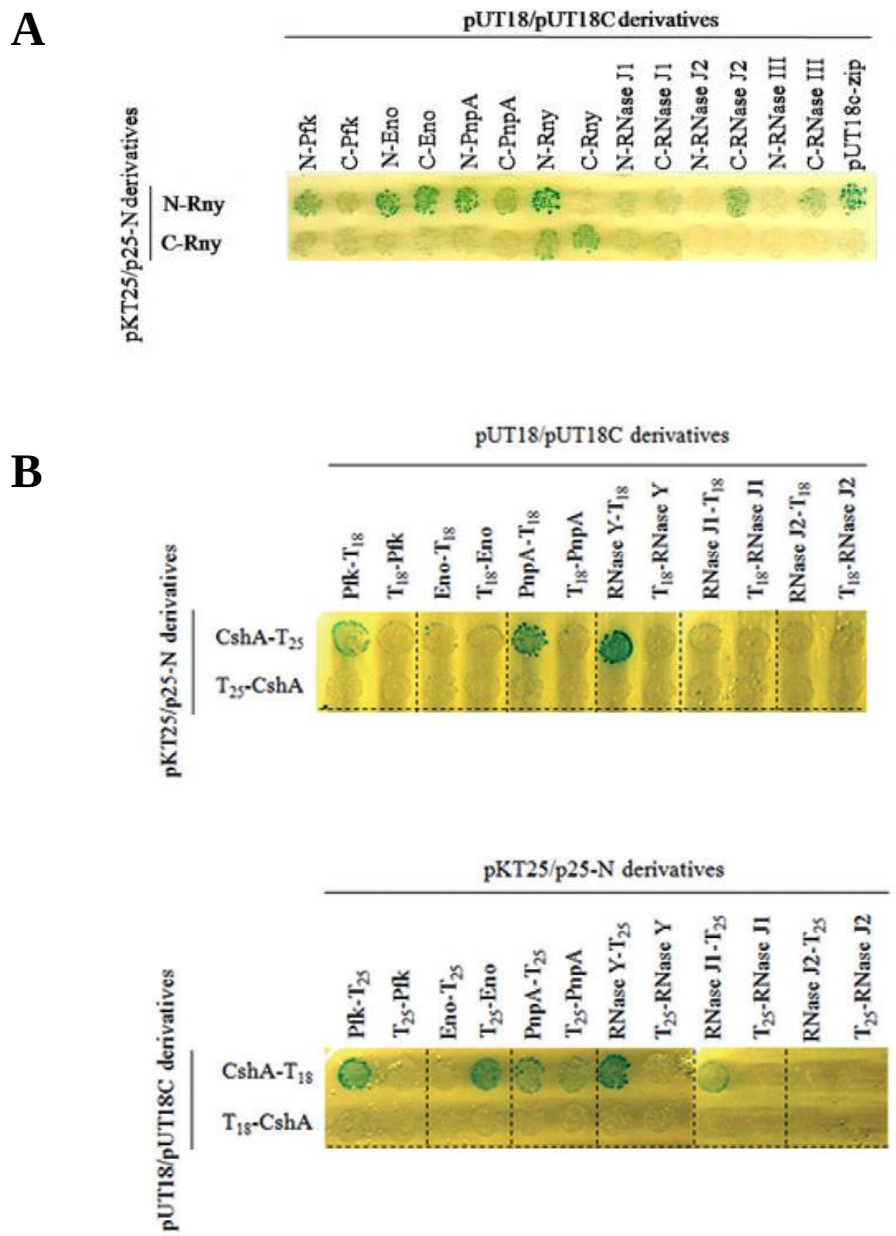


Figure 49 : A. Expérience de double hybride chez *E. coli* montrant les interactions potentiels entre la RNase Y et des protéines du métabolisme des sucres (Phosphofructokinase, Enolase) et du métabolisme des ARN (RNase J1/J2, RNase III, PNPase) (d'après Commichau *et al.*, 2009). **B.** Expérience de double hybride chez *E. coli* montrant les interactions potentielles entre CshA et ses partenaires protéiques potentiels au sein du dégradosome (RNase Y, RNase J, PNPase, phosphofructokinase et énolase). La sous-unité T25 (en haut) ou T18 (en bas) est fusionnée à CshA (d'après Lehnik-habrink *et al.*, 2010).

III.C. Existence d'un dégradosome chez *B. subtilis* ?

i. Un postulat de départ

Au vue des similitudes fonctionnelles avec la RNase E portées par la RNase Y, ainsi que son implication majoritaire dans la dégradation des ARNm, il a été supposé que la RNase Y pourrait potentiellement interagir avec des partenaires cellulaires pour former un dégradosome chez *B. subtilis.*

Plusieurs études, à partir de 2009, ont donc cherché à prouver l'existence de ce dégradosome chez les bactéries à Gram positifs. Les partenaires choisi pour effectuer ces études ont donc été la PNPase, l'énolase et l'homologue de RhlB, l'hélicase d'ARN CshA avec en plus la RNase J mais également la phosphofructokinase, une autre enzyme de la glycolyse. Des expériences de double hybride chez *E. coli* et de co-purification *in vivo,* en présence ou non d'un agent de pontage chimique (formaldéhyde) ont été réalisées.[50,140,146].

ii. Expériences de double hybride chez *E. coli*

Des interactions entre des enzymes du métabolisme de l'ARN (RNase Y, PNPase, RNase J1) et du métabolisme des sucres (énolase, phosphofructokinase) ont d'abord été montrées par double hybride chez *E. coli* (Figure 49A). Les auteurs ont utilisé l'adénylate cyclase qui via un gène rapporteur permet de détecter des interactions entre deux partenaires (annexes 1). Dans cette première étude, on constate que la RNase Y interagit avec l'énolase, la PNPase et la phosphofructokinase uniquement lorsque la sous-unités T18 de l'adénylate cyclase est fusionnée à l'extrémité N-terminale de la RNase Y (figure 49A)[50]. Ce qui laisse penser que l'interaction avec ces partenaires se ferait via le domaine C-terminale de la RNase Y (430-520) dont la fonction est inconnue, car lorsque la sous-unité de l'adénylate cyclase est en C-terminale il n'y aucune interaction, probablement parce qu'elle doit gêner les contactes protéine -protéine.

Puis, il a été montré par cette même technique que l'hélicase d'ARN CshA interagirait avec la RNase Y et la PNPase (Figure 49B) au sein du dégradosome de *B. subtilis.* Ici, l'interaction avec la RNase Y est observée uniquement lorsque la sous-unité de l'adénylate cyclase (T18 ou T25) est fusionnée à l'extrémité C-terminale de la RNase Y. Il est donc possible, dans le cas de CshA, que l'interaction se fasse au niveau du domaine intrinsèquement désordonnée de la RNase Y[146].

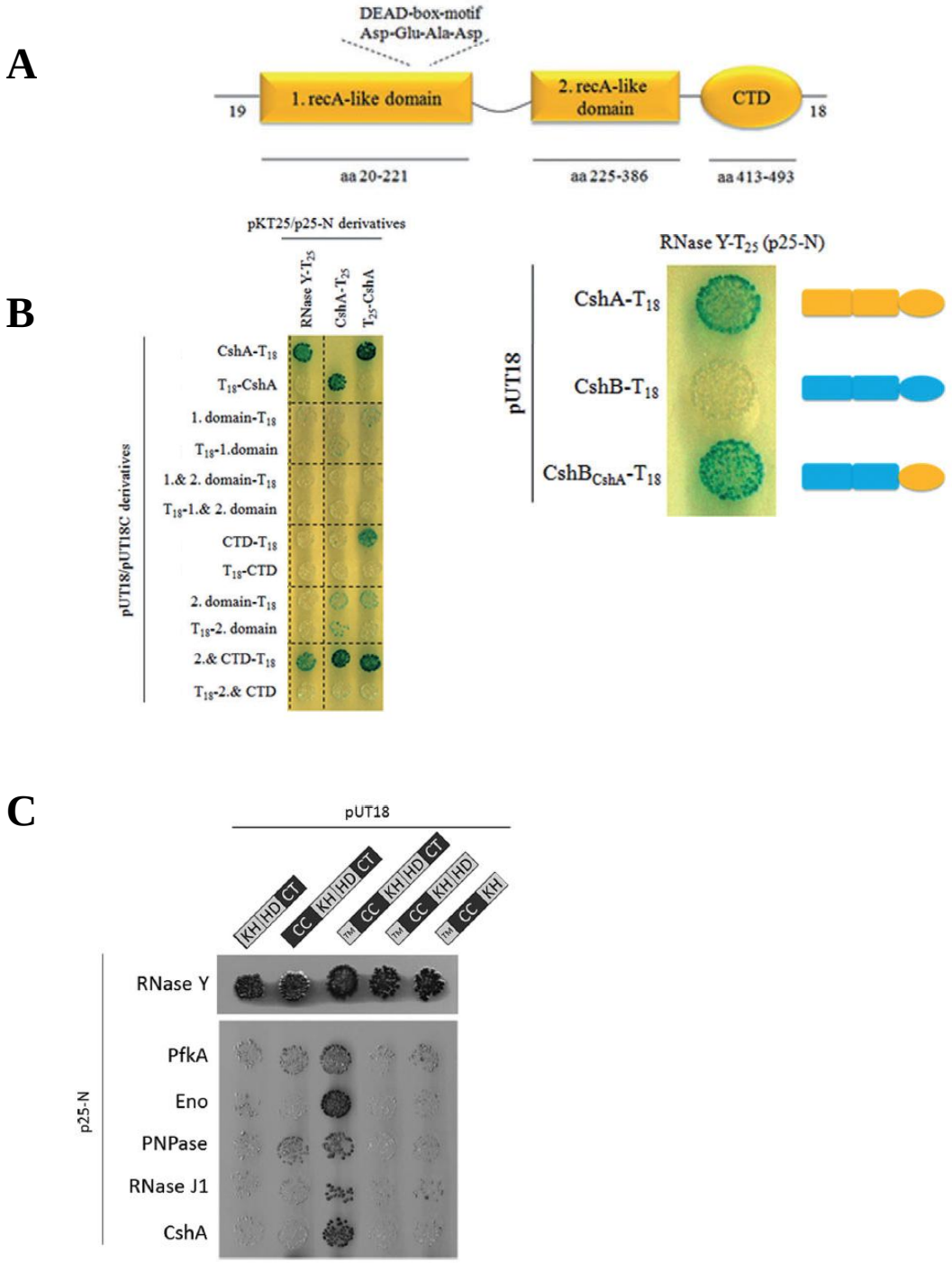


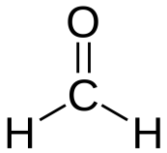
Figure 50: A. Organisation des domaines de l'hélicase CshA de *B. subtilis* **B.** Expréiences de double-hybride chez *E. coli* montrant l'interaction des diverses constructions de CshA avec la RNase Y (a gauche) et l'interaction de la chimère CshB portant le domaine CTD de CshA avec la RNase Y (à droite) (d'après Lehnik-habrink et al., 2010) **C.** Expérience de double-hybride chez *E. coli* montrant les interactions des différentes constructions de la RNase Y déletés d'un ou plusieurs domaines avec les partenaires potentiels du dégradosome de *B. subtilis* (d'après Lehnik-habrink *et al.*, 2011)

Dans cette même étude sur CshA, il a été montré que CshA interagirait avec la RNase Y via son domaine RecA n°2 (225-386) et son domaine C-terminal. CshA est constitué de deux domaines RecA (n°1 et n°2) et un domaine C-terminal. Le domaine RecA n°1 portant le motif DEAD permettant de lier l'ATP. (figure 50A). Les auteurs ont réalisés des constructions différentes de CshA dans lesquelles certains domaines sont délétés. Cinq constructions ont été réalisées : RecA n°1 seul / RecA n°2 seul / RecA n°1 + RecA n°2 / CTD seul / RecA n°2 + CTD. L'interaction avec la RNase Y a été testée pour l'ensemble des constructions. Seule la construction avec les domaines RecA n°2 + CTD interagit avec la RNase Y (figure 50B, gauche).

Ensuite, les auteurs ont également utilisé CshB une autre hélicase DEAD ATP-dépendante. Cette hélicase n'a aucune implication au sein du dégradosome. Ils ont donc fusionné le domaine CTD (413-493) de CshA à l'extrémité C-terminale de CshB. Par double-hybride, on constate que la chimère CshB-CTD (CshA) interagit avec la RNase Y (figure 50B, droite). Ceci confirme que c'est le domaine C-terminale de CshA qui permet l'interaction avec la RNase Y.

Une étude similaire à celle faite sur CshA a été réalisée sur la RNase Y. Diverses constructions de la RNase Y ont été clonés dans lequelles ils manque un ou plusieurs domaines[140]. L'interaction de ces constructions avec les différents partenaires a ensuite été testée également par double hybride. Cette étude a révélé que seule la forme entière de la RNase Y possède la capacité d'interagir avec les différents partenaires impliqués dans la formation du dégradosome et qu'en particulier la présence de l'hélice α (TM) est nécessaire (figure 50C). On peut donc supposer que la localisation membranaire de la RNase Y joue un rôle important dans la formation de ces interactions.

A



B

Co-purification *in vivo* avec CshA
(étiquette Strep en C-ter)

strain		crude extract	-FA	+FA
GP1029	α-PnpA*			
GP1028	α-RNase J1*			
GP1031	α-RNase Y*			
GP1026	α-Eno			
GP1027	α-PfkA*			
GP1026	α-GltC			
GP1032	α-YtsJ*			

Co-purification *in vivo* avec RNase Y
(étiquette Strep en C-ter)

strain	α-FLAG	crude extract	co-purification
GP1016	PNPase		
GP1024	RNase J1		
GP1039	RNase J2		
GP1022	CshA		
GP1023	PfkA		
GP1040	Eno		
GP1025	YtsJ		

C

	crude extract	cytosolic fraction	membrane fraction
α-Rny			
α-CshA			
α-CggR			

Figure 51 : A. Formule chimique du formaldéhyde utilisé comme agent chimique de pontage covalent. **B.** Western blot réalisés à partir des co-purifications faites *in vivo* des partenaires en absence ou en présence de formaldéhyde avec comme protéine "appât" CshA (à gauche) ou la RNase Y (à droite) portant une étiquette Streptavidine en C-terminal. Les fractions déposées sur gel SDS-PAGE proviennent de l'élution de la colonne de streptactine, permettant de retenir l'étiquette straptavidine. Dans les deux cas les protéines GltC et YtsJ servent de contrôle (d'après Lehnik-habrink *et al.*, 2010 et d'après Lehnik-habrink et al., 2011). **C.** Western blot des fractions cytosolique et membranaire d'une souche de *B. subtilis* dont la protéine CshA porte une étquette FLAG. Des anticorps anti-FLAG ont été utilisés pour révéler les protéines en Western blot. La protéine CggR qui est une protéine cytoplasmique est utilisée comme contrôle positif (d'après Lehnik-habrink *et al.*, 2010)

iii. Expériences de co-purification in vivo

Des expériences de co-purification *in vivo* ont été réalisées chez *Bacillus subtilis*[140,146]. Ces expériences ont pour but de co-purifier des complexes protéine - protéine. La méthode qui a été utilisé est la méthode appelée SPINE[147]. Dans cette technique, la protéine étudiée, appelée "appât", va porter une étiquette streptavidine permettant de la purifier sur colonne de streptactine. Les protéines partenaires, les « proies », dont on veux savoir si elles interagissent avec la protéine "appât" portent uniquement une étiquette FLAG permettant de les détecter par Western blot avec un anticorps anti-FLAG. Dans cette méthode, seul des interactions binaires sont étudiées (protéine "appât" - protéine "proie"). Les souches sont cultivées jusqu'à DO_{600} = 1,0 puis une partie des cellules est collectés avant d'ajouter un agent chimique de pontage, communément c'est le formaldéhyde qui est utilisé. Cette molécule va permettre de lier chimiquement deux protéines, uniquement si elles sont très proches l'une de l'autre, comme dans le cas d'une interaction. La formule chimique (figure 51A) très simple du formaldéhyde confirme qu'il faut que les deux partenaires soit très proche pour pouvoir être liés chimiquement. Pour l'étude du dégradosome de *B. subtilis*, deux études ont été faites, une dans laquelle la protéine "appât" est la RNase Y (figure 51B, droite) et une autre avec CshA (figure 51B, gauche). Les partenaires étant la PNPase, la RNase J, l'énolase et la phosphofructokinase.

Les interactions de CshA avec les protéines RNase J1, Y, énolase et PfkA ne sont détectées qu'en présence de l'agent chimique. Seule la PNPase interagit avec CshA sans avoir besoin d'agent chimique. Lorsque c'est la RNase Y qui est la protéine "appât", les interactions ne sont détectées qu'en présence de formaldéhyde.

Toujours avec le même principe, une expérience *in vivo* a été réalisé sur la souche de *B. subtilis* dans laquelle CshA porte une étiquette FLAG permettant d'être identifiée par Western blot avec des anticorps anti-FLAG. Des cellules de *B. subtilis* ont été lysées et la fraction membranaire a été séparée de la fraction cytosolique (Figure 51C). On remarque sur le Western blot que la RNase Y et CshA sont toutes les deux localisées dans la fraction membranaire. On peut donc dire que la RNase Y et CshA interagissent *in vivo* et que le complexe est localisé à la membrane via l'hélice α TM de la RNase Y.

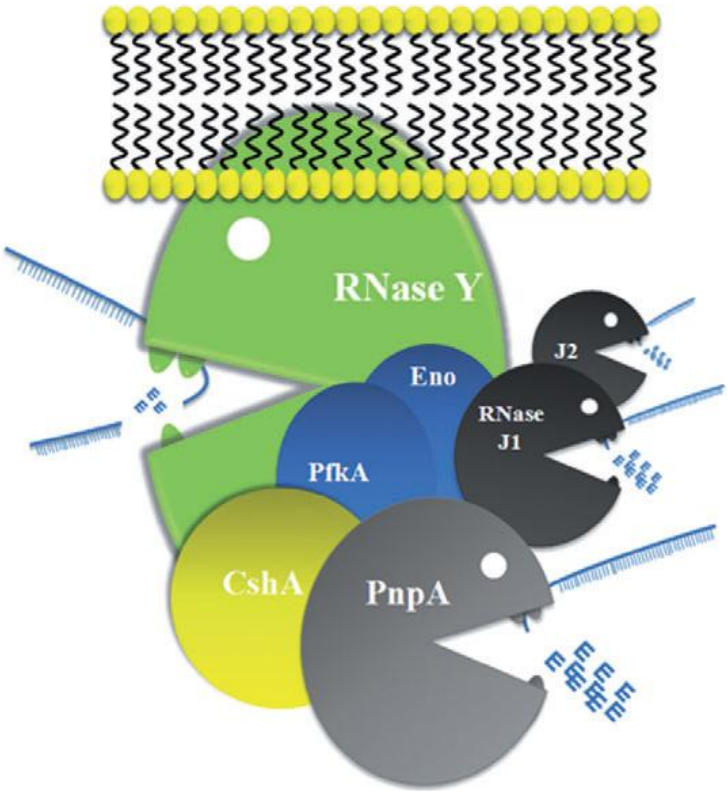


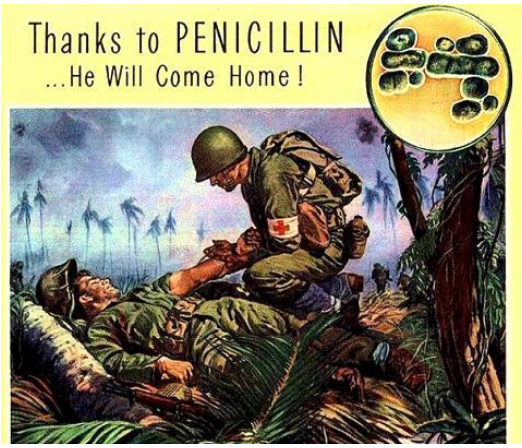
Figure 52: Schéma hypothétique du dégradosome de *B. subtilis* montrant les partenaires interagissant avec la RNase Y au niveau de la membrane cellulaire (d'après Lehnik-habrink *et al.*, 2010)

iv. Le dégradosome de Bacillus subtilis éxiste-t-il ?

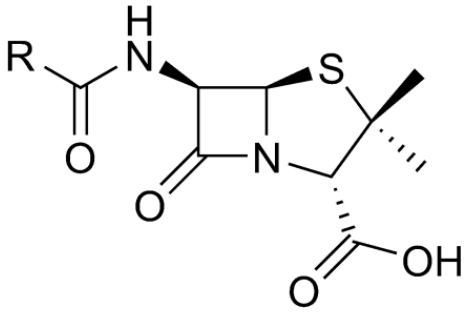
A l'inverse d'*E. coli*, l'existence d'un dégradosome chez *B. subtilis* n'a pas encore été formellement démontrée. En effet, la nécessité d'un pontage chimique pour détecter les interactions par co-purification des partenaires du dégradosome provoque un débat dans la communauté scientifique sur l'existence de ce dégradosome chez les bactéries à Gram positifs. De plus, pour le moment, aucune expérience de co-expression *in vitro* n'a été réalisée avec les protéines de *B. subtilis*. Cette expérience aurait pour objectif de mettre en présence tous les partenaires afin de détecter une interaction. Les expériences de copurification ont été réalisées, jusqu'à présent, uniquement pour des complexes binaires.

A partir de ces données, l'hypothèse émise est que les interactions seraient transitoires, c'est à dire quelles seraient rapides, les protéines interagiraient sur une période de temps très court avec une faible affinité. Cette hypothèse pourrait expliquer l'impossibilité d'isoler un complexe par co-purification sans utiliser un agent chimique de pontage.

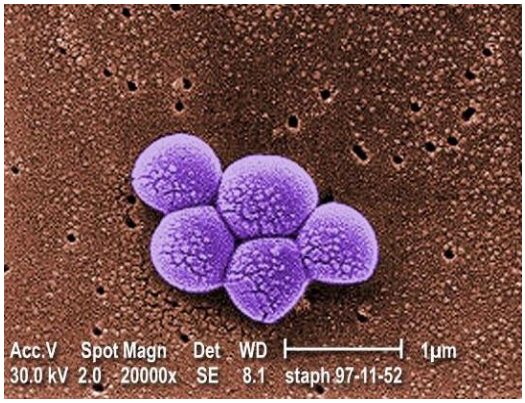
A



B



C



D

E	*Enterococcus faecium* (VRE)
S	*Staphylococcus aureus* (MRSA)
K	*Klebsiella* and *Escherichia coli* producing ESBL
A	*Acinetobacter baumannii*
P	*Pseudomonas aeruginosa*
E	*Enterobacteriaceae*

Figure 53: A. Affiche publicitaire de 1944 publiée dans life Magazine par les laboratoires Schenley. **B.** Structure de la péncilline montrant le noyau β-lactame. **C.** *Staphylococcus aureus* résistant à la méthicilline observé au microscope électronique (d'après le CDC, Centers for Disease Control). **D.** Liste des bactéries multi-résistantes les plus importantes annotées sous le cigle ESKAPE (d'après Lisboa et Nagel, 2011)

III.D. La RNase Y est un facteur de virulence chez *Staphylococcus aureus*

i. Les bactéries font de la résistance

Les maladies infectieuses sont la deuxième cause de mortalité au niveau mondial et la troisième cause de mortalité dans les pays développés au niveau économique[148]. Depuis la découverte fortuite de la pénicilline par Alexander Fleming en 1928[149] et de sa production à grande échelle en 1939, qui permis de soigner de nombreux blessés durant la Seconde guerre mondiale (figure 53A), les antibiotiques sont très largement utilisés à l'heure actuelle dans le traitement des maladies infectieuses causées par des germes bactériens[7]. Fleming lui même ne pensais pas que sa découverte allais révolutionner le domaine de la médecine et de la santé publique[150].

"*When I woke up just after dawn on September 28, 1928, I certainly didn't plan to revolutionize all medicine by discovering the world's first antibiotic, or bacteria killer. But I suppose that was exactly what I did*."
"I did not invent penicillin. Nature did that. I only discovered it by accident."

- *Sir Alexander Fleming,*

Ces composés ciblent des enzymes impliquées dans des processus cellulaires importants que sont : 1) la réplication et la transcription de l'ADN; 2) la traduction protéique; 3) la synthèse de la paroi bactérienne; 4) le métabolisme du folate[7,148]

Mais après 1940, il y a eu émergence des premières souches résistantes aux antibiotiques. Il y eu alors la découverte d'autres antibiotiques et maintenant il existe de nombreuses classes d'antibiotiques (annexes 3). Mais depuis plusieurs années, il y a l'émergence de bactéries multi-résistantes aux antibiotiques[6]. Ces souches sont résistantes à plusieurs antibiotiques. Ceci est causé par l'utilisation abusive des antibiotiques. En effet, les firmes pharmaceutiques envoient la moitié de leurs production en antibiotiques aux industries agro-alimentaires[148]. Les producteurs de porc, de volaille ou encore les exploitations laitières utilise de façon excessive les antibiotiques sur leurs animaux, même quand ces derniers ne sont infectés par aucuns germes. Cette utilisation abusive permet de sélectionner des bactéries résistantes qui finissent par trouver un chemin dans la population humaine et rendre des maladies traitables en maladies incurables[148].

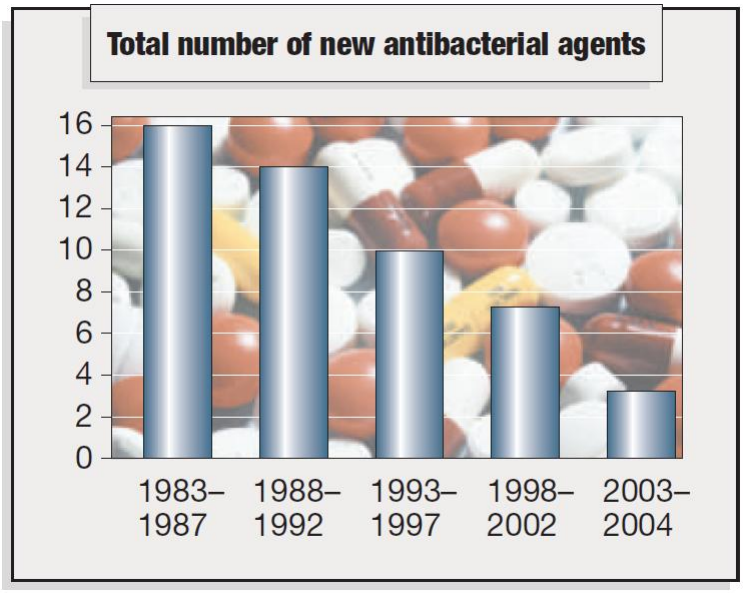


Figure 54: Nombre de nouveaux composés anti-microbiens découvert entre 1983 et 2004 (d'après Nathan, 2004).

Par exemple, pour les β-lactames, la classe principale d'antibiotique dont fait partie la pénicilline, nous en sommes à la quatrième génération d'antibiotiques. Malgré cela les souches multi-résistantes persistent et la plupart des antibiotiques utilisés à l'heure actuelle sont des dérivés de molécules découvertes il y a 50 ans et ont perdu leurs efficacité contre ces pathogènes. En plus de cela, depuis les années 80, la découverte de nouveaux composés anti-microbiens ne cesse de diminuer[148] (figure 54)

De nombreux pathogènes bactériens développent donc des résistances aux antibiotiques actuels. L'un des plus résistants est le Staphylocoque doré (*Staphylococcus aureus*) qui fait parti des pathogènes multi-résistants annotés sous le cigle ESKAPE (figure 53D). En 1940, les infections aux staphylocoque doré étaient traitées avec la première génération de pénicilline[7]. Après un an, la résistance est apparue chez *S. aureus*. En effet, le Staphylocoque doré était devenu capable de produire une pénicillinase permettant de dégrader la pénicilline G. Cette résistance aux β-lactames a donc poussé le développement et l'introduction de la méthicilline, un antibiotique résistant aux pénicillinases. On a eu alors l'apparition du SARM (*Staphylococcus aureus* résistant à la méthicilline) devenu très courant. En 1986, la vancomycine est devenu la barrière contre le SARM. Mais le SARM a alors évolué en SARV (*Staphylococcus aureus* résistant à la vancomycine). Depuis 1990, des combinaisons d'antibiotiques sont donc utilisés pour lutter contre ces bactéries résistantes[7].

Avec l'émergence de ces souches multi-résistantes comme *Staphylococcus aureus,* la nécessité de découvrir de nouvelles cibles thérapeutiques devient primordiale pour lutter efficacement contre ces infections[7].

Tableau 14: Liste des RNases essentielles (E) et non-essentielles (NE) présentes chez les bactéries du groupe ESKAPE (d'après Eidem *et al.*, 2012)

(A) Gram-Negative[a]						
Enzyme	*Escherichia coli*[24,25,47–58]	*Acinetobacter baumannii*[59b]	*Klebsiella pneumoniae*	*Pseudomonas aeruginosa*[59]	*Enterobacter cloacae*	Human
Oligoribonuclease	E	E (63%)	P (93%)	NE (67%)	P (96%)	REX2 (49%)
PNPase	NE	E (61%)	P (96%)	E (65%)	P (96%)	Exosome core[c] (25–33%)
RNase I	NE	P (14%)	P (70%)	—	P (79%)	RNase T2 (25%)
RNase II	NE	—	P (90%)	—	P (80%)	Dis3 (26%) Dis3L (25%)
RNase III	NE	NE (58%)	P (97%)	E (57%)	P (98%)	Drosha (30%) RNase III (30%)
RNase BN/Z	NE	P (31%)	P (69%)	NE[d] (29%)	P (77%)	tRNase Z/ELAC1 (35%)
RNase D	NE	NE (25%)	NE (38%)	NE (38%)	P (82%)	PM-Scl100/Rrp6 (26%)
RNase E	E	E (61%)	P (74%)	NE (63%)	P (77%)	No known ortholog
RNase G	NE	P (54%)	P (97%)	NE (58%)	P (96%)	No known ortholog
RNase HI	NE	E[e] (49%)	P (91%)	E (62%)	P (92%)	RNase H1 (37%)
RNase HII	NE	NE (65%)	P (89%)	NE (71%)	P (91%)	RNase H2, subunit A (30%)
RNase P (RnpA)	E	P (39%)	P (97%)	NE (47%)	P (97%)	No known ortholog[f]
RNase PH	NE	E (67%)	P (96%)	NE (72%)	P (96%)	Exosome core[c] (25-28%)
RNase R	NE	NE (42%)	P (93%)	NE (53%)	P (93%)	Dis3 (29%) Dis3L (26%)
RNase T	NE	NE (54%)	P (86%)	E (63%)	P (89%)	No known ortholog
(B) Gram-Positive[g]						
Enzyme	*Staphylococcus aureus*[38,60]	*Enterococcus faecium*[61]				Human
PNPase	NE	P (63%)				Exosome core[c] (25-33%)
RNase III	NE	P (53%)				Drosha (36%) RNase III (31%)
RNase BN/Z	NE	P (53%)				tRNase Z/ELAC1 (38%)
RNase HI	NE	P (35%)				RNase H1 (35%)
RNase HII	NE	P (48%)				RNase H2, subunit A (26%)
RNase HIII	NE	P (41%)				RNase H2, subunit A (22%)
RNase J1	E	P[h] (63%)				No known ortholog
RNase J2	E	NE[h] (35%)				No known ortholog
RNase M5	NE	P (51%)				No known ortholog
RNase P (RnpA)	E	P (47%)				No known ortholog[f]
RNase R	E	P (48%)				Dis3 (31%) Dis3L (26%)
RNase Y	NE	P (63%)				No known ortholog

E, essential; NE, nonessential; P, present, essentiality unknown.
[a]Protein percent identity calculated in comparison to *E. coli* RNase.
[b]Deletion studies performed in closely related *Acinetobacter baylyi*.
[c]Exosome core components Rrp41, Rrp42, Mtr3, OIP2/Rrp43, Rrp46, and PM-Scl75/Rrp45.
[d]Putative metal-dependent hydrolase.
[e]DnaQ.
[f]Human RNase P subunits Rpp14, Rpp20, Rpp21, Rpp25, Rpp29, Rpp30, Rpp38, Rpp40, Pop1, and Pop5 have no significant similarity to *E. coli* or *S. aureus* RnpA proteins.
[g]Protein percent identity calculated in comparison to *S. aureus* RNase.
[h]Metallo-beta-lactamase superfamily protein.

ii. De nouvelles cibles thérapeutiques

Les bactéries à Gram positif et négatif comprennent de nombreux pathogènes, en particulier humains. Malgré la quatrième génération de molécules anti-bactériennes, celles-ci ciblent toujours les mêmes voies essentielles chez la bactérie[151].

Chez les bactéries, la transcription est couplée à la traduction. Par contre, chez les cellules eucaryotes, la compartimentalisation fait que la transcription et la maturation a lieu dans le noyau et que la traduction se fait dans le cytoplasme. Sans surprise, la machinerie de dégradation des ARNm chez les procaryotes et chez les eucaryotes font appels à des enzymes très différentes. Ces différences peuvent donc être exploitées pour la recherche de nouvelles cibles et donc de nouveaux antibiotiques. En effet, nous avons vu que la dégradation des ARNm est unprocessus cellulaire extrêmement important pour l'adaptation de la bactérie a son environnement[151].

Chez les bactéries pathogènes certaines RNases bactériennes sont essentielles, comme la RNase E, la RNase J, la RNase R ou en encore la RNase P, ces deux dernières étant présentes à la fois chez les bactéries à Gram positifs et négatifs (Tableau 14). De plus, la RNase E est très conservée chez les bactéries à Gram négatifs et il n'y a pas d'homologie significative avec des protéines humaines[73]. Pour trouver des cibles potentielles, il faut que ces cibles soient essentielles. Mais l'essentialité d'un gène est très difficile à définir. En effet, traditionnellement, les tests pour déterminer quelles enzymes sont essentielles se font dans des milieux riches et fortement oxygéné. Mais ces conditions (*in vitro*) sont très différentes de celles rencontrées chez un patient infecté (*in vivo*). Le milieu peut différer chez l'hôte au niveau de l'oxygène, du pH, de la source de carbone, et le système immunitaire va également se défendre pour éliminer le pathogène[151].

La bactérie va donc devoir s'adapter a ce nouvelle environnement en exprimant certains gènes, différents de ceux exprimés en culture de laboratoire. Un gène peut donc être essentiels dans les conditions de laboratoire et ne plus l'être dans les conditions d'une infection *in vivo*, et inversement, un gène peut ne pas être déterminé comme essentiel dans les conditions de laboratoire mais le devenir *in vivo*. Nous avons vu précédemment que la RNase Y n'était pas essentielle chez *B. subtilis* et il en est de même pour la RNase J, mais nous avons remarqué que la physiologie de la cellule est très affectée lorsque la RNase Y n'assure plus ses fonctions. La RNase Y est également présent chez le Staphylocoque doré, alors, quelle est l'influence de l'absence de cette RNase sur cette bactérie pathogène ?

iii. La RNase Y de Staphylococcus aureus

Chez le Staphylocoque doré, la RNase Y est appelée CvfA pour Conserved virulence factor A. En effet, chez *S. aureus*, la RNase Y, ou CvfA, est un facteur de virulence. La virulence d'un organisme pathogène désigne l'intensité de son pouvoir pathogène, c'est une mesure de la sévérité de la maladie dont il est responsable. Cette virulence est mesurée par la diminution de la valeur sélective qui reflète la capacité d'un organisme a survivre et à se reproduire. Un facteur de virulence est donc une protéine qui va jouer un rôle dans la virulence du pathogène en question.

Par la suite, nous n'allons pas dire CvfA mais RNase Y, pour éviter les confusions et pour plus de clarté.

Tout d'abord, le staphylocoque doré pénètre dans l'organisme via une brêche dans l'épiderme de l'hôte. Les symptômes d'une infection par le staphylocoque doré sont des infections des tissus de l'épiderme. La peau est rouge, tuméfiée et inflammée. Cette infection peut aussi se présenter sous forme de furoncles, kystes ou autre abcès. Mais *S. aureus* peut également causé des infections urinaires, des septicémie ou même des pneumonies. Toutes les parties du corps peuvent être potentiellement infectées et dans les cas les plus graves ces infections peuvent entrainer la mort du patient.

Tous ces symptômes sont dus à l'expression combinée de différents facteurs de virulence que sont les protéines de surface cellulaire pour l'adhésion à l'hôte, les hémolysines provoquant la lyse des globules rouges, des protéases et nucléases pour détruire les tissus de l'hôte. L'expression de ces facteurs de virulence est sous le contrôle de facteurs de transcription comme *agr, sarA, srrAB* et *arlRS*[137].

La protéine codée par le gène *agr* se va réguler positivement l'expression des facteurs de virulence et il a été montré que la RNase Y va promouvoir l'expression du gène *agr*. La RNase Y est donc impliqué indirectement dans la régulation de l'expression des facteurs de virulence chez *S. aureus*[137].

Plusieurs études ont donc été réalisées sur la RNase Y de *S. aureus*[137,152]. Tout d'abord, il a été montré que la RNase Y de *S. aureus* possède une activité phosphodiestérase, cette activité est responsable du clivage des nucléotides cycliques (AMPc, GMPc,). Chez les bactéries, les nucléotides cycliques sont très importants dans la régulation de nombreux processus cellulaires, dont la virulence[137].

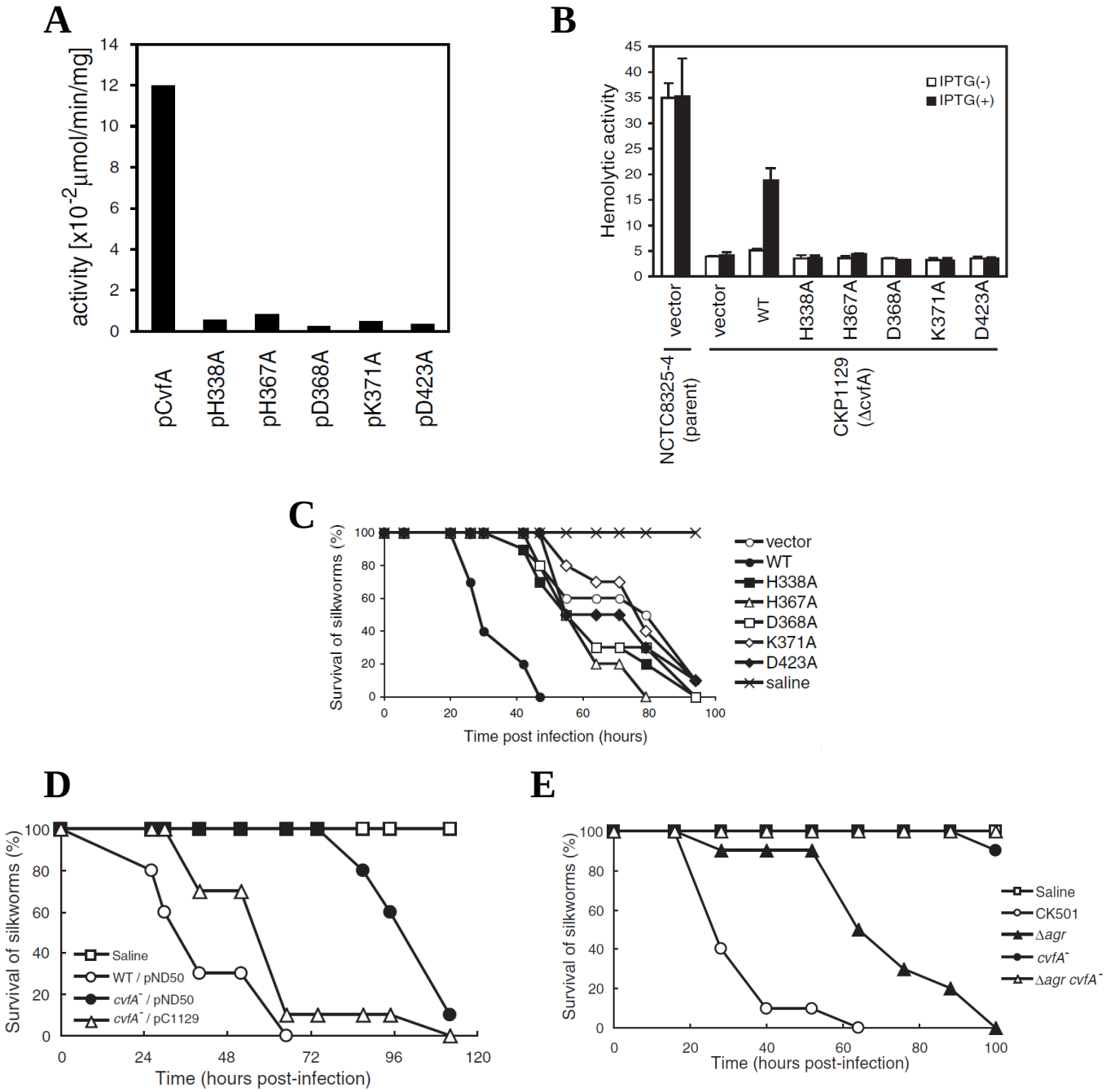


Figure 56 : A. Activité phosphodiestérase mesuré in vitro sur le substrat bis-pNPP pour la RNase Y sauvage et sur les diférents mutants du domaine HD (H338A; H367A;D368A;K371A; D423A) **B.** Activité hémolytique de S. aureus sauvage et de la souche délétée en RNase Y transformé avec le plasmide contenant les divers mutant de la RNase Y. **C.** Virulence chez le verre à soie de S. aureus sauvage et des mutants de la RNase Y (d'après Nagata *et al.*, 2008). **D.** Virulence chez le verre à soie de S. aureus sauvage, Δrny, Δrny + plasmide (rny). **E.** Virulence chez le verre à soie de S. aureus sauvage, Δrny, Δagr et Δrny/Δagr (d'après Kaito *et al.*, 2005)

Ces mêmes auteurs ont également réalisés des mutants de la RNase Y, ces mutations ont été faites dans le domaine HD portant l'activité catalytique. Ils ont tout d'abord vérifié que ces mutants ne possédaient plus d'activité phosphodiestérase en mesurant la dégradation d'un composé, le Bis(p-nitrophenyl) phosphate (bis-pNPP), et effectivement l'ensemble des mutants ne possède plus d'activité phosphodiestérase *in vitro* (figure 56A)[137].

Les auteurs ont ensuite regardé l'effet de ces mutations sur l'activité hémolytique de *S. aureus* ainsi que sur sa virulence. Pour cela, le gène codant la RNase Y a été délété chez une souche de *S. aureus*. Les gènes mutés au niveau du domaine HD de la RNase Y ont été clonés individuellement dans un vecteur puis insérés chez la souche délété en RNase Y. Le gène muté étant sous le contrôle d'un promoteur inductible à l'IPTG[137].

L'activité hémolytique de ces mutants a été mesurée sur des globules rouges infectés par les différentes souches contenant les mutants de la RNase Y. Ils ont constaté que les mutants avaient une très faible activité hémolytique, 5 fois inférieure à celle de la souche sauvage (figure 56B). De plus, la virulence de ces mutants a également été mesurée. Pour cela, les auteurs ont utilisé comme modèle le ver à soie qu'il ont infecté de la même manière que pour les globules rouges. Ils ont constaté que la durée de vie des vers à soie étaient augmentée dans le cas ou ils étaient infectés avec les mutants de la RNase Y en comparaison de la souche sauvage (figure 56C).

Une autre étude a également été réalisée dans laquelle le gène de la RNase Y est également délété chez une souche de *S. aureus*[152]. Le gène de la RNase Y a également été cloné dans un plasmide dans le but de voir si l'insertion de ce plasmide chez la souche délétée permet de rétablir le phénotype sauvage. Ce qui a été le cas. En effet, toujours chez le ver à soie, les souches délétés en RNase Y sont moins virulentes et provoquent une augmentation de la durée de vie des vers à soie infectés par cette souche (figure 56D). Dans cette même étude, les auteurs ont également déletés le gène *agr* responsable de la régulation de l'expression des facteurs de virulence[153]. Ils ont constaté que la déletion du gène *agr* ainsi que du gène de la RNAse Y, au sein de la même souche, a un effet beaucoup plus drastique sur la virulence de *S. aureus*. En déletant ces deux gènes, la virulence des souches est plus faible que lorsque un seul gène est délété (*agr* ou *rny*) (figure 56E).

En conclusion, il est vrai que la RNase Y n'est pas une RNase essentielle, mais lorsque cette enzyme est délétée chez *S. aureus* on constate une diminution de la virulence. De plus, cette chute de virulence de *S. aureus* est d'autant plus importante lorsque le gène *agr* est délété en plus de celui de la RNase Y.

De plus, plusieurs RNases sont éssentielles chez *S. aureus* (tableau 15), comme la RNase J, la RNase P et la RNase R. En combinant cela, la production d'antibiotiques contre les RNases essentielles mais également contre les RNases dont l'absence chez *S. aureus* cause une chute de la virulence, comme la RNase Y, serait un bon moyen de lutte contre les souches multi-résistantes. Dans ce cas-ci, des multi-thérapies seraient également mises en place.

L'étude structurale de ces RNases est donc primordiale pour que, dans un futur espérons le proche, nous puissions trouver de nouveaux antibiotiques afin de lutter contre ces germes bactériens résistants à un grand nombre d'anti-microbiens.

RÉSULTATS

Introduction des articles

Le but de mon travail a consisté à étudier structuralement la RNaseY, qui n'a aucun homologue structural connu. La RNase Y est constituée d'une hélice α transmembranaire (TM), d'un domaine intrinsèquement désordonné (ID), d'une région catalytique avec les domaines KH et HD et d'une région C-terminale (C-ter) (figure 1A, annexe). Nous avons essayé de cristalliser différentes construction dans le but de déterminer leur structure par diffraction des rayons X. Nous avons d'abord cloné et exprimé la RNase Y sans l'hélice TM (Δ1-23) provenant de deux organismes thermophiles, *Geobacillus stearothermophilus* et *Thermus thermophilus*. Ces protéines portaient une étiquette hexahistdine, respectivement en C-terminal et N-termainl, pour faciliter la purification sur colonne de chromatographie d'affinité nickel.

La protéine de *T. thermophilus* était protéolysée au cours de la purification. Nous n'avons pas réussi, malgré l'utilisation d'anti-protéases, à obtenir une protéine homogène pour des essais de cristallisation. Un alignement des séquences des trois RNases Y (*B. subtilis*, *T. thermophilus* et *G. stearothermophilus*) a permis de mettre en évidence que la protéine de *T. thermophilus* possède un domaine ID plus long (+ 60 acides aminés) que celui de la RNase Y de *B. subtilis* et *G. stearothermophilus.* Cela a été confirmé par le prédicteur de désordre PONDR qui a déterminé la présence d'un domaine non-structuré pour la RNase Y de *T. thermophilus* entre les acides-aminés 30 et 260 (230 résidus). Nous pensons que c'est la raison pour laquelle la RNase Y de *T. thermophilus* est très sensible aux protéases d'*E.coli*.

Ensuite, la protéine de *G. stearothermophilus* forme un oligomère soluble de haut poids moléculaire dont la taille augmente avec la concentration en protéine. Les tests de cristallisation de cet oligomère n'ont pas conduit à la formation de cristaux. La purification de la RNase Y (Δ1-23) de *B. subtilis* par chromatographie d'exclusion stérique a permis d'isoler une forme dimérique et une forme oligomèrique soluble, que nous avons obtenus sous forme homogène. Mais les essais de cristallisation n'ont également pas abouti. Nous avons donc ensuite fusionné une étiquette maltose binding protein (MBP) à l'extrémité N-terminale de la RNase Y (Δ1-23) de chaque organisme en utilisant un plasmide codant pour une MBP optimisée pour la cristallisation[154]. En effet, cette MBP possède des mutations de surface afin de remplacer des acides aminés flexibles (Arg, Lys, Asp, Glu, Gln et Asn) par des alanines afin de diminuer l'entropie du système et favoriser les contacts cristallins. D'autre part, cette MBP non-clivable, est fusionnée à la protéine cible par un court linker de trois acides aminés

(AAA) permettant de réduire la flexibilité entre les deux protéines et favoriser la cristallisation. Pour les trois protéines, les constructions MPB ont conduit aux mêmes résultats que pour les étiquettes histidines = la construction de *T. thermophilus* est protéolysée, celle de *G. stearothermophilus* forme des oligomères et celle de *B. subtilis* ne cristallise pas.

Suite à ces résultats négatifs au niveau de la cristallisation, nous avons décidé de faire des constructions de la RNase Y de *B. subtilis* sans son domaine ID, car se fut la seule protéine formant un dimère pouvant être purifié.

Nous avons réalisé deux constructions de la RNase Y: une sans les domaines TM et ID et une sans les domaines TM, ID et C-ter (figure 1C et 1D - annexe). Ces constructions ont été conçues à partir d'une analyse *in silico* avec le logiciel Hydrophobic Cluster Analysis (HCA) qui permet de déterminer la présence de clusters hydrophobes au sein de la séquence d'une protéine et donc de délimiter les zones structurées (hélices α et feuillets β) (figure 1B, article 2)[155]. Ces deux constructions ont été clonées dans le plasmide pGEX-6P2 afin d'introduire une étiquette glutathion-S-transférase (GST) clivable en N-terminal. Ces constructions se sont avérées non solubles (Figure 2B - annexe), sauf en co-expression avec des protéines chaperonnes (DnaK-DnaJ-GrpE). Mais dans ce dernier cas, il n'a pas été possible de séparer la RNase Y des chaperonnes.

Article 1: Characterization of an Fab fragment that promotes ordering of the intrinsically disordered domain of *B. subtilis* RNase Y

Les divers essais de cristallisation de la RNase Y n'ayant pas abouti, nous avons décidé de produire un anticorps dirigé contre la RNase Y. En effet, les anticorps sont souvent utilisés comme outil pour la cristallisation des protéines[156,157]. En particulier, ils ont beaucoup été utilisés pour cristalliser des protéines membranaires, ou des protéines virales car ils permettent de masquer les régions hydrophobes et favorisent donc les contacts cristallins. Récemment, des anticorps ont également été utilisés pour étudier des protéines intrinsèquement désordonnées[124,158–161]. Seul le fragment reconnaissant l'antigène, appelé "Fragment antigen-binding" (Fab), est utilisé pour la co-cristallisation (Figure 4, annexe). L'anticorps, produit par la Société Abmart® (Chine), reconnait un peptide de 12 acides aminés situés dans la région ID de la RNase Y de *B. subtilis*. Nous avons donc produit le Fab, par protéolyse ménagée à la papaïne de l'anticorps. Les essais de cristallisation avec le dimère de RNase Y n'ont pas abouti. Nous avons cependant cristallisé et résolu la structure du Fab en complexe avec le peptide synthétique qui a servi à produire l'anticorps. Les constantes d'affinité pour les complexes Fab-peptide et Fab-BsRNase Y sont respectivement de 355 pM et 122 nM. L'affinité pour le peptide est donc 350 fois supérieure à celle pour le dimère. On peut donc supposer que le peptide est enfoui dans le dimère et que la région N-terminale participe potentiellement à la dimérisation de la protéine. Cependant, une analyse en SAXS des complexes Fab-dimère (BsRNase Y) et Fab-Nter (BsRNase Y) montre que la liaison du Fab permet de structurer le domaine N-terminal. Notre conclusion est que l'anticorps pourrait mimer un des partenaires physiologiques de la RNase Y et induire la dissociation du dimère de Nter-BsRNase Y et/ou des changements de conformation du domaine N-terminal. Ceci suggère que le domaine N-terminal pourrait être impliqué dans la formation d'un complexe de dégradation des ARNm chez *Bacillus subtilis*.

Article 2: The N-terminal domain of *B. subtilis* RNase Y is a molten-globule

Les protéines ID sont très présentes au sein du monde vivant et assurent dans les cellules eucaryotes des fonctions physiologiques importantes, comme la signalisation intracellulaire ou encore la formation de complexes macromoléculaires comme le spliceosome[100]. Il y a très peu de protéines ID caractérisées chez les bactéries. La RNase E d'*E. coli* possède un domaine intrinsèquement désordonné en C-terminal permettant la formation d'un complexe de dégradation des ARNm[75,83,85,89,98]. Lorsque ce domaine est délété, l'enzyme conserve son activité endonucléase pour la maturation des ARNt et ARNr mais elle n'est alors plus capable de cliver efficacement les ARNm[36,162].

Tout comme la RNase E, la RNase Y possède un domaine prédit pour être intrinsèquement désordonné, mais en N-terminal. Aucune expérience n'avait encore confirmé cette prédiction[140]. Nous avons donc cloné et purifié le domaine ID de la RNase Y (acides aminés 24 à 192). Le SEC-MALLS indique également que ce domaine est en équilibre entre une forme monomérique et une forme dimérique. En collaboration avec Carine Van-Heijenoort et Nadine Assrir du laboratoire de chimie et biologie structurale de l'ICSN, une analyse en résonnance magnétique nucléaire (RMN) et en dichroïsme circulaire (CD) de ce domaine a été réalisée. L'analyse du spectre RMN du domaine N-terminal marqué à l'azote ^{15}N indique qu'il y a une centaine de résidus, sur les 176, qui sont déstructurés ou qui forment des hélices α. L'analyse par CD de ce domaine a révélé une structuration majoritaire en hélice α. Cette analyse est en accord avec le programme COILS qui prédit que 120 acides aminés seraient structurés en domaine coiled-coil sur les 162 prédits pour être non-structurés par le programme PONDR. De plus, une analyse en diffusion des rayons X aux petits angles montre que le domaine N-terminal n'est pas complètement déstructuré. En conclusion, la combinaison des différentes méthodes biophysiques (RMN, CD, SAXS) montrent que le domaine N-terminal de la RNase Y possède majoritairement des éléments de structures secondaires de type hélice α et environ 70 acides aminés pleinement déstructurés et qu'il ne forme pas de structure tertiaire. On peut donc dire que le domaine N-ter est un globule fondu. Il reste à déterminer la fonction de ce domaine et à savoir en particulier s'il interagit avec ses partenaires potentiels au sein du dégradosome.

Article 3: The elongated dimer of *Bacillus subtilis* RNase Y assembles as a globular oligomer of high molecular weight

Avant de débuter ce projet, une étude biochimique réalisée en 2011 dans l'équipe de Jörg Stülke avait montré qu'une construction de la RNase Y, sans l'hélice TM, purifiée par chromatographie d'exclusion stérique, se présentait sous deux formes : une forme de très haut poids moléculaire et une forme soluble de masse apparente 280 kDa (masse théorique d'un dimère de RNase Y = 114,2 kda). Une analyse par ultracentrifugation analytique avait révélé que celle-ci correspondait à un dimère (~100 kDa) de forme allongée[140]. L'oligomère de haut poids moléculaire n'avait fait l'objet d'aucune étude approfondie.

Lors de nos purifications de la construction Δ1-23 de la RNase Y portant une étiquette hexahistidine en position C-terminale, nous nous sommes aperçus que la quantité de l'oligomère formé était dépendante de la concentration en protéine et qu'au cours du temps, à haute concentration en protéine, la proportion de la forme dimérique diminuait au profit de la forme de haut poids moléculaire (Figure 3, annexe). Une analyse en SEC-MALLS a permis de déterminer les masses absolues du dimère et de l'oligomère. Le dimère a une masse molaire de 105 kDa et un rayon hydrodynamique de 60 Å alors que l'oligomère soluble une masse de 980 kDa et un rayon hydrodynamique de 130 Å.

Des tests d'activité sur les deux formes de la RNase Y, dimère et oligomère, ont été réalisés dans l'équipe d'Harald Putzer. Ces tests ont révélé que les deux formes ont une activité endonucléase sur le transcrit *yitJ*. Les formes dimérique et oligomérique de la RNase Y ont aussi fait l'objet d'une analyse en diffusion des rayons X aux petits angles (SAXS). Cette analyse révèle que le dimère a une forme allongée (Rg = ~60 Å) et ne présente pas de partie déstructurée et que l'oligomère possède une forme globulaire avec un rayon de giration de 100 Å. Ces résultats sont en accords avec les analyses respectives de ces deux formes en SEC-MALLS.

Nous avons alors décidé d'approfondir l'analyse de la forme oligomérique en solution. En collaboration avec l'équipe de Catherine Venien-Bryan à l'UPMC, nous avons commencé une étude en microscopie électronique de l'oligomère. Des images en coloration négative ont révélé que cette forme était très homogène et formait des particules compactes de rayon compris entre 100 et 150 Å possédant un tunnel. Le rayon des particules observé en coloration négative est en accord avec le Rh obtenu en SEC-MALLS et le Rg obtenu en SAXS. Nous avons pu générer un modèle initial de cet oligomère qui confirme une forme en anneau avec un tunnel au centre. Ce modèle sera prochainement affiné à l'aide d'images prises à 0° et 45°

("tiltées") qui devraient permettre de visualiser d'autres orientations des particules. Les images collectées en cryo-microscopie pourront alors être traitées afin d'obtenir un modèle de plus haute résolution.

Annexe de l'introduction des articles

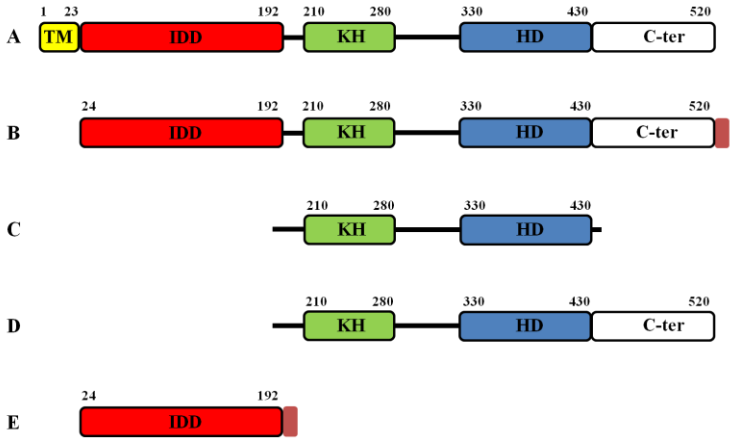
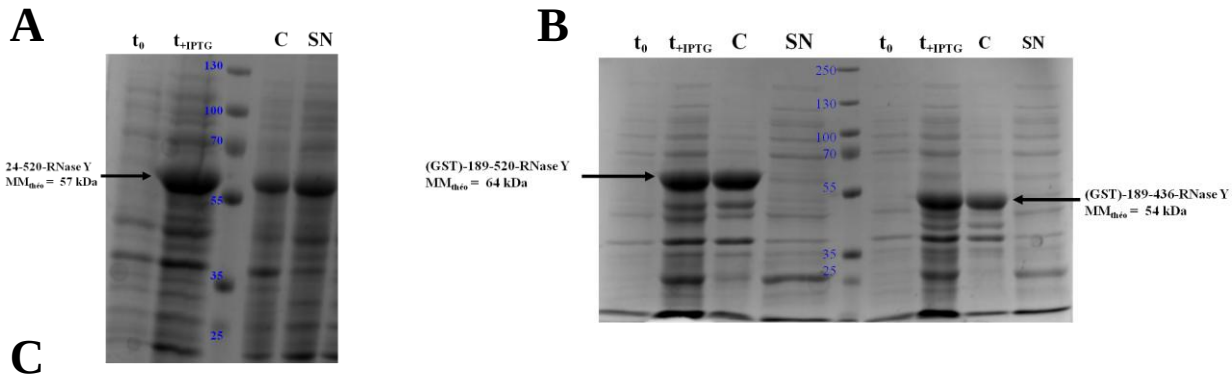


Figure 1: Constructions de la RNase Y réalisées au cours du projet. A. Schéma à l'échelle de l'organisation des domaines de la RNase Y. **B.** RNase Y sans son hélice α TM avec une étiquette hexahistidine en C-terminal. **C.** RNase Y sans son hélice α, son domaine ID et son domaine C-terminal de fonction inconnue. **D.** RNase Y sans son hélice α et son domaine ID. **E.** Domaine ID de la RNase Y avec une étiquette hexahistidine en C-terminale. (Légende: TM = hélice α transmembranire; IDD = domaine intrinsèquement désordonné; KH = domaine KH; HD = domaine HS; C-ter = domaine C-terminal)



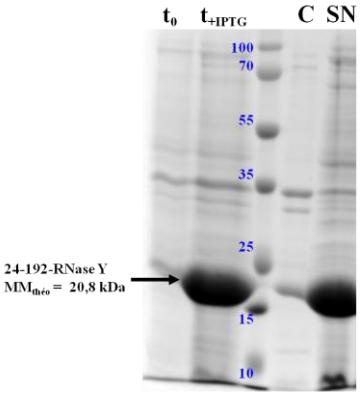


Figure 2: Expression et solubilité des constructions de la RNaseY. Toutes ces constructions ont été exprimées dans la souche d'*E. coli* BL21 (DE3) Star codonPlus. **A.** Solubilité de la RNase Y (Δ1-23). **B.** Solubilté des constructions courtes, 189-520 (à gauche) et 189-436 (à droite) en absence de chaperonne. C. Solubilité du domaine ID (24-192) de la RNase Y (t0 = avant induction à l'IPTG; t+IPTG = après induction à l'IPTG; C= culot; SN = surnageant).

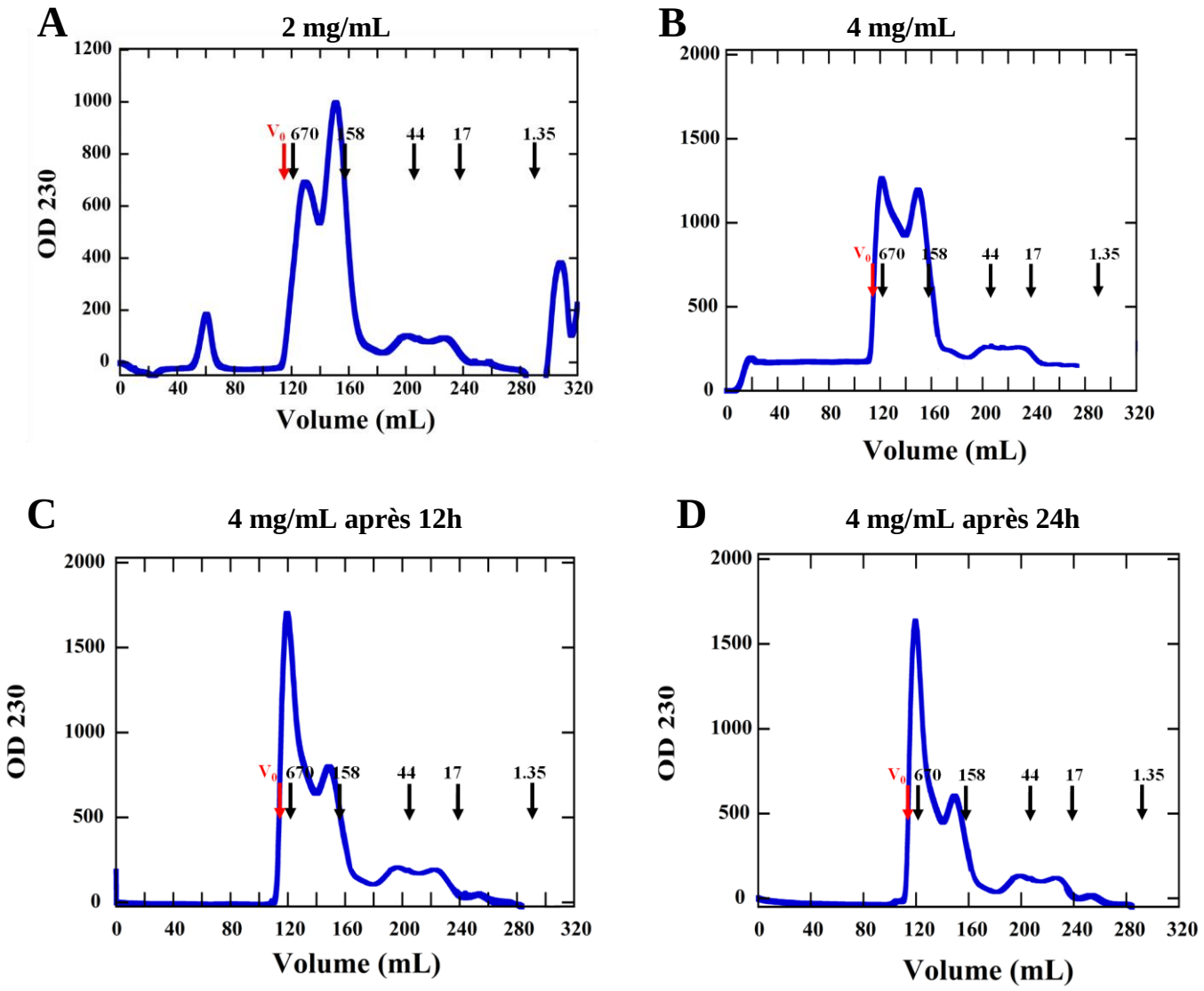


Figure 3: Formation de la forme oligomérique de haut poids moléculaire au cours du temps. Purification de la RNase Y sur colonne de chromatographie d'exclusion stérique (Superdex 200 26/60 PG) dans 20mM HEPES pH 7,5, 500mM NaCl, 10% glycérol. **A.** Injection de 10mL à 2 mg/mL. **B.** Injection de 10mL à 4 mg/mL, **C.** après 12 heures, **D.** après 24 heures.

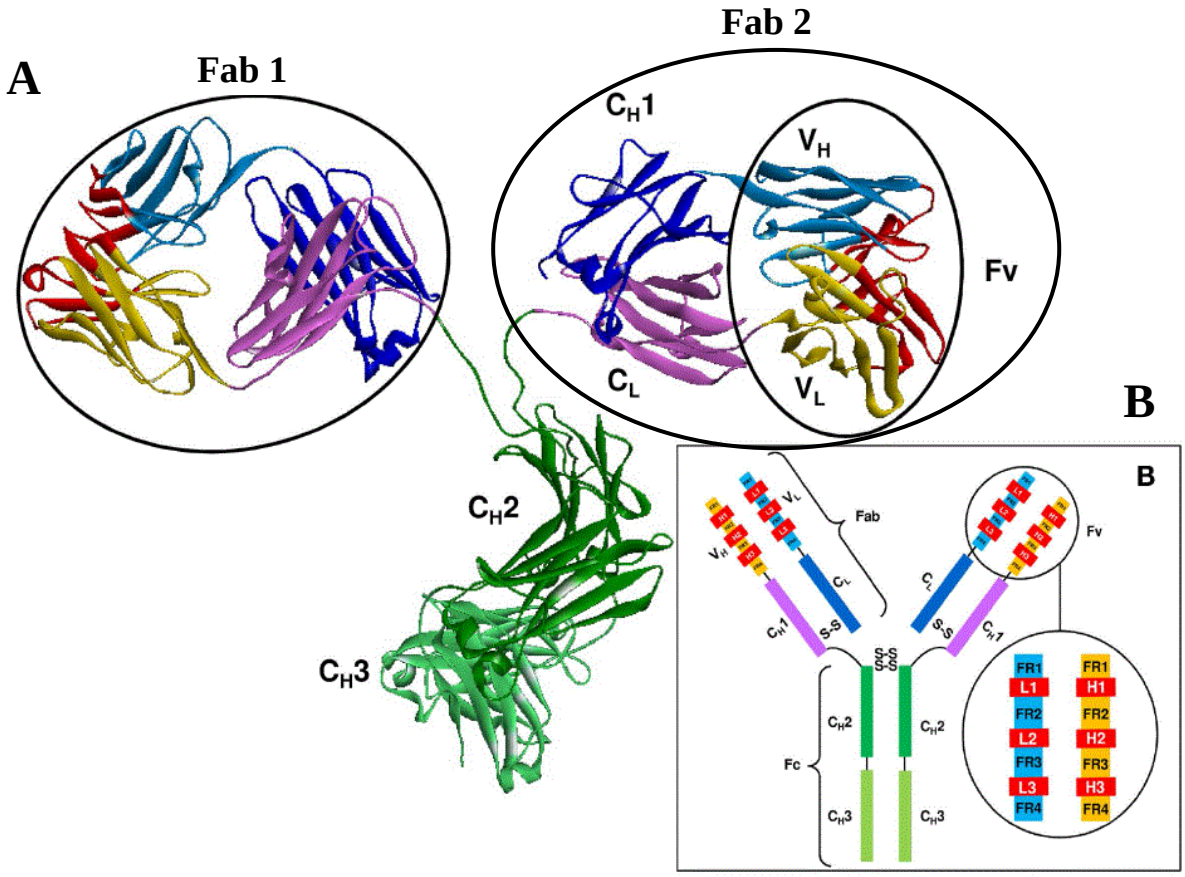


Figure 4: A. Structure cristallographique d'un anticorps entier (PDB : 1IGT). B) Structure schématique d'un anticorps constitué de deux chaines lourdes (H), formées chacune des domaines CH1, CH2, CH3 et VH, ainsi que de deux chaines légères (L), formées par les domaines CL et VL. Chaque anticorps est constitué de deux Fab capables de lier l'antigène. **B.** Les boucles hypervariables ou CDR (Complementary determining regions) permettant la liaison à l'antigène sont présentes dans la partie Fv (VH/VL) du Fab. Ces boucles (trois boucles de la chaine légère L1, L2 et L3 et trois boucles de la chaine lourde H1, H2 et H3) sont représentées en rouge. (D'après Sela-Culang *et al*., 2013)

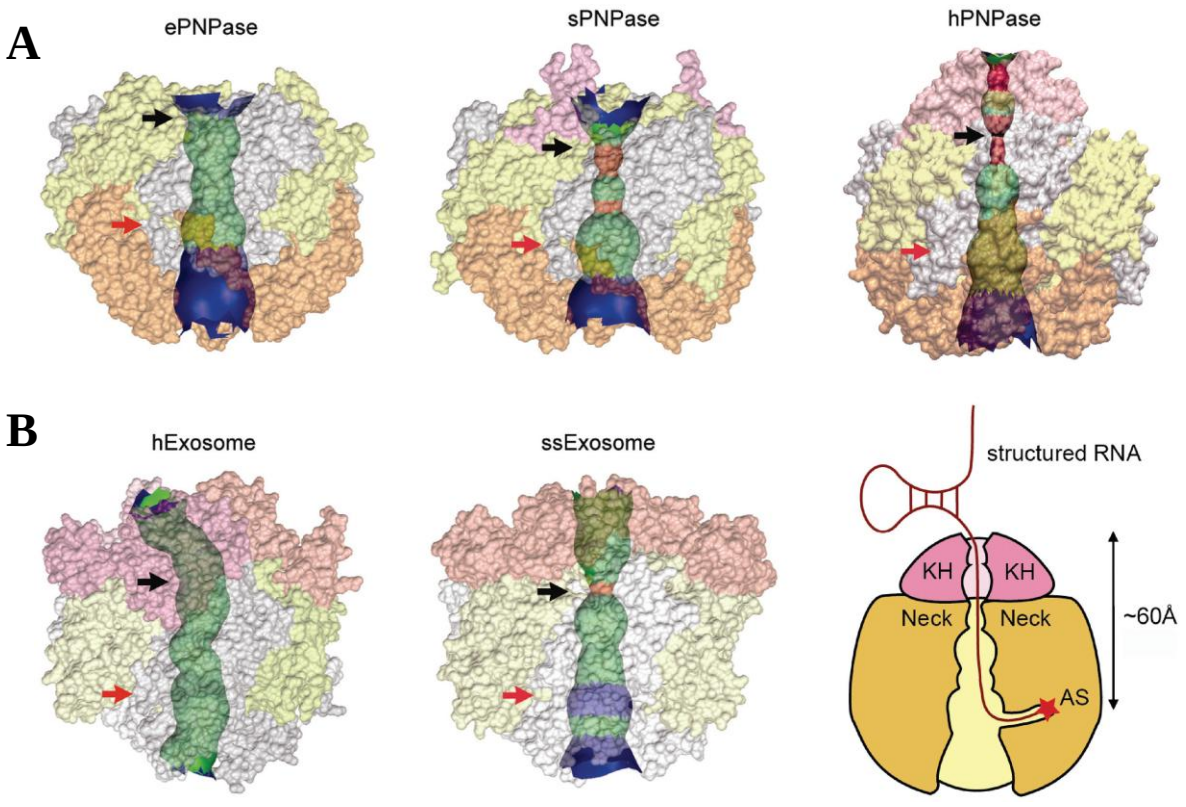


Figure 5: Tunnel central de liaison de l'ARN chez la PNPase et l'exosome. **A.**Comparaison du tunnel de liaison de l'ARN entre les différentes structures cristallographiques des PNPases (PNPase d'*E. coli* = ePNPase; PNPase de *S. antibioticus* = sPNPase; PNPase humaine = hPNPase). Les régions du tunnel ayant un diamètre inférieur à 5 Å, entre 10-20 Å et entre 20-30 Å sont respectivement indiquées en rouge, vert et bleu. Le site actif est indiqué par une flêche noire et la zone d'entrée de l'ARN est indiquée par une flêche rouge. **B**.Comparaison du tunnel de liaison de l'ARN entre les différentes structures cristallographiques des exosomes (Exosome humain = hExosome; Exosome de *S. solfataricus* = ssExosome). Ces structures montrent que le tunnel est plus étroit pour l'exosome de *S. solfataricus* que pour celui de l'homme. **C.** Modèle de la fixation d'un ARN chez la PNPase humaine. (d'après Liang Lin *et al.*, 2011)

ARTICLES

Article 1

Characterization of an Fab fragment that promotes ordering of the intrinsically disordered domain of *B. subtilis* RNase Y

Pierre Hardouin[1], Charles Bou-Nader[1], Soumaya Laalami[2], Harald Putzer[2], Béatrice Golinelli-Pimpaneau[1*]

[1] Laboratoire de Chimie des Processus Biologiques, UMR 8229 CNRS, Collège de France, Université Pierre et Marie Curie, 11 Place Marcelin Berthelot, 75231 Paris cedex 05, France
[2] CNRS UMR8261 - Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

**Corresponding author:*
beatrice.golinelli@college-de-france.fr

Although RNase Y acts as the key enzyme initiating mRNA decay in *Bacillus subtilis* and likely in many other Gram positive bacteria, its three-dimensional structure remains unknown. An antibody, belonging to the rare IgG2b, λx isotype, was raised against a 12-residue conserved peptide from the N-terminal non-catalytic domain of *Bacillus subtilis* RNase Y (Nter-BsRNaseY) that is predicted to be intrinsically disordered. Small angle X-ray scattering indicates that Nter-BsRNaseY gets ordered upon binding the Fab (fragment antigen binding) of this antibody. The dissociation constants of the Fab/Nter-BsRNaseY and the Fab/peptide complexes are similar, suggesting that the peptide has a main contribution to the affinity of Nter-BsRNaseY for the Fab. The crystal structure of the Fab in complex with the antigen peptide shows that the peptide adopts an elongated U-shaped conformation, in which the unique hydrophobic residue of the peptide, Leu6, is completely buried. Hence, the Fab/peptide complex may mimic the interaction of a microdomain of Nter-BsRNaseY with one of its cellular partners within the degradosome complex.

Keywords V lambda x chain, antibody, Fab/peptide complex, crystal structure, RNase Y, intrinsically disordered domain

Introduction

The processing and degradation of mRNA is a key element in the control of prokaryotic gene expression [1-3]. In *E. coli*, RNase E has a pivotal role in these processes [4] but *Bacillus subtilis* and many other Gram positive bacteria have a very different set of enzymes directing mRNA metabolism [5-7]. Indeed, in *B. subtilis,* a novel endoribonuclease called RNase Y (BsRNaseY) shares extensive functional homologies with RNase E despite an absence of sequence similarity [8]. Although not essential for cell viability, the crucial role of RNase Y in initiating mRNA clivage is highlighted by deletion experiments of the *rny* gene, which leads to defects in cell growth, cell morphology and hypersensitivity to antibiotics [9]. The half-life of mRNA is increased two-fold when RNase Y is depleted [8]. Indeed, the abundance of a large number of transcripts is affected by decreasing RNase Y expression [8, 10-14].

A structure prediction analysis revealed the probable architecture of RNase Y [12]: a transmembrane (TM) region anchoring the protein to the membrane [15], formed by the 1-24 first amino acids, followed by an intrinsically disordered domain (IDD), a catalytic domain with KH [16] and HD motifs [17] and a C-terminal region of unknown function. The predicted non-structured IDD region is hypothesized to adopt various conformations to enable the protein to interact with several cellular partners (RNA or proteins), as typical for intrinsically disordered proteins [18-22].

Only a few *in vitro* studies of BsRNaseY have been reported, including activity tests [8] and ultracentrifugation experiments that showed that the main oligomeric form is a dimer [12]. To get more information about the three dimensional structure of RNase Y, we raised an antibody against a small region of the IDD. Not only could this antibody stabilize the conformation of RNase Y, which would favor crystallization of the complex, but it could also serve to get insights into the structure/function of the IDD.
We report here the characterization of the interaction between BsRNaseY and an antibody raised against a 12-residue peptide from the IDD domain using ELISA (Enzyme Linked ImmunoSorbent Assay) and Size Exclusion Chromatography coupled with Small Angle X-ray Scattering (SEC-SAXS). We also report the crystal structure of the fragment antigen binding (Fab) RY79-90 in complex with the peptide used to raise the antibof-dy. Our results show that the antibody can bind BsRNaseY but is more specific for its N-terminal intrinsically disordered domain (Nter-BsRNaseY). The comparison of the low resolution structures obtained by SAXS of Nter-BsRNaseY, alone or in complex with Fab RY79-90, indicates that this domain gets more ordered upon Fab binding. We reveal the fold adopted by the epitope of the intrinsically disordered region of BsRNaseY upon binding with RY79-90 and discuss how the detailed interactions of the Fab/peptide complex observed in the crystal structure could reflect the binding of a microdomain of Nter-BsRNaseY and a cellular partner.

RESULTS AND DISCUSSION

RY79-90 is an IgG2b, λx antibody induced against a 12-residue peptide of the IDD

After having identified peptides from BsRNase Y most susceptible to yield mouse monoclonal antibodies, six monoclonal antibodies were produced. Among them, one antibody induced against the 12-residue peptide (ERRNELQKQENR) covering residues 79 to 90, which belongs to a conserved region of the IDD of BsRNaseY (Fig. 1A), was selected because it had the highest affinity for BsRNaseY, as determined by Western-blot and non-denaturing SDS PAGE (Fig. 1B &C). The antibody, named RY79-90, has a heavy chain belonging to the IgG2b isotype and utilizes a rare Vλx light chain that is observed in only 0.5% of antibody sequences [23] (Figure S1). The amino acid sequences of Vλx chains are only 30-33% identical to that of other Vλ or Vκ light chains. They generally display very few somatic mutations [24-26] and no junctional diversity caused by the rearrangement of the Vλx and Jλ2 segments [23]. A characteristic of the Vλx class is that CDRL3, which is entirely encoded by the Vλx gene fragment, is four amino acids longer than in all other light chains (Figure S1). It has been suggested that this unusual light chain confers specificity to the antibodies that contain it [27]. To our knowledge, there are only two known crystal structures of Fab (fragment antigen binding) bearing a λx light chain: Fab 13F6-1-2 (PDB code 2QHR) [24] and Fv (variable fragment) MW1 (PDB code 2OTU and 2OTW)[26], both solved in complex with their peptide antigen.

Sequence comparison of the light chains of Fab RY79-90, Fab 13F6-1-2 and the Vλx germline gene indicates that Fab RY79-90 underwent only a few maturation affinity mutations, which translate into Glu at position L80 instead of Pro (Figure S1). Interestingly, the heavy chain isotype distribution of Vλx-bearing antibodies in normal mouse serum, upon immunization with myelin basic protein, showed a biased H chain isotype towards g2a and 2b isotypes [27], as is the case for Fab RY79-90. The IgG2b heavy chain of Fab RY79-90 is a combination of the V segment of the Igh-VGAM3.8 VH9 family, a D segment that could not be identified because it is too short and the JH3 segment. There are 10 amino acid mutations compared to the germline gene in the V segment, including one in CDRH1 and two in CDRH2. CDRH3 is exceptionally short, with only three amino acids.

IgG RY79-90 has a higher affinity for the inducing peptide and Nter-BsRNaseY than for BsRNaseY

To know if the conformation of the peptide in interaction with the antibody is representative of BsRNaseY in its active conformation in solution, the affinity of IgG RY79-90 for the inducing peptide, Nter-BsRNaseY and BsRNaseY was determined by a competitive ELISA test [28]. The dissociation constants for the complexes with BsRNaseY, Nter-BsRNase Y and the inducing peptide were respectively 122.6 ± 6.6 nM, 500 pM and 355 ± 39 pM, respectively (Figure 2). Thus, the affinity of IgG RY79-90 for the peptide is 350 times stronger than for BsRNaseY, whereas the K_Ds for the complexes of the antibody with the peptide or Nter-BsRNaseY are in the same order of magnitude, suggesting that the interaction of the antibody with the peptide is close to that with Nter-BsRNaseY. For further studying structurally the antibody/Nter-BsRNaseY interaction, a Fab fragment of IgG RY79-90 was generated by papain digestion and reduction with 2-mercaptoethanol, and purified.

Fab binding promotes ordering of Nter-BsRNaseY

To ensure that the Fab recognizes not only a peptide from BsRNaseY but a longer region within the IDD of the native protein, SEC-SAXS was used to monitor any structural change of Nter-BsRNaseY upon Fab binding. The scattering intensity profiles I(q) *vs* q (where I(q) and q are the scattering intensity and vector) of NterBsRNaseY, Fab 79-90 and the NterBsRNaseY/Fab 79-90 complex show that all samples were monodisperse (Figure S2A). The gyration radius Rg that monitors the extent of macromolecular elongation was calculated from the slope of the linear part of the Guinier plots, ln(I(q)) *vs* q^2 at very small angles (Figure S2B, C & D) [29]. Rg values of 2.79 ±0.05 nm, 5.63 nm ± 0.05 and 4.64 ± 0.15 nm were

estimated for Fab RY79-90, Nter-BsRNaseY and Fab/Nter-BsRNaseY, respectively. The Kratky plot, which is a representation of the scattering intensity I(q). q^2 as a function of the scattering vector q [30], was used for a qualitative analysis of the scattering profiles and to assess the degree of compactness or folding/unfolding of the Fab/Nter-BsRNaseY complex [31]. The normalized Kratky plot of Nter-BsRNaseY had no maximum and a plateau at high q values (Figure 3A, blue curve), which indicated disorder. Moreover, the probability distance distribution function p(r), which is a real-space representation of the distribution of all interatomic distances within the molecule [32], exhibited an asymmetric profile with a maximum at small distances characteristic of an elongated molecule (Figure 3B, blue curve) and the maximum diameter of the molecule D_{max} was determined to be 20 nm. In contrast, the normalized Kratky plot of the Fab alone (Figure 3A, black curve) was bell-shaped with a well-defined maximum, and its p(r) function was rather narrow and symmetric (Figure 3B, black curve), with maximal dimension D_{max} of 9.69 nm, indicating a folded globular structure. Interestingly, the normalized Kratky plot of the Fab/Nter-BsRNaseY mixture (Figure 3A, red curve) is intermediate between the plots of the proteins alone, indicating the formation of a stable complex that is less disordered and more compact than Nter-BsRNaseY. The p(r) distribution (Figure 3B, red curve) is characteristic of a flexible multidomain protein with maximal dimension D_{max} of 13.23 nm rather than a protein with a compact conformation. Altogether, the SAXS data indicated that Nter-BsRNaseY gets a more ordered conformation upon Fab binding.

Overview of the Fab RY79-90 structure

Co-crystallization trials of Fab RY79-90 with Nter-BsRNaseY remained unsuccessful. Yet, the crystal structure of Fab RY79-90 in complex with the inducing peptide was solved at 2.6 Å resolution by molecular replacement using Fab 13F6-1-2 as model (Table 1). It contains two molecules in the asymmetric unit (Figure S3A). The Fab displays a large elbow angle of 138.5° and 133.1° between the Fv fragment and the CL/CH dimer, for molecules LH and AB, respectively, in agreement with the light chain belonging to the λ-class. The peptide-binding pocket of Fab RY79-90 is a shallow cleft, roughly 50 Å by 44 Å wide and 16 Å deep, at the upper part of the VL/VH interface (Figure S3B). The peptide exhibits a standard type 1 β-turn implicating an intramolecular hydrogen bond between the Asn4 O and Gln7 N atoms, and thus adopts a U-shape extended conformation. All CDRs except L2 contribute to the combining site (Figure 4A).

Antibody-peptide interactions

8 out of the 12 residues of the peptide are in contact with the antibody and display clear electron density, whereas electron density was ill-defined for Glu1, Arg3, Gln9, Arg12 that were oriented towards the solvent (Figure 4C). The high specificity and binding affinity of the antibody/peptide interaction is achieved through numerous side-chain interactions. Interestingly, despite its very short length, CDRH3 is involved in several main interactions with the peptide, including a bidentate ionic interaction. Three ionic interactions occur between the side chains of ArgH94 (CDRH3) and Glu5, AspH52 (CDRH2) and Arg3, and GluL34 (CDRL1) and Lys8. Hydrogen bonds are formed between the side chains of ThrH96 (CDRH3) and Glu5, AsnH35 (CDRH1) and Asn4, ThrL93 (CDRL3) and Gln7, ThrL32 (CDRL1) and Glu10. There are additional hydrogen bonds between the side chain of Asn11 and the main chains oxygen atoms of GlyL91 (CDRL3) and oxygen and nitrogen atoms of ThrL32 (CDRL1). Interestingly, the only hydrophobic residue of the peptide, Leu6, is completely buried at the bottom of the combining site. Its side chain is stacked between GlyH95 (CDRH3) and TyL96 (CDRL3). Other peptide side chains are involved in stacking

interactions: Arg2 with TyH53 and ThrH54 of CDRH2, Glu5 with TyrH32 (CDRH1) and Asn11 with TyrL31 (CDRL1).

Comparison of Fab RY79-90 to Fab13F6-1-2 and Fv MW1

Considering the rarity of Vλx usage, it is interesting to know whether the other monoclonal antibodies having Vλx light chains share similar antigen specificity. In addition to the structure of Fab reported in this study, two other Fab crystal structures of antibodies bearing a Vλx light chain have been determined, also in complex with a peptide. The structure of Fab 13F6-1-2, an antibody specific for Ebola viral surface glycoprotein GP was solved in complex with the 11-residue peptide VEQHHRRTDND [24]. Its CDRH3 loop is very long compared to Fab 79-90, which explains that its peptide binds more on the surface of the combining site compared to the peptide of Fab 79-90, which is much more buried (Figure S4A). Antibody MW1 was directed against a poly-glutamine sequence relevant for Huntington's disease. The structure of Fv MW1 solved in complex with the $GQ_{10}G$ peptide, shows that the latter adopts an extended coil-like structure and is less buried in the antibody combining site than the BsRNaseY peptide [26] (Figure S4B).

The specificity of naturally-occurring Vλx-bearing antibodies for myelin basic protein (MBP) could be localized to a particular epitope encompassing residues 25–34 (RHGFKPRHRD) of the MBP molecule [27]. Interestingly, the binding of a recombinant form of Vλx alone to MBP revealed an important contribution of the Vλx light chain. Yet, no crystal structure of Vλx-bearing anti-MBP antibodies is available to further get insight into the molecular basis for such specificity. More structures of Vλx-bearing antibodies are need to know if the light chain isotype could be linked to the nature of the antigen, which up to now was always a peptide.

CONCLUSION

Areas of residual structure in unfolded proteins can be critical to forming the initial site of interaction with a binding partner [33], and delineation of those sites is required for the design of drugs and antibodies that interfere with binding. Moreover, several monoclonal antibodies combining sites have been shown to act as stable molds of *in vivo* conformation of important segments of intrinsically disordered proteins [34]. Here, we used an antibody directed against a peptide located in a conserved region of the IDD of BsRNaseY as a tool to question whether this region could become ordered upon binding a cellular partner. The peptide used to generate RY79-90 lies in a conserved, mostly positively charged sequence region of Nter-RNaseY that could be a binding site for RNA or protein partners contributing to the degradosome complex (Figure 1). Indeed, like in *E. coli* [35, 36], only microdomains of the intrinsically structured region of RNase Y are thought to interact with protein partners to form the *B. subtilis* degradosome complex. If the conserved peptide used to induce RY79-90 belongs to such microdomain, the peptide/Fab RY79-90 interaction could mimic the interaction between the microdomain and a cellular protein partner.

We revealed that the fold adopted by the epitope of the intrinsically disordered region upon binding RY79-90 is a U-shaped extended conformation. Leu6, the unique hydrophobic residue of the peptide, is sandwiched between CDRL3 and CDRH3 at the bottom of the combining site, along with the formation of a β-turn due to hydrogen bonds between Asn4 and Gln7. Interestingly, such a β-turn structure has previously been observed in the complex of a peptide from the intrinsically disordered Tat protein with an Fab fragment, with the burying of a tryptophan residue central to the epitope [37]. In fact, intrinsically disordered

proteins have generally the propensity to expose their few hydrophobic residues, probably to favor interactions with their biological partner.
The concept of conformation specific antibodies that recognize a compact fold of the antigen requires rigidity of the antibody-combining site. We did not solve the structure of Fab 13F6-1-2 alone but, as its CDRH3 loop is extremely short, it is probable that no large conformational changes of the antibody would occur upon peptide binding because these are usually caused by the inherent high flexibility of CDRH3 [38]. In addition, the fact that Fab RY79-90 is able to stabilize Nter-BsRNaseY by ordering it suggests that a similar structural organization of Nter-BsRNaseY might occur *in vivo* upon binding its cellular partners. Therefore, Fab RY79-90 likely acts as a molecular imprint of a specific structural state of Nter-BsRNaseY and the Fab RY79-90/peptide complex could mimic the interaction of a microdomain of Nter-BsRNaseY with one of its cellular partners within the degradosome complex. In future studies, Fab13F6-1-2 could be used as a tool to induce a specific fold to disordered Nter-BsRNaseY.

MATERIALS & METHODS

Production, purification and characterization of the Fab fragment

The sequence of *B. subtilis* RNase Y was analyzed by the Abmart Company (Shanghai, China) using the SEAL™ algorithm designed for identifying protein fragments most susceptible to yield monoclonal antibodies. Among the six antigen peptides identified for immunization optimization, peptide 79-90 (ERRNELQKQENR) was chosen for monoclonal antibody production in mice ascitic fluid. The IgG (IgG2b, λ) was produced and purified on a Protein A affinity column (Abmart) and stored in 50 mM Tris, 100 mM Glycine pH 9, 0.5% ProClin 300 (Sigma-Aldrich). The Fab was generated by papain digestion of the antibody at 37°C in 20 mM sodium phosphate, 274 mM NaCl, 5.4 mM KCl pH 7.4, 1.25 mM EDTA, 1.25 mM 2-mercaptoethanol using a 0.5 % papain to antibody ratio (w/w) for 5 hours. Undigested IgG and Fc fragment were removed by HiTrapQ anion exchange chromatography. The Fab purity was analyzed on an HPLC BioSec-3 column, dialyzed against 200 mM NaCl, 5% glycerol and concentrated to 6.5 mg/ml.

Sequence determination

Sequencing of the light and heavy chains was performed by ProMab Biotechnologies Inc, Richmond, CA (USA). Total RNA was extracted from 100,000 hybridoma cells and, using PCR methods and a mixture of constant region sequences, the scFv region was amplified for cloning and sequencing. The sequence of the variable light chain could not be determined by this method, probably because of its very rare Vλx isotype. Molecular replacement using various antibodies of known structure identified antibody Fab13F6-1-2 (PDB code 2QHR) as the best model for both the Fv and the CL-CH1 dimer [24]. Mass spectrometry analysis of the lower molecular mass protein band of Fab on a SDS-PAGE gel under reducing conditions confirmed that the amino acid sequence of Fab light chain shares only 3 amino acid sequence differences with Fab13F6-1-2.

Expression and purification of Nter-BsRNaseY and BsRNaseY

The coding sequence of the BsRNaseY gene starting with amino acid K26 (thus lacking the N-terminal transmembrane domain) was PCR-amplified from chromosomal DNA of strain SSB1002, a wild-type *B. subtilis* laboratory stock strain derived from strain 168, using oligonucleotides HP1182 (5' TGATTCACTCATATGAAAACCATTGCCGAAGCGAAAATTGCG) and HP116 (5' AGCTAGGATCCTTAGTGATGATGATGATGGTGTTTTGCATACTCTACGGCTCGAGTC). The 1.5 kb PCR fragment containing the 6xHis-tag sequence at the 3' end was inserted as an NdeI-BamHI fragment into the T7 expression vector pKYB1 plasmid (New England Biolabs) cleaved with NdeI and BglII. Recombinant plasmids from *E. coli* JM109 were then transformed into *E. coli* BL21(*DE3*) star codon$^+$ competent cells. One colony was used to inoculate 200 ml of Luria Broth medium supplemented with kanamycine (25 µg/ml) and chloramphenicol (30 µg/ml).

For the Nter-BsRNaseY construct, the sequence of the BsRNaseY gene corresponding to amino acid V24 to N192 appended with a C-terminal hexahistidine tag was PCR-amplified from chromosomal DNA of strain SSB1002, a wild-type *B. subtilis* laboratory stock strain derived from strain 168, using the Sequence Ligation Independent Cloning method [39]. The PCR fragment containing the 6xHis-tag sequence at the 3' end was inserted as an NdeI-BamHI fragment into the T7 expression vector pET22b plasmid (New England Biolabs). The plasmid was then transformed into *E. coli* BL21(*DE3*) star codon$^+$ competent cells. One colony was used to inoculate 200 ml of Luria Broth medium supplemented with ampicillin (50 µg/ml) and chloramphenicol (30 µg/ml).

The precultures grown overnight at 37°C at 180 rpm (0.5%) were used to inoculate 2 liters of LB medium supplemented with the same antibiotics. Cultures were grown at 37°C to an OD_{600} of 0.8. Expression was then induced at 23°C by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM and incubation was extended overnight at 23°C. After centrifugation for 10 min at 5000 rpm and 4°C, cells were stored at −80°C.

Puification of Nter-BsRNaseY and BsRNase Y

Cells were resuspended in 1.5 ml/g of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol (buffer A) containing 20 µl/ml cOmplete™ Protease Inhibitor Cocktail (Roche), 10 mM β-mercaptoethanol, 2 µg/ml of RNase A (Thermofischer), 1 mg/ml lysozyme, 0.25% Tween 20, and disrupted by sonication. Cells debris were removed by ultracentrifugation at 100,000g for 1h at 4°C. The supernatant was loaded on a 15 mL Ni-NTA Superflow column (Qiagen) equilibrated in buffer A containing 20 mM imidazole, 0.25% Tween 20. The column was washed with 70 mM imidazole and eluted with 200 mM imidazole in buffer A. Fractions containing the proteins were then loaded at 2.5 mg/ml (BsRNaseY) or 4.5 mg/ml (Nter-BsRNaseY) on a Superdex 200 26/60 PG column (GE Healthcare) in buffer A. Protein elution was followed by the optical density at 230 nm because of the rarity or absence of aromatic residues in the BsRNaseY and NterBsRNaseY sequences. After concentration to 5 mg/ml, using Amicon concentrators (30 kD cutoff, Milllipore), proteins were aliquoted, frozen in liquid nitrogen and stored at -80°C.

Crystallization of the Fab79-90/peptide complex

Fab79-90 (5 mg/ml) was mixed with the peptide in a 1:1.2 ratio. Crystals were grown in 16 % PEG 4000, 0.2 M $MgCl_2$, 0.1 M Tris pH 8.5, then flash cooled in a nitrogen stream at 100K in the same solution containing 20% glycerol.

X-Ray Data Collection and Structure Determination

Diffraction data were recorded on PX1 beamline at the SOLEIL synchrotron. Data were processed with XDS (Table 1). The structure was solved by molecular replacement with PHASER using as models the Fv and the CL-CH1 dimer of Fab 13F6-1-2 (PDB code 2QHR), which belongs to the IgG2a, λ class. The atomic model was refined alternating cycles of model reconstruction with *COOT* and refinement with *BUSTER*. The final refinement statistics are given in Table 1. Figures were drawn with *PYMOL*. ThrL51 in CDRL2, which lies in the disallowed region of the Ramachandran plot, belongs to a γ turn, as commonly observed in all antibody structures.

Determination of the dissociation constants for the Fab/BsRNaseY, Fab/Nter-BsRNaseY, and Fab/peptide complexes

The K_ds were estimated by a competitive ELISA assay, as described previously [28]. In this method, the antibody at constant concentration is first incubated in solution with the antigen at various concentrations until the equilibrium is reached. Then, the amount of unbound antibody is measured by a classical indirect ELISA test.

First, calibration curves were set up in duplicate to ensure that the antibody concentration used is in the linear range of the ELISA response and only a small fraction of the total free antibody in solution is retained on the plate. To this end, 96-well plates were coated with 0.5-4 nM BsRNase Y, 5-40 nM Nter-BsRNaseY or 2-16 nM peptide in PBS buffer pH 7.4 and incubated overnight at 4°C. Then, the plates were saturated with 250 µL of 3% BSA, 3% Milk in PBS pH 7.4, 0.05% Tween 20 at 37°C for one hour. After three washes with 20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 0.05% Tween 20, antibody at various concentrations (0.5 to 60 nM) was dispensed into the wells. After incubation for one hour at 25°C, the wells were washed three times and 100 µL of peroxidase-labeled anti-mouse IgG antibody (Sigma Aldrich®) at a dilution of 1:4000 in 20 mM Hepes pH 7.5, 150 mM NaCl, 10% glycerol and 10 mg/mL BSA was added to each well. After 1 hour incubation at 25°C, plates were washed five times and 100 µL of tetramethylbenzidine (Sigma Aldrich®) was added to each well. After incubation for 10 min in the dark, the reaction was stopped by adding 1 M sulfuric acid, and absorbance at 450 nm, which should be proportional to the concentration of free antibody, was measured.

Second, the concentration of free antibody in a series of solutions preincubated with various concentrations of the different antigens was determined in triplicate. The binding equilibria were carried out in 20 mM Hepes 7.5, 150 mM NaCl, 10% glycerol, 10 mg/mL BSA overnight at 25°C with the antibody at a fixed concentration (7.6 nM, 1.5 nM and 3.8 nM for the complexes with BsRNaseY, Nter-BsRNaseY or the peptide, respectively), and 31.25, 62.5, 125, 250, 500, 1000 nM BsRNaseY, 0.025, 0.05, 0.1, 0.2, 0.4 et 0.8 nM Nter-BsRNaseY or 0.031, 0.063, 0.125, 0.250, 0.5, 1 nM peptide. Antibody–antigen mixtures were transferred into a 96-well plate previously coated overnight with 2 nM BsRNase Y, 40 nM Nter-BsRNaseY or 16 nM peptide in PBS buffer pH 7.4, saturated and washed as described above. After incubation at 25°C for 30 min for the antibody–BsRNaseY complex or 10 min for the antibody–Nter-BsRNaseY complex and antibody–peptide complexes, the concentration of free antibody was determined by the indirect ELISA test described above. The results were plotted using a Scatchard plot as described [28], the slope being equal to $-1/K_D$ and the y-axis intercept point to $1/K_D$.

SAXS data collection and analysis

SAXS data were collected on beam line SWING of the SOLEIL synchrotron. The Fab (in 200 mM NaCl, 5% glycerol buffer) and Nter-BsRNAseY (20 mg/ml in 20 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol) were injected at 20 mg/ml onto an BioSEC3 column (Agilent) equilibrated in 20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol. The complex sample (2Fab:Nter-BsRNaseY ratio) was injected after incubation for 30 min. Scattering data were recorded on the eluted peak and processed using FOXTROT [40]. The gyration radius was calculated by the Guinier approximation using PRIMUS from the ATSAS suite [41]. The GNOM program from the ATSAS suite was used to calculate the pair distance distribution function P(r) and D_{max}. *Ab initio* shape reconstruction of BsRNAseY dimer was performed by 20 independent DAMMIF [42] runs averaged using DAMAVER [43] and further refined with DAMMIN [44].

Acknowledgements.

This work was supported by the French State Program 'Investissements d'Avenir' (Grants "LABEX DYNAMO", ANR-11-LABX-0011), and has benefited from the facilities and expertise of the Macromolecular Interaction Platform of I2BC. We thank SOLEIL for provision of synchrotron radiation facilities, Pierre Legrand for assistance in using beamline PROXIMA 1.

References

1. Laalami, S., Zig, L. & Putzer, H. (2014). Initiation of mRNA decay in bacteria. *Cell Mol Life Sci* 71, 1799-1828.
2. Hui, M. P., Foley, P. L. & Belasco, J. G. (2014). Messenger RNA degradation in bacterial cells. *Annu Rev Genet* 48, 537-559.
3. Mohanty, B. K. & Kushner, S. R. (2016). Regulation of mRNA Decay in Bacteria. *Annu Rev Microbiol* 70, 25-44.
4. Mackie, G. A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* 11, 45-57.
5. Bechhofer, D. H. (2011). *Bacillus subtilis* mRNA decay: new parts in the toolkit. *Wiley Interdiscip Rev RNA* 2, 387-394.
6. Lehnik-Habrink, M., Lewis, R. J., Mader, U. & Stulke, J. (2012). RNA degradation in *Bacillus subtilis:* an interplay of essential endo- and exoribonucleases. *Mol Microbiol* 84, 1005-1017.
7. Durand, S., Tomasini, A., Braun, F., Condon, C. & Romby, P. (2015). sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiol Rev* 39, 316-330.
8. Shahbabian, K., Jamalli, A., Zig, L. & Putzer, H. (2009). RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J* 28, 3523-3533.
9. Figaro, S., Durand, S., Gilet, L., Cayet, N., Sachse, M. & Condon, C. (2013). *Bacillus subtilis* mutants with knockouts of the genes encoding ribonucleases RNase Y and RNase J1 are viable, with major defects in cell morphology, sporulation, and competence. *J Bacteriol* 195, 2340-2348.
10. Yao, S. & Bechhofer, D. H. (2010). Initiation of decay of *Bacillus subtilis* rpsO mRNA by endoribonuclease RNase Y. *J Bacteriol* 192, 3279-3286.
11. Bruscella, P., Shahbabian, K., Laalami, S. & Putzer, H. (2011). RNase Y is responsible for uncoupling the expression of translation factor IF3 from that of the ribosomal proteins L35 and L20 in *Bacillus subtilis*. *Mol Microbiol* 81, 1526-1541.
12. Lehnik-Habrink, M., Newman, J., Rothe, F. M., Solovyova, A. S., Rodrigues, C., Herzberg, C., *et al.* (2011). RNase Y in *Bacillus subtilis*: a Natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J Bacteriol* 193, 5431-5441.

13. Durand, S., Gilet, L., Bessieres, P., Nicolas, P. & Condon, C. (2012). Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs. *PLoS Genet* 8, e1002520.
14. Laalami, S., Bessieres, P., Rocca, A., Zig, L., Nicolas, P. & Putzer, H. (2013). *Bacillus subtilis* RNase Y activity in vivo analysed by tiling microarrays. *PLoS One* 8, e54062.
15. Hunt, A., Rawlins, J. P., Thomaides, H. B. & Errington, J. (2006). Functional analysis of 11 putative essential genes in *Bacillus subtilis*. *Microbiology* 152, 2895-2907.
16. Grishin, N. V. (2001). KH domain: one motif, two folds. *Nucleic Acids Res* 29, 638-643.
17. Aravind, L. & Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci* 23, 469-472.
18. Tompa, P. (2012). On the supertertiary structure of proteins. *Nat Chem Biol* 8, 597-600.
19. Liu, Z. & Huang, Y. (2014). Advantages of proteins being disordered. *Protein Sci* 23, 539-550.
20. Varadi, M., Vranken, W., Guharoy, M. & Tompa, P. (2015). Computational approaches for inferring the functions of intrinsically disordered proteins. *Front Mol Biosci* 2.
21. Basu, S. & Bahadur, R. P. (2016). A structural perspective of RNA recognition by intrinsically disordered proteins. *Cell Mol Life Sci* 73, 4075-4084.
22. van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., *et al.* (2014). Classification of intrinsically disordered regions and proteins. *Chem Rev* 114, 6589-6631.
23. Sanchez, P., Marche, P. N., Rueff-Juy, D. & Cazenave, P. A. (1990). Mouse V lambda x gene sequence generates no junctional diversity and is conserved in mammalian species. *J Immunol* 144, 2816-2820.
24. Lee, J. E., Kuehne, A., Abelson, D. M., Fusco, M. L., Hart, M. K. & Saphire, E. O. (2008). Complex of a protective antibody with its Ebola virus GP peptide epitope: unusual features of a V lambda x light chain. *J Mol Biol* 375, 202-216.
25. Maier, C. C., Galin, F. S., Jarpe, M. A., Jackson, P., Krishna, N. R., Gautam, A. M., *et al.* (1994). A V lambda x-bearing monoclonal antibody with similar specificity and sequence to encephalitogenic T cell receptors. *J Immunol* 153, 1132-1140.
26. Li, P., Huey-Tubman, K. E., Gao, T., Li, X., West, A. P., Jr., Bennett, M. J., *et al.* (2007). The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model. *Nat Struct Mol Biol* 14, 381-387.
27. Galin, F. S., Maier, C. C., Zhou, S. R., Whitaker, J. N. & Blalock, J. E. (1996). Murine V lambda x and V lambda x-containing antibodies bind human myelin basic protein. *J Clin Invest* 97, 486-492.
28. Friguet, B., Chaffotte, A. F., Djavadi-Ohaniance, L. & Goldberg, M. E. (1985). Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. *J Immunol Methods* 77, 305-319.
29. Koch, M. H., Vachette, P. & Svergun, D. I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* 36, 147-227.
30. Kikhney, A. G. & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 589, 2570-2577.
31. Receveur-Brechot, V., Bourhis, J. M., Uversky, V. N., Canard, B. & Longhi, S. (2006). Assessing protein disorder and induced folding. *Proteins* 62, 24-45.
32. Svergun, D. I. & Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 66, 1735-1782.
33. Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 11, 739-756.
34. Skrabana, R., Dvorsky, R., Sevcik, J. & Novak, M. (2010). Monoclonal antibody MN423 as a stable mold facilitates structure determination of disordered tau protein. *J Struct Biol* 171, 74-81.
35. Callaghan, A. J., Aurikko, J. P., Ilag, L. L., Gunter Grossmann, J., Chandran, V., Kuhnel, K., *et al.* (2004). Studies of the RNA degradosome-organizing domain of the *Escherichia coli* ribonuclease RNase E. *J Mol Biol* 340, 965-979.

36. Marcaida, M. J., DePristo, M. A., Chandran, V., Carpousis, A. J. & Luisi, B. F. (2006). The RNA degradosome: life in the fast lane of adaptive molecular evolution. *Trends Biochem Sci* 31, 359-365.
37. Serrière, J., Dugua, J. M., Bossus, M., Verrier, B., Haser, R., Gouet, P., *et al.* (2011). Fab'-induced folding of antigenic N-terminal peptides from intrinsically disordered HIV-1 Tat revealed by X-ray crystallography. *J Mol Biol* 405, 33-42.
38. Sela-Culang, I., Alon, S. & Ofran, Y. (2012). A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *J Immunol* 189, 4890-4899.
39. Li, M. Z. & Elledge, S. J. (2012). SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol Biol* 852, 51-59.
40. David, G. & Pérez, J. (2009). Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J Appl Cryst* 42, 892-900.
41. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J Appl Cryst* 25, 495-503.
42. Franke, D. & Svergun, D. I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst* 42, 342-346.
43. Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of ab-initio shape determination in small-angle scattering. *J Appl Cryst* 42, 342-346.
44. Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76, 2879-2886.
45. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539.
46. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305-308.

Figure legends

Figure 1: Selection of a monoclonal antibody induced against a 12-residue peptide of Nter-BsRNase Y. A Alignment of sequences of the N-terminal domain of RNases Y from *B. subtilis, Staphylococcus aureus, Streptococcus pyogenes, Listeria monocytogenes, Clostridium perfringens, Lactoccus lactis, Borrelia burgdorferi, Leadbetterella byssophila* was performed with Clustal Omega [45] and rendered with *ESPript* [46]. **B** Denaturing immunoblotting gel **1 :** 10 µg crude extract; **2-4 :** BsRNaseY after Nickel chromatography :10 ng (**2**), 100 ng (**3**) and 1µg (**4**). **C** Formation of the BsRNaseY/Fab RY79-90 complex. BsRNaseY was incubated with Fab 79-90 at different ratios and deposited on a 4% native retardation gel. **1** : Fab alone (2 µg); **2** : BsRNaseY dimer alone (2 µg) **3** :Fab/BsRNaseY mixture. 0.1 (**3**), 0.25 (**4**), 0.5 (**5**), 0.75 (**6**), 1 (**7**), 2.5 (**8**), 5 (**9**), 7.5 (**10**), 10 (**11**), 12.5 (**12**) and 15µM (**13**) Fab79-90.

Figure 2: **Determination of the dissociation constants (K_D) for the complexes of IgG RY79-90 with BsRNaseY (A) or the inducing peptide (B)**. The Scatchard plot representation ($Y= 1/K_D (1 - X)$ with X and Y defined below) was used to fit the data and determine the dissociation constant for each complex. $X = (A_0 – A)/A_0$ and $Y = ((A_0-A)/A_0)/[C – C_{Ac}*((A_0-A)/A_0)]$. A_0 is the measured absorbance for antibody alone, A is the absorbance measured for each mixture, C is the initial antigen concentration and C_{Ab} is the fixed antibody concentration.

Figure 3: SAXS analysis of the Nter-BsRNAseY complex.
A Normalized Kratky plots of Fab RY79-90, Nter-BsRNAseY, and their complex shown in black, blue and red, respectively. **B** Pair distance distribution function p(r) of Fab RY79-90, Nter-BsRNAseY, and their complex shown in black, blue and red, respectively.

Figure 4: **Recognition of the inducing peptide by Fab RY79-90.**
(**A)** Location of the peptide at the VL/VH interface in the active site of Fab RY79-90. The ligand is in yellow and CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, and CDRH3 are colored violet-purple, pink, magenta, orange, red and brown, respectively. **B** Details of the interactions between Fab RY79-90 and the inducing peptide. The residues that directly interact with the peptide are represented as sticks and colored according to the CDR. H-bonding interactions are indicated as dashes. **C** Stereoview of Fo-Fc electron density map, calculated by omitting the peptides and a 1.8-Å spherical region around them, is contoured at the level of 2.0 σ and superimposed on the active-site structure.

Table 1: Data collection and refinement statistics[a]

	Fab79-90/peptide complex
Data collection	
Wavelength (Å)	0.9793
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	98.68, 130.45, 72.05
α, β, γ (°)	90, 90, 90
Resolution (Å) [a]	50-2.60 (2.75-2.60)
R_{merge} [a]	0.096 (4.76)
R_{pim} [a]	0.030 (1.44)
$I/\sigma(I)$ [a]	13.2 (0.9)
$CC_{1/2}$ [a]	0.986 (0.463)
Completeness (%)[a]	99.8 (99.7)
Redundancy [a]	8.7 (8.6)
Refinement	
Resolution (Å)	46.1-2.6 (2.69-2.6)
No. reflections	
R_{work} / R_{free}	0.222/0.257 (0.232/0.302)
No. atoms	
Protein	6512
Peptide	222
Water	225
B factors (Å^2)	
Protein	84.9
Peptide	89.2
Water	71.7
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.19

[a] Values in parentheses are for highest-resolution shell.

Figure 1:

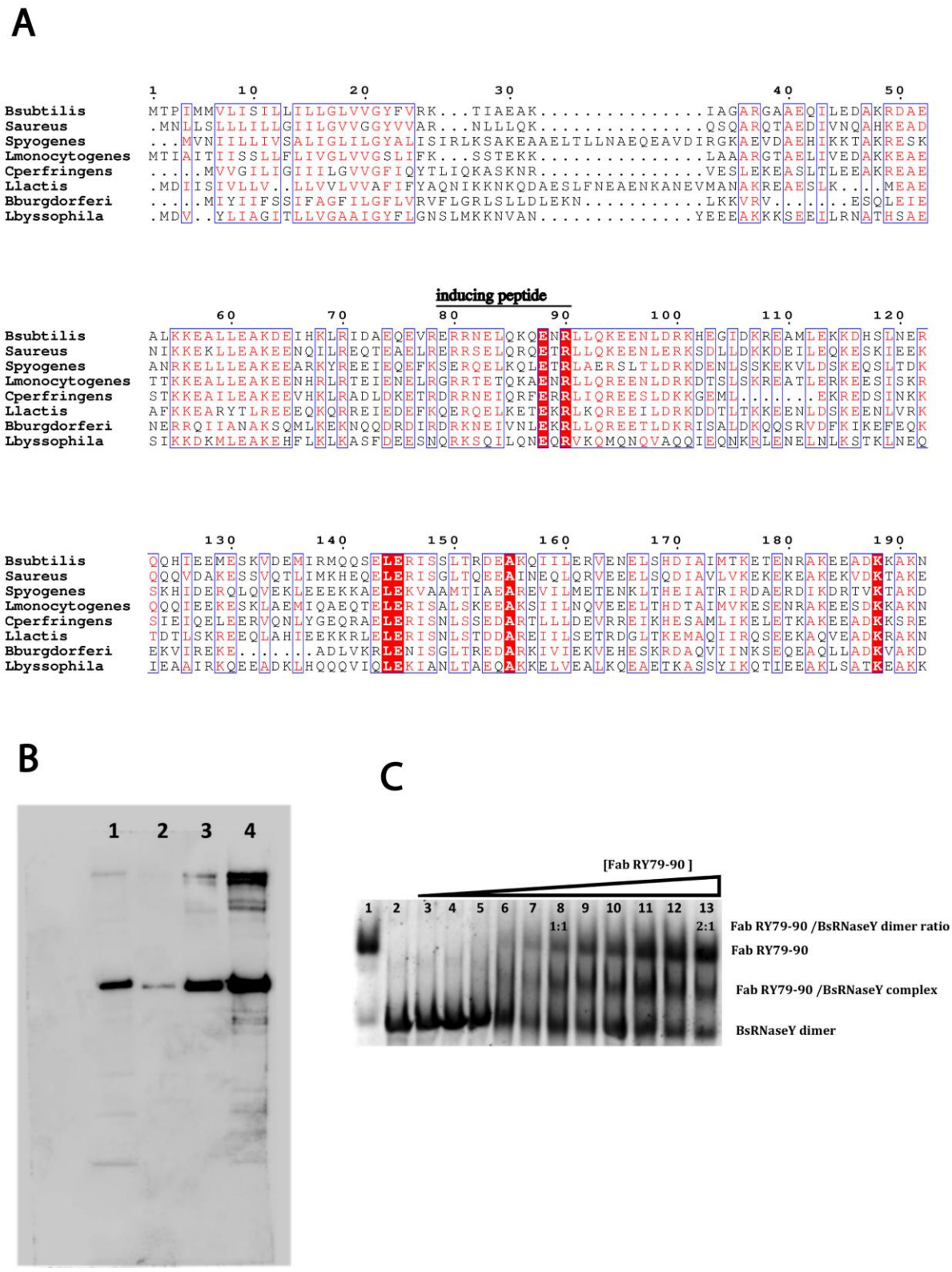

A
inducing peptide
Bsubtilis
Saureus
Spyogenes
Lmonocytogenes
Cperfringens
Llactis
Bburgdorferi
Lbyssophila
B
1 2 3 4
C
[Fab RY79-90]
1 2 3 4 5 6 7 8 9 10 11 12 13
1:1
2:1
Fab RY79-90 /BsRNaseY dimer ratio
Fab RY79-90
Fab RY79-90 /BsRNaseY complex
BsRNaseY dimer

Figure 2

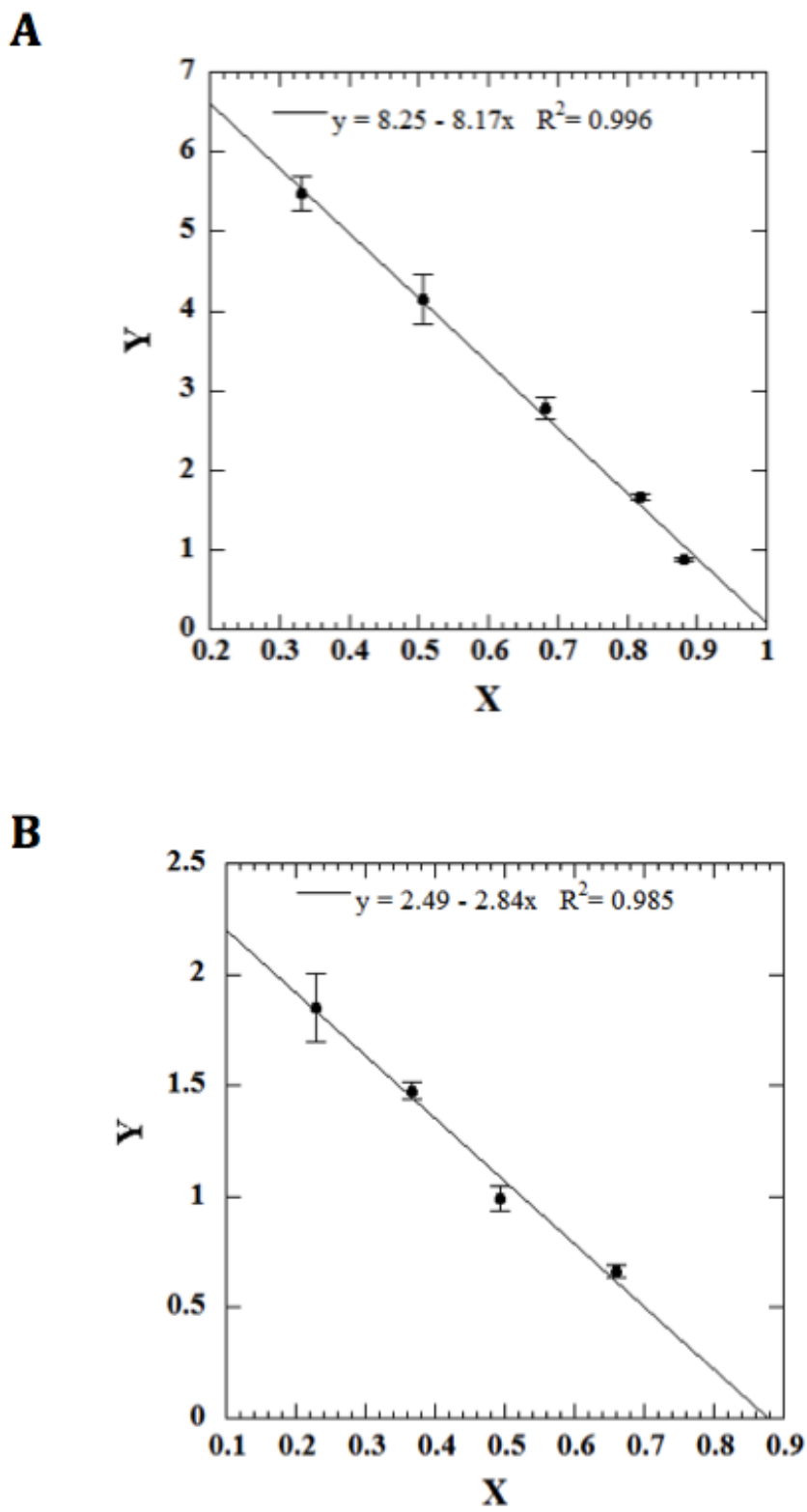

A
y = 8.25 - 8.17x R2= 0.996
Y
X
0
1
2
3
4
5
6
7
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
B
y = 2.49 - 2.84x R2= 0.985
Y
X
0
0.5
1
1.5
2
2.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

Figure 3

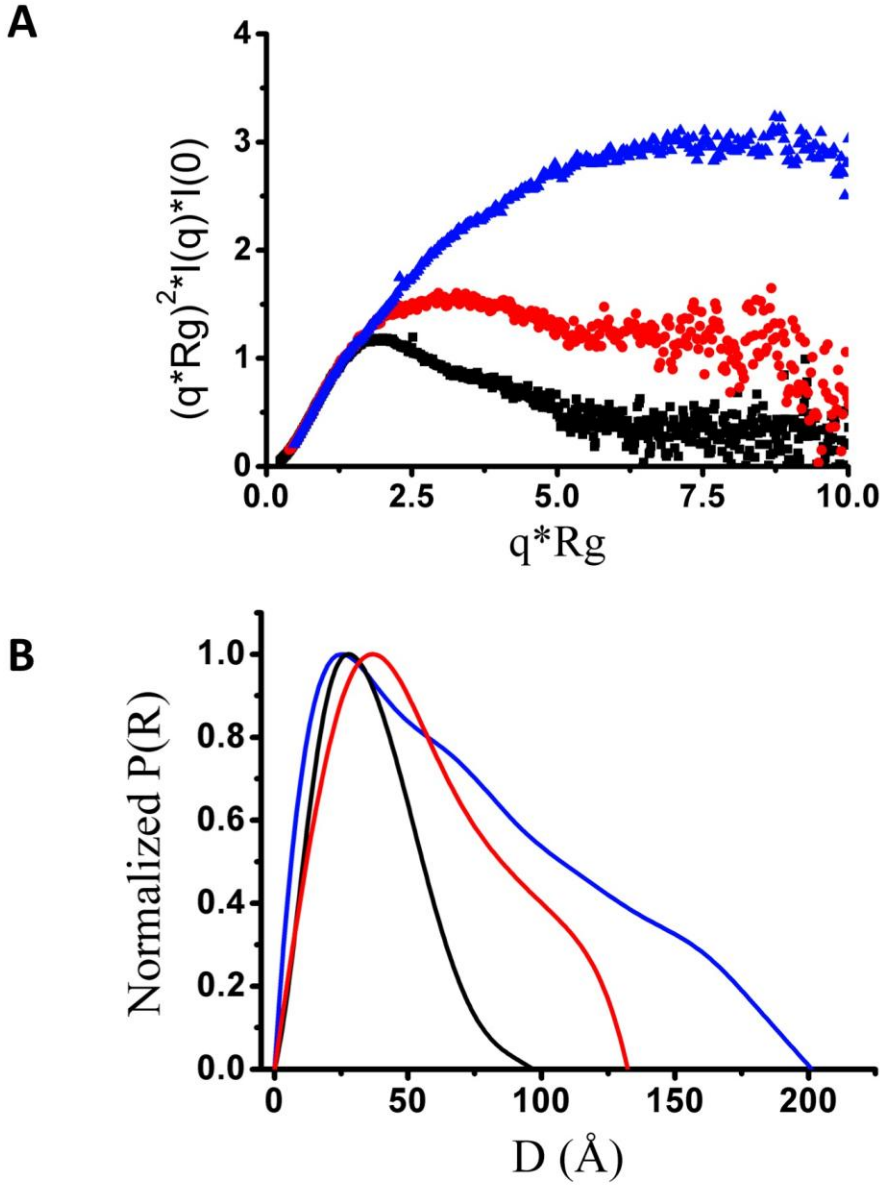
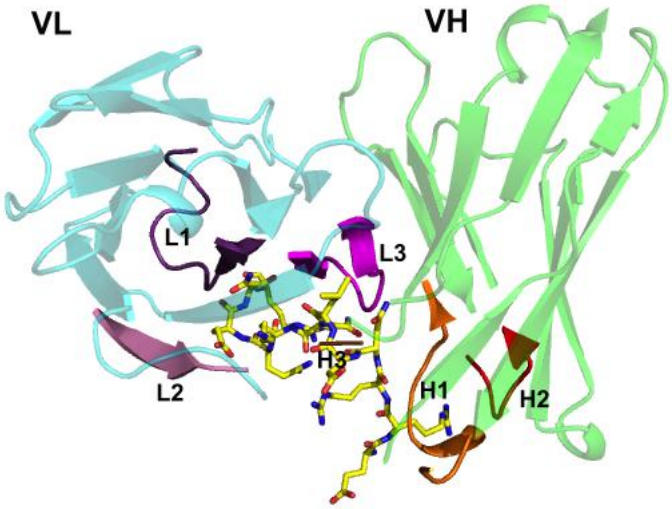
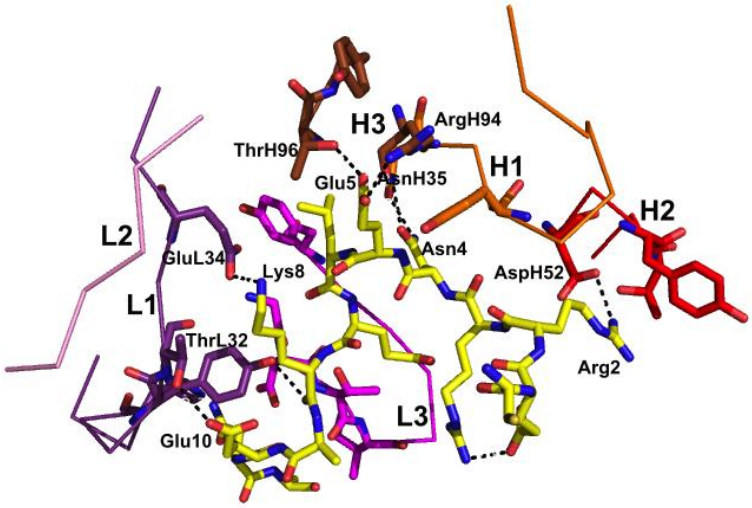


Figure 4

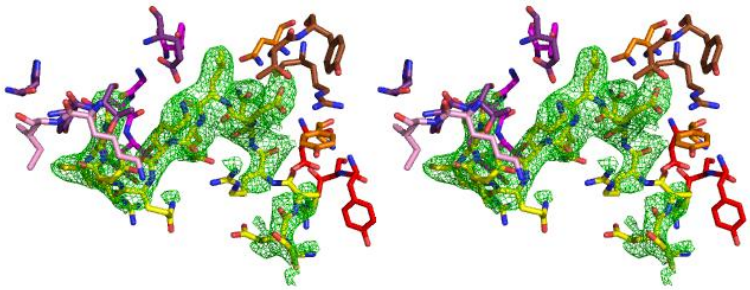
A


VL
VH
L1
L3
H3
L2
H1
H2

B


H3
ArgH94
ThrH96
H1
Glu5
AsnH35
H2
L2
GluL34
Asn4
Lys8
AspH52
L1
ThrL32
Arg2
L3
Glu10

C



Supplementary Figures

Figure S1: Amino acid sequences of the VL and VH domains of Fab79-90 and Fab13F6-1-2 and their likely corresponding germline genes. Dots denote sequence identity to Fab79-90. H denotes the heavy chain and L the light chain. Sequence numbering and CDR definitions (indicated in bold) follow Kabat nomenclature. The likely germline sequences are indicated below the sequence of the antibodies.

```
                                          CDRL1
              1       10        20     27a 30        40        50
LRY79-90      QLVLTQSSSASFSLGASAKLTCTLSSQHSTYTIEWYQQQPLKPPKYVMEL
L13F6-1-2     ........................R..................R.....
Vλx       GSFS.................................................

             CDRL2                                        CDRL3
             54abcd    60        70        80        90   95ab
LRY79-90   KKDGSHSTGDGIPDRFSGSSSGADRYLSISNIQEEDEAIYICGVGDTIKEQ
L13F6-1-2  ................................P.................
Vλx        ................................P.................

           cd  100   106a
LRY79-90   FVYVFGGGTKVTVL
L13F6-1-2  ..............
   Vλx     ..
   Jλ2       ............


                                          CDRH1
           1       10        20        30        40        5052a
HRY79-90   QVQLQESGPELKKPGETVKISCKVSGYPFTNYGMNWVKQAPRKVLKWMGWIDT
H13F6-1-2  E..VV...GG.V...GSL.L..AA..FA.SS.D.S..R.T.E.R.E.VAY.SR
IGHV156    .I..VQ.................A...T.............G.G.......N.

           CDRH2                                        CDRH3

                 60        70        8082abc      90       100
HRY79-90   YTGDPTYADDFKGRFAFSLDTSASTAYLQINNLKNEDTATYFCARGT-----
H13F6-1-2  GG.YTY.P.TV....TI.R.NAKN.L...MSS..S....M.Y.S.HIYYGSS
IGHV156    ...E...............E.........................

                   110
HRY79-90   ------YWGQGTLVTVSS
H13F6-1-2  HYYAMD......S.....
JH3           WFA............
```

Figure S2 : Analysis of Fab RY79-90, Nter-BsRNaseY and their complex by SEC-SAXS. A X-ray scattering curves of Fab alone, Nter-BsRNaseY and their complex, in black, blue and red, respectively. **B** Guinier plot of Fab alone. **C** Guinier plot of Nter-BsRNaseY alone. **D** Guinier plot of the Fab/Nter-BsRNaseY complex.

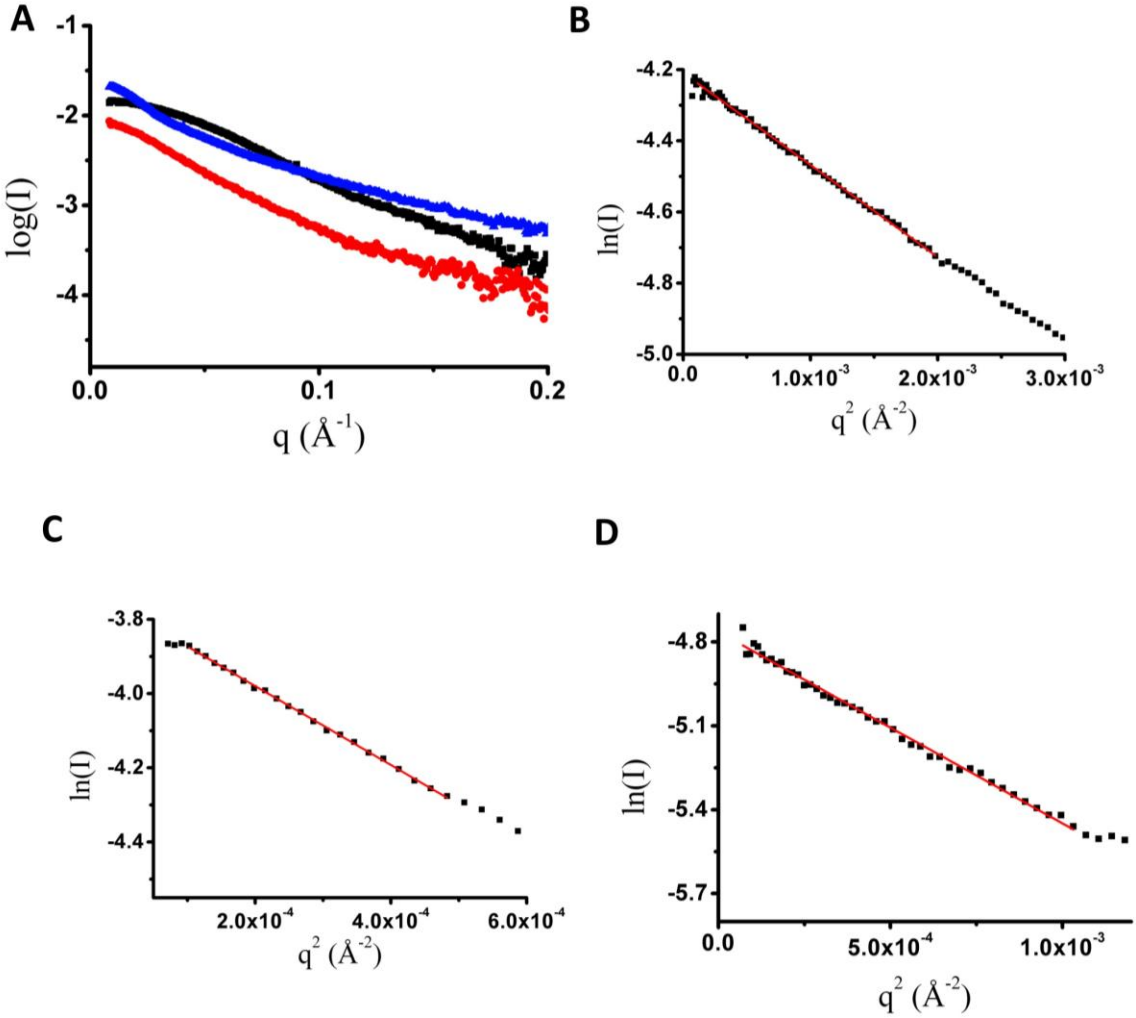
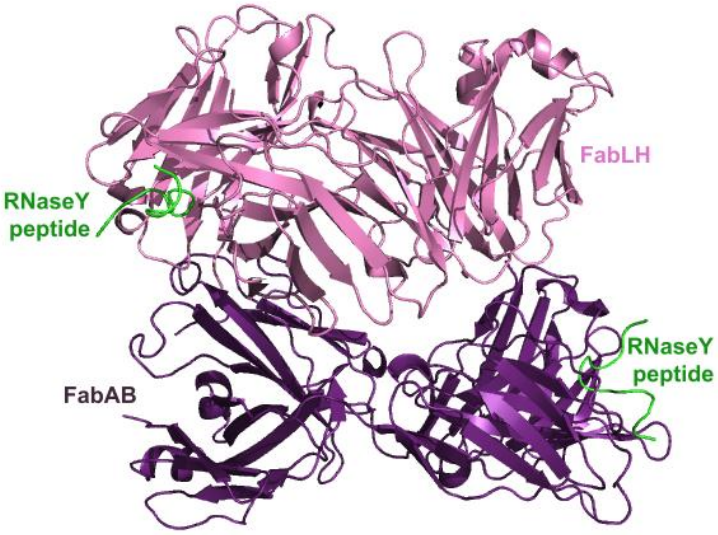


Figure S3: **Overall view of the Fab RY79-90/peptide complex**. **A** Overall view of the two Fab molecules in the asymmetric unit (colored pink and magenta) with the peptides drawn as green coils. **B** Overall view of the peptide in the combining site of Fab RY79-90. The molecular surfaces of the VL and VH chains are shown in cyan and green, respectively.

A



B

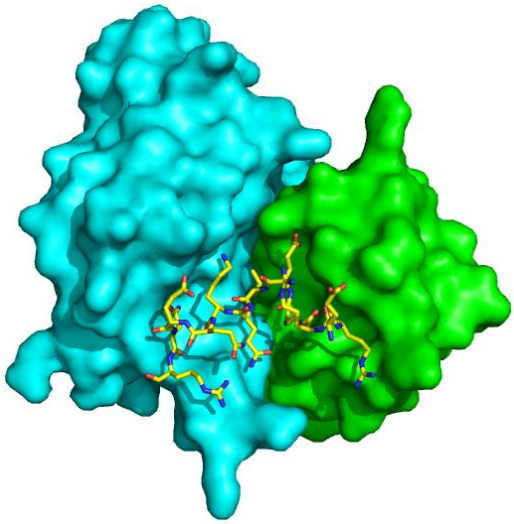
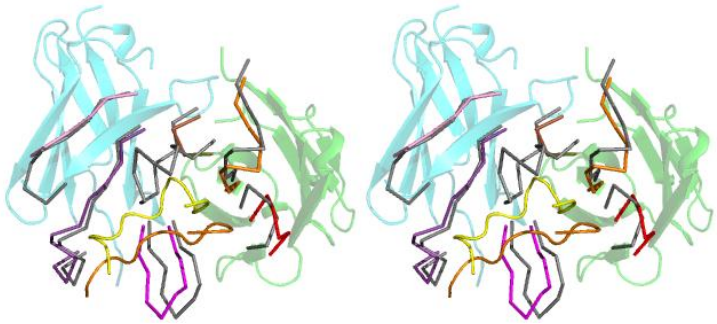
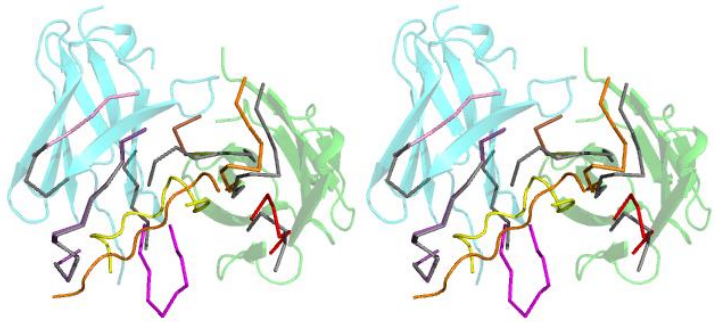


Figure S4: **Comparison of structures of antibodies with vλx light chain in complex with peptide**. **A** Stereo view of the superposition of the complexes of Fab RY79-90 and Fab 13F6-1-2 with their respective peptides. The colors for Fab RY79-90 are as in Figure 4 and the CDR loops of Fab 13F6-1-2 are in grey, with the peptide shown as an orange cartoon. **B** Stereo view of the superposition of the complexes of Fab RY79-90 and Fv MW1 (PDB code 2OTU) with their respective peptides. The colors for Fab RY79-90 are as in Figure 4 and the CDR loops of v MW1 are in grey, with the peptide shown as an orange cartoon.

A



B



Article 2

The N-terminal domain of *B. subtilis* RNase Y is a molten globule

Pierre Hardouin[1], Charles Bou Nader[1], Nadine Assrir[2], Carine van Heijenoort[2], Christophe Velours[3], Béatrice Golinelli-Pimpaneau[1*]

[1] Laboratoire de Chimie des Processus Biologiques, UMR 8229 CNRS, Collège de France, Université Pierre et Marie Curie, 11 Place Marcelin Berthelot, 75231 Paris cedex 05, France

[2] Structural Chemistry and Biology Team, Institut de Chimie des Substances Naturelles, CNRS UPR 2301, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France

[3] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, Avenue de la Terrasse, 91198 Gif-sur-Yvette cedex, France

**Corresponding author:*
beatrice.golinelli@college-de-france.fr

Although RNase Y acts as the key enzyme initiating mRNA decay in many Gram positive bacteria, its three-dimensional structure remains unknown. Bioinformatic analysis had predicted that its N-terminal domain (NTD) would be intrinsically disordered. Here we show that the NTD can be produced as a stand-alone protein and provide experimental evidence by SAXS, CD and NMR that it undergoes conformational changes between a monomeric and dimeric forms that can be described as molten globules with some regions organized in α-helices.

Keywords RNase Y, intrinsically disordered domain, molten globule, SAXS, NMR, CD

Introduction

In bacteria, gene expression in response to different environmental conditions is controlled by the fluctuation of transcript abundance [1-3]. The rapid mRNA turnover rate in bacteria (average half-life of several minutes) relies on the synthesis and degradation rates of each mRNA. RNases, responsible for RNA degradation, can be either very specific for their RNA substrate, or involved in bulk RNA turnover. The major RNases involved in bulk RNA turnover generally form multi-protein complexes for efficient and regulated degradation that are called RNA degradosomes [4]. Whereas most RNA synthesis machinery is similar in Gram-positive and Gram-negative bacteria, the machinery for bulk RNA degradation is quite different. The well-studied hydrolytic endoribonuclease RNase E plays a major role in mRNA degradation and RNA processing in *Escherichia coli* [5], whereas RNase Y is one of the key enzymes directing RNA metabolism in *Bacillus subtilis* [6] as well as other low-GC Gram positive species [7], including many pathogens [8-10]. In particular, RNase Y was shown to be necessary for the full virulence of medically important Gram-positive bacterial pathogens [8,11]. Studying the function and structure of RNase Y is not only crucial to understand RNA metabolism in the many eubacteria that, unlike *E. coli*, do not rely on RNase E-based strategies of RNA degradation/maturation, but also to consider whether RNase Y could be targeted for the development of new antimicrobial agents in the future since RNase Y orthologs do not exist in eukaryotic cells.

Whereas RNase Y and RNase E share many functions [12], they display only low sequence similarity. RNase E (118 kDa) is composed of two major domains: the N-terminal domain that bears the endoribonuclease activity [13-15] and the C-terminal scaffolding domain that organizes the RNA degradosome [4,16-18]. This latter domain is intrinsically unstructured and unlikely to be extensively folded within the degradosome except for a few microdomains mostly involved in interactions with other components in the degradosome [19,20]. Indeed, although most of the C-terminal domain sequence is not much conserved among RNase E orthologs, some parts are more conserved than others: the first microdomain, called the membrane targeting sequence, forms an amphipathic helix that interacts with the membrane [21]. Other conserved microdomains ranging from 20 to 70 residues in size are binding sites for RNA substrates and other protein components of the degradosome [18].

In *E. coli*, RNase E forms the core of the degradosome, which in its minimal version includes polynucleotide phosphorylase (PNPase), the ATP dependent DEAD-box RhlB RNA helicase (RhlB), and the glycolytic enzyme enolase [22]. Enolase is recruited into the degradosome through interactions with a small recognition motif. The crystal structures of enolase bound to its cognate site from RNase E (residues 833 to 850 at 1.6 Å resolution [23] or residues 823-850 at 1.9 Å resolution [24]) show that a single molecule of the RNase E peptide binds asymmetrically in a conserved cleft at the interface of the enolase dimer and suggest that enolase may help to organize an adjacent conserved RNA-binding motif in RNase E. Crystal structures of *E. coli* PNPase complexed with the recognition site from RNase E (residues 1021-1061) show that the homotrimeric PNPase binds RNase E on the periphery of its ring-like architecture via a sheet–strand interaction [25]. Finally, the location of the interaction sites between RNase E and RhlB were analyzed using limited protease digestion, domain cross-linking and homology modeling [26]. These data indicated that RhlB recognizes a segment of RNase E that is no greater than 64 residues and it was proposed that their interaction stimulates the unwinding and ATPase activities of RhlB and that RhlB gains proximity to two RNA-binding sites on RNase E, with which it cooperates to unwind RNA. Although resident components of the complex appear not to be exchanged under normal growth conditions once they are engaged in the complex [27,28], the degradosome is believed to be a dynamic complex, capable of exchanging components under certain conditions. Altogether, these studies suggest that the carboxy-terminal domain of RNase E acts as a flexible tether of the degradosome components.

Contrary to RNase E, the catalytic domain of RNaseY is located at the C-terminus [12]. The amino acids composition of the N-terminal domain of RNase Y from *Bacillus subtilis* (BsRNaseY) indicates that this domain is mostly charged and polar (70%), does not contain any aromatic residue and displays only a few hydrophobic residues (Figure 1A). An *in silico* study was previously performed [12]. The *PONDR_FIT* [29] and *metaPrDOS* [30] programs indicated a high disorder propensity spanning around 100 residues with a sharp decrease at about residue 190, which is slightly different from the *IUPRED* prediction (Figure 1A). Charge-hydropathy analysis [31] also clearly showed that the N-terminal domain of BsRNaseY has a distribution characteristic of intrinsically disordered

proteins. In addition, the COILS program [32] predicted the presence of a coiled-coil domain with the potential to form a leucine zipper, spanning residues 30 to 150. Altogether, this region is predicted to form a flexible coiled-coil-like structure that is also significantly disordered and able to adopt many different conformations [12]. Furthermore, the interaction of the isolated domains of RNase Y was analyzed by monitoring the restoration of adenylate cyclase activity upon interaction of the partners fused to the N- and C-terminal domains of *Bordetella pertussis* adenylate cyclase [12]. This study revealed that only the transmembrane and the putative coiled-coil domains showed self-interactions, suggesting that they were major contributors to the oligomerization of RNase Y.
RNase Y has been proposed to participate, together with RNases J1 and J2, polynucleotide phosphorylase PNPase, helicase CshA and the glycolytic enzymes 6-phosphofructokinase and enolase in a multiprotein complex [33-37], similar to the RNase E-based degradosome in *E. coli*. It is likely that the N-terminal domain of RNase Y could be the key element for assembling a degradosome complex in *B. subtilis*, similarly to the C-terminal domain of RNase E in *E. coli*, and that it adopts a more ordered structure upon binding to its interaction partners. It has been shown that the interaction with the membrane through the N-terminal transmembrane region is required for the interaction of RNase Y with the other components in the degradosome *in vivo* [12].
As no more information about the structure of the N-terminal domain of BsRNaseY has been reported, we decided to express and purify it as a stand-alone domain to analyze experimentally its behavior in solution and try to gain information about its function.

RESULTS AND DISCUSSION

The secondary structure of BsRNAse Y was appraised by the HCA (hydrophobic cluster analysis) method (Figure 1B) [38]. The HCA plot offers the best correspondence between the hydrophobic clusters (Figure 1B, in green) and the regular secondary structures. Examination of the HCA plot allows us to easily identify globular regions from non-globular ones and, in globular regions, to identify secondary structures. This method confirms the previous bioinformatic prediction that the N-terminal domain of RNase Y should be intrinsically disordered and that it ends around residue 190 [12].

Nter-BsRNaseY diffuses as a non-globular molecule.

The sequence corresponding to residues 24 to 192 (Nter-BsRNaseY) was then expressed as a highly soluble form and purified (Figure 2A). To evaluate the homogeneity of the Nter-BsRNaseY preparation and try to get a first experimental estimate of the size of the molecule, the protein was analyzed by Size Exclusion Chromatography (SEC) on a Superdex 200 column (Figure 2B). The complex was eluted as a reasonably symmetrical peak and had a larger Stokes radius than a globular protein of equal mass. Indeed, the deduced average partition coefficient *K*av derived from the standard curve corresponds to an apparent molecular mass of 131 kDa, whereas the theoretical molecular weight of the monomer is 20.6 kDa. Thus, the SEC data indicate that Nter-BsRNaseY diffuses as a non-globular molecule. The elongated nature of the protein would cause it to elute earlier in size exclusion, with a consequent overestimation of its apparent molecular mass.

The monomeric and dimeric forms of Nter-BsRNaseY are in rapid equilibrium

The absolute molar mass of Nter-BsRNaseY in solution was then determined by Multiangle Laser Light Scattering coupled with Size Exclusion chromatography (SEC-MALLS), a technique that does not rely on the calibration of standards and assumptions of the globular state of the protein [39,40]. The chromatogram contains a major species eluting at 12.3 ml on a Superdex 200 increase column (Figure 2C). The close superposition of the LS and RI (refractive index) signals indicates a highly monodisperse sample. The measured hydrodynamic radius Rh was 4.33 ± 0.3 nm and the molar mass 33.1 ± 1.1 kDa (Table 1). This mass value, intermediate between the expected value for a monomer and a dimer, suggests that the monomer and another oligomeric form of Nter-BsRNaseY are in rapid equilibrium. To confirm that this mass corresponds to Nter-BsRNaseY and not a protein contaminant, the sample was analyzed by MALDI-MS (Supplementary Figure 1). The spectrum shows a major peak with m/z=21,108 Da, 10,475 Da and 7028 Da, for the +1, +2 and +3 charged species, respectively, which gives on average a mass of 21,048 Da, a value very close to the theoretical mass of Nter-BsRNaseY. To know whether the dimer assembly is formed by hydrophobic or electrostatic

interactions, the SEC-MALLS analysis of Nter-BsRNaseY was performed in the presence of 2M NaCl. As shown in Figure 2D, the measured mass of 40.66 ± 1.2 kDa corresponds to the dimer, indicating that dimerization operates via hydrophobic interactions that are favored at high salt concentration.

Nter-BsRNaseY displays various extended conformations

SAXS is a method complementary to crystallography and NMR that is widely used for studying the low-resolution structure of macromolecules in solution [41]. This method can provide insight into tertiary structure of biological macromolecules but also enables to distinguish ordered and disordered proteins [42-44]. In particular, SAXS is a well-suited tool to assess disorder in proteins and get insight into their structure ensembles [43,45,46]. The Guinier plot of SAXS data shows a nice concordance between the experimental points and the Guinier equation, which confirms the monodispersity of Nter-BsRNaseY in solution (Supplementary Figure 2A). The gyration radius Rg that monitors the extent of macromolecular elongation was estimated from the slope of the linear part of the Guinier plot at very small angles to be 5.63 nm (Supplementary Figure 2A) [47] and the maximum diameter of the molecule D_{max} was determined to be 20 nm (Table 1). The Rg and D_{max} values are much larger than the corresponding values expected for a spherical protein of n=169 residues (Rg ~ $3n^{1/3}$ = 16.6 Å and D_{max} ~2.6 x Rg = 43.1Å), indicating an elongated shape of Nter-BsRNaseY in solution. The Kratky plot, which is a representation of the scattering intensity I(q). q^2 as a function of the scattering vector q [45], was used for a qualitative analysis of the protein scattering profile. The normalized Ktratky plot has no maximum, and it plateaus at high q values, indicating the presence of intrinsic disorder (Figure 3A). The probability distance distribution function p(r), which is a real-space representation of the distribution of all interatomic distances within the molecule [41], exhibits an asymmetric profile with a maximum at small distances (r ~ 25 Å) characteristic of an ensemble of elongated molecules (Figure 3B).

Nter-BsRNaseY is a flexible molten globule

NMR has emerged as a preeminent method to investigate the molecular structure of IDPs, despite the fact that their spectra are generally characterized by extremely low chemical shift dispersion [48]. To examine whether Nter-BsRNaseY is an ID domain, we recorded the (^{1}H, ^{15}N)-HSQC spectrum of N^{15}-labeled Nter-BsRNaseY (Figure 4, Supplementary Figure 3B), which provides correlations between nitrogen atoms and amide protons. The spectrum should contain a peak for each amino acid except proline over the 6.5-10 ppm ^{1}H range if the protein adopts a tertiary structure. Around a hundred narrow peaks, out of the 176 expected peaks, gather in the 8–8.5 ppm ^{1}H region. The narrow chemical shift dispersion in the ^{1}H dimension of the ^{1}H-^{15}N HSQC spectrum is characteristic of disordered or α-helices-forming residues. Other residues are either absent in the spectrum or have broad peaks signature, indicating that a conformational exchange between more or less folded conformers or oligomerization takes place.

The SEC-MALLS results indicated that addition of salts shifts the oligomer equilibrium towards the dimeric form. Upon addition of 2M NaCl, a large number of peaks were broadened beyond detection in the (^{1}H, ^{15}N)-HSQC spectrum (Supplementary Figure 3C). This broadening is compatible with the formation of a larger system, probably a dimer, in agreement with the SEC-MALLS data. The remaining peaks, which mostly superimpose with the narrow peaks observed in the spectrum recorded at 200 mM NaCl, indicate that disordered segments are still present in the oligomeric form.

To confirm that the native protein contains some ordered structural elements, its spectrum was compared to that of a denatured sample. The number of peaks observed under denaturing conditions (6 M urea) was characteristic of a completely disordered protein (Supplementary Figure 3A). Thus, by comparison, the spectrum of the native protein is not that of a completely disordered monomer (Supplementary Figure 3B). Out of the 3 glycines present in Nter-BsRNaseY, Gly36 and Gly39 are located within the region potentially forming a coiled-coil and Gly105 is present in the region predicted to be intrinsically disordered (Figure 1). One of the glycine peaks of the (^{1}H, ^{15}N)-HSQC spectrum remains unchanged upon urea and salt titration, showing that its electronic environment does not change significantly (Supplementary Figure 3, top). This is compatible with a glycine in a highly solvent exposed, disordered region. On the contrary, the intensities of the two other glycine peaks vary significantly for the three conditions, with narrow resonances in urea that broaden in a 40 mM MES,

200 mM NaCl buffer and broaden even more in the presence of 2M NaCl. This behavior is compatible with the formation of a partially structured segment encompassing these two glycine residues, and possibly with the stabilization of a dimeric structure upon salt increase.

Nter-BsRNaseY contains mainly α-helices

To confirm that Nter-BsRNAseY behaves as a molten globule [49], its secondary structure was analyzed by near-UV circular dichroism (CD) spectroscopy (Figure 5). The CD spectrum clearly shows that the protein contains helices. A more qualitative analysis using the DICHROWEB algorithm indicates a content of around 50% of helices and 15% of disordered regions.

CONCLUSION

IDPs frequently function in protein interaction networks that regulate crucial cellular signaling pathways [50]. A common mechanism of action for natively unfolded proteins involves partial or complete folding upon interaction with a binding partner [51,52] and areas of residual structure in unfolded proteins can be critical to forming the initial site of interaction with a binding partner [52]. Disordered proteins can provide large and versatile surface area, enabling multiple interactions with other molecules [53]. Although DNA is one of the most common targets bound by unfolded proteins and many nucleic acid-binding proteins contain flexible extensions and loops thought to contribute to their chaperone activity, these domains have been little studied owing to their weak sequence conservation and conformational disorder [54,55].

Although natively unfolded proteins are much less abundant among archaea and eubacteria (5%) than eukaryotes [56], a recent report has addressed the function of the bacterial Hfq protein that contains an ID domain [57]. The RNA chaperone Hfq is a major regulator of bacterial stress responses and virulence. It is required to protect small noncoding RNAs (sRNAs) and facilitate their interactions with mRNAs. It was shown that its intrinsically disordered C-terminal domain is not needed to stimulate RNA base pairing but rather contributes to the chaperone activity by accelerating the release of dsRNA. By altering the dynamics of Hfq–RNA interactions, the C-terminal domain of Hfq maintains normal sRNA levels in the cell and tunes the efficiency of specific sRNA regulatory pathways.

The ID domain of RNAse Y may have a related function in modulating RNA interactions within the degradosome complex for the regulation of mRNAs degradation. We have shown here that Nter-BsRNaseY can be isolated as a stand-alone protein and is not a completely disordered monomer but rather a molten globule organized as helices with unstructured regions, which is distributed between a monomeric and dimeric state. Coiled-coil domains are well known to be involved in the self-interaction of many proteins [58]. Accordingly, BsRNAseY was previously predicted to form dimers with a high tendency to oligomerize and it was suggested that oligomerization is mediated by the N-terminal domain [12]. The formation of higher oligomeric structures extends significantly the potential interaction interface for binding other proteins. In order to gain more precise insights into the structural and dynamic characteristic of the domain, free and bound to its partners, the next step will be the assignment of the backbone resonances of [^{15}N, ^{13}C, ^{2}H]-Nter-BsRNAseY and the investigation of the evolution of the spectra upon binding its potential protein partners. These studies will help to reveal the oligomerization state of the ID domain of RNase Y within the degradosome complex and understand its function.

MATERIALS & METHODS

Overexpression of Nter-BsRNaseY

The sequence of the BsRNaseY gene corresponding to amino acid V24 to N192 (thus lacking the N-terminal transmembrane domain) appended with a C-terminal hexahistidine tag was PCR-amplified from chromosomal DNA of strain SSB1002, a wild-type *B. subtilis* laboratory stock strain derived from strain 168, using the Sequence Ligation Independent Cloning method [59]. The PCR fragment containing the 6xHis-tag sequence at the 3' end was inserted as an NdeI-BamHI fragment into the T7 expression vector pET22b plasmid (New England Biolabs). The plasmid was then transformed into *E. coli* BL21(*DE3*) star codon$^+$ competent cells. One colony was used to inoculate 200 ml of Luria Broth medium supplemented with ampicillin (50 µg/ml) and chloramphenicol (30 µg/ml). This preculture grown overnight at 37°C at 180 rpm (0.5%) was used to inoculate 2 liters of LB medium supplemented with the same antibiotics. Cultures were grown at 37°C to an OD_{600} of 0.8. Expression was then induced at 23°C by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM and incubation was extended overnight at 23°C. Cells were centrifuged for 10 min at 5000 rpm and 4°C.

Purification of Nter-BsRNaseY

Cells were resuspended in 1.5 ml/g of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol (buffer A) containing 20 µl/ml cOmplete™ Protease Inhibitor Cocktail (Roche), 10 mM β-mercaptoethanol, 2 µg/ml of RNase A (Thermofischer), 1 mg/ml lysozyme, 0.25% Tween 20, and disrupted by sonication. Cells debris were removed by ultracentrifugation at 100,000g for 1h at 4°C. The supernatant was loaded on a 15 mL Ni-NTA Superflow column (Qiagen) equilibrated in buffer A containing 20 mM imidazole, 0.25% Tween 20. The column was washed with 70 mM imidazole and eluted with 200 mM imidazole in buffer A. Fractions containing the protein were then loaded at 3.5-5 mg/mL, in two successive steps, on a Superdex 200 26/60 PG column (GE Healthcare) in buffer A. Protein elution was followed by the optical density at 230 nm due to the absence of aromatic residues in the Nter-BsRNaseY sequence.

Culture and purification of ^{15}N-labeled Nter-BsRNaseY

A preculture in 200mL of LB medium, supplemented with ampicillin (50 µg/ml) and chloramphenicol (30 µg/ml) and grown overnight at 37°C at 180 rpm, was used (0.5%) to inoculate one liter of minimal medium (M9) containing ^{15}N-labeled NH_4Cl and supplemented with 2 mM Mg_2SO_4, 20% Glucose, 10 µM $CaCl_2$, 5 mg/mL thiamine and antibiotics as in the preculture. Cultures were grown at 37°C to an OD_{600} of 0.8. Expression was then induced at 23°C by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM and incubation was extended overnight at 23°C. Purification by affinity chromatography was done as previously described, except that 40 mM MES pH 6.8, 200 mM NaCl (buffer B) was used for SEC. After analysis for purity on a Superdex 200 10/300 GL increase column, unlabeled and ^{15}N-labeled Nter-BsRNaseY were finally concentrated to 10 mg/ml (480 µM) using Amicon concentrators (30 kD cutoff, Milllipore), then aliquoted, frozen in liquid nitrogen and stored at -80°C.

Static Light Scattering

^{15}N-labeled Nter-BsRNase Y was analyzed using a SEC-MALLS/UV/RI/quasi-elastic light scattering Wyatt instrument. 100 µl at 10.5 mg/ml was injected, and separation was performed using a Shimadzu liquid chromatography system (Prominence) equipped with a Superdex 200 10/300GL increase column before light-scattering measurements. Isocratic elution was performed at a flow rate of 0.5 ml/min in 40 mM MES, pH 6.8, 200 mM NaCl, at 25°C. Light scattering was monitored using a mini DAWNTM TREOS system equipped with a quasi-elastic light scattering module (Wyatt Technology, Santa Barbara, CA) and refractometer (OptilabrEX, Wyatt Technology). The concentration of the molecule in solution was determined using a specific refractive index increment (dn/dc) of 0.168 ml/g at 658 nm. The absolute refractive index of the solution was measured using the refractive index detector. The light scattering detector was calibrated with toluene. The detector was controlled and the data were analyzed using the ASTRA V software (version 5.3.4.20; Wyatt Technology).

SAXS data collection and analysis

SAXS data were collected on beam line SWING of the SOLEIL synchrotron. The protein sample (20 mg/ml in 20 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol) was injected onto an BioSEC3 column (Agilent). Scattering data were recorded on the eluted peak and processed using FOXTROT [60]. The gyration radius was calculated by the Guinier approximation using PRIMUS from the ATSAS suite [61]. The GNOM program from the ATSAS suite was used to calculate the pair distance distribution function P(r) and D_{max}.

Circular dichroism (CD) spectroscopy

CD spectra were recorded using a JASCO J-810 CD spectropolarimeter equipped with a sample cell
temperature control unit (PFD 423S/L Peltier). All measurements were collected at 25°C in a 0.1 mm path-length quartz cuve over the 180-260 nm spectral range. The spectra consisted of an average of five scans acquired at a speed of 50 nm/min, with a time response of 1s for a 1.0 nm band-width. Non-denatured Nter-BsRNaseY (190 µM) was examined in 40 mM MES pH 6.8, 200 mM NaCl, 5 % D_2O.). For each spectrum, the buffer contribution was subtracted and the signal normalized to protein concentration for converting the results to molar ellipticity per residue units (MER). The DICHROWEB algorithmn [62] was used for CD spectrum deconvolution using the CDSSTR method with reference data set 7 [63].

NMR experiments

NMR samples of ^{15}N-labeled (190 mM) Nter-BsRNAseY were prepared in MES buffer (40 mM, pH 6.8, 200 mM NaCl) in H_2O/D_2O (9.5:0.5) for the reference condition. MES buffer was supplemented by 6 M urea or NaCl was increased to 2 M for experiments at high urea and high saline concentrations, respectively. NMR spectra were collected at 298K on a Bruker AVANCE III HD 950 MHz equipped with TCI cryoprobes. ^{1}H-^{15}N correlation experiments were collected using the SOFASTHMQC pulse sequences [64]. Data processing and analysis were performed using Topspin ® 3.5 and CCPNMR softwares [65].

Acknowledgements

This work was supported by the French State Program 'Investissements d'Avenir' (Grants "LABEX DYNAMO", ANR-11-LABX-0011 and ANR-11-BSV8-026), and has benefited from the facilities and expertise of the Macromolecular Interaction Platform of I2BC. We thank SOLEIL for provision of synchrotron radiation facilities, TGIR-RMN-THC FR3050 CNRS for access to the NMR spectrometers, Aurélien Agneray for maintenance of beamline SWING and the IMAGIF platform for cloning facilities, Harald Putzer for fruitful discussions.

References

1. Laalami S, Putzer H: mRNA degradation and maturation in prokaryotes: the global players. *Biomol Concepts* 2011, 2:491-506.
2. Laalami S, Zig L, Putzer H: Initiation of mRNA decay in bacteria. *Cell Mol Life Sci* 2014, 71:1799-1828.
3. Mohanty BK, Kushner SR: Regulation of mRNA Decay in Bacteria. *Annu Rev Microbiol* 2016, 70:25-44.
4. Gorna MW, Carpousis AJ, Luisi BF: From conformational chaos to robust regulation: the structure and function of the multi-enzyme RNA degradosome. *Q Rev Biophys* 2012, 45:105-145.
5. Mackie GA: RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* 2013, 11:45-57.
6. Shahbabian K, Jamalli A, Zig L, Putzer H: RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J* 2009, 28:3523-3533.
7. Durand S, Tomasini A, Braun F, Condon C, Romby P: sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiol Rev* 2015, 39:316-330.
8. Kaito C, Kurokawa K, Matsumoto Y, Terao Y, Kawabata S, Hamada S, Sekimizu K: Silkworm pathogenic bacteria infection model for identification of novel virulence genes. *Mol Microbiol* 2005, 56:934-944.
9. Bonnin RA, Bouloc P: RNA Degradation in *Staphylococcus aureus*: Diversity of Ribonucleases and Their Impact. *Int J Genomics* 2015, 2015:395753.
10. Obana N, Nakamura K, Nomura N: Role of RNase Y in *Clostridium perfringens* mRNA Decay and Processing. *J Bacteriol* 2017, 199.
11. Kang SO, Caparon MG, Cho KH: Virulence gene regulation by CvfA, a putative RNase: the CvfA-enolase complex in *Streptococcus pyogenes* links nutritional stress, growth-phase control, and virulence gene expression. *Infect Immun* 2010, 78:2754-2767.
12. Lehnik-Habrink M, Newman J, Rothe FM, Solovyova AS, Rodrigues C, Herzberg C, Commichau FM, Lewis RJ, Stulke J: RNase Y in *Bacillus subtilis*: a Natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J Bacteriol* 2011, 193:5431-5441.
13. Callaghan AJ, Grossmann JG, Redko YU, Ilag LL, Moncrieffe MC, Symmons MF, Robinson CV, McDowall KJ, Luisi BF: Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain. *Biochemistry* 2003, 42:13848-13855.
14. Callaghan AJ, Marcaida MJ, Stead JA, McDowall KJ, Scott WG, Luisi BF: Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* 2005, 437:1187-1191.
15. Koslover DJ, Callaghan AJ, Marcaida MJ, Garman EF, Martick M, Scott WG, Luisi BF: The crystal structure of the *Escherichia coli* RNase E apoprotein and a mechanism for RNA degradation. *Structure* 2008, 16:1238-1244.
16. Vanzo NF, Li YS, Py B, Blum E, Higgins CF, Raynal LC, Krisch HM, Carpousis AJ: Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev* 1998, 12:2770-2781.
17. Dominguez-Malfavon L, Islas LD, Luisi BF, Garcia-Villegas R, Garcia-Mena J: The assembly and distribution in vivo of the *Escherichia coli RNA* degradosome. *Biochimie* 2013, 95:2034-2041.

18. Ait-Bara S, Carpousis AJ, Quentin Y: RNase E in the gamma-Proteobacteria: conservation of intrinsically disordered noncatalytic region and molecular evolution of microdomains. *Mol Genet Genomics* 2015, 290:847-862.
19. Callaghan AJ, Aurikko JP, Ilag LL, Gunter Grossmann J, Chandran V, Kuhnel K, Poljak L, Carpousis AJ, Robinson CV, Symmons MF, et al.: Studies of the RNA degradosome-organizing domain of the *Escherichia coli* ribonuclease RNase E. *J Mol Biol* 2004, 340:965-979.
20. Marcaida MJ, DePristo MA, Chandran V, Carpousis AJ, Luisi BF: The RNA degradosome: life in the fast lane of adaptive molecular evolution. *Trends Biochem Sci* 2006, 31:359-365.
21. Khemici V, Poljak L, Luisi BF, Carpousis AJ: The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol Microbiol* 2008, 70:799-813.
22. Carpousis AJ: The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* 2007, 61:71-87.
23. Chandran V, Luisi BF: Recognition of enolase in the *Escherichia coli* RNA degradosome. *J Mol Biol* 2006, 358:8-15.
24. Nurmohamed S, McKay AR, Robinson CV, Luisi BF: Molecular recognition between *Escherichia coli* enolase and ribonuclease E. *Acta Crystallogr D Biol Crystallogr* 2010, 66:1036-1040.
25. Nurmohamed S, Vaidialingam B, Callaghan AJ, Luisi BF: Crystal structure of *Escherichia coli* polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradosome assembly. *J Mol Biol* 2009, 389:17-33.
26. Chandran V, Poljak L, Vanzo NF, Leroy A, Miguel RN, Fernandez-Recio J, Parkinson J, Burns C, Carpousis AJ, Luisi BF: Recognition and cooperation between the ATP-dependent RNA helicase RhlB and ribonuclease RNase E. *J Mol Biol* 2007, 367:113-132.
27. Coburn GA, Miao X, Briant DJ, Mackie GA: Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. *Genes Dev* 1999, 13:2594-2603.
28. Worrall JA, Gorna M, Crump NT, Phillips LG, Tuck AC, Price AJ, Bavro VN, Luisi BF: Reconstitution and analysis of the multienzyme *Escherichia coli* RNA degradosome. *J Mol Biol* 2008, 382:870-883.
29. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN: PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta* 2010, 1804:996-1010.
30. Ishida T, Kinoshita K: Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics* 2008, 24:1344-1348.
31. Uversky VN, Gillespie JR, Fink AL: Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 2000, 41:415-427.
32. Lupas A, Van Dyke M, Stock J: Predicting coiled coils from protein sequences. *Science* 1991, 252:1162-1164.
33. Lehnik-Habrink M, Pfortner H, Rempeters L, Pietack N, Herzberg C, Stulke J: The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol Microbiol* 2010, 77:958-971.
34. Lehnik-Habrink M, Lewis RJ, Mader U, Stulke J: RNA degradation in *Bacillus subtilis:* an interplay of essential endo- and exoribonucleases. *Mol Microbiol* 2012, 84:1005-1017.
35. Salvo E, Alabi S, Liu B, Schlessinger A, Bechhofer DH: Interaction of *Bacillus subtilis* Polynucleotide Phosphorylase and RNase Y: Structural Mapping and effect on mRNA turnover. *J Biol Chem* 2016, 291:6655-6663.
36. Cascante-Estepa N, Gunka K, Stulke J: Localization of Components of the RNA-Degrading Machine in *Bacillus subtilis*. *Front Microbiol* 2016, 7:1492.
37. Cho KH: The Structure and Function of the Gram-Positive Bacterial RNA Degradosome. *Front Microbiol* 2017, 8:154.
38. Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP: Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell Mol Life Sci* 1997, 53:621-645.
39. Folta-Stogniew E: Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors. *Methods Mol Biol* 2006, 328:97-112.

40. Sahin E, Roberts CJ: Size-exclusion chromatography with multi-angle light scattering for elucidating protein aggregation mechanisms. *Methods Mol Biol* 2012, 899:403-423.
41. Svergun DI, Koch MHJ: Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 2003, 66:1735-1782.
42. Receveur-Brechot V, Bourhis JM, Uversky VN, Canard B, Longhi S: Assessing protein disorder and induced folding. *Proteins* 2006, 62:24-45.
43. Bernado P, Svergun DI: Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol Biosyst* 2012, 8:151-167.
44. Receveur-Brechot V, Durand D: How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr Protein Pept Sci* 2012, 13:55-75.
45. Kikhney AG, Svergun DI: A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 2015, 589:2570-2577.
46. Bernado P, Svergun D: Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. In *Instrumental analysis of intrinsically disordered proteins: assessing structure and conformation* Edited by Uversky VN LSe: John Wiley & Sons, Inc.; 2010:451-475.
47. Koch MH, Vachette P, Svergun DI: Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* 2003, 36:147-227.
48. Brutscher B, Felli IC, Gil-Caballero S, Hosek T, Kummerle R, Piai A, Pierattelli R, Solyom Z: NMR Methods for the Study of Instrinsically Disordered Proteins Structure, Dynamics, and Interactions: General Overview and Practical Guidelines. *Adv Exp Med Biol* 2015, 870:49-122.
49. Naiyer A, Hassan MI, Islam A, Sundd M, Ahmad F: Structural characterization of MG and pre-MG states of proteins by MD simulations, NMR, and other techniques. *J Biomol Struct Dyn* 2015, 33:2267-2284.
50. Babu MM: The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem Soc Trans* 2016, 44:1185-1200.
51. Fink AL: Natively unfolded proteins. *Curr Opin Struct Biol* 2005, 15:35-41.
52. Uversky VN: Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 2002, 11:739-756.
53. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R: Extended disordered proteins: targeting function with less scaffold. *Trends Biochem Sci* 2003, 28:81-85.
54. Tompa P, Csermely P: The role of structural disorder in the function of RNA and protein chaperones. *FASEB J* 2004, 18:1169-1175.
55. Charlier C, Bouvignies G, Pelupessy P, Walrant A, Marquant R, Kozlov M, De Ioannes P, Bolik-Coulon N, Sagan S, Cortes P, et al.: Structure and Dynamics of an Intrinsically Disordered Protein Region That Partially Folds upon Binding by Chemical-Exchange NMR. *J Am Chem Soc* 2017, 139:12219-12227.
56. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT: Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 2004, 337:635-645.
57. Santiago-Frangos A, Kavita K, Schu DJ, Gottesman S, Woodson SA: C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proc Natl Acad Sci U S A* 2016, 113:E6089-E6096.
58. Burkhard P, Stetefeld J, Strelkov SV: Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol* 2001, 11:82-88.
59. Li MZ, Elledge SJ: SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol Biol* 2012, 852:51-59.
60. David G, Pérez J: Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J Appl Cryst* 2009, 42:892-900.
61. Svergun DI: Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J Appl Cryst* 1992, 25:495-503.
62. Whitmore L, Wallace BA: Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 2008, 89:392-400.

63. Sreerama N, Woody RW: Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 2000, 287:252-260.
64. Schanda P, Brutscher B: Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* 2005, 127:8014-8015.
65. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED: The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 2005, 59:687-696.
66. Dosztanyi Z, Csizmok V, Tompa P, Simon I: IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005, 21:3433-3434.

Table 1. Biophysical parameters of Nter-BsRNaseY

Method	Parameter	Nter-BsRNaseY
SEC-MALLS	Molecular mass (kDa)	33.1 ± 1.1
	Rh (nm)	4.3 ± 0.3
SAXS	Rg (nm)	5.63 ± 0.05
	Dmax (nm)	20.0

Figure legends

Figure 1: Secondary structure analysis of BsRNaseY. A Bioinformatics analysis of Nter-BsRNAseY amino acid sequence (Top). The amino acid sequence of Nter-BsRNAseY (buttom) is shown with the disorder propensity, as calculated by *IUPRED* [66]. Glycine are represented by purple bars or in bold. The most flexible regions are highlighted in green. **B** Hydrophobic Cluster Analysis plot of BsRNaseY. This method is based on the use of a bidimensional HCA plot, in which the protein sequence is written as an α-helix displayed on a cylinder with 3.6 residue per turn. The cylinder is cut along the horizontal axis, unrolled and repeated to recreate the full environment of each amino acid. Hydrophobic amino acids (colored green) gather into clusters (encircled), whose form is generally indicative of the type of secondary structure: vertical clusters are often associated to β-strands and horizontal ones to α-helices. Special symbols are used for proline (red star), threonine and serine (square and dotted square) and glycine (black diamond).

Figure 2: **Molecular mass determination of Nter-BsRNaseY. A** Purification of Nter-BsRNaseY analyzed on a 12% SDS-PAGE gel. 1 and 7: molecular weight standards commercial standards (BioRad n° 151-1901): Tyroglobulin (670 kDa), γ-globulin (158 kDa), Ovalbumin (44 kDa), Myoglobin (17 kDa), and Vitamin B12 (1,36 kDa); 2 and 3: cells before and after IPTG induction; 4: Pellet; 5: Supernatant; 6: after Nickel affinity column; 8: after first size exclusion chromatography. **B** Chromatographic profile of Nter-BsRNaseY on a Superdex 200 26/60 PG column with the pooled fractions delimited by the two green arrows. The void volume V_0 and elution volumes of molecular weight standards (in kDa) are indicated by red and black arrows, respectively. **C & D** SEC-MALLS analysis of Nter-BsRNaseY in 40 mM MES pH 6.2 and either 200 mM NaCl (**C**) or 2M NaCl (**D**). Dashed and continuous curves, representing respectively the refractometer signal (dRI) and light scattering signals (LS), are plotted as function of the elution volume. The molar masses calculation curve is in orange.

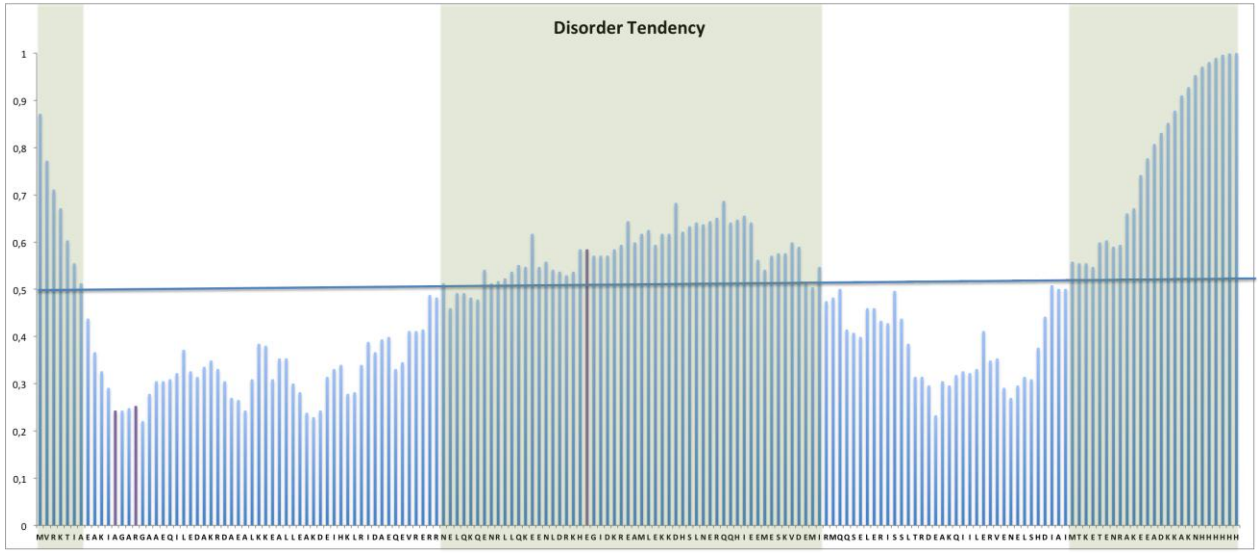
Figure 3: SAXS analysis of the Nter-BsRNAseY. A Normalized Kratky plot of Nter-BsRNAseY. **B** Pair distance distribution function p(r) of BsRNAseY dimer, indicating an elongated species.

Figure 4: **^{1}H, ^{15}N-HSQC Spectrum of Nter-BsRNAseY** (190 µM) in MES 40 mM pH 6.2, 200 mM NaCl, 10 % D_2O collected at 298K on a Bruker AVANCE III HD 950 MHz spectrometer equipped with TCI cryoprobes.

Figure 5: **CD Spectrum of Nter-RNAseY**. The CD Spectrum of the native protein (190 µM) was recorded in 40 mM MES pH 6.2, 200 mM NaCl, 10 % D_2O. Plots with crosses and dots correspond to the experimental data or data fitted by the CDSSTR method, respectively.

Figure 1

A



24 30 40 50 60 70 80 90 100

MVRKTIAEAKIA**G**AR**G**AAEQILEDAKRDAEALKKEALLEAKDEIHKLRIDAEQEVRERRNELQKQENRLLQKEENLDRKHE**G**IDKR

110 120 130 140 150 160 170 180 190

EAMLEKKDHSLNERQQHIEEMESKVDEMIRMQQSELERISSLTRDEAKQIILERVENELSHDIAIMTKETENRAKEEADKKAKNHHHHHH

B

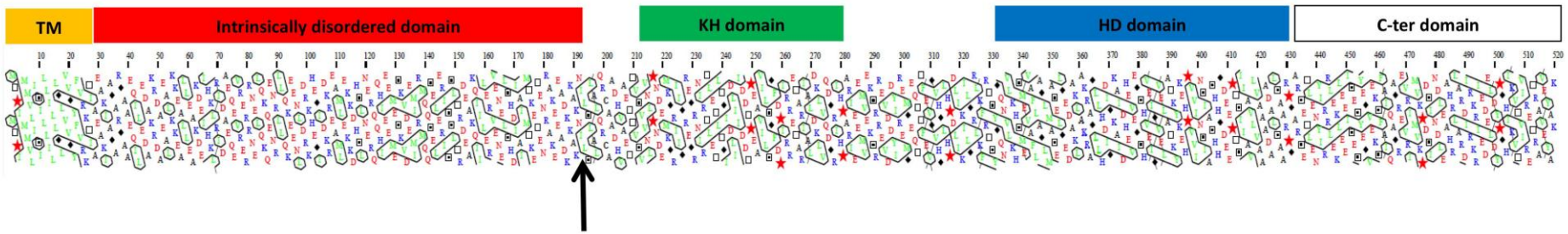


Figure 2

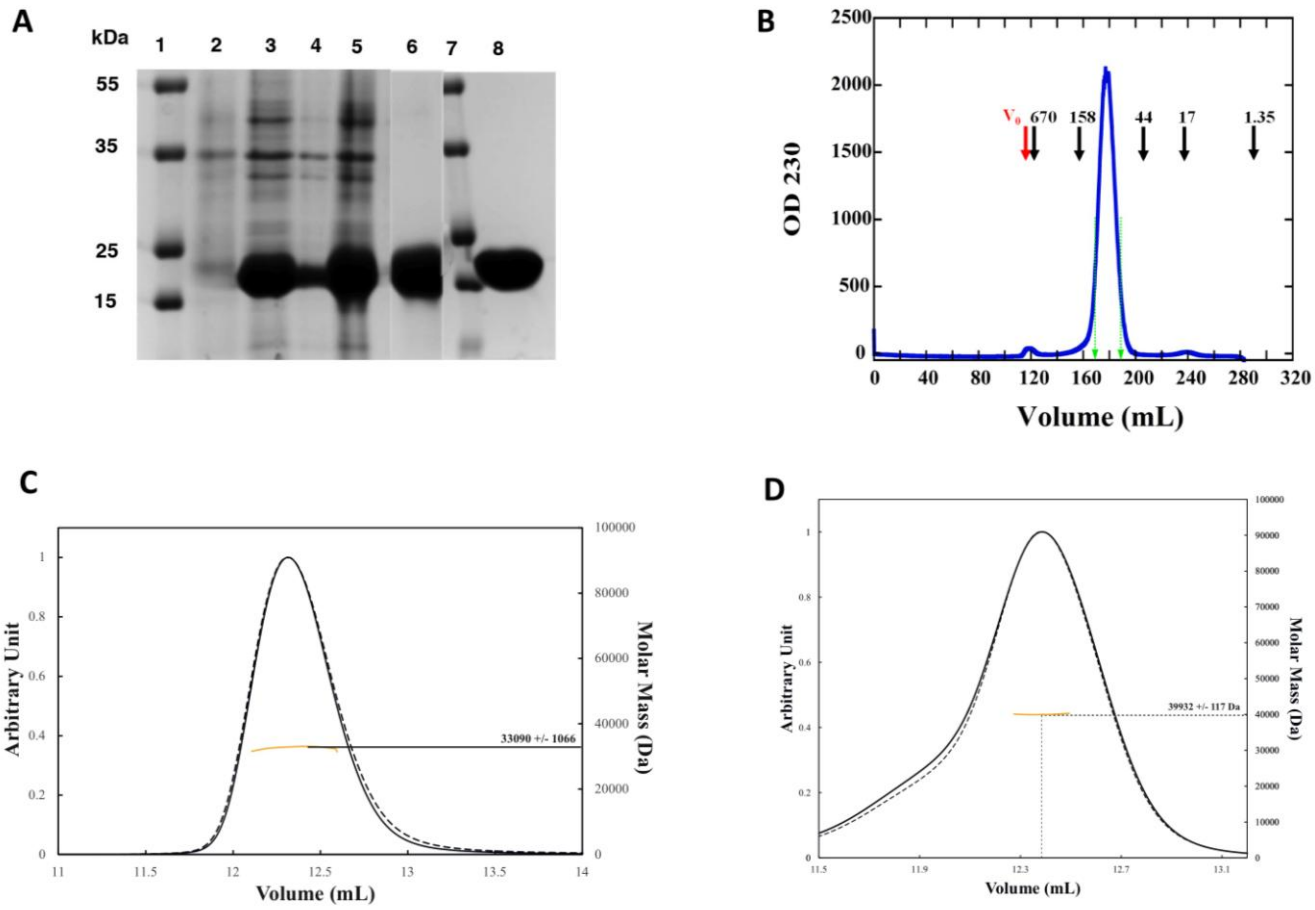

A
kDa
1 2 3 4 5 6 7 8
55
35
25
15
B
OD 230
Volume (mL)
V0 670 158 44 17 1.35
C
Arbitrary Unit
Molar Mass (Da)
Volume (mL)
33090 +/- 1066
D
Arbitrary Unit
Molar Mass (Da)
Volume (mL)
39932 +/- 117 Da

Figure 3

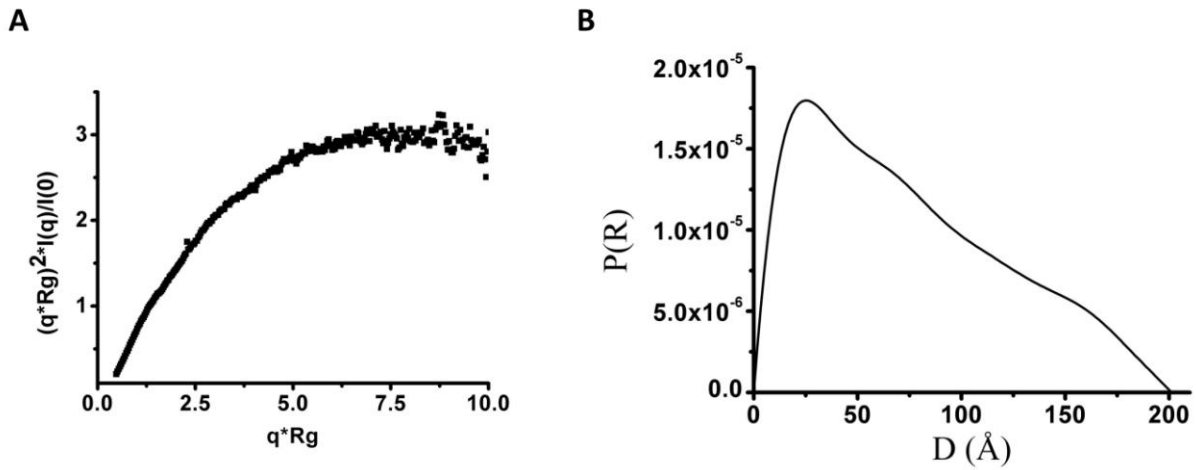

A
(q*Rg)2*I(q)/I(0)
3
2
1
0.0
2.5
5.0
7.5
10.0
q*Rg
B
2.0x10-5
1.5x10-5
1.0x10-5
5.0x10-6
0.0
P(R)
0
50
100
150
200
D (Å)

Figure 4

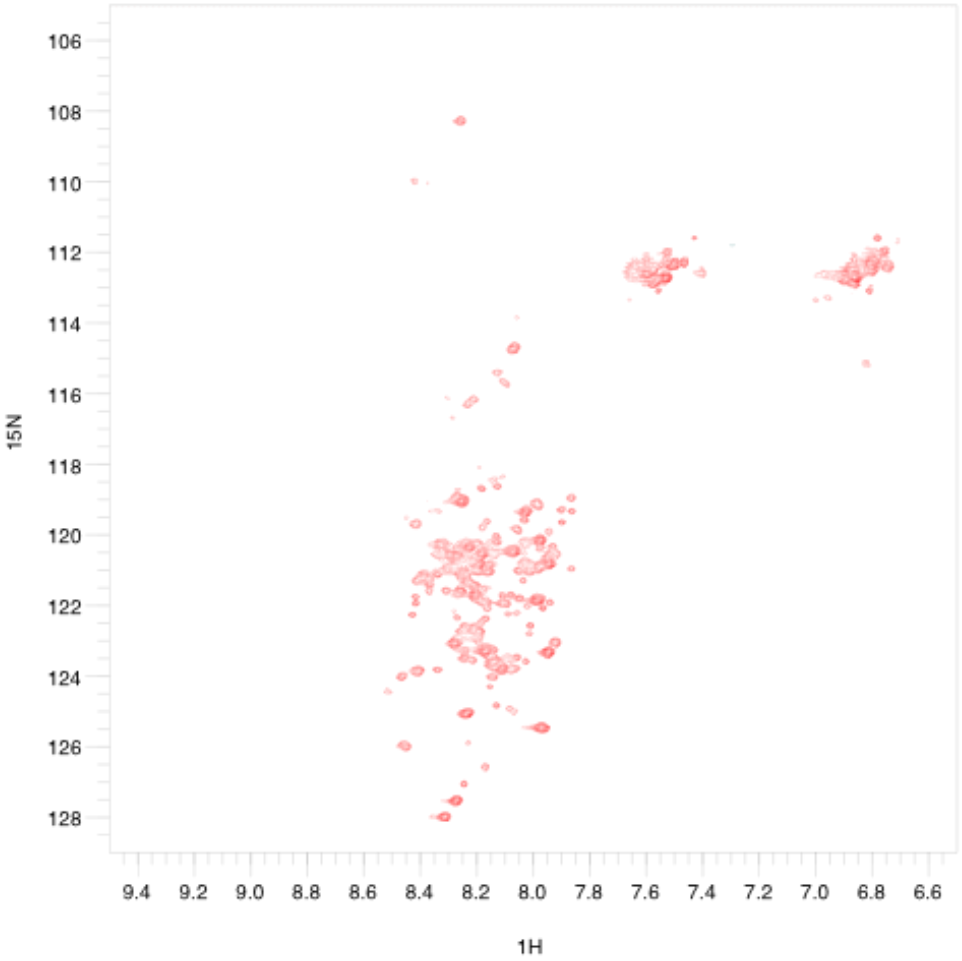
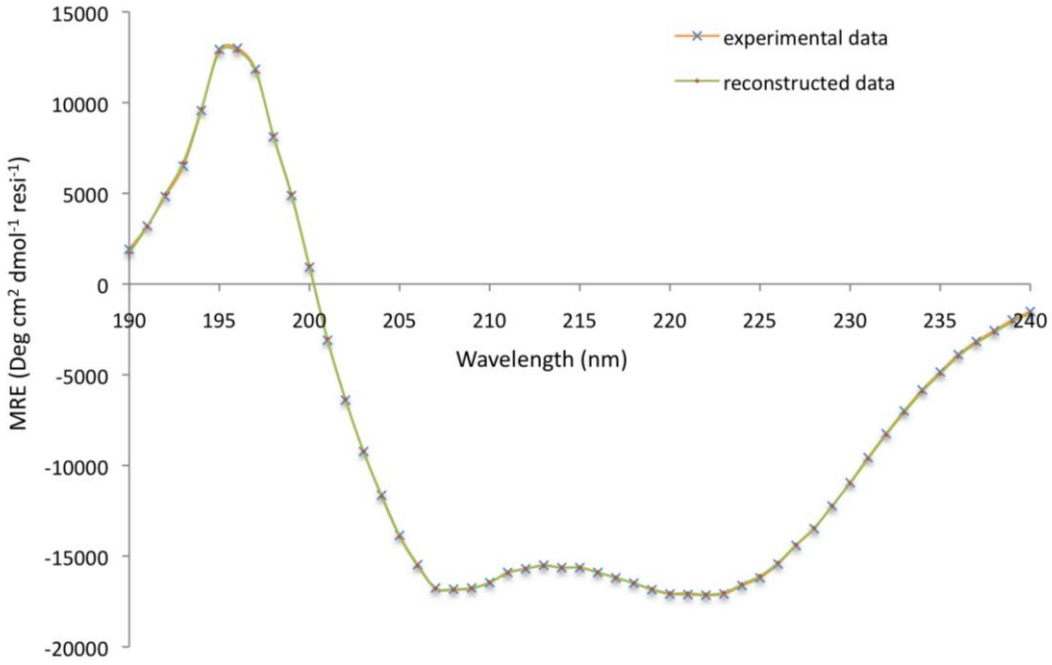
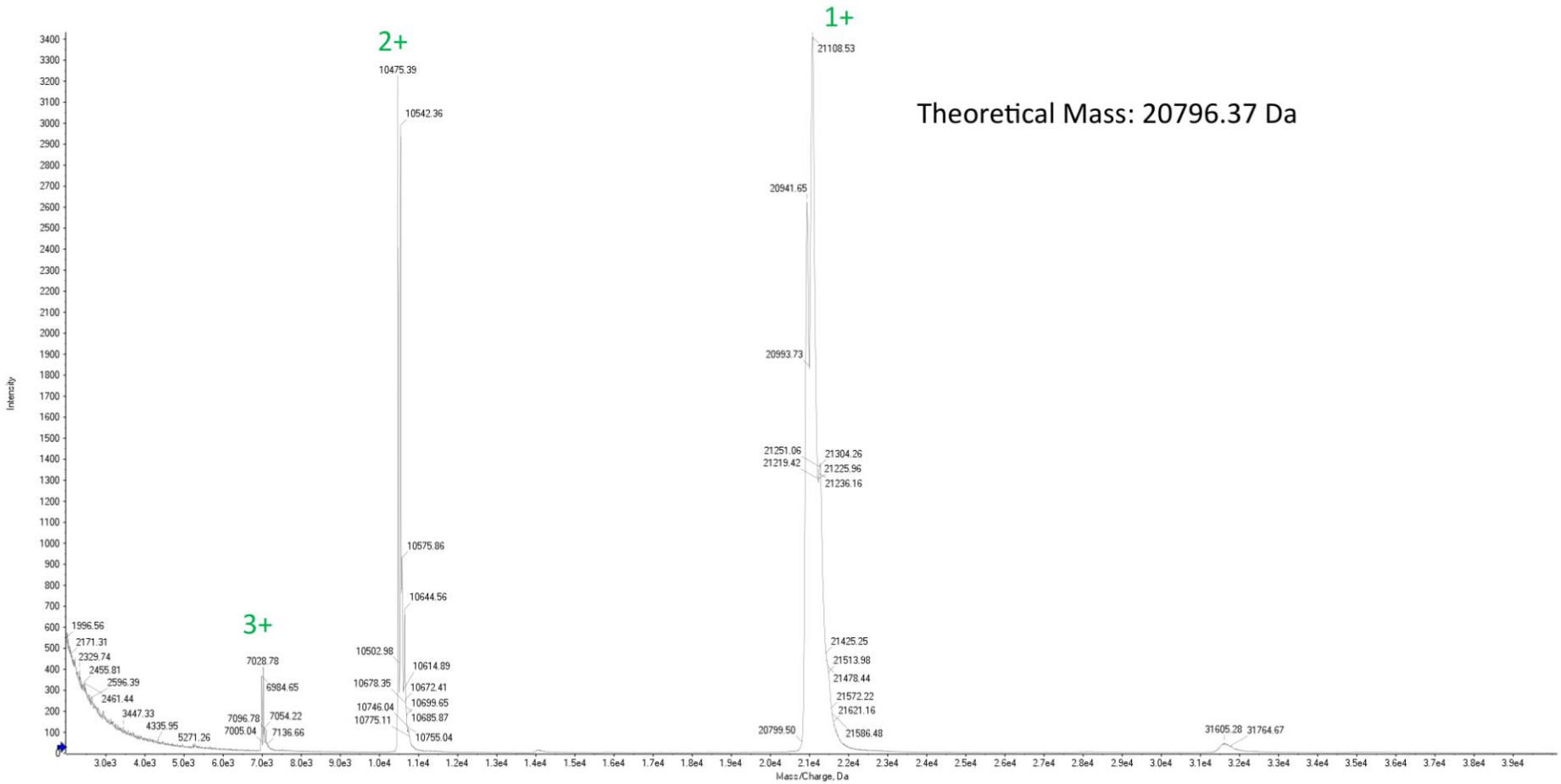

106
108
110
112
114
116
118
120
122
124
126
128
15N
9.4
9.2
9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
6.6
1H

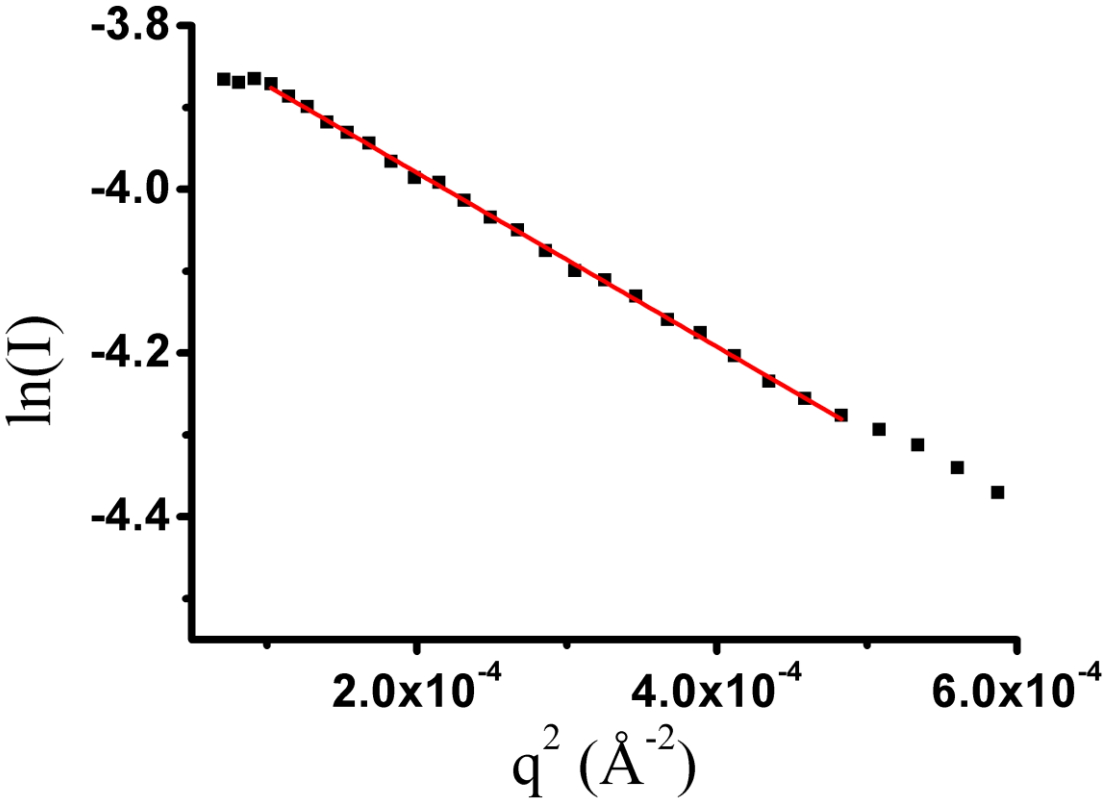
Figure 5


experimental data
reconstructed data
MRE (Deg cm² dmol⁻¹ resi⁻¹)
Wavelength (nm)
15000
10000
5000
0
-5000
-10000
-15000
-20000
190
195
200
205
210
215
220
225
230
235
240

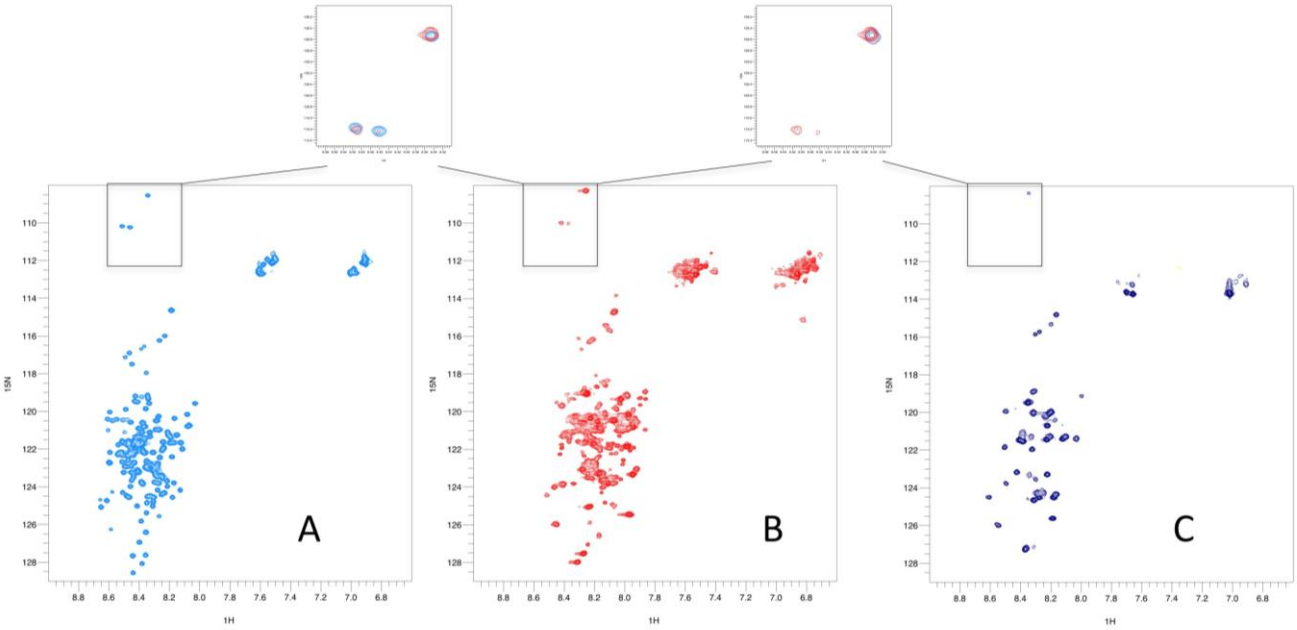
Supplementary Figure 1: Analysis of ^{15}N-labelled Nter-BsRNaseY by MALDI-MS



Supplementary Figure 2: Guinier plot of Nter-BsRNAseY.



Supplementary Figure 3: (^{1}H,^{15}N)-HSQC spectra recorded at 950 MHz of Nter-BsRNaseY (190 µM) in MES 40mM pH 6.8, 5% D_2O in various salt conditions. A 200 mM NaCl, 6M urea. **B** 200 mM NaCl (reference). **C** 2M NaCl. The top panels represent a superposition of the Glycine regions between spectra A and B (left), and spectra C and B (right).



Article 3

The elongated dimer of *Bacillus subtilis* RNase Y assembles as a globular oligomer of high molecular weight.

Pierre Hardouin[1], Charles Bou-Nader[1], Christophe Velours[2], Saravuth Ngo[3], Soumaya Laalami[3], Harald Putzer[3], Catherine Venien-Bryan[4], Béatrice Golinelli-Pimpaneau[1*]
[1] Laboratoire de Chimie des Processus Biologiques, UMR 8229 CNRS, Collège de France, Université Pierre et Marie Curie, 11 Place Marcelin Berthelot, 75231 Paris cedex 05, France
[2] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS Université Paris-Sud, Université Paris-Saclay, Avenue de la Terrasse, 91198 Gif-sur-Yvette cedex, France
[3] Institut de Biologie Physicochimique, PSL Research University, Chimie-ParisTech, Institut de Recherche de Chimie-Paris, 11 rue Pierre et Marie Curie, 75005 Paris, France
[4] Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, UMR 7590, CNRS, UPMC, IRD, MNHN, 75005 Paris, France

**Corresponding author:*
beatrice.golinelli@college-de-france.fr

Abstract
The instability of mRNA is crucial for the control of prokaryotic gene expression. The recently discovered endoribonuclease Y is an important player defining RNA metabolism in many Gram positive bacteria. We have purified two catalytically active oligomeric forms of *Bacillus subtilis* RNase Y: a dimer and a soluble oligomer of high molecular weight. Characterization of these two forms by various biophysical methods indicates that the dimer, which adopts an elongated form, can associate to form an 8-membered ring complex of ~1 M Da with an inner tunnel that could be visualized by electron microscopy.
Keywords: RNase Y, SEC-MALLS, SAXS, electron microscopy, oligomer

Introduction
Riboregulation in bacteria is an important aspect of gene expression control, especially in response to stress [1-3]. RNase Y is an endoribonuclease that acts as the key enzyme initiating mRNA decay in the Gram positive bacterium *Bacillus subtilis*. It is the only enzyme known, besides the RNA decay initiating RNase E from *Escherichia coli*, to significantly affect global mRNA half-life when depleted [4]. It is present in most low-GC Gram positive species [5], where it plays important roles in mRNA metabolism [2,6-8]. In *B. subtilis*, RNase Y likely initiates the decay of more than 1500 mRNAs [2,9-11]. In particular, RNase Y is crucial for the turnover of all S-adenosylmethionine dependent riboswitches in *B. subtilis* [4] and has been implicated in the decay initiation of the *rpsO* mRNA (encoding r-protein S15) [12]. Moreover, this nuclease also affects the expression of the *infC-rpmI-rplT* ribosomal operon (encoding translation factor IF3 and the ribosomal proteins L35, L20) [13]. Following the rate-limiting initial endonucleolytic cleavage, the resulting cleavage products are then degraded exonucleolytically by 3' and 5' exoribonucleases [2].
Bacillus subtilis RNase Y (BsRNaseY) shares important functional features with *E. coli* RNase E, despite both enzymes sharing no sequence similarity whatsoever [2]: 1) The cleavage specificity appears to be similar, consisting in a single stranded AU rich sequence close to a mRNA secondary structure. 2) The initial cleavage rate of RNase Y is stimulated 20-fold by a monophosphorylated 5'-end of the substrate [4]. 3) BsRNaseY is localized at the cell periphery [14], consistent with the presence of a predicted N-terminal transmembrane domain. By comparison, *E. coli* RNase E contains a peptide with the propensity to form an amphipathic α-helix that is necessary and sufficient for membrane binding [15].
RNase Y was the first discovered member of a new class of ribonucleases that constitute a subgroup within the superfamily of HD (His-Asp-containing phosphohydrolase) proteins [16] that combine the metal-chelating HD domain with the RNA-binding KH (ribonucleoprotein K homology) domain. It

has been shown that the conserved His and Asp signature residues of the HD domain are essential for the ribonuclease activity [4]. Very few structures of proteins containing both a KH and HD domains have been reported [17]. The potential interactions of RNase Y with other proteins could modulate the activity and/or specificity of RNase Y under varying metabolic conditions. In this context, it is likely that RNase Y interacts with its RNA and protein partners via its N-terminal domain, which is predicted to be intrinsically disordered [18]. In fact, RNase Y has been proposed to participate, together with RNases J1 and J2, polynucleotide phosphorylase PNPase and the glycolytic enzymes 6-phosphofructokinase and enolase in a multiprotein complex [19-23] similar to the RNase E-based degradosome in *E. coli*. In addition, helicase CshA has been shown to bind BsRNaseY and was also suggested to belong to the *B. subtilis* degradosome complex [24]. However, the recruitment of RNase J1/J2 and possibly other enzymes into the RNase Y-based assembly appears to be highly dynamic as some of the proposed interactions have not been observed in other studies [25,26] and a RNase Y-based degradosome has never been isolated without crosslinking.
A full comprehension of the mechanistic details of the catalytic activity, including the 5'-end dependence of RNase Y, requires information that can only be provided by 3D structure determination. Conversely, the knowledge of how RNase Y functions at the atomic level would open the way for a detailed mutational analysis of the enzyme. Since the crystallization of BsRNaseY was not successful so far, we studied the conformation and structure of BsRNaseY by various biophysical techniques, including Multi-Angle Laser Light Scattering coupled with Size Exclusion Chromatography (SEC-MALLS) and SEC-coupled Small Angle X-ray Scattering (SEC-SAXS). Two catalytically active forms were obtained and the higher molecular weight form was further characterized by electron microscopy combined to image analysis of negatively stained specimen. Altogether, these studies indicate that the elongated dimeric form of BsRNAse Y is able to organize as a compact 8-membered ring 1 MDa complex with a radius of 10-13 nm containing an inner tunnel.

Results

Two oligomeric forms of BsRNaseY can be separated

An RNase Y construct lacking the N-terminal transmembrane domain had previously been expressed, purified and analyzed by ultracentrifugation. Although a significant proportion of the protein seemed to be in an aggregated form, RNase Y formed a mixture of mostly dimers, but also some tetramers [18]. Here, we cloned and expressed in *E. coli* BsRNaseY, lacking the first 25 amino acids predicted to be inserted into the membrane [14,18] and appended with a C-terminal hexahistidine tag, as described in Material and Methods. The protein was purified by Nickel affinity (Fig. 1A) and size exclusion chromatography (Fig. 1B). At this stage, two oligomeric forms were separated and each one was loaded onto a second gel filtration column for further purification (Fig. 1C and D). As previously reported [18], the lower molecular weight form corresponds to a non-globular dimeric form. Indeed, as SEC fractionates proteins on the basis of their Stokes radius and not their molecular weight, the calculation of the molar mass using the eluted volume of molecular weight standards that assumes a globular form of the investigated protein cannot be used here. The higher molecular weight oligomer was conveniently named "oligomer" in the following. Both forms possessed similar ribonuclease activity for the *yitJ* riboswitch RNA substrate [4] (Supplementary Figure 1).

The dimer adopts an extended conformation and the soluble oligomer of high molecular weight forms a compact assembly

The oligomerization states of BsRNaseY in solution were further investigated by SEC-MALLS, a technique that does not rely on the calibration of standards and assumptions of the globular state of the protein [27,28]. Static light scattering, which measures the intensity of the scattered light as a function of angle, can be used to determine the absolute molecular mass using the Zimm formalism. The chromatograms of the dimer (theoretical mass of 57.1 kDa) and oligomer each contain a major species eluting at 11.1 ml and 9.5 ml, respectively, on a Superdex 200 and Superose 6 10/300 GL increase columns, respectively (Fig. 2). A molar mass of 105.3 kDa ± 1.0 kDa was measured for the main peak of the dimeric form, in reasonable agreement with the theoretical mass of 114.2 kDa expected for the BsRNAseY dimer (Table 1). The RI (refraction index) signal superimposed onto the LS (scattered

light) signal, which confirms that the dimer sample is monodisperse (Fig. 2A). The polydispersity at the peak summit of 0.268 mg/ml was low (1.027), and the hydrodynamic radius Rh value was 5.9 ± 0.3 nm. The mass and Rh values indicate a highly elongated form for the dimer, with a friction coefficient f/f0 of 1.8. The oligomer sample was contaminated by a species of higher molecular weight and strong absorbance, eluting at 8.5 ml on a Superose 6 10/300GL increase column (Fig. 2B), that could not be eliminated by SEC. Yet, the molar mass measured for the main peak of the oligomer, which was defined by the summit of the MALLS and RI curves, was 987.0 ± 21.7 kDa with a polydispersity of 1.009 at the peak plateau of 0.024 mg/ml, and a Rh value of 12-14 nm (Table 1, Fig. 2B). A spherical particle with this molar mass is expected to have a radius of 6.6 nm [29], which indicates that the oligomer behaves as an elongated particule in solution. Assuming a partial specific volume of 0.73 cm^3/g [29], the oligomer is calculated to occupy a volume of around 1,200,000 $Å^3$, similar to that of the 50 S ribosomal subunit [30].

To further examine the shape of the BsRNAseY dimer and oligomer in solution and assess whether the purified oligomer was an amorphous aggregate or ordered assembly, we recorded small-angle X-ray scattering data [31,32] of the eluting peaks from a BioSEC3 column (Fig. 3) [33]. The scattering intensity profiles I(q) *vs* q, where I(q) and q are the scattering intensity and vector, show that both samples were monodisperse (Supplementary Figure 2). The gyration radius Rg that monitors the extent of macromolecular elongation was calculated from the slope of the linear part of the Guinier plots, ln(I(q)) *vs* q^2 at very small angles (Supplementary Figure 2B & C) [34]. Rg values of 5.98 nm and of 10.15 nm were estimated for the dimer and oligomer, respectively (Table 1).
The Kratky plots $I(q).q^2$ *vs* q can be used to qualitatively identify disordered states and distinguish them from globular proteins [35]. The normalized Kratky plot of the dimer is characteristic of a folded protein with an extended conformation (Fig 3A, red curve)[32]. More structural information can be learned from the pair distance distribution function p(r), which is a real-space representation of the distribution of all interatomic distances within the molecule [31]. The p(r) function of the dimer displays a broad asymmetric shape with a clear maximum at small distances, which confirms the presence of a compact elongated species (Fig. 3B) [31,32]. The estimated maximum distance between atoms D_{max} of 22 nm corresponds to the radius of the cross-section of the dimer [31].
The normalized Kratky plot of the oligomer (Fig. 3A, black) is bell-shaped with a well-defined maximum and a normalized value near 1.104 at $q.R_g$ of 1.732 indicative of a folded globular structure [36]. The p(r) distribution (Fig. 3C) is also characteristic of a globular particule with maximal dimension D_{max} of 41.9 nm. Thus, the oligomer adopts a more globular conformation than the dimer.

Structural models of the BsRNaseY dimer and oligomer

To visualize the spatial organization of the BsRNAseY dimer, we performed *ab initio* shape reconstructions based on the SAXS p(r) function. As seen in Fig. 3D & E, the generated model for the dimer shows an extended molecule, 22.6 nm long and 6.8 nm wide, comprising two main domains, with χ^2 of 0.675 between the model curve and the experimental data.
To further assess the homogeneity of the oligomer sample and gain insights into its overall structure, we performed negative staining electron microscopy (EM). The micrographs show a homogenous sample of particles with globular form, which adsorb to the carbon support film in a limited number of preferred orientations (Fig. 4A). To examine in more details the particles from the electron micrographs, 7500 particles selected by hand were subjected to a reference free alignment and class assignment procedure in EMAN2 [37]. From the classification of the particle images into 32 classes, we kept the five best class-averages representing different views of the particles (Fig. 4B). The negative staining method creates stain clouds that surround the molecules, producing a strong contrast between the background and the particle [38]. The oligomer particles show a clear ring-like shape with an average radius of 10-15 nm and a central tunnel. In some classes (Fig. 4B, classes 0, 1, 5, 31; Fig. 4C), an eight-membered ring shape for the oligomer is observed, which suggests C8 symmetry for the particles. Construction of an initial model gave a close match between the 5 best class averages and the final projection classes (Fig. 4C).

Discussion

BsRNAseY was previously predicted to form dimers with a high tendency to oligomerize [18]. Here, we pursued the biochemical and biophysical analysis of the BsRNAseY dimer and provide the first characterization of a higher molecular weight form of BsRNAseY.

RNAseY was predicted to contain an intrinsically disordered domain (IDD) at the N-terminus, which was assumed to mediate oligomerization since a bacterial adenylate cyclase two-hybrid assay indicated that this domain showed the strongest self-associations [18]. However, our SAXS data indicate that the normalized Kratky plot of the dimer is not characteristic of an intrinsically disordered protein (IDP) nor an IDD-containing protein (Fig. 3A). In contrast, the SAXS data indicate that the dimeric form is folded into an extended conformation. The elongated shape of the dimer is confirmed by the non-globular behavior of the dimer during SEC-chromatography as well as the friction coefficient calculated from SEC-MALLS data (1.8). A slightly different f/f0 coefficient of 1.4 had previously been determined from sedimentation velocity analytical ultracentrifugation experiments [18]. Rh, as measured by dynamic light scattering, is defined as the radius of an equivalent hard sphere diffusing at the same rate as the molecule under observation, whereas Rg is defined as the mass weighted average distance from the core of a molecule to each mass element in the molecule. Shape information is provided by the Rg/Rh ratio. When molecules deviate from globular to non-spherical or elongated structures, Rg/Rh tend to values upwards of 0.775 [39]. This is the case for the dimer with a Rg/Rh value close to 1 (Table 1). Altogether, these results confirm that the BsRNAseY dimer adopts an elongated conformation and show that the N-terminal domain within the dimer is less disordered than predicted, as a result of intra- or inter-molecular interactions.

The oligomer, which could be separated from the dimer during BsRNaseY purification, appears to form a globular soluble complex. The three different methods that have been used to estimate the radius of the oligomer molecule, SEC-MALLS, SAXS and EM, converge to a value of 10-13 nm (Table 1). The Rg/Rh ratio of 0.779 for the oligomer is close to the 0.775 value characteristic for a globular protein. Single particle analysis of the oligomer complex observed by negative staining EM and construction of an initial model suggest that the oligomer has C8 symmetry and contains an inner tunnel. Oligomerization of the dimer as an octamer would give a molecular mass of 917 kDa, in close agreement with the value of 987 kDa for the oligomer obtained by SEC-MALLS. The discrepancy between the SAXS and EM models could be indicative of a different behavior of the studied particule in solution or on the EM grids, as previously reported in other cases [40].

The possible existence of an octameric ring assembly of BsRNaseY is reminiscent of the PNPase and exosome [41,42] structures. PNPases and exosomes are complexes containing 3' to 5' exoribonucleases that, like RNase Y, have important roles in processing decay and quality control of various RNA molecules [43,44]. The crystal structures of bacterial PNPases show a trimeric architecture with six RNase PH domains assembled into a hexameric ring-like conformation containing a central channel for RNA binding and degradation [45-47]. Human PNPase has the same architecture and is capped with a trimeric KH pore [48]. Interestingly, it has been reported that chloroplast PNPase elutes as a ~ 600 kDa homo-multimer enzyme complex that is distinct from a degradosome complex [49]. Current models for RNA degradation by PNPases propose that the RNA substrates are bound by the S1 domain and/or the KH domain and are further threaded into the central channel [48]. Similarly, the archaeal and human exosome cores contain nine proteins with six RNase PH-like proteins forming a hexameric ring-like structure and three S1/KH domain-containing proteins forming a pore at the top. Entrapment of the single-stranded RNA substrate in the central ~ 110 Å long and ~ 10-20 Å wide channel explains the processive degradation of extended RNAs [41,42]. A similar function of RNA binding and processing may possibly be attributed to the inner hole present in the BsRNaseY oligomer, which displays larger dimensions with a diameter of ~ 60 Å. This model fits to an exoendonuclease and could also explain the endonuclease activity because the dimensions of the hole are sufficient for RNA to be threaded through the hole until it reaches the endonucleolytic cleavage site.

Because the BsRNaseY oligomer sample was homogenous, a vitrified sample is amenable to structure determination by cryo-EM. A higher resolution model is currently being constructed using this technique to confirm whether the architecture of BsRNaseY is similar to that of other RNA-degrading complexes.

Material and methods

Cloning and overexpression of BsRNaseY.

The coding sequence of the BsRNaseY gene, starting with amino acid K26 (thus lacking the N-terminal transmembrane domain), was PCR-amplified from chromosomal DNA of strain SSB1002, a wild-type *B. subtilis* laboratory stock strain derived from strain 168, using oligonucleotides HP1182 (5' TGATTCACTCATATGAAAACCATTGCCGAAGCGAAAATTGCG) and HP116 (5' AGCTAGGATCCTTAGTGATGATGATGATGGTGTTTTGCATACTCTACGGCTCGAGTC). The 1.5 kb PCR fragment containing the 6xHis-tag sequence at the 3' end was inserted as an NdeI-BamHI fragment into the T7 expression vector pKYB1 plasmid (New England Biolabs) cleaved with NdeI and BglII. Recombinant plasmids from *E. coli* JM109 were then transformed into *E. coli* BL21(*DE3*) star codon$^+$ competent cells. One colony was used to inoculate 200 ml of Luria Broth medium supplemented with kanamycine (25 µg/ml) and chloramphenicol (30 µg/ml). This preculture (0.5%), grown overnight at 37°C at 180 rpm, was used to inoculate 2 x 3 liters of LB medium supplemented with the same antibiotics. Cultures were grown at 37°C to an OD_{600} of 0.8. Expression was then induced at 23°C by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM and incubation was extended overnight at 20°C. After centrifugation for 10 min at 5000 rpm and 4°C, cells were stored at −80°C.

Purification of two oligomeric forms of BsRNaseY.

Cells were resuspended in 1.5 ml/g of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol (buffer A) containing 20 µl/ml cOmplete™ Protease Inhibitor Cocktail (Roche), 10 mM β-mercaptoethanol, 2 µg/ml of RNase A (Thermofischer), 1 mg/ml lysozyme, 0.25% Tween 20, and disrupted by sonication. Cells debris were removed by ultracentrifugation at 100,000g for 1h at 4°C. The supernatant was loaded on a 15 mL Ni-NTA Superflow column (Qiagen) equilibrated in buffer A containing 20 mM imidazole, 0.25% Tween 20. The column was washed with 70 mM imidazole and eluted with 200 mM imidazole in buffer A. Fractions containing the protein were then loaded at 4.5 mg/ml on a Superdex 200 26/60 PG column (GE Healthcare) in buffer A. The two main peaks (called "dimer" and "oligomer") were further purified individually by SEC. Protein elution was followed by the optical density at 230 nm because of the rarity of hydrophobic residues in the BsRNaseY sequence. The dimer was concentrated and loaded at 2.5 mg/mL onto a Superdex 200 26/60 PG column in buffer A. The oligomer was concentrated to 3 mg/ml and loaded onto a Superose 6 10/300 GL columns in buffer A. The dimer and oligomer were finally concentrated to 5 mg/ml, using Amicon concentrators (30 kD cutoff, Milllipore). Before being aliquoted, frozen in liquid nitrogen and stored at -80°C, the dimer and the oligomer were analyzed for purity on a Superdex 200 10/300 GL increase or a Superose 6 10/300 GL columns in buffer A, respectively.

Endonuclease activity of BsRNaseY dimer and oligomer.

The catalytic activity of each form was tested with the *yitJ* mRNA substrate as described previously [4].

Negative staining electron microscopy.

For analysis using negative staining electron microscopy, the high molecular weight oligomer was freshly repurified on a Superose 6 10/300 GL gel filtration column in 20 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol and concentrated at 150 µg/ml. Samples were prepared using copper mashed grids covered with a thin film of carbon and made hydrophilic by glow discharge. A 3 µl drop of the protein diluted 1/3 in buffer without glycerol was deposited and left to adsorb on the grid for about 1 min. The drop was then blotted and washed twice on ultrapure water to avoid artefacts related to buffer composition. Finally, a drop of 2% uranyl acetate was added and left to adsorb for 30 sec, blotted and left dry before observation at the microscope. 40 images were collected at a nominal magnification of x 20,000 (image pixel size of 4.9 Å) and 1 sec exposure using a JEOL 2100 (200 kV LaB6) equipped with a camera Gatan US 1000 (2k x 2k). 7500 particles were peaked by hand and data were analyzed with EMAN2.2.

Static Light Scattering.
The two oligomeric forms of RNase Y were analyzed using a SEC-MALLS/UV/RI/quasi-elastic light scattering Wyatt instrument. 100 µl of the dimer at 5.6 mg/ml or the oligomer at 2 mg/ml were injected, and separation was performed using a Shimadzu liquid chromatography system (Prominence) equipped with a Superdex 200 10/300GL increase column (dimer) or a Superose 6 10/300GL increase column (oligomer) before light-scattering measurements. Isocratic elution was performed at a flow rate of 0.5 ml/min in 20 mM HEPES, pH 8.0, 500 mM NaCl, at 17°C. Light scattering was monitored using a mini-DAWN TREOS system equipped with a quasi-elastic light scattering module (Wyatt Technology, Santa Barbara, CA) and refractometer (Optilab T-rEX, Wyatt Technology). The concentration of the molecule in solution was determined using a specific refractive index increment (dn/dc) of 0.183 ml/g at 658 nm. The absolute refractive index of the solution was measured using the refractive index detector. The light scattering detector was calibrated with toluene. The detector was controlled and the data were analyzed using the ASTRA 6.1 software (Wyatt Technology).

SAXS data collection and analysis
SAXS data were collected on beam line SWING of the SOLEIL synchrotron [33]. The protein samples (20 mg/ml in 20 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol) were injected onto a BioSEC3 column (Agilent) and data were recorded on the eluted peak. The scattering data were processed using FOXTROT [33]. The gyration radius was calculated by the Guinier approximation using PRIMUS from the ATSAS suite. The GNOM program from the ATSAS suite [50] was used to calculate the pair distance distribution function P(r) and D_{max}. *Ab initio* shape reconstructions of the BsRNAseY dimer and oligomer were performed by 20 independent DAMMIF [51] runs averaged using DAMAVER [52] and further refined with DAMMIN [53].

ACKNOWLEDGEMENTS
This work was supported by the French State Program ‘Investissements d’Avenir’ (Grants “LABEX DYNAMO”, ANR-11-LABX-0011 and ANR-11-BSV8-026), and has benefited from the facilities and expertise of the Macromolecular Interaction Platform of I2BC. We thank SOLEIL for provision of synchrotron radiation facilities, Aurélien Agneray for maintenance of SWING beamline, Ludovic Pecqueur for participating in the initial stage of the project, Aude Pouclet for purifying the proteins and picking particles on the EM micrographs, Jean-Michel Guignier for taking EM pictures and Leonardo Feletto for advices about running the EMAN2 program.

REFERENCES

1. Laalami S, Putzer H: **mRNA degradation and maturation in prokaryotes: the global players**. *Biomol Concepts* 2011, **2**:491-506.
2. Laalami S, Zig L, Putzer H: **Initiation of mRNA decay in bacteria**. *Cell Mol Life Sci* 2014, **71**:1799-1828.
3. Mohanty BK, Kushner SR: **Regulation of mRNA Decay in Bacteria**. *Annu Rev Microbiol* 2016, **70**:25-44.
4. Shahbabian K, Jamalli A, Zig L, Putzer H: **RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis***. *EMBO J* 2009, **28**:3523-3533.
5. Durand S, Tomasini A, Braun F, Condon C, Romby P: **sRNA and mRNA turnover in Gram-positive bacteria**. *FEMS Microbiol Rev* 2015, **39**:316-330.
6. Condon C, Bechhofer DH: **Regulated RNA stability in the Gram positives**. *Curr Opin Microbiol* 2011, **14**:148-154.
7. Khemici V, Prados J, Linder P, Redder P: **Decay-Initiating Endoribonucleolytic Cleavage by RNase Y Is Kept under Tight Control via Sequence Preference and Sub-cellular Localisation**. *PLoS Genet* 2015, **11**:e1005577.
8. Obana N, Nakamura K, Nomura N: **Role of RNase Y in *Clostridium perfringens* mRNA Decay and Processing**. *J Bacteriol* 2017, **199**.
9. Lehnik-Habrink M, Schaffer M, Mader U, Diethmaier C, Herzberg C, Stulke J: **RNA processing in *Bacillus subtilis:* identification of targets of the essential RNase Y**. *Mol Microbiol* 2011, **81**:1459-1473.
10. Durand S, Gilet L, Bessieres P, Nicolas P, Condon C: **Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs**. *PLoS Genet* 2012, **8**:e1002520.
11. Laalami S, Bessieres P, Rocca A, Zig L, Nicolas P, Putzer H: ***Bacillus subtilis* RNase Y activity in vivo analysed by tiling microarrays**. *PLoS One* 2013, **8**:e54062.
12. Yao S, Bechhofer DH: **Initiation of decay of *Bacillus subtilis* rpsO mRNA by endoribonuclease RNase Y**. *J Bacteriol* 2010, **192**:3279-3286.
13. Bruscella P, Shahbabian K, Laalami S, Putzer H: **RNase Y is responsible for uncoupling the expression of translation factor IF3 from that of the ribosomal proteins L35 and L20 in *Bacillus subtilis***. *Mol Microbiol* 2011, **81**:1526-1541.
14. Hunt A, Rawlins JP, Thomaides HB, Errington J: **Functional analysis of 11 putative essential genes in *Bacillus subtilis***. *Microbiology* 2006, **152**:2895-2907.
15. Khemici V, Poljak L, Luisi BF, Carpousis AJ: **The RNase E of *Escherichia coli* is a membrane-binding protein**. *Mol Microbiol* 2008, **70**:799-813.
16. Aravind L, Koonin EV: **The HD domain defines a new superfamily of metal-dependent phosphohydrolases**. *Trends Biochem Sci* 1998, **23**:469-472.
17. Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, et al.: **HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase**. *Nature* 2011, **480**:379-382.
18. Lehnik-Habrink M, Newman J, Rothe FM, Solovyova AS, Rodrigues C, Herzberg C, Commichau FM, Lewis RJ, Stulke J: **RNase Y in *Bacillus subtilis*: a Natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli***. *J Bacteriol* 2011, **193**:5431-5441.
19. Commichau FM, Rothe FM, Herzberg C, Wagner E, Hellwig D, Lehnik-Habrink M, Hammer E, Volker U, Stulke J: **Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing**. *Mol Cell Proteomics* 2009, **8**:1350-1360.
20. Lehnik-Habrink M, Lewis RJ, Mader U, Stulke J: **RNA degradation in *Bacillus subtilis:* an interplay of essential endo- and exoribonucleases**. *Mol Microbiol* 2012, **84**:1005-1017.
21. Cascante-Estepa N, Gunka K, Stulke J: **Localization of Components of the RNA-Degrading Machine in *Bacillus subtilis***. *Front Microbiol* 2016, **7**:1492.

22. Salvo E, Alabi S, Liu B, Schlessinger A, Bechhofer DH: **Interaction of *Bacillus subtilis* Polynucleotide Phosphorylase and RNase Y: Structural Mapping and effect on mRNA turnover**. *J Biol Chem* 2016, **291**:6655-6663.
23. Cho KH: **The Structure and Function of the Gram-Positive Bacterial RNA Degradosome**. *Front Microbiol* 2017, **8**:154.
24. Lehnik-Habrink M, Pfortner H, Rempeters L, Pietack N, Herzberg C, Stulke J: **The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex**. *Mol Microbiol* 2010, **77**:958-971.
25. Mathy N, Hebert A, Mervelet P, Benard L, Dorleans A, de la Sierra-Gallay IL, Noirot P, Putzer H, Condon C: ***Bacillus subtilis* ribonucleases J1 and J2 form a complex with altered enzyme behaviour**. *Mol Microbiol* 2010, **75**:489-498.
26. Newman JA, Hewitt L, Rodrigues C, Solovyova AS, Harwood CR, Lewis RJ: **Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome**. *J Mol Biol* 2012, **416**:121-136.
27. Folta-Stogniew E: **Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors**. *Methods Mol Biol* 2006, **328**:97-112.
28. Sahin E, Roberts CJ: **Size-exclusion chromatography with multi-angle light scattering for elucidating protein aggregation mechanisms**. *Methods Mol Biol* 2012, **899**:403-423.
29. Erickson HP: **Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy**. *Biol Proced Online* 2009, **11**:32-51.
30. Voss NR, Gerstein M: **Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly**. *J Mol Biol* 2005, **346**:477-492.
31. Svergun DI, Koch MHJ: **Small-angle scattering studies of biological macromolecules in solution**. *Rep. Prog. Phys.* 2003, **66**:1735-1782.
32. Bernado P, Svergun DI: **Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering**. *Mol Biosyst* 2012, **8**:151-167.
33. David G, Pérez J: **Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline**. *J Appl Cryst* 2009, **42**:892-900.
34. Koch MH, Vachette P, Svergun DI: **Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution**. *Q Rev Biophys* 2003, **36**:147-227.
35. Kikhney AG, Svergun DI: **A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins**. *FEBS Lett* 2015, **589**:2570-2577.
36. Receveur-Brechot V, Durand D: **How random are intrinsically disordered proteins? A small angle scattering perspective**. *Curr Protein Pept Sci* 2012, **13**:55-75.
37. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ: **EMAN2: an extensible image processing suite for electron microscopy**. *J Struct Biol* 2007, **157**:38-46.
38. Ohi M, Li Y, Cheng Y, Walz T: **Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy**. *Biol Proced Online* 2004, **6**:23-34.
39. Burchard W: In *Light Scattering from Polymers*. Edited by Springer-Verlag: Cantow, H.-J. et al., Eds.; 1983:1-124.
40. Dias J, Renault L, Perez J, Mirande M: **Small-angle X-ray solution scattering study of the multi-aminoacyl-tRNA synthetase complex reveals an elongated and multi-armed particle**. *J Biol Chem* 2013, **288**:23979-23989.
41. Lorentzen E, Dziembowski A, Lindner D, Seraphin B, Conti E: **RNA channelling by the archaeal exosome**. *EMBO Rep* 2007, **8**:470-476.
42. Malet H, Topf M, Clare DK, Ebert J, Bonneau F, Basquin J, Drazkowska K, Tomecki R, Dziembowski A, Conti E, et al.: **RNA channelling by the eukaryotic exosome**. *EMBO Rep* 2010, **11**:936-942.
43. Briani F, Carzaniga T, Deho G: **Regulation and functions of bacterial PNPase**. *Wiley Interdiscip Rev RNA* 2016, **7**:241-258.
44. Houseley J, LaCava J, Tollervey D: **RNA-quality control by the exosome**. *Nat Rev Mol Cell Biol* 2006, **7**:529-539.

45. Symmons MF, Jones GH, Luisi BF: **A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation**. *Structure* 2000, **8**:1215-1226.
46. Shi Z, Yang WZ, Lin-Chao S, Chak KF, Yuan HS: **Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation**. *RNA* 2008, **14**:2361-2371.
47. Hardwick SW, Gubbey T, Hug I, Jenal U, Luisi BF: **Crystal structure of *Caulobacter crescentus* polynucleotide phosphorylase reveals a mechanism of RNA substrate channelling and RNA degradosome assembly**. *Open Biol* 2012, **2**:120028.
48. Lin CL, Wang YT, Yang WZ, Hsiao YY, Yuan HS: **Crystal structure of human polynucleotide phosphorylase: insights into its domain function in RNA binding and degradation**. *Nucleic Acids Res* 2012, **40**:4146-4157.
49. Baginsky S, Shteiman-Kotler A, Liveanu V, Yehudai-Resheff S, Bellaoui M, Settlage RE, Shabanowitz J, Hunt DF, Schuster G, Gruissem W: **Chloroplast PNPase exists as a homo-multimer enzyme complex that is distinct from the Escherichia coli degradosome**. *RNA* 2001, **7**:1464-1475.
50. Svergun DI: **Determination of the regularization parameter in indirect-transform methods using perceptual criteria** *J Appl Cryst* 1992, **25**:495-503.
51. Franke D, Svergun DI: **DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering**. *J Appl Cryst* 2009, **42**:342-346.
52. Volkov VV, Svergun DI: **Uniqueness of ab-initio shape determination in small-angle scattering**. *J Appl Cryst* 2003, **42**:342-346.
53. Svergun DI: **Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing**. *Biophys J* 1999, **76**:2879-2886.

Table 1. Biophysical parameters of the BsRNaseY dimer and oligomer.

Method	Parameter	Dimer	Oligomer
SEC-MALS	Molecular mass (kDa)	105.3 ± 1	987.0 ± 22
	Rh (nm)	5.9 ± 0.3	13 ± 1
	Polydispersity	1.027	1.009
SAXS	Rg (nm)	5.98 ± 0.42	10.15 ± 0.24
	Dmax (nm)	22.0	41.9
EM	Particle symmetry Molecular mass (kDa)	-	Octamer 916.8
	Particle radius (nm)	-	10.0-15.0

Figures legends

Figure 1: Purification of the RNAseY dimer and oligomer. A Purification of BsRNaseY analyzed on a 12% SDS-PAGE gel. 1 and 2: cells before and after IPTG induction; 3: Pellet; 4: Supernatant; 5: Nickel affinity column; 6: First size exclusion chromatography; 7: molecular weight standards. The position of the BsRNaseY monomer is indicated in red. **B** Separation of two main oligomeric forms of BsRNase Y, corresponding to the fractions limited by the dashed green arrows, on a Superdex 200 26/60 PG column (1: oligomer, 2: dimer). The void volume V_0 and the elution volumes of molecular weight standards (in kDa) are indicated by red and black arrows, respectively. **C** Analysis of the dimer homogeneity on a Superdex 200 10/300 GL increase column. **D** Analysis of the oligomer homogeneity on a Superose 6 10/300 GL column.

Figure 2: SEC-MALLS analysis of the BsRNAseY dimer (A) and oligomer (B). SEC-MALLS elution profiles of the dimer and oligomer. Dashed traces represent the signal recorded from the refractive index detector (RI) plotted as a function of the elution volume and dotted lines correspond to the scattered light recorded on the detector at 90°. The red trace indicates the molecular mass.

Figure 3: SAXS analysis of the RNAseY dimer and oligomer. A Normalized Kratky plot of BsRNAseY dimer and oligomer shown in black and red, respectively. **B** Pair distance distribution function p(r) of BsRNAseY dimer, indicating an elongated species. **C** Pair distance distribution function p(r) of BsRNAseY oligomer, indicating a compact species. **D** Superposition of the experimental (black) and theoretical scattering curves (red) of the *ab initio* shape model of the BsRNAseY dimer. **E** *Ab initio* model of the BsRNaseY dimer.

Figure 4: Structure of the BsRNaseY oligomer examined by single-particle electron microscopy. A Negative-stain electron micrograph of the BsRNaseY oligomer particles. A field of particles adsorbed onto a glow-discharged carbon grid was stained with 2% (w/v) uranyl acetate and processed for imaging. Particles are white on a dark background. Boxes (size 50 nm x 50 nm) used for particles picking are indicated in blue. **B** Classification of the BsRNaseY oligomer particles into 32 2D class averages. **C** Adjacent pairs of images showing the 5 representative 2D class averages and the projections of the initial model. The octameric symmetry of one class is highlighted by red circles. **D** Initial model of the BsRNaseY oligomer.

Figure 1

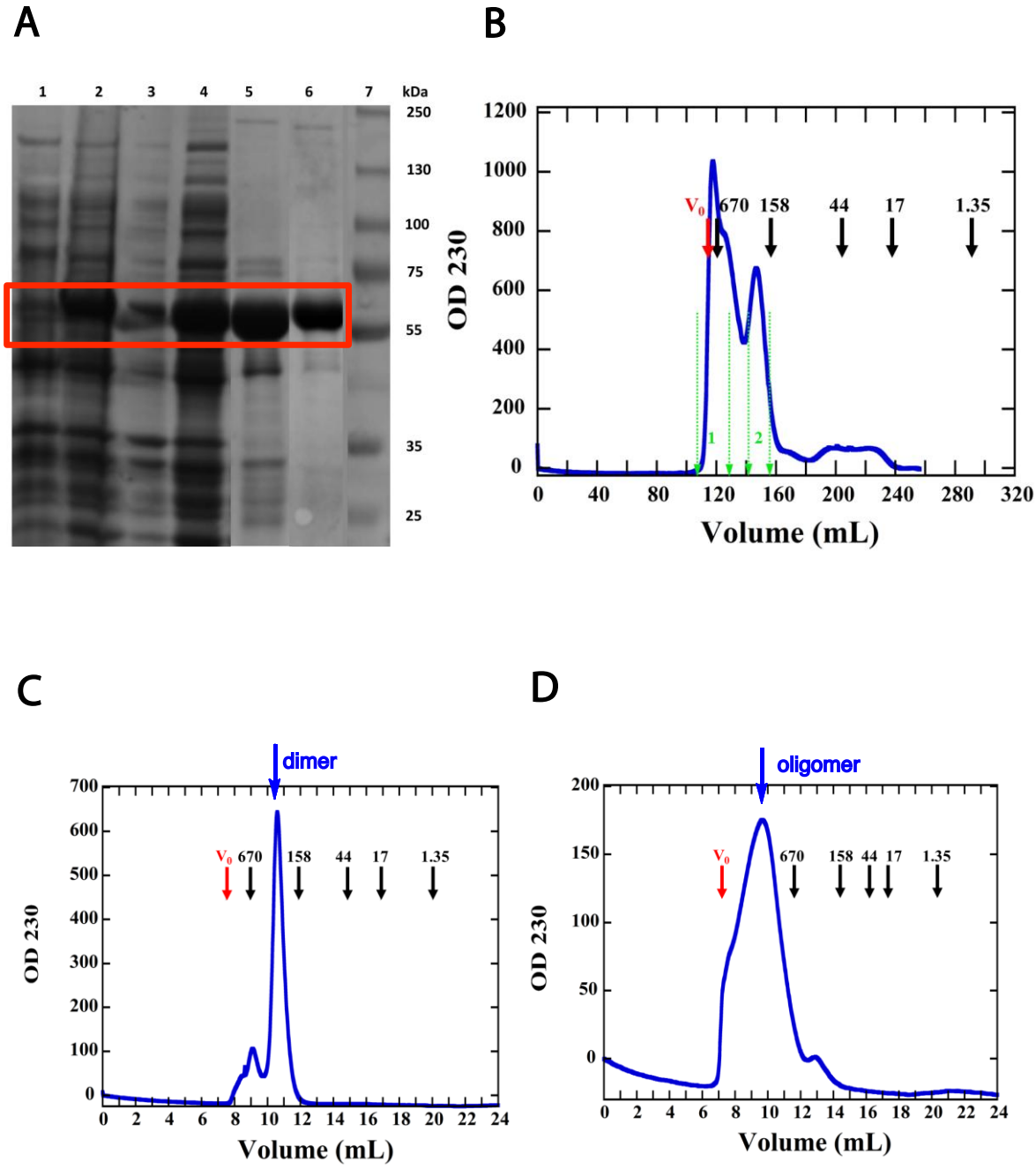
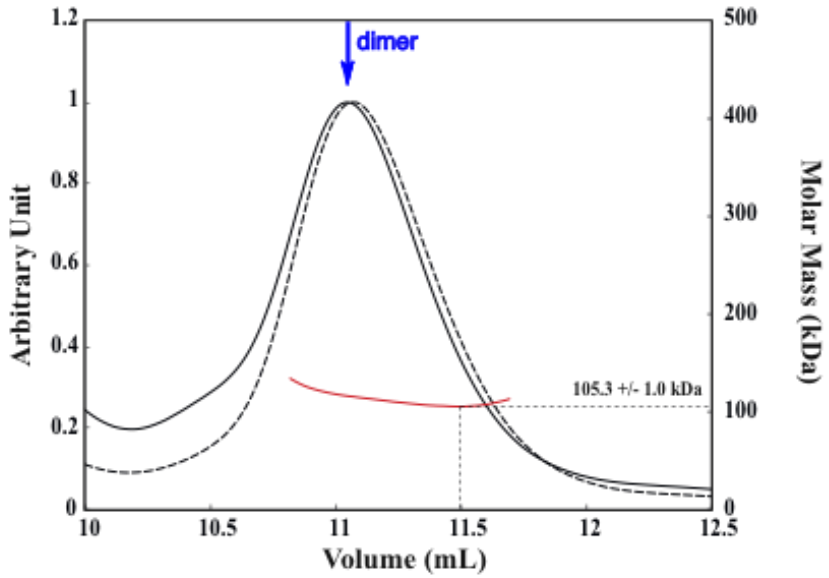

A
1
2
3
4
5
6
7
kDa
250
130
100
75
55
35
25
B
OD 230
Volume (mL)
V0
670
158
44
17
1.35
1
2
C
dimer
OD 230
Volume (mL)
V0
670
158
44
17
1.35
D
oligomer
OD 230
Volume (mL)
V0
670
158
44
17
1.35

Figure 2

A


dimer
105.3 +/- 1.0 kDa
Arbitrary Unit
Molar Mass (kDa)
Volume (mL)
1.2
1
0.8
0.6
0.4
0.2
0
500
400
300
200
100
0
10
10.5
11
11.5
12
12.5

B

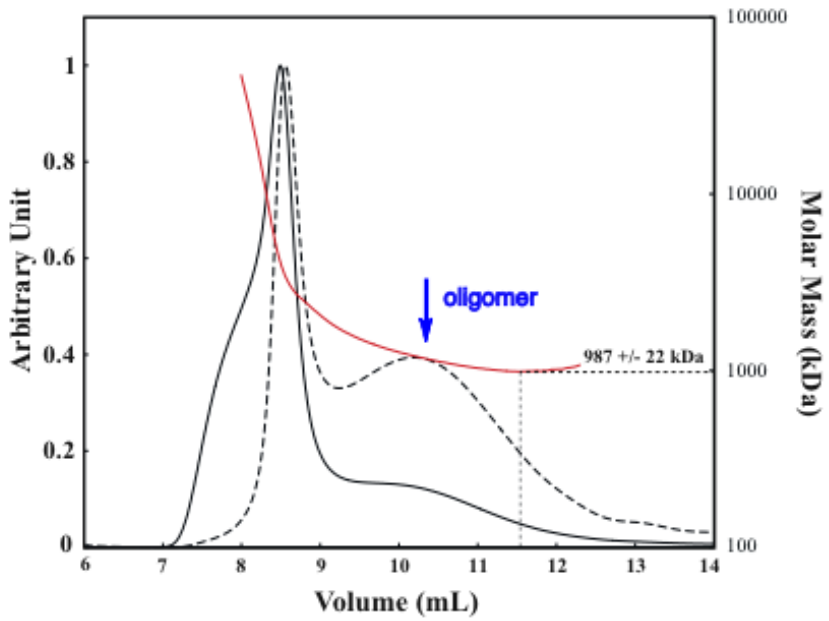

oligomer
987 +/- 22 kDa
Arbitrary Unit
Molar Mass (kDa)
Volume (mL)
1
0.8
0.6
0.4
0.2
0
100000
10000
1000
100
6
7
8
9
10
11
12
13
14

Figure 3

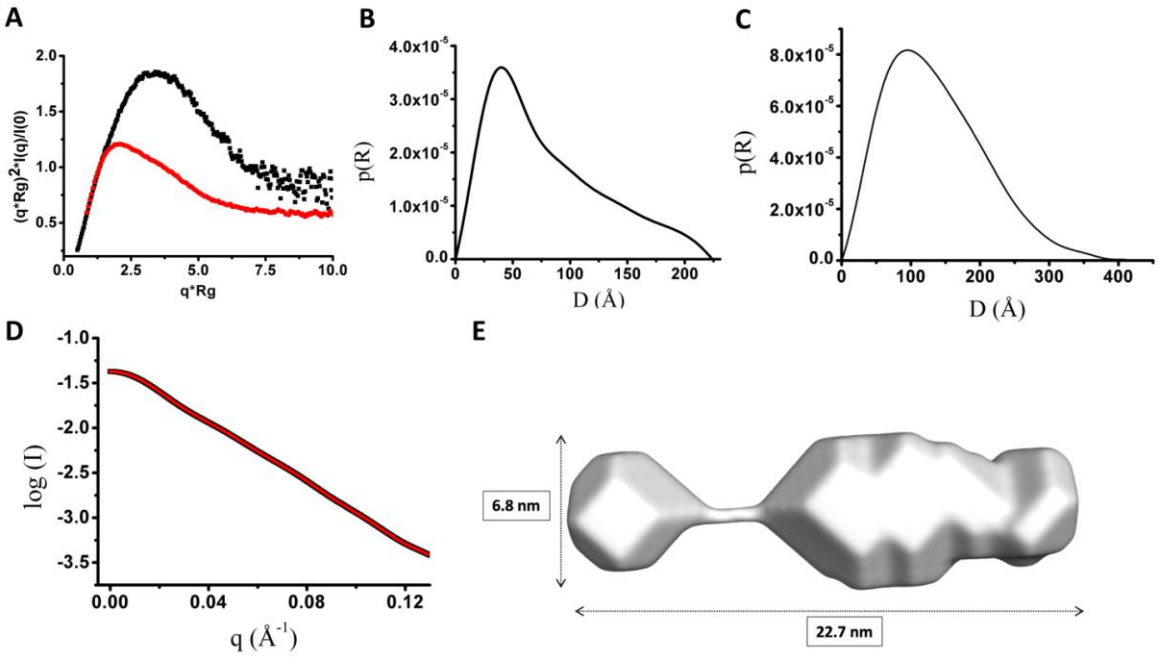
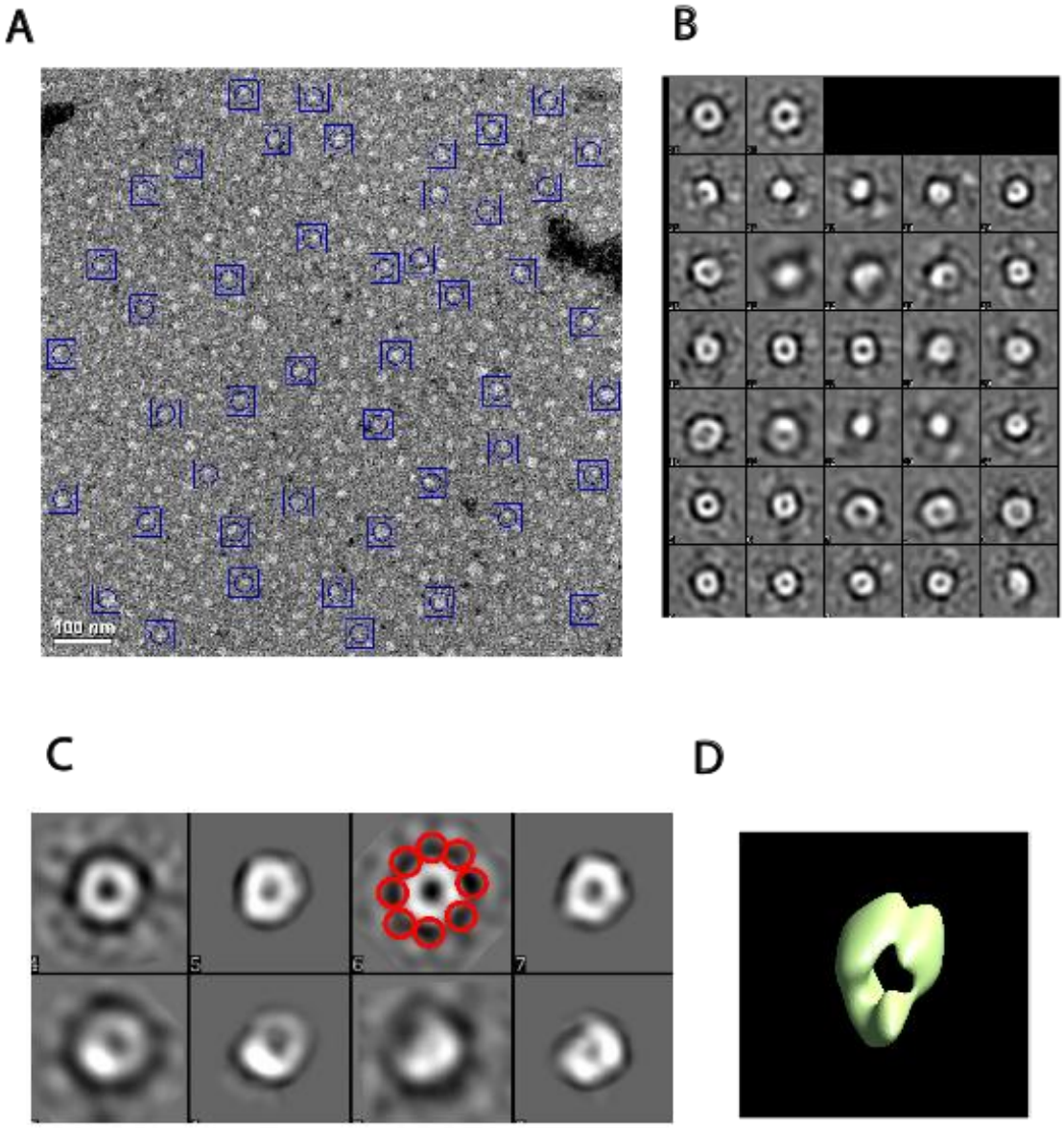
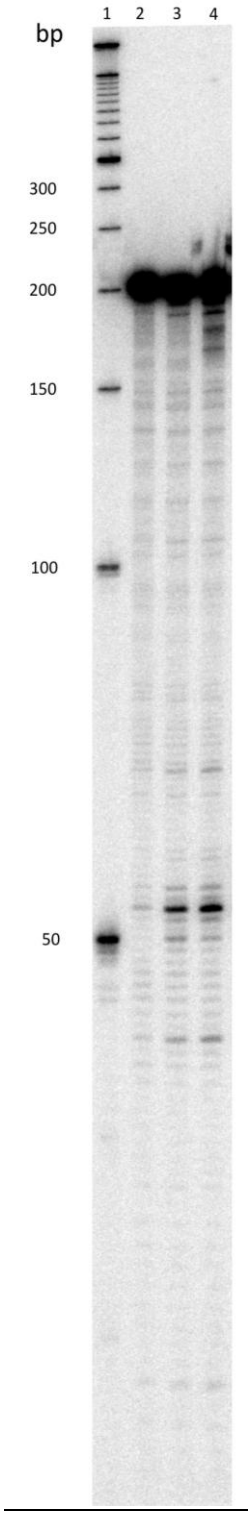


Figure 4

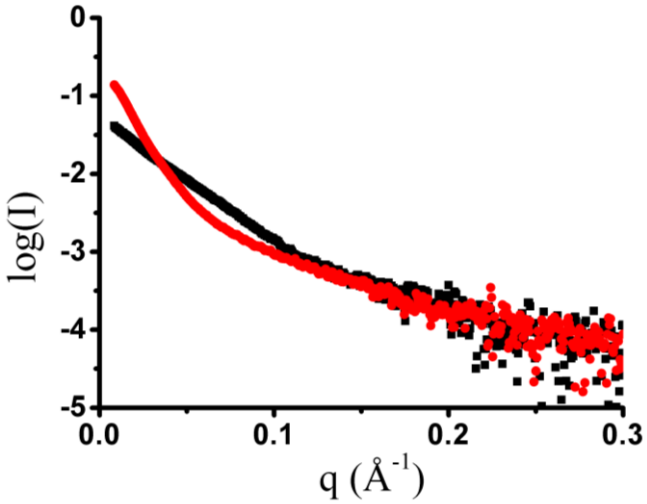


Supplementary Figure 1: Activity tests of the BsRNaseY dimer and oligomer. Endonuclease activity test of BsRNAseY (10 pmoles) using the 5' end-labelled monophosphorylated *yitJ* mRNA transcript (~200 bp). Cleavage of the mRNA substrate produces a major 52 nt fragment as well as other additional products. 1: 5'end-labelled 50 bp DNA ladder used as a size marker (Invitrogen); 2: control without enzyme; 3: Dimer; 4: Oligomer.

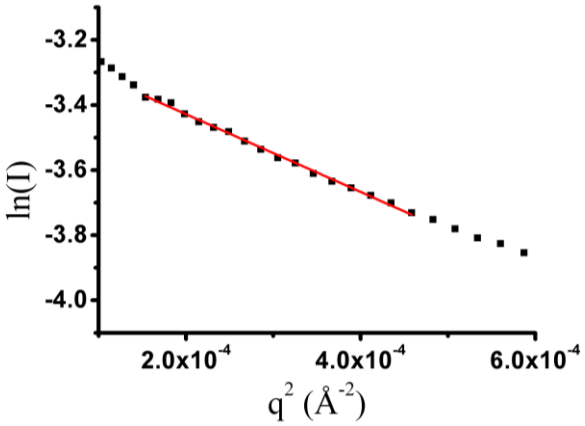


Supplementary Figure 2: Analysis of BsRNY by SEC-SAXS. A) X-ray scattering curves of the BsRNAseY dimer and oligomer, in red and black, respectively. B) Guinier plot of the BsRNAseY dimer. C) Guinier plot of the BsRNAseY oligomer.

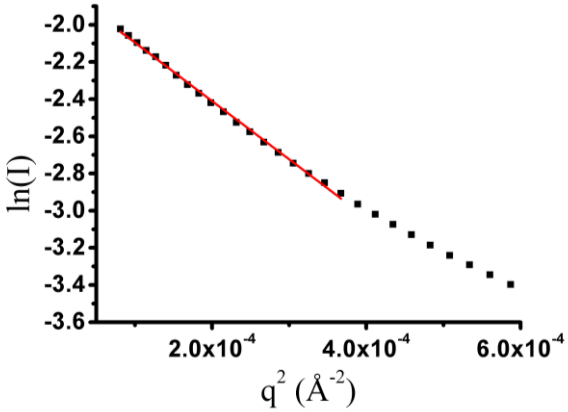
A



B



C



DISCUSSION et PERSPECTIVES

Mon projet d'étude structurale de la RNase Y était initialement basé sur la cristallographie des rayons X. Comme les différentes constructions réalisées n'ont pas cristallisées, je me suis tourné vers d'autres techniques biophysiques pour obtenir des informations structurales. Ainsi, les SEC-MALLS et le SAXS ont permis d'obtenir les masses molaires exactes des constructions et d'examiner l'état de structuration de ces formes. Les expériences de RMN devraient pouvoir être poursuivies sur le domaine N-terminal marqué (^{15}N, ^{13}C, ^{2}H) de manière à identifier les zones en hélice de ce domaine et celles en interaction avec les protéines partenaires du dégradosome. La cryo-microscopie électronique devrait pouvoir résoudre la structure de la forme oligomérique en solution à basse résolution. Les premiers résultats de microscopie électronique en coloration négative ont montré que la RNase Y de *B. subtilis* peut adopter une organisation oligomérique sous forme d'un anneau constitué probablement de 8 dimères et possédant un tunnel central. D'autres complexes de dégradation de l'ARN sont organisés de la même manière. Par exemple, la polynucléotide phosphorylase (PNPase) est une 3'-exonucléase dont la structure a été résolue chez de nombreux organismes vivants dont l'homme[31,86–88]. Cette enzyme est organisée sous forme d'un trimère formant un tunnel au centre duquel l'ARN se fixe pour y être dégradé. De même, l'exosome, un complexe multiprotéique de dégradation des ARN à activité exonucléase, présent chez les eucaryotes mais également chez les archées, a une organisation en anneau avec un tunnel central[44,163]. Les complexes PNPase et exosome ont une activité exonucléase alors que la RNase Y possède une activité endonucléase.

Les questions qui se posent alors sont les suivantes : Est-ce que cette organisation oligomérique de la RNase Y existe *in vivo* ? Si oui, pourquoi la cellule a choisi cette organisation oligomérique pour la RNase Y ? Enfin, comment expliquer l'activité endonucléase à partir d'une organisation oligomérique en anneau de la RNase Y ?

Voici quelques éléments de réponse. Tout d'abord, les structures cristallographiques de la PNPase et de l'exosome montrent un tunnel dont le diamètre est d'environ 10-20 Å[31,87], ne permettant d'accommoder qu'une seule molécule d'ARN, ce qui explique l'activité exonucléase de ces enzymes (figure 5, annexe). Le tunnel observé à la surface de l'oligomère de la RNase Y sur les images de coloration négative possède un diamètre beaucoup plus gros, d'environ 50 à 65 Å. C'est peut être la grande dimension de ce tunnel qui pourrait expliquer que la RNase Y puisse lier l'ARNm dans différentes conformations afin d'effectuer le clivage endonucléolytique. Ensuite, la RNase Y est localisée à la membrane[131] et l'absence de compartimentalisation dans la bactérie favorise peut être la formation de complexes de dégradation de façon à optimiser le renouvellement des acides nucléiques. De même, comme

nous l'avons mentionné dans l'introduction, la localisation de la RNase E à la membrane permet d'éviter une dégradation trop rapide de l'ARNm afin qu'il soit diffusé dans le cytoplasme après avoir été transcrit au niveau du nucléoïde. Après traduction, l'ARNm serait acheminé vers la membrane par des protéines liant l'ARN, comme la protéine Hfq, pour y être dégradé.

D'autre part, au sein du laboratoire d'Harald Putzer, il a été observé par microscopie à fluorescence que la RNase Y s'associe et se dissocie pour former des auto-assemblages de protéine à la membrane *in vivo*. Pour le moment, on ne comprend pas encore le rôle de ces *foci* de RNase Ydans la cellule mais il est possible que cette organisation de la RNase Y observée *in vivo* corresponde à celle observée *in vitro* en coloration négative.

Nous savons que la forme oligomérique de la RNase Y est active *in vitro*, maintenant il serait également intéressant de voir combien de molécule d'ARN sont liées par oligomère et savoir si l'ARNm se lie bien dans le tunnel. Cette étude est possible en utilisant des ARN marqués avec des nanobilles d'or[44]. En effet, par microscopie électronique, les molécules d'ARN seraient alors visibles en coloration négative.

D'après les résultats que nous avons obtenus en RMN, CD et SAXS concernant le domaine N-terminal de la RNase Y et des études montrant l'existence potentielle d'un dégradosome chez *B. subtilis* les questions suivantes se posent: Est-ce que ce domaine est bien impliqué dans des interactions avec des partenaires pour former un dégradosome chez *B. subtilis*, similaire à celui présent chez *E. coli* impliquant la RNase E ? Est-ce que ce domaine interagit avec l'ARN, comme c'est le cas pour le domaine ID de la RNase E ? Est-ce que l'hélicase CshA, la PNPase et l'énolase sont impliquées dans le dégradosome de *B. subtilis* ?
La RNase Y est deux fois plus courte que la RNase E (520 résidus pour la RNase Y contre 1061 pour la RNase E) ainsi que son domaine ID (160 résidus pour la RNase Y contre 560 résidus pour la RNase E). Ces deux domaines sont prédits pour être déstructurés par le logiciel PONDR. De même que pour la RNase Y, la RNase E possèderait un domaine coiled-coil entre les résidus 633-662 et 685-712, dont une partie interagit avec l'ARN (RBD)[83]. Il a été montré que le site d'interaction de l'hélicase RhlB se situe entre les deux domaines d'interaction avec l'ARN et juste après le domaine prédit en coiled-coil[75,83]. Nous savons qu'aucune structuration du site d'interaction de RhlB sur la RNase E n'a lieu lorsque RhlB se lie, via son domaine C-terminal. Cela a été montré par une étude faite en CD[83]. Notre étude en RMN et CD montre que le domaine N-terminal de la RNase Y possède une région déstructurée d'environ 70 acides aminés et une région organisée en hélice α.

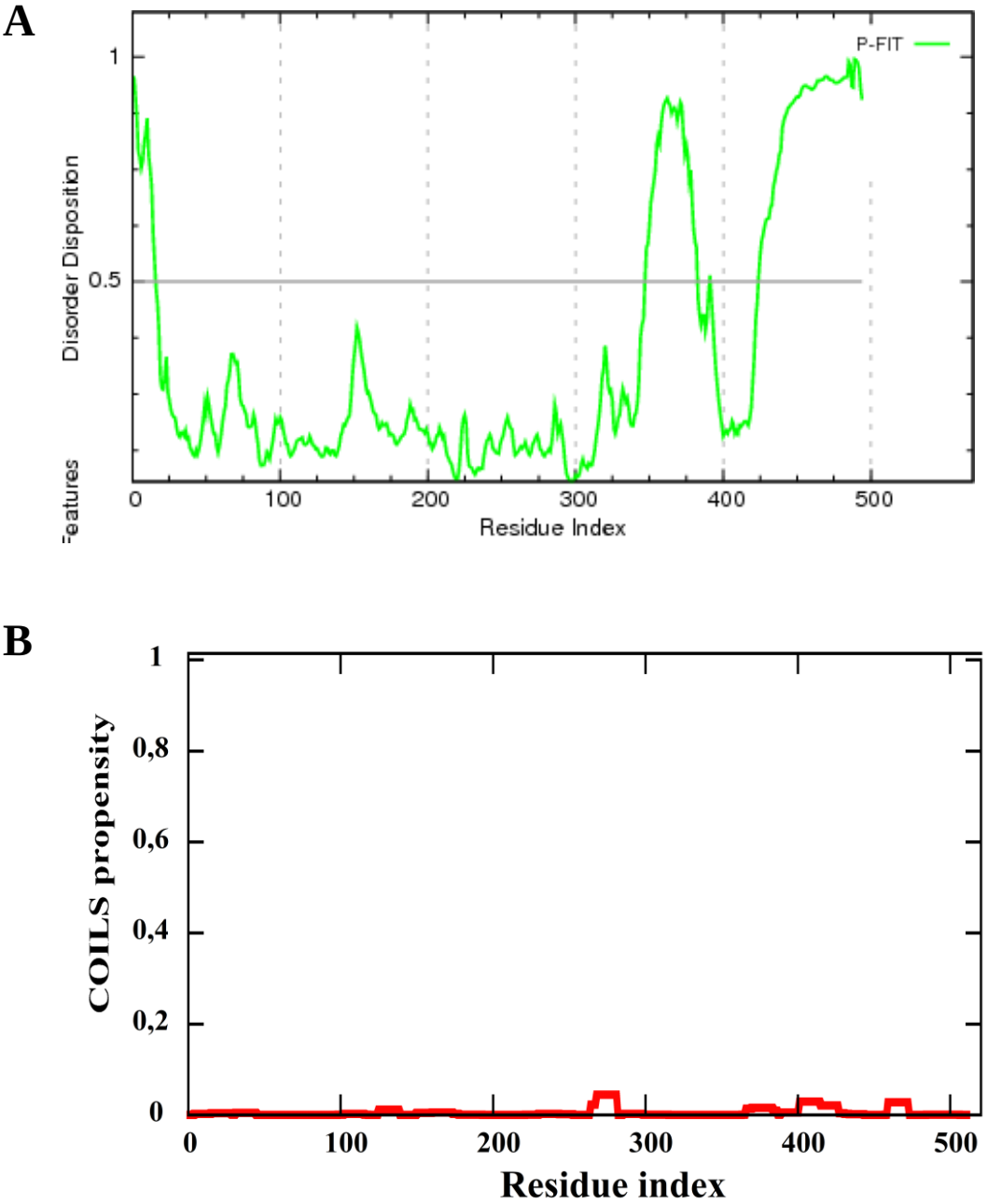


Figure 1: A. Prédiction du désordre au sein de la séquence de CshA de *B. subtilis* par le programme PONDR. **B.** Prédiction de structure de type coiled-coil au sein de la séquence de CshA par le programme COILS.

Le domaine C-terminal de l'hélicase CshA ainsi qu'une partie du domaine RecA n°2 sont prédits, via le serveur PONDR, pour être déstructurés (figure 1A) et nous savons que ces deux domaines sont impliqués dans l'interaction avec la RNase Y[146]. Par contre, le programme COILS ne prédit aucune structuration en domaine coiled-coil au sein de la séquence de CshA (figure 2B). Néanmoins, il est probable que CshA interagisse avec le domaine coiled-coil de la RNase Y via son domaine C-terminal prédit pour être déstructuré.
Il est donc intéressant de tester l'interaction de cette hélicase avec le domaine N-terminal de la RNase Y en RMN. Cette étude utilisant le domaine N-terminal triplement marqué (^{15}N, ^{13}C, D_2O) est en cours. Elle permettra de mettre en évidence une interaction entre les deux protéines, même faible. Si cette interaction existe, la comparaison des spectres RMN du domaine N-terminal seul et en présence d'hélicase permettra de déterminer quels acides aminés du domaine N-terminal de la RNase Y sont impliqués dans cette interaction.
Ensuite, comme le domaine prédit en coiled-coil de la RNase E interagit avec l'ARN[83], il serait intéressant également de regarder si le domaine N-terminal de la RNase Y interagit avec l'ARN. Nous pourrions tout d'abord déterminer s'il y a interaction avec l'ARN par gel retard.

Si l'ARN présente une affinité pour le domaine N-terminal seul, nous étudierons cette interaction du domaine N-terminal de la RNase Y marqué ^{13}C, ^{15}N, ^{2}H avec l'ARN en RMN afin de déterminer les acides aminés de la RNase Y impliqués dans cette interaction.

Pour terminer, il est connu que les domaines KH sont des domaines très flexibles[138]. En effet la structure du domaine KH dans des cristaux de PNPase n'a pu être résolue car la densité électronique pour ce domaine n'était pas visible, ce qui prouve sa flexibilité[31]. Le motif conservé des domaines KH interagissant avec l'ARN (VIGxxGxxI) forme une boucle flexible lorsqu'il n'est pas en interaction avec l'ARN[138]. Cette flexibilité est probablement une des raisons pour laquelle les deux constructions courtes de la RNase Y que nous avons réalisées (189-436 et 189-520) ne sont pas solubles lorsqu'elles sont exprimées en absence de chaperonnes(figures 1, annexe).

Enfin, durant ce projet, des essais de protéolyse ménagée ont été réalisés avec la protéine de *B. subtilis* (Δ1-23) portant une étiquette MBP (figure 2). La visualisation des produits après protéolyse ménagée à la chymotrypsine montre une série de bandes très proches les unes des autres entre 70 et 55 kDa et une bande nette à 25 kDa sur gel SDS-PAGE. Après passage sur résine d'amylose permettant de retenir la MBP, seule la protéine de poids moléculaire 25 kDa n'était pas retenue sur la colonne. Sachant que l'étiquette MBP est située en N-terminal, la bande d'environ 25kDa pourrait correspondre aux régions HD et Cter de la RNase Y et la série de bandes entre 70 et 55 kDa pourrait correspondre a différents

clivages de la région N-terminale comprenant le domaine N-terminal et le domaine KH.Il serait donc intéressant de refaire une construction de la RNase Y ne contenant que les domaines HD et C-terminal afin de faire des essais de cristallisation.

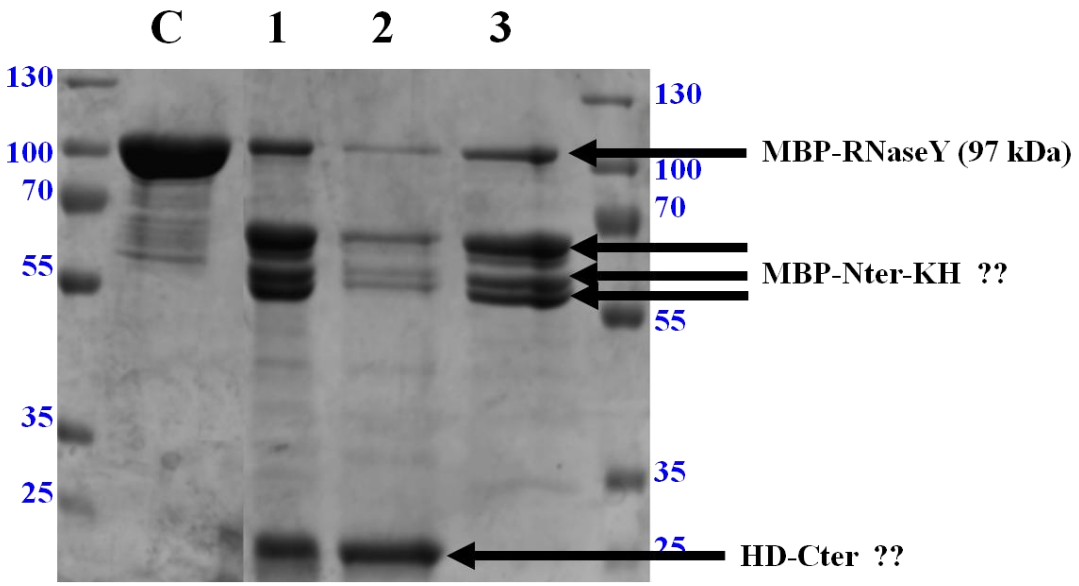


Figure 2: Protéolyse ménagée à la chymotrypsine de la RNase Y (Δ1-23) portant une étiquette MBP en N-terminal. Après digestion pendant 80 min à 25°C avec un ratio protéase:RNase Y de 1:2000, l'échantillon est chargé sur une résine d'amylose. (C= contrôle, protéine non digérée; puits 1= après protéolyse et avant passage sur la résine; puits 2 = fraction non retenue sur la résine; puits 3 = fraction éluée de la résine avec du maltose)

Masses théoriques calculées avec Protparam :

MBP-RNase Y (Δ1-23) = 97 kDa

HD-Cter = ~25-26 kDa

MBP-Nter = ~60,5 kDa

MBP-Nter-KH = ~71 kDa

Enfin, nous avons vu dans la partie III que des expériences de double hybride montrent que la PNPase semble interagir avec le domaine C-terminal de la RNase Y[50]. Une construction HD-Cter de la RNase Y marquée ^{15}N, ^{13}C, ^{2}H pourrait servir à étudier l'interaction de ce domaine avec la PNPase par RMN.

La RNase Y semble donc avoir un mécanisme de reconnaissance de l'ARN et de catalyse complexe. Les expériences obtenues suggèrent que ce domaine N-terminal soit impliqué dans la reconnaissance de l'ARNm ainsi que de l'hélicase CshA et que la PNPase interagirait avec la partie C-terminal de la RNase Y, le domaine N-terminal de la RNase Y étant beaucoup plus court que le domaine C-terminal de la RNase E, il est difficile d'imaginer que l'ensemble des partenaires puisse interagir en même temps pour former le dégradosome, sauf si ces interactions sont transitoires. En effet, si les interactions sont de faibles affinités mais très spécifiques, on peut imaginer que les partenaires puissent venir chacun leur tour interagir avec le domaine N-terminal, assurer leurs fonction au niveau de l'ARN, puis se dissocier afin de laisser la place au partenaire suivant pour continuer la dégradation de l'ARN. L'ARNm viendrait se fixer au niveau de la RNase Y, puis CshA viendrait interagir avec la RNase Y pour déplier les structures secondaires de l'ARN et permettre le clivage endonucléolytique par la RNase Y. CshA laisserait ensuite la place à la PNPase pour venir dégrader les fragments en 3'. Les hypothèses demandent à être verifiées par l'expérience

Bibliographie

1. Woese, C. R. There must be a prokaryote somewhere: microbiology's search for itself. *Microbiol. Rev.* **58,** 1–9 (1994).
2. Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci.* **95,** 6578–6583 (1998).
3. Lederberg, J. & Tatum, E. L. Gene recombination in Escherichia coli. *Nature.* **158**, 558 (1946)
4. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R. & Kolter, R. Sticking together: building a biofilm the Bacillus subtilis way. *Nat. Rev. Microbiol.* **11,** 157–168 (2013).
5. Bacillus subtilis Systems Biology: Applications of -Omics Techniques to the Study of Endospore Formation. in *The Bacterial Spore: from Molecules to Systems* (eds. Driks, A. & Eichenberger, P.) 129–144 (American Society of Microbiology, 2016). doi:10.1128/microbiolspec.TBS-0019-2013
6. Fischbach, M. A. & Walsh, C. T. Antibiotics for emerging pathogens. *Science* **325,** 1089–1093 (2009).
7. Walsh, C. Where will new antibiotics come from? *Nat. Rev. Microbiol.* **1,** 65–70 (2003).
8. Coates, A., Hu, Y., Bax, R. & Page, C. The future challenges facing the development of new antimicrobial drugs. *Nat. Rev. Drug Discov.* **1,** 895–910 (2002).
9. Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4,** 44 (2004).
10. Chaney, S. G. & Boyer, P. D. Incorporation of water oxygens into intracellular nucleotides and RNA: II. Predominantly hydrolytic RNA turnover in Escherichia coli. *J. Mol. Biol.* **64,** 581–591 (1972).

11. Duffy, J. J. , Chaney, S. G. & Boyer, P. D. Incorporation of water oxygens into intracellular nucleotides and RNA: I. Predominantly Non-hydrolytic RNA Turnover in *Bacillus subtilis. J. Mol. Biol.* **64**, 565-579 (1972)

12. Deutscher, M. P. & Reuven, N. B. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in Escherichia coli and Bacillus subtilis. *Proc. Natl. Acad. Sci.* **88,** 3277–3280 (1991).

13. Yang, W. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* **44,** 1–93 (2011).

14. Zuo, Y., Wang, Y. & Malhotra, A. Crystal Structure of Escherichia coli RNase D, an Exoribonuclease Involved in Structured RNA Processing. *Structure* **13,** 973–984 (2005).

15. Grossman, D. & van Hoof, A. RNase II structure completes group portrait of 3′ exoribonucleases. *Nat. Struct. Mol. Biol.* **13,** 760–761 (2006).

16. Frazão, C. *et al.* Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* **443,** 110–114 (2006).

17. Schmier, B. J., Seetharaman, J., Deutscher, M. P., Hunt, J. F. & Malhotra, A. The Structure and Enzymatic Properties of a Novel RNase II Family Enzyme from Deinococcus radiodurans. *J. Mol. Biol.* **415,** 547–559 (2012).

18. Gan, J. *et al.* Structural Insight into the Mechanism of Double-Stranded RNA Processing by Ribonuclease III. *Cell* **124,** 355–366 (2006).

19. de la Sierra-Gallay, I. L., Zig, L., Jamalli, A. & Putzer, H. Structural insights into the dual activity of RNase J. *Nat. Struct. Mol. Biol.* **15,** 206–212 (2008).

20. Court, D. L. *et al.* RNase III: Genetics and Function; Structure and Mechanism*. *Annu. Rev. Genet.* **47,** 405–431 (2013).

21. Kazantsev, A. V. *et al.* Crystal structure of a bacterial ribonuclease P RNA. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 13392–13397 (2005).

22. Harlow, L. S. Crystal structure of the phosphorolytic exoribonuclease RNase PH from Bacillus subtilis and implications for its quaternary structure and tRNA binding. *Protein Sci.* **13,** 668–677 (2004).

23. Vincent, H. A. & Deutscher, M. P. Insights into How RNase R Degrades Structured RNA: Analysis of the Nuclease Domain. *J. Mol. Biol.* **387,** 570–583 (2009).

24. Hsiao, Y.-Y. *et al.* Structural basis for RNA trimming by RNase T in stable RNA 3′-end maturation. *Nat. Chem. Biol.* **7,** 236–243 (2011).

25. Pellegrini, O., Li de la Sierra-Gallay, I., Piton, J., Gilet, L. & Condon, C. Activation of tRNA Maturation by Downstream Uracil Residues in B. subtilis. *Structure* **20,** 1769–1777 (2012).

26. Callaghan, A. J. *et al.* Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. *Nature* **437,** 1187–1191 (2005).

27. Shahbabian, K., Jamalli, A., Zig, L. & Putzer, H. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in Bacillus subtilis. *EMBO J.* **28,** 3523–3533 (2009).

28. Laalami, S., Zig, L. & Putzer, H. Initiation of mRNA decay in bacteria. *Cell. Mol. Life Sci.* **71,** 1799–1828 (2014).

29. Durand, S., Tomasini, A., Braun, F., Condon, C. & Romby, P. sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiol. Rev.* **39,** 316–330 (2015).

30. Portnoy, V., Palnizky, G., Yehudai-Resheff, S., Glaser, F. & Schuster, G. Analysis of the human polynucleotide phosphorylase (PNPase) reveals differences in RNA binding and response to phosphate compared to its bacterial and chloroplast counterparts. *RNA* **14,** 297–309 (2007).

31. Lin, C. L., Wang, Y.-T., Yang, W.-Z., Hsiao, Y.-Y. & Yuan, H. S. Crystal structure of human polynucleotide phosphorylase: insights into its domain function in RNA binding and degradation. *Nucleic Acids Res.* **40,** 4146–4157 (2012).

32. Gegenheimer, P., Watson, N. & Apirion, D. Multiple pathways for primary processing of ribosomal RNA in Escherichia coli. *J. Biol. Chem.* **252,** 3064–3073 (1977).

33. Robertson, D. Hugh., Webster, E. Robert. & Zinder D. Norton. Purification and Properties of Ribonuclease III from Escherichia coli. *J. Biol. Chem*. **243**, 82-91, (1968)

34. Kuwano, M. *et al.* Gene affecting longevity of messenger RNA: A mutant ofEscherichia coli with altered mRNA stability. *Mol. Gen. Genet. MGG* **154,** 279–285 (1977).

35. Misra, T. K. & Apirion, D. RNase E, an RNA processing enzyme from Escherichia coli. *J. Biol. Chem.* **254,** 11154–11159 (1979).

36. Ehretsmann, C. P., Carpousis, A. J. & Krisch, H. M. Specificity of Escherichia coli endoribonuclease RNase E: in vivo and in vitro analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes Dev.* **6,** 149–159 (1992).

37. Ow, M. C. & Kushner, S. R. Initiation of tRNA maturation by RNase E is essential for cell viability in E. coli. *Genes Dev.* **16,** 1102–1115 (2002).

38. Li, Z. & Deutscher, M. P. RNase E plays an essential role in the maturation of Escherichia coli tRNA precursors. *Rna* **8,** 97–109 (2002).

39. Kunst F. et al., The complete genome sequence of the Gram-positive bacterium Bacillus subtilis, *Nature*, **390**, 249-256 (1997)

40. Blattner, F. R. *et al.* The complete genome sequence of Escherichia coli K-12. *science* **277,** 1453–1462 (1997).

41. Condon, C., Putzer, H., & Grunberg-Manago M. Processing of the leader mRNA plays a major role in the induction of thrS expression following threonine starvation in Bacillus subtilis. *Proc. Natl. Acad. Sci.* **93**, 6992-6997 (1996)

42. Condon, C., Putzer, H., Luo, D. & Grunberg-Manago, M. Processing of the Bacillus subtilis thrS leader mRNA is RNase E-dependent in Escherichia coli. *J. Mol. Biol.* **268,** 235–242 (1997).

43. Deikus, G., Condon, C. & Bechhofer, D. H. Role of Bacillus subtilis RNase J1 Endonuclease and 5'-Exonuclease Activities in trp Leader RNA Turnover. *J. Biol. Chem.* **283,** 17158–17167 (2008).

44. Malet, H. *et al.* RNA channelling by the eukaryotic exosome. *EMBO Rep.* **11,** 936–942 (2010).

45. Britton, R. A. *et al.* Maturation of the 5' end of Bacillus subtilis 16S rRNA by the essential ribonuclease YkqC/RNase J1. *Mol. Microbiol.* **63,** 127–138 (2007).

46. Mathy, N. *et al.* 5′-to-3′ Exoribonuclease Activity in Bacteria: Role of RNase J1 in rRNA Maturation and 5′ Stability of mRNA. *Cell* **129,** 681–692 (2007).

47. Even, S. Ribonucleases J1 and J2: two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. *Nucleic Acids Res.* **33,** 2141–2152 (2005).

48. Mathy, N. *et al. Bacillus subtilis* ribonucleases J1 and J2 form a complex with altered enzyme behaviour. *Mol. Microbiol.* **75,** 489–498 (2010).

49. Kim, J. N. & Breaker, R. R. Purine sensing by riboswitches. *Biol. Cell* **100,** 1–11 (2008).

50. Commichau, F. M. *et al.* Novel Activities of Glycolytic Enzymes in Bacillus subtilis INTERACTIONS WITH ESSENTIAL PROTEINS INVOLVED IN mRNA PROCESSING. *Mol. Cell. Proteomics* **8,** 1350–1360 (2009).

51. Lehnik-Habrink, M. *et al.* RNA processing in Bacillus subtilis: identification of targets of the essential RNase Y: Targets of the essential RNase Y in Bacillus subtilis. *Mol. Microbiol.* **81,** 1459–1473 (2011).

52. Durand, S., Gilet, L., Bessières, P., Nicolas, P. & Condon, C. Three Essential Ribonucleases—RNase Y, J1, and III—Control the Abundance of a Majority of Bacillus subtilis mRNAs. *PLoS Genet.* **8,** e1002520 (2012).

53. Laalami, S. *et al.* Bacillus subtilis RNase Y Activity In Vivo Analysed by Tiling Microarrays. *PLoS ONE* **8,** e54062 (2013).

54. Park, C., Schultz, L. W. & Raines, R. T. Contribution of the Active Site Histidine Residues of Ribonuclease A to Nucleic Acid Binding †. *Biochemistry* **40,** 4949–4956 (2001).

55. Palermo, G. *et al.* Catalytic Metal Ions and Enzymatic Processing of DNA and RNA. *Acc. Chem. Res.* **48,** 220–228 (2015).

56. Laalami, S. & Putzer, H. mRNA degradation and maturation in prokaryotes: the global players. *Biomol. Concepts* **2,** (2011).

57. Condon, C. & Bechhofer, D. H. Regulated RNA stability in the Gram positives. *Curr. Opin. Microbiol.* **14,** 148–154 (2011).

58. Deutscher, M. P. Degradation of Stable RNA in Bacteria. *J. Biol. Chem.* **278,** 45041–45044 (2003).

59. Deutscher, M. P. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.* **34,** 659–666 (2006).

60. Motorin, Y. & Helm, M. tRNA Stabilization by Modified Nucleotides. *Biochemistry* **49,** 4934–4944 (2010).

61. Deana, A. & Belasco, J. G. Lost in translation: the influence of ribosomes on bacterial mRNA decay. Genes & Development **19**, 2526-2533 (2005)

62. Condon, C. RNA Processing and Degradation in Bacillus subtilis. *Microbiol. Mol. Biol. Rev.* **67,** 157–174 (2003).

63. Richards, J. *et al.* An RNA Pyrophosphohydrolase Triggers 5′-Exonucleolytic Degradation of mRNA in Bacillus subtilis. *Mol. Cell* **43,** 940–949 (2011).

64. Hsieh, P. -k., Richards, J., Liu, Q. & Belasco, J. G. Specificity of RppH-dependent RNA degradation in Bacillus subtilis. *Proc. Natl. Acad. Sci.* **110,** 8864–8869 (2013).

65. Py, P. et al., A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. *Letters to Nature.* **381**, 169-172 (1996)

66. Mackie, G. A. RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.* **11,** 45–57 (2012).

67. Piton, J. *et al.* Bacillus subtilis RNA deprotection enzyme RppH recognizes guanosine in the second position of its substrates. *Proc. Natl. Acad. Sci.* **110,** 8858–8863 (2013).

68. Hakim, J. A., Koo, H., van Elsas, J. D., Trevors, J. T. & Bej, A. K. CRISPR–Cas system: a new paradigm for bacterial stress response through genome rearrangement. *Stress Environ. Regul. Gene Expr. Adapt. Bact. 2 Vol. Set* **2,** (2016).

69. Geissmann, T., Marzi, S. & Romby, P. The role of mRNA structure in translational control in bacteria. *RNA Biol.* **6,** 153–160 (2009).

70. Naville, M. & Gautheret, D. Transcription attenuation in bacteria: theme and variations. *Brief. Funct. Genomic. Proteomic.* **8,** 482–492 (2009).

71. Arraiano, C. M. *et al.* The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* **34,** 883–923 (2010).

72. Jahn, N. & Brantl, S. One antitoxin—two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res.* **41,** 9870–9880 (2013).

73. Aït-Bara, S. & Carpousis, A. J. RNA degradosomes in bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs: RNA degradosome evolution. *Mol. Microbiol.* **97,** 1021–1135 (2015).

74. Aït-Bara, S., Carpousis, A. J. & Quentin, Y. RNase E in the γ-Proteobacteria: conservation of intrinsically disordered noncatalytic region and molecular evolution of microdomains. *Mol. Genet. Genomics* **290,** 847–862 (2015).

75. Vanzo, N. F. *et al.* Ribonuclease E organizes the protein interactions in the Escherichia coli RNA degradosome. *Genes Dev.* **12,** 2770–2781 (1998).

76. Koslover, D. J. *et al.* The Crystal Structure of the Escherichia coli RNase E Apoprotein and a Mechanism for RNA Degradation. *Structure* **16,** 1238–1244 (2008).

77. Newman, J. A. *et al.* Unusual, Dual Endo- and Exonuclease Activity in the Degradosome Explained by Crystal Structure Analysis of RNase J1. *Structure* **19,** 1241–1251 (2011).

78. Nagarajan, V. K., Jones, C. I., Newbury, S. F. & Green, P. J. XRN 5′→3′ exoribonucleases: Structure, mechanisms and functions. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1829,** 590–603 (2013).

79. Dorléans, A. *et al.* Molecular Basis for the Recognition and Cleavage of RNA by the Bifunctional 5′–3′ Exo/Endoribonuclease RNase J. *Structure* **19,** 1252–1261 (2011).

80. Palzkill, T. Metallo-β-lactamase structure and function: Metallo-β-lactamase structure and function. *Ann. N. Y. Acad. Sci.* **1277,** 91–104 (2013).

81. Py, B., Causton, H., Mudd, E. A. & Higgins, C. F. A protein complex mediating mRNA degradation in Escherichia coli. *Mol. Microbiol.* **14,** 717–729 (1994).

82. Carpousis, A. J. et al., Copurification of E. coli RNAase E and PNPase : Evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell.* **76**, 889-900 (1994)

83. Callaghan, A. J. *et al.* Studies of the RNA Degradosome-organizing Domain of the Escherichia coli Ribonuclease RNase E. *J. Mol. Biol.* **340,** 965–979 (2004).

84. Kühnel, K. & Luisi, B. F. Crystal structure of the Escherichia coli RNA degradosome component enolase. *J. Mol. Biol.* **313,** 583–592 (2001).

85. Chandran, V. & Luisi, B. F. Recognition of Enolase in the Escherichia coli RNA Degradosome. *J. Mol. Biol.* **358,** 8–15 (2006).

86. Symmons, M. F., Jones, G. H. & Luisi, B. F. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* **8,** 1215–1226 (2000).

87. Shi, Z., Yang, W.-Z., Lin-Chao, S., Chak, K.-F. & Yuan, H. S. Crystal structure of Escherichia coli PNPase: Central channel residues are involved in processive RNA degradation. *RNA* **14,** 2361–2371 (2008).

88. Hardwick, S. W., Gubbey, T., Hug, I., Jenal, U. & Luisi, B. F. Crystal structure of Caulobacter crescentus polynucleotide phosphorylase reveals a mechanism of RNA substrate channelling and RNA degradosome assembly. *Open Biol.* **2,** 120028–120028 (2012).

89. Nurmohamed, S., Vaidialingam, B., Callaghan, A. J. & Luisi, B. F. Crystal Structure of Escherichia coli Polynucleotide Phosphorylase Core Bound to RNase E, RNA and Manganese: Implications for Catalytic Mechanism and RNA Degradosome Assembly. *J. Mol. Biol.* **389,** 17–33 (2009).

90. Lykke-Andersen, S., Brodersen, D. E. & Jensen, T. H. Origins and activities of the eukaryotic exosome. *J. Cell Sci.* **122,** 1487–1494 (2009).

91. Hardwick, S. W. & Luisi, B. F. Rarely at rest: RNA helicases and their busy contributions to RNA degradation, regulation and quality control. *RNA Biol.* **10,** 56–70 (2013).

92. Linder, P. & Jankowsky, E. From unwinding to clamping — the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* **12,** 505–516 (2011).

93. Chandran, V. *et al.* Recognition and Cooperation Between the ATP-dependent RNA Helicase RhlB and Ribonuclease RNase E. *J. Mol. Biol.* **367,** 113–132 (2007).

94. Khemici, V., Poljak, L., Luisi, B. F. & Carpousis, A. J. The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol. Microbiol.* (2008). doi:10.1111/j.1365-2958.2008.06454.x

95. Taghbalout. A, & Rothfield. L., RNase E and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. *PNAS.* **104**, 1667-1672 (2007)

96. Taghbalout, A. & Rothfield, L. RNaseE and RNA Helicase B Play Central Roles in the Cytoskeletal Organization of the RNA Degradosome. *J. Biol. Chem.* **283,** 13850–13855 (2008).

97. Bandyra, K. J., Bouvier, M., Carpousis, A. J. & Luisi, B. F. The social fabric of the RNA degradosome. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1829,** 514–522 (2013).

98. Worrall, J. A. R. *et al.* Reconstitution and Analysis of the Multienzyme Escherichia coli RNA Degradosome. *J. Mol. Biol.* **382,** 870–883 (2008).

99. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293,** 321–331 (1999).

100. van der Lee, R. *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **114,** 6589–6631 (2014).

101. Kissinger, R. et al., Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Letters to Nature*. **378**, 641-644 (1995)

102. Huber, R. & Bennett, W. S. Functional significance of flexibility in proteins. *Biopolymers* **22,** 261–279 (1983).

103. Pande, V. S. & Rokhsar, D. S. Is the molten globule a third phase of proteins? *Proc. Natl. Acad. Sci.* **95,** 1490–1494 (1998).

104. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1804,** 1231–1264 (2010).

105. Receveur-Bréchot, V., Bourhis, J.-M., Uversky, V. N., Canard, B. & Longhi, S. Assessing protein disorder and induced folding. *Proteins Struct. Funct. Bioinforma.* **62,** 24–45 (2005).

106. Receveur-Bréchot, V. & Durand, D. How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.* **13,** 55–75 (2012).

107. Kosol, S., Contreras-Martos, S., Cedeño, C. & Tompa, P. Structural Characterization of Intrinsically Disordered Proteins by NMR Spectroscopy. *Molecules* **18,** 10802–10828 (2013).

108. Konrat, R. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J. Magn. Reson.* **241,** 74–85 (2014).

109. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27,** 527–533 (2002).

110. Campen, A. *et al.* TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.* **15,** 956–963 (2008).

111. Theillet, F.-X. *et al.* The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord. Proteins* **1,** e24360 (2013).

112. Mathura, V. *et al.* The protein non-folding problem: amino acid determinants of intrinsic order and disorder. in *Pacific Symposium on Biocomputing* **6,** 89–100 (2001).

113. GARNER, E., YOUNG, J. & TAKAYAMA, S. The protein non-folding problem: amino acid determinants of intrinsic order and disorder. *Biocomput. 2001* 89 (2000).

114. Uversky, V. N. What does it mean to be natively unfolded? *Eur. J. Biochem.* **269,** 2–12 (2002).

115. Meng, F., Uversky, V. N. & Kurgan, L. Comprehensive review of methods for prediction of intrinsic disorder and its molecular functions. *Cell. Mol. Life Sci.* (2017). doi:10.1007/s00018-017-2555-4

116. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. & Uversky, V. N. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1804,** 996–1010 (2010).

117. Wintjens, R. *et al.* ^{1}H NMR Study on the Binding of Pin1 Trp-Trp Domain with Phosphothreonine Peptides. *J. Biol. Chem.* **276,** 25150–25156 (2001).

118. Russo, A. A. *et al.*, Crystal structure of the p27 cyclin-dependent-kinase inhibitor bound to the cyclinA-Cdk2 complex. *Nature.* **382**, 325-331 (1996)

119. Guharoy, M., Pauwels, K. & Tompa, P. SnapShot: Intrinsic Structural Disorder. *Cell* **161,** 1230–1230.e1 (2015).

120. Acuner Ozbabacan, S. E., Engin, H. B., Gursoy, A. & Keskin, O. Transient protein-protein interactions. *Protein Eng. Des. Sel.* **24,** 635–648 (2011).

121. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* **18,** 1233–1243 (2010).

122. Oldfield, C. J. & Dunker, A. K. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu. Rev. Biochem.* **83,** 553–584 (2014).

123. Dunker, A. K. et al., Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*. **19**, 26-59 (2001)

124. Serrière, J. *et al.* Fab′-Induced Folding of Antigenic N-Terminal Peptides from Intrinsically Disordered HIV-1 Tat Revealed by X-ray Crystallography. *J. Mol. Biol.* **405,** 33–42 (2011).

125. Xie, H. *et al.* Functional Anthology of Intrinsic Disorder. 1. Biological Processes and Functions of Proteins with Long Disordered Regions. *J. Proteome Res.* **6,** 1882–1898 (2007).

126. Pancsa, R. & Tompa, P. Structural Disorder in Eukaryotes. *PLoS ONE* **7,** e34687 (2012).

127. Dunker, A. K., Romero, P., Obradovic, Z., Garner, E. C. & Brown, C. J. Intrinsic protein disorder in complete genomes. *Genome Inform.* **11,** 161–171 (2000).

128. Schad, E., Tompa, P. & Hegyi, H. The relationship between proteome size, structural disorder and organism complexity. *Genome Biol.* **12,** R120 (2011).

129. Aravind, L. & Koonin, E. V. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* **23,** 469–472 (1998).

130. Kobayashi, K. *et al.* Essential Bacillus subtilis genes. *Proc. Natl. Acad. Sci.* **100,** 4678–4683 (2003).

131. Hunt, A. Functional analysis of 11 putative essential genes in Bacillus subtilis. *Microbiology* **152,** 2895–2907 (2006).

132. Nicholson, A. W. *Ribonucleases, Part A: Functional Roles and Mechanisms of Action*. **341,** (Academic Press, 2001).

133. Goldstone, D. C. *et al.* HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* **480,** 379–382 (2011).

134. Arnold, L. H. *et al.* Phospho-dependent regulation of SAMHD1 oligomerisation couples catalysis and restriction. *PLoS Pathog* **11,** e1005194 (2015).

135. Zimmerman, M. D., Proudfoot, M., Yakunin, A. & Minor, W. Structural Insight into the Mechanism of Substrate Specificity and Catalytic Activity of an HD-Domain Phosphohydrolase: The 5′-Deoxyribonucleotidase YfbR from Escherichia coli. *J. Mol. Biol.* **378,** 215–226 (2008).

136. Vorontsov, I. I. *et al.* Mechanisms of Allosteric Activation and Inhibition of the Deoxyribonucleoside Triphosphate Triphosphohydrolase from *Enterococcus faecalis*. *J. Biol. Chem.* **289,** 2815–2824 (2014).

137. Nagata, M., Kaito, C. & Sekimizu, K. Phosphodiesterase Activity of CvfA Is Required for Virulence in Staphylococcus aureus. *J. Biol. Chem.* **283,** 2176–2184 (2008).

138. Valverde, R., Edwards, L. & Regan, L. Structure and function of KH domains: Structure and function of KH domains. *FEBS J.* **275,** 2712–2726 (2008).

139. Cléry, A. & Allain, F. H.-T. From structure to function of rna binding domains. (2000).

140. Lehnik-Habrink, M. *et al.* RNase Y in Bacillus subtilis: a Natively Disordered Protein That Is the Functional Equivalent of RNase E from Escherichia coli. *J. Bacteriol.* **193,** 5431–5441 (2011).

141. Lupas, A., Van Dyke, M. & Stock, J. Predicting coiled coils from protein sequences. *Science* **252,** 1162 (1991).

142. Bornberg-Bauer, E., Rivals, E. & Vingron, M. Computational approaches to identify leucine zippers. *Nucleic Acids Res.* **26,** 2740–2746 (1998).

143. Karimova, G. et al., A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci.* **95**, 5752-5756 (1998)

144. Vagner, V., Dervyn, E. & Ehrlich, S. D. A vector for systematic gene inactivation in Bacillus subtilis. *Microbiology* **144,** 3097–3104 (1998).

145. Figaro, S. *et al.* Bacillus subtilis Mutants with Knockouts of the Genes Encoding Ribonucleases RNase Y and RNase J1 Are Viable, with Major Defects in Cell Morphology, Sporulation, and Competence. *J. Bacteriol.* **195,** 2340–2348 (2013).

146. Lehnik-Habrink, M. *et al.* The RNA degradosome in Bacillus subtilis: identification of CshA as the major RNA helicase in the multiprotein complex: CshA in the RNA degradosome of Bacillus subtilis. *Mol. Microbiol.* no-no (2010). doi:10.1111/j.1365-2958.2010.07264.x

147. Herzberg, C. *et al.* SPINE: A method for the rapid detection and analysis of protein–protein interactionsin vivo. *PROTEOMICS* **7,** 4032–4035 (2007).

148. Nathan, C. Antibiotics at the crossroads. *Nature* **431,** 899–902 (2004).

149. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. *Br. J. Exp. Pathol.* **10,** 226 (1929).

150. Tan, S. & Tatsumura, Y. Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Med. J.* **56,** 366–367 (2015).

151. Eidem, T. M., Roux, C. M. & Dunman, P. M. RNA decay: a novel therapeutic target in bacteria. *Wiley Interdiscip. Rev. RNA* **3,** 443–454 (2012).

152. Kaito, C. *et al.* Silkworm pathogenic bacteria infection model for identification of novel virulence genes: Silkworm-infection model to identify virulence factors. *Mol. Microbiol.* **56,** 934–944 (2005).

153. Sidote, D. J., Barbieri, C. M., Wu, T. & Stock, A. M. Structure of the Staphylococcus aureus AgrA LytTR Domain Bound to DNA Reveals a Beta Fold with an Unusual Mode of Binding. *Structure* **16,** 727–735 (2008).

154. Moon, A. F., Mueller, G. A., Zhong, X. & Pedersen, L. C. A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci.* NA-NA (2010). doi:10.1002/pro.368

155. Gaboriaud, D. et al., Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEB.* **224**, 149-155 (1987)

156. Hunte, C. & Michel, H. Crystallisation of membrane proteins mediated by antibody fragments. *Curr. Opin. Struct. Biol.* **12,** 503–508 (2002).

157. Griffin, L. & Lawson, A. Antibody fragments as tools in crystallography: Antibody fragments as tools in chrystallography. *Clin. Exp. Immunol.* **165,** 285–291 (2011).

158. Koide, S. Engineering of recombinant crystallization chaperones. *Curr. Opin. Struct. Biol.* **19,** 449–457 (2009).

159. Bukowska, M. A. & Grütter, M. G. New concepts and aids to facilitate crystallization. *Curr. Opin. Struct. Biol.* **23,** 409–416 (2013).

160. Gu, Y. *et al.* Structure of a Novel Shoulder-to-Shoulder p24 Dimer in Complex with the Broad-Spectrum Antibody A10F9 and Its Implication in Capsid Assembly. *PLoS ONE* **8,** e61314 (2013).

161. Skrabana, R., Dvorsky, R., Sevcik, J. & Novak, M. Monoclonal antibody MN423 as a stable mold facilitates structure determination of disordered tau protein. *J. Struct. Biol.* **171,** 74–81 (2010).

162. Bernstein, J. A., Lin, P.-H., Cohen, S. N. & Lin-Chao, S. Global analysis of Escherichia coli RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 2758–2763 (2004).

163. Lorentzen, E., Dziembowski, A., Lindner, D., Seraphin, B. & Conti, E. RNA channelling by the archaeal exosome. *EMBO Rep.* **8,** 470–476 (2007).

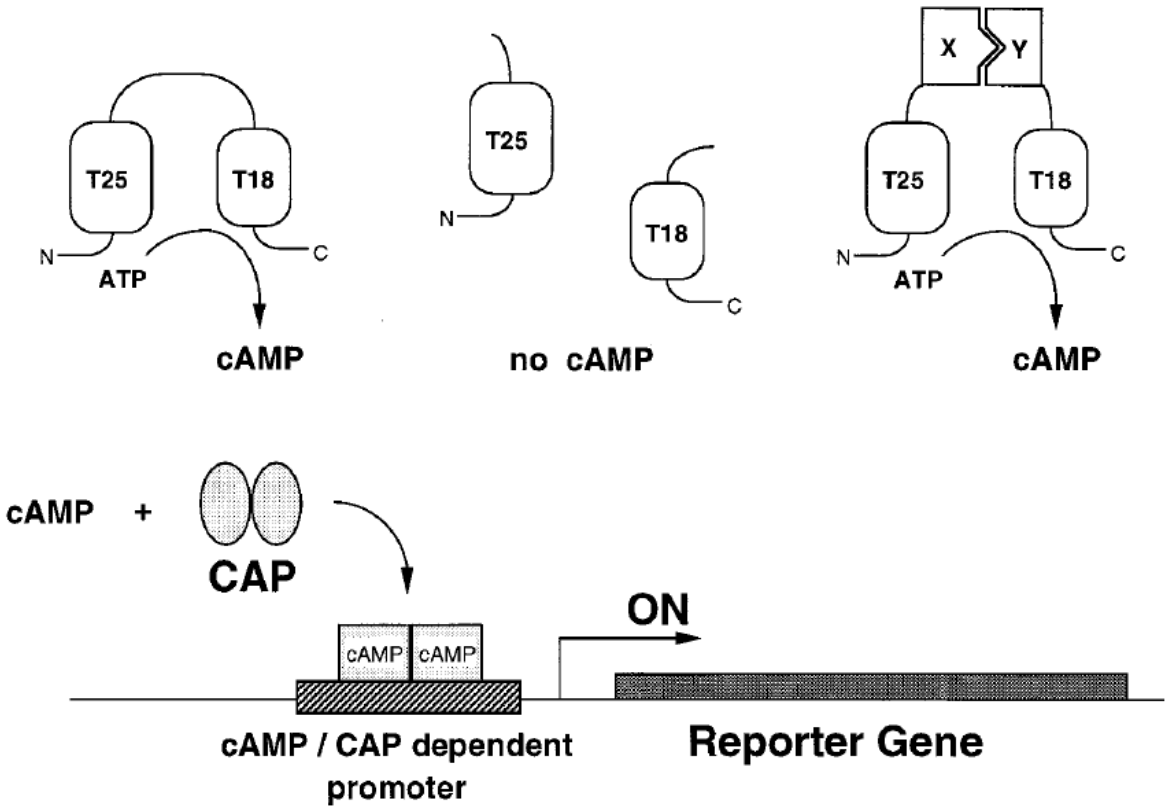
164. Tor, Y. & Fair, R. Antibiotics and Bacterial Resistance in the 21st Century. *Perspect. Med. Chem.* 25 (2014). doi:10.4137/PMC.S14459

Site internet n°1 : https://www.gatesnotes.com/Books/Harvesting-The-Biosphere

ANNEXES

ANNEXE 1

Le principe du double hybride consiste à détecter une interaction entre deux proréines par reconstitution de l'activité adénylate cyclase. L'adénylate cyclase est constituée de deux sous-unités (T25 et T18) . Pour détecter une interaction entre une protéine X et Y, nous allons fusionner par exemple la sous-unité T25 à la protéine X et la sous-unité T18 à la protéine Y. Lorsque les deux sous-unités T18 et T25 sont séparées, on perd l'activité adénylate cyclase. Lors de la co-expression des deux protéines X et Y chez *E. coli* ou chez la levure, si les deux protéines interagissent, cela va reconstituer l'activité adénylate cyclase et induire la transcription d'un gène rapporteur permettant de détecter visuellement sur les colonies de bactéries l'interaction des deux protéines X et Y[143].



ANNEXE 2

Tableau des sites de clivages de la RNase E[36] à gauche et des RNases Y et J à droite[28].

Site	Cleavage sequence	References
RNase E site consensus	G A A U U U A	
9Sa	A$^{\nabla}$A U U U	Ghora and Apirion (1978)
9Sb	A$^{\nabla}$A U A A	Ghora and Apirion (1978)
RNA I	U$^{\nabla}$A U U U	Tomcsányi and Apirion (1985)
T4 gene *32* (−71)	A$^{\nabla}$A U U A	Mudd et al. (1988)
T4 gene *59* (−1340)	G$^{\nabla}$A$^{\nabla}$U U A	Carpousis et al. (1989)
dicF	A A U$^{\nabla}$U U	Faubladier et al. (1990)
gltX	A A$^{\nabla}$U$^{\nabla}$U$^{\nabla}$U	Brun et al. (1990)
gltX	G$^{\nabla}$A U U U	Brun et al. (1990)
f1 D	U A$^{\nabla}$U$^{\nabla}$G$^{\nabla}$U	Kokoska et al. (1990)
f1 E	A A$^{\nabla}$U$^{\nabla}$C U	Kokoska et al. (1990)
f1 F	G A$^{\nabla}$U$^{\nabla}$U U	Kokoska et al. (1990)
f1 G	A A$^{\nabla}$U$^{\nabla}$U U	Kokoska et al. (1990)
T4 gene *32* (+831)	G$^{\nabla}$A U U U	Loayza et al. (1991)
T4 gene *55*/αgt	A$^{\nabla}$A U U C	Loayza et al. (1991)
T2 gene *32* (−76)	A$^{\nabla}$A$^{\nabla}$U$^{\nabla}$U A	Loayza et al. (1991)
rpsO-pnp (M2)	A G$^{\nabla}$U U U	Régnier and Hajnsdorf (1991)
rpsO-pnp (M)	A A$^{\nabla}$G C U	Régnier and Hajnsdorf (1991)
rpsT (301)	G$^{\nabla}$A U C G	Mackie (1991)
ompA A*	A$^{\nabla}$C U U U	Nilsson et al. (1988)
ompA C*	G$^{\nabla}$A U U U	Nilsson et al. (1988)
ompA D*	U$^{\nabla}$A U U U	Nilsson et al. (1988)
10Sa*	G A$^{\nabla}$U$^{\nabla}$U$^{\nabla}$A	Chauhan and Apirion (1989)
10Sb*	$^{\nabla}$G A U U U	Gurevitz et al. (1983)
10.1S*	C A$^{\nabla}$A A A	Pragai and Apirion (1982)

RNase Y	
Bs *yitJ* C1*	(--)GACACGAAAAUUU^CAUAU CCG(--)
Bs *yitJ* C2	(--)GAGACA^AAAUCACUGAC(--)
Bs *gapA*	(--)CAAAGAA^GU(--)
Bs *infC* C1	(--)TATTG^TGTAGAATAGT
Bs *infC* C2	(--)TGACCGTAC^ATTTTTATTGA
Sa *sae*	TATACAACTAT^TAAATCCCATAA
RNases J1 and J2	
Bs *thrS* leader**	GAUUCCG^UUUAUUC
Bs *thrZ* leader	CCACGGG^UUAAUCA
Bs *trp* leader	CAUUAUG^U^U^UAUUC
Bs *ilv-leu*	GAGAACA^GGUACA
Bs scRNA	AUCAUCA^AAUUUUC

ANNEXE 3

Liste des différentes classes d'antibiotiques existants à l'heure actuelle[164]

CLASS	CLINICALLY INTRODUCED SINCE 2000
Sulfonamides	None
β-lactams	Biapenem , ceftaroline, doripenem , ertapenem
Aminoglycosides	None
Amphenicols	None
Macrolides	Telithromycin
Tetracyclines	Tigecycline
Rifamycins	Rifaximin
Glycopeptides	Telavancin
Quinolones	Balafloxacin, gemifloxacin, pazufloxacin, prulifloxacin
Streptogramins	None
Polymyxins	None
Oxazolidinones	Linezolid
Lipopeptides	Daptomycin
Pleuromutilins	Retapmulin
Macrolactones	Fidaxomicin
Diarylquinolines	Bedaquiline

ANNEXE 4

Liste des nucléases eucaryotes et procaryotes et de leurs fonctions dans la cellule

Process	Function	Enzyme	Substrate	Section	Nucleophile
DNA replication	Proofreading	3′ Exo and DnaQ	ssDNA	6.1.1	H_2O
	Primer removal	5′ Exo, FEN1; Dna2; RNase H	DNA with a 5′ RNA flap	6.4.1; 6.3.3; 6.2.1	H_2O
Rolling-circle replication	Nicking	Rep endonuclease	ssDNA	7.3.1	Tyr
DNA replication, transcription, recombination	Topoisomerization	Type IA; II; Type IB	DNA	8.1.1; 8.1.2; 10.2.1	Tyr
Site-specific recombination	Retroviral integration and V(D)J recombination transposition	Tn5, Hermes, mariner, RAG and integrase	Transposon, phage DNA and plasmid	6.2.2	H_2O; 3′-OH
	Transposition, gene inversion, phage integration and plasmid segregation	Cre, Flp and Int recombinase	Two DNA-specific sites	10.2.2	Tyr
		Resolvase and invertase	Two DNA-specific sites	8.2	Ser
	ICEs and ssDNA-mediated transposition	TraI, TrwC; IS605 TnpA	ssDNA	7.3.2; 7.3.3	Tyr
	Homing endonuclease	LAGLIDAG	DNA	6.10	H_2O
		HNH and His-Cys	DNA	7.1.1	H_2O
		GIY-YIG	DNA	7.2	H_2O
DNA homologous recombination	Processing broken DNA ends	RecBCD; RecE, λ Exo	DNA ends	6.3.2; 6.3.3	H_2O
	HJ resolution	RuvC; Endo I; Mus81; Gen1; RusA; EndoVII	Four-way DNA junction	6.2.5; 6.3.4; 6.3.5; 6.4.2; 6.14; 7.1.2	H_2O
	Meiosis breaks	Spo11	DNA	8.1.1	Tyr
DNA repair	Incision in base excision repair (BER)	Ape1/Exo III; FEN1; Endo IV	Modified bases	6.7.2; 6.4.1; 6.16.1	H_2O
	Incision in nucleotide excision repair (NER)	UvrC (N&C); XPF; XPG	DNA lesions	6.2.6 & 7.2; 6.3.5; 6.4.2	H_2O
	Incision and removal in mismatch repair (MMR)	3′ ExoI and X; MutH; Vsr; 5′ ExoI; RecJ	Mismatched base pairs	6.1.2; 6.3.1; 6.3.6; 6.4.3; 6.5	H_2O
	Double-strand break repair (DSBR)	Wrn; Mre11; Artemis; Dna2	DNA ends	6.1.4; 6.9; 6.8.3; 6.3	H_2O

RNA processing	tRNA 5′ processing	RNase P	tRNA	6.15.3	H_2O
	tRNA 3′ processing	RNase D; T; Z	tRNA	6.1.5; 6.1.6; 6.8.1	H_2O
	tRNA splicing	Endonuclease	tRNA	9.5	2′-OH
	mRNA splicing	Spliceosome	mRNA	6.15.2	H_2O, 3′-OH
	mRNA end processing and degradation	RNase D (Prp6); PARN; Rai; Rat1; PIN; RNase E/G; II; PH; CPSF-73	RNA	6.1.6; 6.1.7; 6.3.7; 6.4.4; 6.4.5; 6.7.3; 6.11; 6.13; 6.8.2	H_2O; PO_4
	rRNA maturation	RNase T; ERI-1; RNase M5; RNase E/G; J1; group I/II intron	RNA	6.1.5; 6.1.8; 6.6; 6.7.3; 6.8.1, 6.15.1/2	H_2O; 2′ and 3′-OH
RNA interference	Dicer	RNase III	dsRNA	6.12	H_2O
	Slicer	Argonaute	RNA	6.2.4	H_2O
Cell death	DNA decay	NucA, EndoG, CAD; Nuc, DNase II	DNA and ssDNA	7.1.4; 10.1.1	H_2O; His
	RNA decay	TREX1 and TREX2; RNase H2; Crn4	RNA and RNA/DNA hybrid	6.1.3; 6.2.3; 6.1.8	H_2O
Nutrition source	DNA digestion and scavenging	DNase I; nuclease P1/S1; Serratia	ssDNA and dsDNA	6.7.1; 6.16.2; 7.1.4	H_2O
	RNA digestion	RNase T1; A; T2	RNA	9.1; 9.3; 9.4	2′-OH; H_2O;
Defense	Restriction/modification	EcoRI; KpnI; BfiI; halfpipe	DNA	6.3.1; 7.1.3; 10.1.2; 10.3	H_2O; Tyr
	Microbial toxins	Restrictocin; colicin E3/4/6; colicin D/E5	RNA	9.1.3; 9.1.4; 9.2	2′-OH
		Colicin E7-9	DNA	7.1.4.3	H_2O
	CRISPR	Cas1; Cas2; Cas6	DNA and RNA	6.17.2; 6.17.3; 9.6	H_2O

Abstract :

Structural study of RNase Y, an endoribonuclease involved in messenger RNA degradation in *Bacillus subtilis*

Keywords : Bacteria, mRNA, ribonuclease, oligomer, intrinsically diordered protein, SAXS, RMN, electron microscopy.

Messenger RNA decay and processing in bacteria involve a set of various ribonucleases (RNases). In *Bacillus subtilis*, a newly discovered endoribonuclease called RNase Y was shown to play a central role in mRNA decay by catalyzing the initial endonucleolytic cleavage reaction. Its action releases several RNA pieces that are more sensitive to degradation by exoribonucleases. Interestingly, although RNase Y presents a functional homology with *E. coli* RNase E, it does not bear sequence similarity with it and no known structural homolog has been found yet.

As *Bacillus subtilis* is a model of Gram positive bacteria, which includes several human pathogens (such as *Listeria, Staphylococcus, Streptococcus, Clostridium* species). Several bacterial species have developed antibiotics resistance like Methicillin-resistant *Staphylococcus aureus* (SARM) and RNases have recently been perceive as new therapeutic targets for new antimicrobial compounds determination. So understanding the mechanism and structure/function of this enzyme may help to design specific inhibitors with antibiotic properties.

The main project consists in producing and crystallizing various constructs of the protein, with the goal to solve their structure by X-ray crystallography. But all crystallization trials had never allowed to get crystals. RNase Y contains an intrinsically disordered domain at the N-terminus, which represents an important part (~30% of the sequence), which is a challenge for crystallography[3]. So we decided to orientate our experiments in specific methods which permit to study this kind of protein with an intrinsically disordered region (IDR). These methods are small angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR), circular dichroism (CD) and electron microscopy.

By combination of these different approaches, it was found that RNase Y seems to be organized in a oligomeric form with a ring shape. This oligomer is homogenous in microscopy, globular and structured in SAXS. The intrinsically disordered domain is not completely disordered as we can conclude with the SAXS analysis. It contains some secondary structural elements (α helix) but it has no tertiary structure as we can conclude with CD and NMR analysis respectively. SAXS analysis of the RNase Y dimeric form show that this dimeric form is structured but have a very elongated shape. *In vitro* activity tests on both oligomeric and dimeric form reveal that this two form are active. Oligomer cryomicroscopy pictures collected will allow us to get a structure with a better resolution. At the moment, this oligomer seems to be composed by eight RNase Y dimer.

Some questions need to get an answer now : Does this oligomeric organization exist *in vivo* ? How we can explain an endonuclease activity with this RNase Y structural organization ? Does this ID domain interacts with protein partners in the cell like *E. coli* RNase E which can interacts with partners to form the degradosome complex for mRNA degradation ? Does this ID domain can also interacts with RNA like the RNase E ID domain ?

Résumé :

Résumé :

Étude structurale de la RNase Y, une endoribonucléase impliquée dans la dégradation des ARNm chez *Bacillus subtilis*

Mots clés : Bactérie, ARNm, ribonucléase, oligomère, protéine intrinsèquement désordonnée, SAXS, RMN, microscopie électronique

Le processus de maturation et de dégradation des ARNm chez les bactéries implique différentes ribonucléases (RNases). Une endoribonucléase a été découverte en 2009 chez *Bacillus subtilis* appelée RNase Y. Cette RNase joue un rôle centrale en catalysant la réaction de clivage initiatrice, un clivage endonucléolytique. Les fragments d'ARN issus du clivage par la RNase Y deviennent alors plus sensibles à la dégradation, en 5' et 3', par les exonucléases. La RNase Y possède les mêmes caractéristiques fonctionnelles que la RNase E d' *E. coli*, il s'agit de son homologue fonctionnel, en effet ces deux enzymes ne partagent aucune homologie de séquence et la RNase Y ne possède aucun homologue structural connu.

Bacillus subtilis est la bactérie modèle chez les Gram positif qui contiennent également de nombreux pathogènes pour l'homme comme par exemple les espèces *Listeria, Staphylococcus, Streptococcus, Clostridium.* Certaines de ces bactéries ont développé des résistances aux antibiotiques et les RNases ont récemment été perçus comme de nouvelles cibles thérapeutiques pour la détermination de nouveaux composés antimicrobiens. Obtenir des informations structurales sur la RNase Y est donc d'une grande importance pour le futur de la recherche de nouveaux antibiotiques.

L'objectif principal de ce projet a été de produire et cristalliser différentes constructions de la RNase Y afin de résoudre sa structure par diffraction des rayons X. Mais les essais de cristallisation de cette protéine n'ont pas abouti. La RNase Y possède un domaine intrinsèquement désordonné (ID) en N-terminal (~30% de la séquence de la RNase Y), ce type de domaine est en général très difficile à cristalliser. Nous avons donc décidé d'orienter nos expérimentations vers des techniques permettant d'étudier ce type de protéine, comme la diffusion des rayons X aux petits angles (SAXS), la résonance magnétique nucléaire (RMN), le dichroïsme circulaire (CD) et la microscopie électronique.

En combinant ces différentes techniques, il s'est avéré que la RNase Y semble s'organiser sous forme d'un oligomère ayant une forme en anneau. Cet oligomère est homogène en microscopie, globulaire et structuré en SAXS. Le domaine ID lui n'est pas complètement déstructuré comme le révèle son analyse en SAXS. Il contient des éléments de structure secondaire de type hélice α et ne possède pas de structure tertiaire, comme le concluent respectivement les analyses en CD et RMN. L'analyse en SAXS de la forme dimérique de la RNase Y révèle que le dimère est structuré mais qu'il possède une forme allongée. Les tests d'activité *in vitro* sur la forme dimérique et oligomérique montre que ces deux formes sont activent. Les images collectées en cryo-microscopie nous permettrons d'obtenir une structure de meilleure résolution de l'oligomère, qui semble être constitué de 8 dimères de RNase Y.

Les questions qui se posent maintenant sont les suivantes: Est-ce que cette forme oligomérique est présente *in vivo* ? Comment expliquer une activité endonucléase par une telle organisation structurale en oligomère de la RNase Y ? Est-ce que le domaine ID interagit avec des partenaires protéiques comme c'est le cas de la RNase E chez *E. coli* qui forme un complexe de dégradation des ARNm appelé dégradosome ? Est-ce que ce domaine ID est également capable d'interagir avec l'ARN comme le domaine ID de la RNase E ?