

Université de Montréal
Faculté des études supérieures et postdoctorales

Étude structuro-fonctionnelle de SecP43 et caractérisation de ses interactions avec SepSecS et l'ARNt^{Sec} *in vitro*

Département de biochimie et médecine moléculaire
Faculté de médecine

Mémoire présenté à la faculté des études supérieures en vue de l'obtention du
grade de M. Sc. en biochimie
Option biologie structurale

Ce mémoire a été évalué par le jury suivant :

Président : Serguei Chteinberg
Directeur Christian Baron et Jurgen Sygusch codirecteur
Membre du jury: Nazzareno D'Avanzo

Décembre 2013

Présenté par
Arnold Dopgwa Puemi

Abstract

Selenoproteins are proteins containing selenium in the form of the 21st amino acid, selenocysteine. Selenocysteine (**Sec**) is directly synthesized onto its cognate tRNA (tRNA^{[Ser]Sec} or tRNA^{Sec}) and inserted into selenoproteins co-translationally with the help of various *cis*- and *trans*-acting factors. Among those factors, SecP43 has been reported to possibly play an essential role in the methylation at the 2'-hydroxyribosyl moiety in the wobble position (Um34) of Sec-tRNA^{[Ser]Sec} and consequently reduce the expression of glutathione peroxidase 1. SecP43 also called tRNA^{Sec}-associated protein has also been reported to interact in with SepSecS and tRNA^{Sec} *in vivo* and the targeted removal of one of these proteins affected the binding of the other to the Sec-tRNA^{Sec}.

The initial aim of the project was to solve the structure of SecP43 by means of x-ray crystallography. Secondly, we were interested in characterizing the interaction of the latter with some of the components of the selenocysteine insertion machinery. These factors are SepSecS and tRNA^{Sec}. We were able to optimize the expression and the purification of soluble form of the human homologue of SecP43 and of SepSecS by using an adapted auto-induction protocol. This was a major challenge considering that full length SecP43 has not been expressed and purified to date. We did not succeed in crystallizing SecP43. Our failure to crystallize SecP43 is probably due to the fact that it is a partially folded protein as we were able to demonstrate by SAXS (*Small Angle X-ray Scattering*). The SecP43 envelope calculated by SAXS displayed a rod-shape like structure. In order to enhance the stability of SecP43 required for crystallization, binding affinity studies were conducted to characterize the interaction between SecP43, tRNA^{Sec} and SepSecS. We did not detect an interaction between SecP43 and tRNA^{Sec} by using EMSA (*Electrophoretic Mobility Shift Assay*) and gel filtration. We also could not detect an interaction between SecP43 and SepSecS using a cross-linking assay. In contrast, the tRNA^{Sec}/SepSecS interaction was demonstrated by EMSA and the addition of SecP43 seemed to reduce the binding affinity. Therefore, SecP43 might induce a conformational change in SepSecS in the presence of tRNA^{Sec}.

Keywords: Selenocysteine, SecP43, SepSecS, tRNA^{Sec}.

Résumé

Les sélénoprotéines sont des protéines contenant du sélénium sous la forme du 21^{ème} acide aminé, la sélénocystéine. Sélénocystéine (**Sec**) est synthétisé sur son ARNt (ARNt^{[Ser]Sec} ou ARNt^{Sec}) et inséré dans les sélénoprotéines de manière co-translationnelle, avec l'aide de divers facteurs agissant en *cis* et en *trans*. Parmi ces facteurs, SecP43 joue possiblement un rôle dans la méthylation du groupement 2'-hydroxyribosyl dans la position d'oscillation (Um34) du Sec-ARNt^{[Ser]Sec} et par conséquent, l'inhibition de SecP43 réduit l'expression de la glutathion peroxydase 1. L'interaction SecP43/ARNt^{Sec}/SepSecS a été démontrée *in vivo* et le retrait ciblé d'une de ces protéines affecte la liaison de l'autre à Sec-ARNt^{Sec}.

L'objectif initial du projet était de résoudre la structure de SecP43 par cristallographie. En second lieu, nous avons voulu caractériser l'interaction de ce dernier avec certaines composantes de la machinerie d'insertion de la sélénocystéine. Ces facteurs sont notamment SepSecS et ARNt^{Sec}. Nous avons optimisé la surexpression et la purification des homologues humains de SecP43 et SepSecS en utilisant un protocole d'auto-induction. Ce fut un défi majeur étant donné que SecP43 n'avait pas encore été surexprimée et purifiée à ce jour. Nous n'avons pas réussi à cristalliser SecP43. Cet échec s'explique probablement par le fait que SecP43 est une protéine dynamique et partiellement repliée comme nous avons pu le démontrer par SAXS (*Small Angle X-ray Scattering*). L'enveloppe de SecP43 calculée par SAXS présente une structure filiforme et non globulaire. Afin d'améliorer la stabilité de SecP43 requise pour la cristallisation, des études d'affinité de liaison ont été conduites pour caractériser les interactions entre SecP43, ARNt^{Sec} et SepSecS. Aucune interaction substantielle n'a pu être démontrée entre SecP43 et ARNt^{Sec} par retard sur gel ou par chromatographie d'exclusion stérique. Le même constat fut fait pour l'interaction SecP43/SepSecS. En revanche, l'interaction ARNt^{Sec}/SepSecS a été démontrée par EMSA (*Electrophoretic Mobility Shift Assay*) et l'addition de SecP43 semble réduire l'affinité de liaison de ladite interaction. Par conséquent, SecP43 induirait un changement conformationnel de SepSecS en présence du ARNt^{Sec}.

Mots clés : Sélénocystéine, SecP43, SepSecS, ARNt^{Sec}

Table des matières

Abstract.....	i
Résumé.....	ii
Liste des tableaux.....	viii
Liste des figures.....	ix
Liste des sigles, liste des abréviations.....	xii
Dédicace.....	xiv
Remerciements.....	xv

Chapitre 1. Introduction1

1. Le sélénium et ses propriétés chimiques et biologiques.....1

1.1. Le sélénium.....	1
1.1.1. La sélénocystéine.....	5
1.1.2. Les sélénoprotéines.....	6
1.2. Biosynthèse des sélénoprotéines.....	10
1.2.1. Biosynthèse de la sélénocystéine.....	11
1.2.1.1. ARNt ^{Sec}	11
1.2.1.2. Conversion de la sérine à la phosphosérine (O-phosphoserine- ARNt ^{Sec} Kinase/PSTK).....	14
1.2.1.3. Conversion de la phosphosérine à la sélénocystéine.....	15
1.2.1.3.1. Génération du donneur de sélénium (SPS1/2).....	15
1.2.1.3.2. Conversion de Sep-ARNt ^{Sec} à Sec-ARNt ^{Sec} (SecS).....	16
1.3. Incorporation de la sélénocystéine (Sec).....	17

1.3.1. Facteurs agissant en <i>Cis</i>	18
1.3.1.1. <i>Selenocysteine Incorporation Sequence (SECIS)</i>	18
1.3.1.1. a. Localisation dans l'ARNm	18
1.3.1.1. b. Structure secondaire.....	19
1.3.1.2. <i>Selenocysteine codon Redefinition Element (SRE)</i>	22
1.3.2. Facteurs agissant en <i>Trans</i>	23
1.3.2.1. EFSec.....	23
1.3.2.2. SBP2.....	25
1.3.2.3. RpL30 (<i>Ribosomal protein L30</i>).....	27
1.4. SecP43.....	28
1.5. Autres protéines.....	31
1.6. Modèles d'incorporation de Sec.....	32
1.7. Objectifs et grandes lignes de ce mémoire.....	34
Chapitre 2. Matériels et Méthodes	35
2. Méthodes.....	35
2.1. Souches bactériennes décrites dans le mémoire.....	35
2.2. Tampons divers.....	35
2.3. Culture cellulaire.....	37
2.4. Transfection des cellules HEK293-6E en suspension.....	37
2.5. Clonage de SecP43 et de SepSecS.....	37
2.6. Expression des protéines SecP43 et SepSecS.....	39

2.6.1. Induction à l'IPTG (Isopropyl β -D-1-thiogalactopyranoside).....	39
2.6.2. Auto induction.....	39
2.7. Purification des protéines SecP43 et SepSecS.....	40
2.8. Analyse de SecP43 par spectrométrie de masse.....	41
2.9. Analyse structurale de SecP43.....	42
2.9.1. Essais de cristallisation.....	42
2.9.2. Mesure et analyse des données de BIOSAXS (<i>Small Angle x-rays scattering</i>).....	42
2.10. Synthèse <i>in vitro</i> et purification de l'ARNt ^{Sec}	44
2.11. Interaction SecP43-SepSecS-ARNt ^{Sec}	45
2.11.1. Essais de réticulation entre SecP43 et SepSecS.....	45
2.11.2. Essais de retard de mobilité sur gel d'électrophorèse.....	45
Chapitre 3. Résultats	47
3.1. Optimisation de l'expression de SecP43.....	47
3.1.1. Expression de SecP43 dans les cellules humaines HEK 293.....	47
3.2. Expression de SecP43 dans la souche BL21 DE3 d' <i>E.coli</i>	49
3.2.1. Expression dans le milieu LB.....	49
3.2.2. Milieu LB ; effet du vecteur pBB535.....	50
3.2.3 Expression dans le milieu d'auto-induction.....	51
3.3. Purification de SecP43.....	55
3.3.1. Chromatographie d'affinité IMAC-Histrap.....	55

3.3.1.2. Étape de polissage par filtration sur gel.....	56
3.3.1.3. Étude la protéine purifiée par spectrométrie de masse.....	57
3.4. Purification de SecP43 tronquée.....	58
3.5. Vérification de la masse et de la séquence par spectrométrie de masse.....	60
3.6. Caractérisation structurelle de SecP43.....	62
3.6.1. Essais de cristallisation.....	62
3.6.2. Caractérisation structurelle de SecP43 par BIOSAXS (<i>Biological-Small Angle X-rays Scattering</i>).....	65
3.6.2.1. Description de la technique de BIOSAXS.....	65
3.6.2.2. Collecte des données de diffraction de SecP43 tronquée par BIOSAXS.....	66
3.6.3. Validation de la collecte des données de diffraction de SecP43 tronquée...68	
3.6.4. Fonction de distribution des distances $p(r)$	74
3.6.5. Fusion des courbes de diffraction des échantillons SecP43-Tris.....	75
3.6.6. Repliement de SecP43 (<i>Kratky plot normalisé</i>).....	76
3.7. Calcul des enveloppes	78
3.8. Études des interactions SecP43/ARNt ^{Sec} /SepSecS.....	81
3.8.1. Analyse des interactions SecP43/ARNt ^{Sec} par chromatographie d'exclusion stérique.....	81
3.8.2. Analyse des interactions SecP43/ARNt ^{Sec} /SepSecS par réticulation au formaldéhyde.....	85
3.8.3. Analyse des interactions SecP43/ARNt ^{Sec} /SepSecS par retard sur gel.....	88

Chapitre 4. Discussion	91
4.1 Mise en contexte.....	91
4.2. Surexpression et purification de SecP43.....	93
4.3. Caractérisation structurelle de SecP43.....	94
4.4. Interaction SecP43/SepSecS/ARNt ^{Sec}	95
4.5. Fonction de SecP43.....	97
4.6. Conclusion et perspective.....	99
Références.....	101
Annexe.....	113

Liste des Tableaux

Tableau I : Sélénoprotéines eucaryotes identifiées et leurs fonctions.....	9
Tableau II : Exemples de sélénoprotéines impliquées dans les maladies humaines.....	10
Tableau III : Souches bactériennes obtenues par transformation de la souche BL21(DE3).....	35
Tableau IV : Séquences des amorces.....	36
Tableau V : Conditions des réactions de synthèse de l'ARNt ^{Sec}	44
Tableau VI : Collection de données et paramètres dérivés des données de dispersion.....	67
Tableau VII : Paramètres des courbes de diffraction unidimensionnelles.....	72
Tableau VIII : Peptides générés et détectés par le spectromètre de masse avant l'insertion du codon stop.....	113
Tableaux IX : Peptides générés par la digestion de SecP43 tronquée et détectés par le spectromètre de masse.....	114

Listes des figures

Figure 1. Propriétés réactionnelles du sélénium et du soufre.....	3
Figure 2. Résistance du SecTR à l'inactivation par le radical hydroxyle ($\bullet\text{OH}$).....	4
Figure 3. Structures chimiques de la cystéine et de la sélénocystéine.....	5
Figure 4. Structures secondaires de l'ARNt ^{Sec} et de ses homologues canoniques.....	13
Figure 5. La voie de synthèse de la sélénocystéine.....	15
Figure 6. Modèles des structures secondaires des formes 1 et 2 de l'élément SECIS.....	20
Figure 7. Structures secondaires du SECIS et d'autres RNAs contenant le motif <i>k-turn</i>	21
Figure 8. Les éléments SECIS et SRE des ARNm de SEPNI.....	23
Figure 9. Représentation des facteurs d'élongation spécifiques à la sélénocystéine.....	25
Figure 10. Domaines structuro-fonctionnels de SBP2.....	27
Figure 11. Effet de l'inhibition de SecP43 par siRNA sur la synthèse de Sec-ARNt ^{Sec} Um34.....	30
Figure 12. Modèles d'incorporation de la sélénocystéine.....	33
Figure 13. Test de surexpression transitoire de SecP43 dans les cellules HEK293-6E...	48
Figure 14. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d' <i>E. coli</i>	49
Figure 15. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d' <i>E. coli</i> . Effet du vecteur additionnel pBB535.....	50
Figure 16. Principe du fonctionnement de l'auto-induction.....	53
Figure 17. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d' <i>E. coli</i> . Effet des vecteurs pRARE et pBB535 dans le milieu d'auto-induction 4xLac-8ZYM.....	54
Figure 18. Purification de SecP43 par chromatographie d'affinité (IMAC).....	55
Figure 19. Purification de SecP43 et évaluation de la pureté sur gel SDS-PAGE 10%.....	56

Figure 20. Analyse de la séquence de SecP43 par spectrométrie de masse.....	58
Figure 21. Purification de SecP43 tronquée.....	59
Figure 22. Analyse de la dimérisation de SecP43 par chromatographie d'exclusion stérique.....	60
Figure 23. Analyse de la séquence de SecP43 tronquée par spectrométrie de masse.....	61
Figure 24. Domaines topologiques de SecP43 selon l'algorithme RONN.....	63
Figure 25. Probabilité de cristallisation selon la séquence de SecP43.....	64
Figure 26. Représentation schématique d'une expérience de diffraction à des petits angles (SAS).....	66
Figure 27. Validation de la collecte des données de diffraction par SAXS.....	70
Figure 28. Approximation de Guinier.....	71
Figure 29. Comportement des paramètres R_g et I_0	73
Figure 30. Fonctions de distribution de distances $p(r)$ et effet de la concentration en SecP43.....	75
Figure 31. : Courbe résultant de la fusion des Courbes de diffraction des échantillons SecP43-Tris.....	76
Figure 32. Repliement de SecP43 selon le profil de la courbe Kratky normalisée.....	77
Figure 33. Modèles de basse résolution de SecP43 tronquée.....	79
Figure 34. Superposition des structures atomiques des RRM de SecP43 dans les enveloppes calculées.....	80
Figure 35. Analyse de l'interaction SecP43/ARNt ^{Sec} par chromatographie d'exclusion stérique.....	84
Figure 36. Mécanisme de réticulation des protéines par le formaldéhyde via la formation d'un cation immonium.....	85
Figure 37. Essai de réticulation entre SecP43 et SepSecS.....	86
Figure 38. Essai de réticulation entre SecP43, SepSecS et effet de l'ARNt ^{Sec}	87

Figure 39. Analyse des interactions SecP43/ARNt ^{Sec} et SepSecS/ARNt ^{Sec} par essai de retard sur gel.....	89
Figure 40. Analyse des interactions SecP43/ARNt ^{Sec} et SepSecS/ARNt ^{Sec} par essai de retard sur gel.....	90

Liste des sigles et abréviations

tRNA^{Sec} : ARNt spécifique à la sélénocystéine

bSECIS : *Bacterial Selenocysteine Insertion Sequence*

DIO: Iodothyronine deiodinases

EFSec: Facteur d'élongation spécifique à Sec

EF-Tu: Facteur d'élongation Tu

EMSA : *Electrophoretic Mobility Shift Assay*

GPx-1: Glutathione peroxidase 1

ITC: *Isothermal titration calorimetry*

K-turn: *Kink-turn*

MsrB: *Methione-R-sulfoxide reductase*

mcm5Um34: -5-methylcarboxymethyluridine-2'-O-methylribose

NSEP1: *Nuclease Sensitive Element Binding Protein 1*

PSTK: O-phosphoseryl-tRNA^{Sec} kinase

PHGPx: Phospholipid Hydroperoxide Glutathione peroxidase

RMN: Résonance Magnétique Nucléaire

RPL30: *Ribosomal protein 30*

SAXS: *Small Angle X-ray Scattering*

SBP2: *SECIS Binding Protein 2*

SEC: Sélénocystéine

SECIS: *Selenocysteine Insertion Sequence*

SecS: Selenocysteine synthase

SecTR: Résidu Sec de la thioredoxine réductase

SelA: *Selenocysteine synthase*

SelD: *Selenophosphate synthetase*

SelH,M,N,P,O,T : *Selenoproteins H, M, N, P, O, T*

SPS2: *Selenophosphate synthetase 2*

SRE: *Selenocysteine codon Redefinition Element*

TR: *Thioredoxin reductase*

3'-UTR: *3'-Untranslated region*

Je dédie ce mémoire

À mon père qui a toujours été un modèle de détermination et de persévérance à mes yeux,

À ma maman pour son soutien inconditionnel et ce, même à distance,

À mes oncles Tchaou et Nya pour leur soutien moral et financier,

Et à tous mes ami(e)s pour leurs encouragements lors des périodes d'incertitude.

Remerciements

J'aimerais remercier primordialement mon directeur de recherche, Dr. Christian Baron pour m'avoir accueilli dans son laboratoire et pour la confiance et le soutien qu'il m'a témoigné durant ces années de maîtrise. Mon expérience a été très enrichissante et m'a permis d'acquérir une certaine autonomie dans la conduite d'un projet de recherche. J'ai particulièrement été séduit par la liberté que vous accordez aux étudiants lorsqu'il s'agit d'explorer différentes pistes de recherche. J'en ressors grandi d'un point de vue scientifique et personnel.

Je tiens également à remercier mon co-directeur Dr. Jurgen Sygusch, qui m'a toujours soutenu depuis mes années de stage au baccalauréat. Votre supervision et vos idées généralement multiples ont largement contribuées au développement de mon sens critique et m'ont permis de trouver des alternatives face à des blocages.

Je ne manquerai sans doute pas de remercier le Dr. Pascale Legault et les membres de son laboratoire pour m'avoir donné accès à son équipement dans le cadre des mes travaux.

Ma famille et mes amis m'ont été d'une grande aide durant ces années de maîtrise. Je remercie notamment mes parents pour les sacrifices et les efforts qu'ils ont fournis pour que je puisse me rendre aux études graduées. Merci infiniment au couple Nya qui m'a accueilli comme leur fils à mon arrivée à Montréal. Vous avez facilité mon intégration dans cette belle ville et m'avez apporté tout le soutien dont un fils peut avoir besoin. Je vous en serai éternellement reconnaissant. Merci également à Tchaou qui est a été à la fois un oncle, un ami et un conseiller permanent.

Pour finir, je remercie toute l'équipe du laboratoire du Dr. Baron, Ana Maria, Flore, Jon, Mazhad, Ingrid, Rodolfo, Benoit, Étienne, Charline, Bastien et l'ingénieur Mark. Vous êtes une équipe formidable avec laquelle il fut agréable de travailler.

Chapitre 1. Introduction

1. Le sélénium et ses propriétés chimiques et biologiques

1.1. Le sélénium

Le sélénium, est un métalloïde découvert en 1817 par le chimiste suisse Jacob Berzelius. Ce microélément a longtemps été considéré comme un poison agissant spécialement sur le bétail se nourrissant des plantes accumulatrices de sélénium tel que le type *Astragalus* pendant les périodes sèches. L'établissement du sélénium (Se) comme oligo-élément essentiel remonte aux années 1950 (Berry MJ et al, 1997). La démonstration du rôle de protection contre le stress oxydatif que joue le sélénium alimentaire a suivie au début des années 1970. Ses propriétés d'antioxydant ont largement été attribuées à sa présence au sein de la glutathione peroxydase 1 (GPx-1) cellulaire, une sélénoprotéine. Cependant, les fonctions de cette enzyme prises exclusivement, ne permettaient pas d'expliquer les autres effets associés à la déficience en sélénium (Berry MJ et al, 1997). L'identification d'autres sélénocysteines (**Sec**) de mammifères durant les dernières années a permis d'entrevoir d'autres aspects liés aux fonctions de ce métalloïde. Il s'agit en occurrence des sélénoenzymes appartenant à la famille des oxidoréductases telles que l'iodothyronine deiodinase et la thioredoxine reductase (TR) (Berry MJ et al, 1997; Atkins, J. F et al 2000). Chacune de ces enzymes implique un substrat, en occurrence une molécule oxydante (peroxyde d'hydrogène/ GPx-1) et sont retrouvées dans les organes dont les tissus prolifèrent suite au stress oxydatif (thyroïde/deiodinase). Elles font aussi partie d'un mécanisme antioxydant responsable du maintien des protéines cibles sous leur forme réduite (TR dans le système de la thioredoxine) (Toppo, S et al, 2009 ; Gromer, S et al, 2004). Une hypothèse cohérente suggère que ces enzymes ont évolué de manière à résister à l'inactivation par l'oxydation (Hondal, R. J et al, 2011 ; Ruggles, E. L et al, 2012). Cette hypothèse est incongrue vis-à-vis du fait que le sélénol (Se-H tel que

retrouvé dans la **Sec**) oxydera plus rapidement qu'un thiol (-S-H, retrouvé dans la cystéine) (Cardey, B et al, 2007). Par ailleurs, le potentiel d'oxydoréduction du couple sélénol/diséléride est de loin inférieur à celui du couple thiol/disulfide (Besse, D. et al, 1997). Sur la base de ces propriétés physico-chimiques, des chercheurs dans le domaine déclarèrent : « ...this highly oxidizable amino acid (**Sec**) could be maintained only in anaerobic organisms or in aerobic systems which evolved specific protective mechanisms » (Leinfelder, W et al, 1988). La sélélocystéine (**Sec**) serait, un vestige du monde anaérobique d'il y a deux ou trois milliards d'années (Jukes, T. H, 1990). Toutefois, on retrouve la sélélocystéine dans le site actif des oxydoréductases suscitées qui sont dans leur ensemble impliquées dans des mécanismes antioxydants. Il apparaît donc assez contradictoire qu'un acide aminé si oxydable soit au cœur des mécanismes impliqués dans la lutte contre le stress oxydatif. Cette contradiction est à l'origine du concept intitulé « **paradoxe du sélénium** » associé à la sélélocystéine et au métabolisme de l'oxygène. Des études relatives à ce sujet ont récemment été menées par le groupe du Dr. Robert Joseph Hondal. En effet, la sélélocystéine (RSeH), en raison du caractère nucléophile du séléolate (RSe-M⁺), s'oxyde rapidement en présence d'agents oxydants tels que le peroxyde d'hydrogène. Les oxydes de sélélocystéines qui en résultent à savoir l'acide séléonique (RSeOH) et l'acide séléinique (RSeO₂) peuvent à leur tour être rapidement réduits en sélénol (RSeH) en raison du caractère hautement électrophile de l'atome de sélénium dans ces oxydes. Le soufre dans la cystéine (Cys) peut également être très nucléophile et réagir avec des agents oxydants pour former des oxydes sulfuriques analogues à ceux formés lors de l'oxydation de la sélélocystéine. La différence majeure entre ces deux formes d'oxydes est que l'acide sulfonique (Cys-SO₂) est extrêmement résistant à la réduction chimique (Finlayson, A. J et al, 1979). Sa réduction en thiol (Cys-SH) n'est possible que par l'action de la peroxyredoxine, mais ce mécanisme requiert l'utilisation d'une enzyme de réparation (sulfiredoxine) (Biteau, B et al, 2003). Par ailleurs, tandis que l'acide sulfonique (Cys-SO₂) peut être oxydé en acide sulfonique (Cys-SO₃), l'acide séléinique (Sec-SeO₂) ne peut quant à lui être convertie en acide séléonique (Sec-SeO₃) (Ruggles, E. L et al, 2012). Cette propriété du sélénium, ainsi que sa nature à « **double facette** » (à la fois un bon nucléophile et un bon électrophile) font de lui un élément approprié pour métaboliser les oxydants au sein des enzymes qui

ont besoin d'une résistance accrue à l'inactivation par l'oxydation (Hondal, R. J et al, 2013). La figure 1 montre les paramètres cinétiques mettant en jeu la différence des propriétés réactionnelles intrinsèques du sélénium et du soufre. Par ailleurs, la figure 2 indique que la thioredoxine réductase de mammifères (msec-TR) résiste à l'inactivation par le radical libre ($\bullet\text{OH}$), tandis que son orthologue cys, qui provient de la *D. melanogaster* (DmCysTR), est beaucoup plus fragile. L'étude montre également que l'insertion de la sélénocystéine dans l'orthologue Cys-TR permettrait de sauvegarder son activité suite à un traitement au radical $\bullet\text{OH}$. Cela se traduit par la décroissance logarithmique et non plus linéaire de l'absorbance à 340 nm tel que représenté à la figure 2. Ainsi, s'agissant des réactions impliquées dans le maintien de l'équilibre redox de la cellule, l'avantage de la sélénocystéine comparativement à la cystéine réside en ceci qu'elle confère une meilleure résistance à l'inactivation par l'oxydation dû aux radicaux libres.

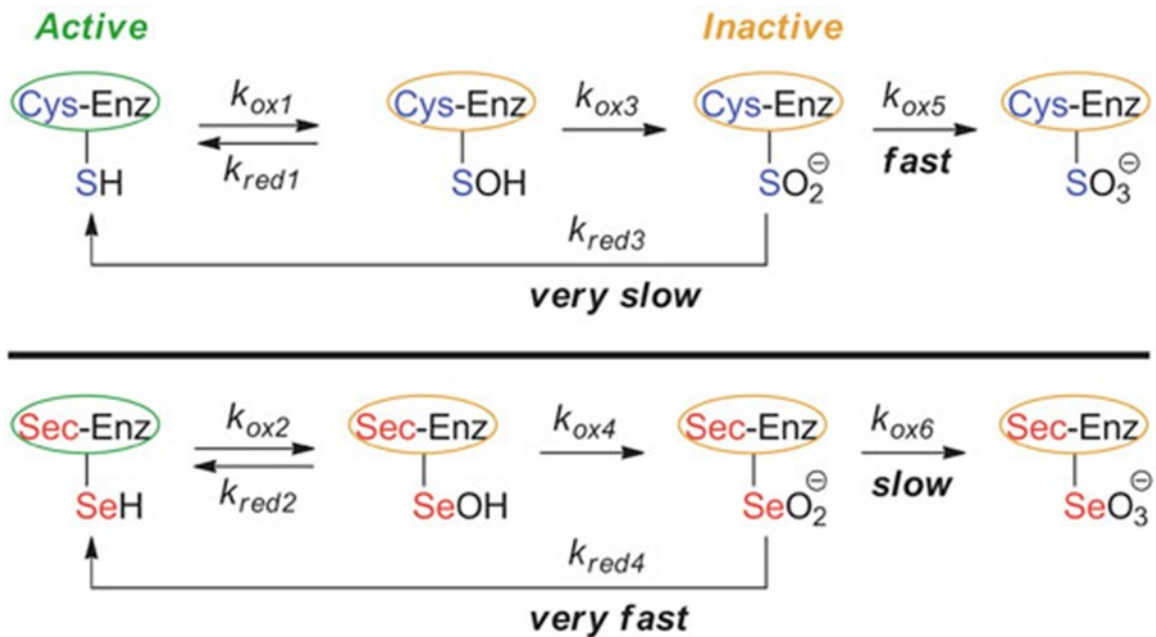


Figure 1. Propriétés réactionnelles du sélénium et du soufre

Les enzymes avec Cys et Sec requièrent respectivement un thiol ou un séléniol réduit, pour être fonctionnellement actives. Chacune de ces enzymes peut être oxydée par le H_2O_2 , pour donner les formes inactives Enz-SOH et Enz-SeOH respectivement, avec des vitesses d'oxydation $k_{ox2} > k_{ox1}$. La réduction

par les thiols restaure l'activité des deux formes inactives avec les vitesses de réduction $k_{red2} > k_{red1}$. L'ajout d'un second équivalent de H_2O_2 aux états d'oxydation $Enz-SOH$ et $Enz-SeOH$ conduit respectivement aux formes $Enz-SO_2^-$ et $Enz-SeO_2^-$. K_{ox4} est présumé supérieur à k_{ox3} quoique ceci n'ait pas été déterminé expérimentalement. Cependant, $k_{red4} \gg k_{red3}$, l'acide sulfonique étant d'une extrême résistance à la réduction. La forme $Enz-SeO_2^-$ résiste davantage à l'oxydation en $Enz-SeO_3^-$ tandis que son homologue $Enz-SO_2^-$ est facilement oxydé en $Enz-SO_3^-$. Dans ce cas, $k_{ox6} \ll k_{ox5}$. (L'image est tirée de l'Article « Chemical Basis for the Use of Selenocysteine », Erik L. Ruggles et al, 2011).

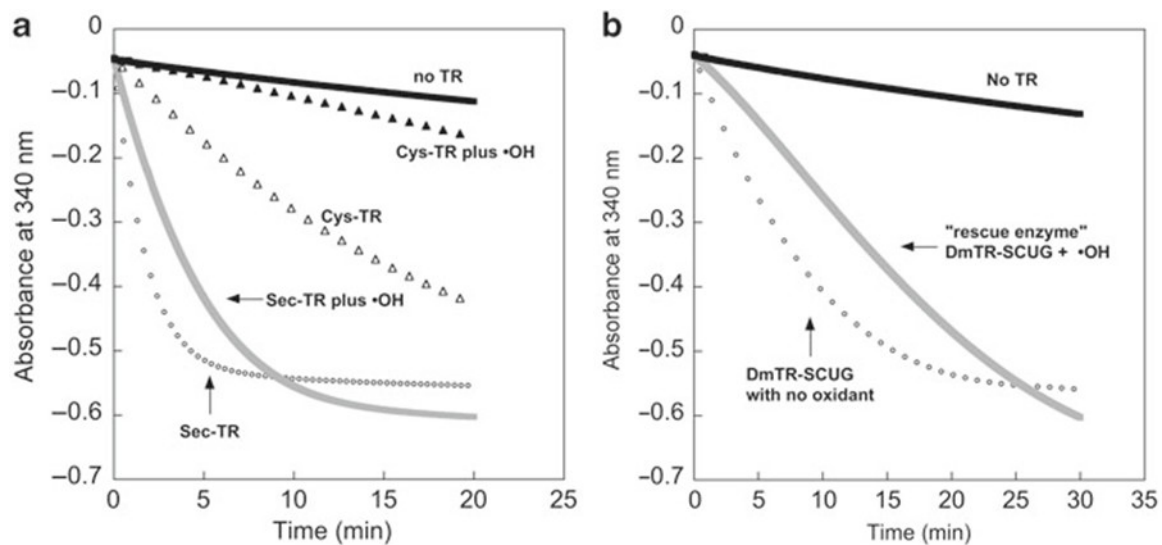


Figure 2. Résistance du SecTR à l'inactivation par le radical hydroxyle ($\bullet OH$)

(a) Exposition de la SecTR-GCUG (GUCG correspond au motif redox) de mammifères au radical $\bullet OH$ généré par $Fe\bullet EDTA/H_2O_2$ *in situ*. L'enzyme est incubée avec (courbe grise) et sans (cercles ouverts) $\bullet OH$ suivi d'une désactivation par étape. Trx est ensuite ajoutée au mélange réactionnel et l'activité est mesurée par la baisse de l'absorbance à 340 nm. La courbe noire représente le contrôle non enzymatique. Tandis que la Sec-TR de mammifères résiste à l'inactivation par $\bullet OH$, son orthologue Cys (DmTR-SCCS) est largement inactivée tel qu'observée en comparant l'activité enzymatique sans $\bullet OH$ (triangles ouverts) à l'activité après le traitement (triangles fermés). La présence du sélénium dans Sec « rescue »-TR (mutant DmTR-SCUG) confère la résistance à l'oxydation par $\bullet OH$ tel que dépeint dans le panneau (b). Comparaison entre DmTR-SCUG avec traitement par $\bullet OH$ (courbe grise) et DmTR-SCCS sans traitement $\bullet OH$ (cercles ouverts). (L'image est tirée de l'Article « Chemical Basis for the Use of Selenocysteine », Erik L. Ruggles et al, 2011).

1.1.1. La sélénocystéine

La sélénocystéine est le 21^{ème} acide aminé du code génétique. Cette forme biologique du sélénium retrouvé chez la plupart des eucaryotes est synthétisée et incorporée de manière co-translationnelle lors la synthèse des protéines correspondantes, les sélénoprotéines (Anton A. Turanov et al, 2013). La structure chimique de la sélénocystéine diffère de celle de la cystéine par la présence d'un atome de sélénium en lieu et place du soufre (figure 3). Bien que le sélénium et le soufre appartiennent à la même famille dans le tableau de la classification périodique des éléments chimiques, la sélénocystéine présente des propriétés chimiques singulières comparativement à la cystéine. En effet, la sélénocystéine a un pKa plus faible de 5,2 versus 8,5 pour la cystéine. Par conséquent, à pH physiologique, celle-ci est déprotonée et présente ainsi une nucléophilie et une réactivité plus accrue. L'activité enzymatique des sélénoenzymes est généralement de 100 à 1000 fois supérieure à celle de leurs homologues cystéiques (Muttenthaler & Alewood, 2008). La sélénocystéine, contrairement aux acides aminés canoniques est encodée par le codon UGA qui est un des trois codons STOP signalant l'arrêt de la traduction. Le mécanisme de « **recodage** » du codon STOP en sélénocystéine et de son incorporation subséquente dans la chaîne polypeptidique en extension sera décrit dans les sections suivantes.

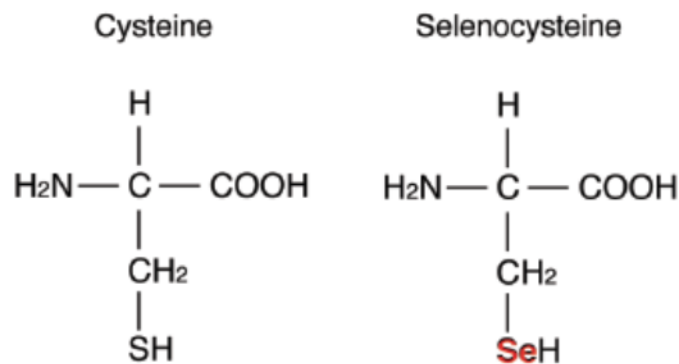


Figure 3. Structures chimiques de la cystéine et de la sélénocystéine. L'atome de soufre (S) de la cystéine est remplacé par l'atome de sélénium en rouge (Se) dans la sélénocystéine.

1.1.2. Les sélénoprotéines

Les sélénoprotéines sont présentes dans les trois domaines de la vie à savoir les bactéries, les archéobactéries et les eucaryotes. Cependant, certains organismes n'utilisent pas la sélénocystéine. Les champignons et les plantes supérieures qui probablement ont perdus la machinerie d'insertion de Sec au cours de l'évolution ne possèdent par conséquent pas de sélénoprotéines. Dans ces organismes, des orthologues contenant la cystéine sont utilisés à la place de la sélénocystéine (Alexey V. Lobanov et al, 2009). Les vertébrés encodent jusqu'à 25-26 sélénoprotéines mais la taille du sélénoprotéome varie considérablement d'un organisme à l'autre. Les organismes aquatiques par exemple possèdent un sélénoprotéome beaucoup plus large que les organismes terrestres. La disponibilité permanente du sélénium dans les eaux marines, aurait permis d'augmenter l'usage de cet élément par diverses oxydoréductases. Le passage de l'habitat aquatique à terrestre aurait engendré la perte ou le remplacement des sélénoprotéines par leurs homologues cystéiques (Lobanov et al, 2007). Tandis que les sélénoprotéines de bactéries et d'archéobactéries sont généralement impliquées dans les voies cataboliques, celles des eucaryotes jouent un rôle majeur dans les voies anaboliques et les processus antioxydant (Herbette et al, 2007).

En fonction de la localisation du résidu **Sec** dans la chaîne polypeptidique, les sélénoprotéines peuvent être classées en deux groupes (tableau I). Le résidu **Sec** des sélénoprotéines du premier groupe se situe près de la région C-terminale de la protéine. Il s'agit par exemple des sélénoprotéines de mammifères K, O, I, R et des thiorédoxines réductases (TRs). D'autres sélénoprotéines de mammifères tels que les sélénoprotéines H, M, T, V, W, Sep15, selenophosphate synthetase 2 (SPS2), les hormones thyroïdiennes deiodinases (DIs) et les glutathione peroxydases (GPXs) possèdent le résidu **Sec** dans le segment N-terminal de la protéine (Kryukov et al, 2003). Certaines sélénoprotéines du second groupe contiennent le motif CXXU (C et U représentant respectivement la cystéine et la sélénocystéine, et X un acide aminé quelconque). Ce motif est similaire au motif CXXC retrouvé dans le site actif des thiorédoxines (Dikiy et al, 2007; Ferguson et al, 2006). De telles signatures séquentielles suggèrent que les sélénoprotéines sont dotées

d'une fonction redox telle qu'indiquée précédemment. Ces réactions d'oxydoréduction jouent un rôle majeur dans la protection des cellules contre le stress oxydatif. Compte tenu de ces observations, il y a de bonnes raisons de penser que les sélénoprotéines dont les fonctions n'ont pas encore été élucidées à ce jour, participent également à des mécanismes redox (Lu & Holmgren, 2009; Novoselov et al, 2007b).

Les sélénoprotéines dont l'activité redox a été établie englobent cinq glutathione peroxydases (GPx), trois thioredoxine réductases (TR), trois iodothyronine deiodinases (DIO) et la selenophosphate synthetase 2 (SPS2).

Les thioredoxine réductases régulent le système thioredoxine qui est impliqué dans plusieurs voies de signalisation cellulaire contrôlant les facteurs de transcription. Ainsi, les thioredoxine réductases sont impliquées dans diverses fonctions cellulaires telles que la prolifération cellulaire, le système de défense redox, et les cascades de signalisation redox dépendantes (Arner, 2009; Lu & Holmgren, 2009).

Première sélénoprotéine de mammifères identifiée en 1973, la glutathione peroxydase 1 (GPx1) fait partie d'une famille d'enzymes formées de sept isoenzymes dont cinq comprennent des sélénocystéines (GPx1, 2, 3, 4 et 6) (Flohe, 2009). Les GPxs réduisent le peroxyde d'hydrogène et les hydroperoxydes organiques et donc protègent les cellules des dommages oxydatifs. Retrouvé abondamment dans le foie et les érythrocytes, la GPx1 prévient les dommages de l'ADN par les radicaux libres. Les polymorphismes de GPx1 sont aussi associés aux risques de cancer. La GPx4 est quant à elle connue sous le nom de phospholipid hydroperoxide GPx (PHGPx) car elle joue un rôle dans l'élimination des peroxydes lipidiques (Flohe, 2009; Gromer et al, 2005).

Les iodothyronine deiodinases (DIOs) clivent spécifiquement les liaisons carbone-iode dans les hormones thyroïdiennes telles que la thyroxine (T4), la 3, 5,3'-tri-iodothyronine (T3) et la 3'3'5' reverse tri-iodothyronine (rT4) qui est moins active que la T3. Ainsi les DIOs régulent l'activité hormonale de la thyroïde. La DIO1 et 2 convertissent la T4 en T3, la DIO 3 convertit T4 en rT3 (revue dans Gromer et al, 2005; Lu & Holmgren, 2009).

La sélénophosphate synthetase 2 (SPS2) est impliquée dans la biosynthèse de la sélénocystéine et sera décrite plus en détails dans les sections suivantes. Le tableau I

résume l'ensemble des sélénoprotéines connues et leurs fonctions respectives. Outre leurs implications dans les mécanismes redox, les sélénocystéines sont impliquées dans le métabolisme des hormones thyroïdiennes, la formation des muscles, la synthèse des sélénocystéines et la maturation du sperme (Rederstorff et al, 2006). Plusieurs troubles pathologiques résultant de l'altération de la structure, de l'activité et de l'expression des sélénoprotéines ont été reportés. La déficience en sélénium, les mutations ou polymorphismes des gènes et des cofacteurs de biosynthèse des sélénoprotéines sont impliqués dans diverses maladies répertoriées dans le tableau II (Bellinger et al, 2009).

Tableau I : Sélénoprotéines eucaryotes identifiées et leurs fonctions

seleno-protein	Function	Schematic representation
DIO1	thyroid hormone maturation (Conversion of T4 to T3 and T4 to rT3)	
DIO2	thyroid hormone maturation (Conversion of T4 to T3)	
DIO3	thyroid hormone catabolism (Conversion of T4 to rT3)	
GPx1	Antioxidant protection	
GPx2	Antioxidant protection	
GPx3	Maintenance of cellular redox status	
GPx4	Detoxification of lipid hydroperoxides	
GPx6	Antioxidant protection	
TR1	Part of the thioredoxin system. Antioxidant defense, redox regulation, cell signaling	
TR2	Part of the thioredoxin system. Antioxidant defense, redox regulation, cell signaling	
TR3	Part of the thioredoxin system. Antioxidant defense, redox regulation, cell signaling	
SPS2	Selenocysteine synthesis	
15kDa	Role in cell apoptosis and mediation of chemopreventive effects of Se	
SelN	Associated with muscular diseases	
SelW	Antioxidant protection Associated with cardiac calcification	
SelP	Plasmatic selenium transport Antioxidant defense	
SelH	Unknown	
SelI	Unknown	
SelJ	Present in fish and sea urchin	
SelK	Unknown	
SelL	Unknown, present in diverse aquatic organisms	
SelM	Unknown	
SelO	Unknown	
SelS	Unknown, involved in cellular redox balance	
SelT	Role in regulation of Ca2+ homeostasis and neuroendocrine secretion	
SelU	Unknown Present in fish, birds, sea urchin, green algae and diatoms only	
SelV	Unknown	
MsrA	Reduction of oxidized methionine residues Selenocysteine in C. reinhardtii only	
MsrB	Reduction of oxidized methionine residues	

(*) La localisation du résidu sélénocystéine (Sec) est mise en évidence par les traits noirs que porte chaque barre rectangulaire (Lobanov et al, 2009).

Tableau II : Exemples de sélénoprotéines impliquées dans les maladies humaines

Disorder	Cause	Selenoprotein or cofactor	Function
Keshan disease	Selenium deficiency/coxsackie B virus	Various, GPX?	??
Kashin-Beck disease	Selenium/iodine deficiency	DIO	Thyroid hormone production
Epilepsy	Selenium deficiency?	GPX?	Oxidative stress
Multiminicore disease	Mutations	SeIN	Calcium signalling
Thyroid dysfunction	Mutations	SBP2	Selenoprotein synthesis
Inflammation responses	Polymorphisms	SeS	Removal of misfolded proteins
Neurodegeneration	??	SeIP, SeIM, ??	Oxidative stress
Cancer	Polymorphisms/ Expression	GPX1, GPX2 and GPX4, SeS, SeIP, Sep15, TRXR	Various
HIV	Virus	TRXR1, GPX?	Viral gene expression, oxidative stress

(Tableau tiré de la revue Bellinger et al, 2009).

1.2. Biosynthèse des sélénoprotéines

Étant donné que la sélénocystéine est encodée par le codon UGA, un des codons STOP du code génétique universel, un mécanisme de reprogrammation, dénommé « **UGA recoding** » est requis. Le mécanisme sous-jacent à cette programmation comporte une machinerie de translation spécialisée impliquant un jeu complexe d'interactions protéine-RNA-ribosome. Bien que ce mécanisme ait été intensément étudié et clairement élucidé chez les eubactéries dès les années 1990 par Böck et ses collaborateurs, il demeurerait insaisissable chez les eucaryotes et archéobactéries (Turanov et al, 2003). Cependant, des progrès majeurs réalisés dans les dernières années ont conduit vers l'élucidation de ce mécanisme chez les eucaryotes. Ce mécanisme sera présenté plus en détail dans la présente section (Allmang et al, 2009; Papp et al, 2007; Squires & Berry, 2008). La synthèse des sélénoprotéines peut être subdivisée en deux étapes à savoir la biosynthèse de la sélénocystéine et son incorporation co-translationnelle dans la sélénoprotéine en cours de synthèse.

1.2.1. Biosynthèse de la sélénocystéine

Contrairement aux acides aminés canoniques, la sélénocystéine n'est pas librement disponible dans la cellule. Sa biosynthèse procède par la conversion de la sérine en sélénocystéine par une série de réactions. Cette synthèse se fait directement sur l'ARNt spécifique à la sélénocystéine, désigné ARNt^{[Ser]^{Sec}} ou ARNt^{Sec} (Frederick P. Bellinger et al, 2009).

1.2.1.1. ARNt^{Sec}

L'ARNt^{Sec} est l'ARN de transfert spécifique à la sélénocystéine et donc contient un anticodon complémentaire au codon UGA. Bien que présentant des différences au niveau de la séquence primaire, des analyses de modélisation computationnelle montrent des similarités au niveau de la structure tridimensionnelle des ARNt^{Sec} des bactéries, eucaryotes et archées (Baron 1993; Sturchler 1993). D'autres études montrent également que la fonction des ARNt^{Sec} est conservée à travers les espèces puisque les ARNt^{Sec} eucaryotiques et archéobactériens fonctionnent dans les systèmes bactériens (Baron et al, 1994).

Initialement considéré comme un accepteur de sérine supprimant le codon UGA, il a été démontré que l'ARNt^{Sec} eucaryotique existe sous une forme conjuguée à la sélénocystéine, sélénocysteyl-ARNt^{Sec} (Hatfield & Portugal, 1970; Lee et al, 1989). Des souris *knock-out* hétérozygotes sont capables de maintenir la synthèse des sélénoprotéines malgré la réduction des niveaux d'ARNt^{Sec}. Ceci indique que l'ARNt^{Sec} n'est pas un facteur limitant pour la synthèse des sélénoprotéines. Par contre, des cas de létalité embryonnaire sont observés chez les souris *knock-out* homozygotes, démontrant ainsi l'importance de la synthèse des sélénoprotéines pour le développement des mammifères (Bosl et al, 1997).

L'ARNt^{Sec} se distingue des autres ARNt cytosoliques d'une part par la présence de certains traits caractéristiques inhabituels au niveau de sa structure secondaire et d'autre part, par des modifications post-transcriptionnelles spécifiques (Allmang & Krol, 2006b). L'ARNt^{Sec} est le plus long ARN de transfert avec notamment 95 nucléotides chez *E.coli*, et 90 nucléotides chez les eucaryotes. Les ARNt canoniques quant à eux sont constitués de 70 à 75 nucléotides (Amberg et al, 1993; Böck, 2006). Bien que les ARNt canoniques possèdent universellement 7 paires de bases dans la branche de fixation de l'acide aminé (*A-stem*), l'ARNt^{Sec} en possède 8 paires (chez les bactéries) ou 9 paires de bases (chez les eucaryotes et archées). Bien que la présence de ces paires de bases supplémentaires ait été clairement établie, le mécanisme d'accommodement d'un tel ARNt au ribosome n'a pas encore totalement été élucidé. En outre, la structure secondaire de l'ARNt^{Sec} présente une branche D (*D-stem*) plus longue qu'une branche D classique. Cette dernière contient notamment 6 paires de bases comparativement à 3-4 paires pour les ARNt^{Sec} canoniques (Ishii et al, 2013). L'ARNt^{Sec} possède donc un bras récepteur (englobe les branches A et T; *A-stem* et *T-stem*) plus étendu. Le bras récepteur des ARNt classiques est formé de 12pb selon la configuration 7 pb *A-stem* et 5 pb *T-stem* tandis que celui de l'ARNt^{Sec} contient 13 pb. Les archéobactéries et les eucaryotes ont une branche-A de 9 pb et une branche-T de 4 pb (modèle 9/4) (figure 4). Les ARNt^{Sec} quant à elles sont configurées selon le modèle 8/5 (figure 4) chez les bactéries, les paires de bases supplémentaires du bras récepteur jouent un rôle déterminant pour le recrutement du facteur d'élongation SelB (Baron & Böck, 1991). Le long bras variable et la base discriminatoire G73 sont des éléments majeurs qui promeuvent la sérylation de l'ARNt^{Sec}, qui est une étape primordiale dans la synthèse de la sélénocystéine (Wu & Gross, 1993, Figure 4).

Des investigations sur les modifications post-transcriptionnelles de l'ARNt^{Sec} du *Xénopus* ont été menées (Diamond et al, 1993; Sturchler et al, 1994). Contrairement aux ARNt canoniques qui contiennent jusqu'à 15-17 bases modifiées, seulement 4 nucléotides sont post-transcriptionnellement modifiés dans l'ARNt^{Sec} des eucaryotes : la pseudouridine (pseudo-U55) et la 1-méthyladenosine (m1A58) situées dans la boucle T, la 6-isopentenyladenosine (i6A37) et le 5-méthylcarboxyméthyluridine-2'-O-méthylribose (mcm5Um34) situés dans la boucle de l'anticodon. Il existe deux formes d'ARNt^{Sec} qui diffèrent par l'état de méthylation du ribose du résidu U34. Il s'agit notamment des

formes mcm5U43 et mcm5Um34. La distribution relative de ces deux formes varie d'une cellule et d'un tissu à l'autre. La méthylation du ribose du résidu U34 requiert la modification préalable des autres bases suscitées et l'intégrité de la structure tertiaire de l'ARNt^{Sec} (Kim et al, 2000). Par ailleurs, la méthylation du ribose du résidu U34 semble être amplifiée par la présence du sélénium (Diamond et al, 1993). Des souris transgéniques surexprimant un gène mutant d'ARNt^{Sec} dépourvu du résidu i6A (rendant par conséquent la méthylation du ribose U34 impossible), présentaient une réduction du niveau d'expression des sélénoprotéines associées au stress tels que la GPx1, GPx3 et SelT (Carlson et al, 2005). Ces résultats suggèrent que les deux isoformes de l'ARNt^{Sec} peuvent avoir des fonctions différentes.

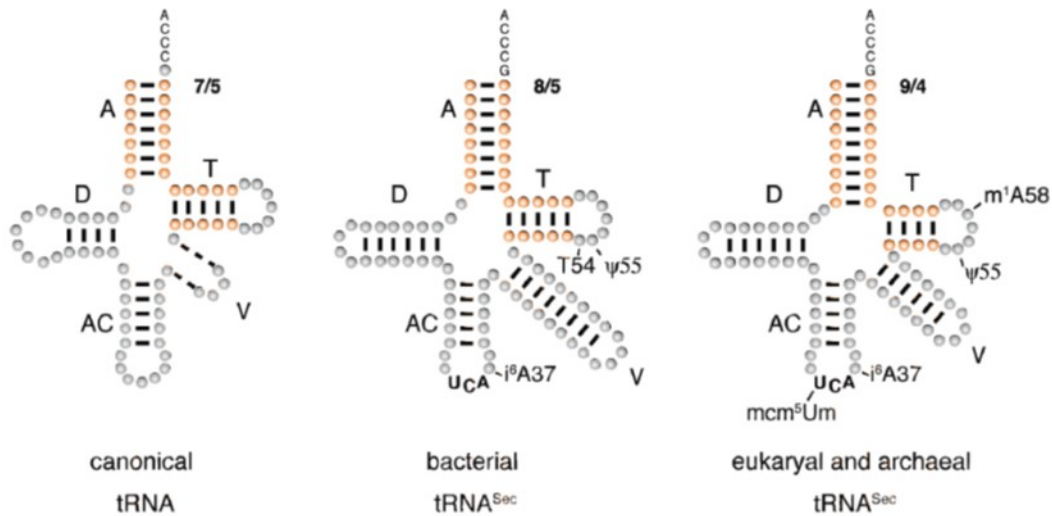


Figure 4. Structures secondaires de l'ARNt^{Sec} et de ses homologues canoniques

Les bras récepteurs sont dépeints en orange : 7/5, 8/5 et 9/4 indiquent le nombre de paires de bases constituant respectivement les branches A et T. i6A37, T54, Ψ55 et m1A58 sont les bases dont les riboses sont modifiés dans ARNt^{Sec} bactériens et eucaryotiques. Les éléments de la structure secondaire sont indiqués par les abréviations suivantes : A- branche de fixation de l'acide aminé, D-Branche D, AC- Branche anticodon, V- bras variable, T- Branche T. la longueur du bras variable varie en fonction des ARNt^{Sec}. La figure est tirée de (Allmang & Krol, 2006b).

1.2.1.2. Conversion de la sérine à la phosphosérine (O-phosphoserine-ARNt^{Sec} Kinase/PSTK)

La synthèse de la sélénocystéine débute par le chargement de la sérine sur l'ARNt^{Sec}. Cette conjugaison de la sérine à l'ARNt^{Sec} est une réaction d'amino-acylation conventionnelle catalysée par la seryl-ARNt synthetase. La conversion du groupement seryl en sélénocystéine a lieu directement sur l'ARNt^{Sec} contrairement aux acides aminés canoniques.

Chez les bactéries, la sélénocystéine synthase (**SeIA**), enzyme liant le cofacteur pyridoxal phosphate, convertit directement le groupement seryl en sélénocystéine en utilisant le monosélénophosphate comme substrat (Böck, 2006) le monosélénophosphate agissant ici comme donneur de sélénium est un produit de la transformation de l'ion séléniure (Se²⁻) par la sélénophosphate synthetase (**SeID**).

Contrairement à la voie de biosynthèse des sélénocystéines chez les bactéries dans laquelle la conversion du Ser-ARNt^{Sec} en Sec-ARNt^{Sec} se fait de manière directe, la biosynthèse des sélénocystéines chez les eucaryotes et les archéobactéries comporte deux étapes qui interviennent suite au chargement de la sérine sur l'ARNt^{Sec}: La phosphorylation du conjugué Ser-ARNt^{Sec} par la O-phosphoseryl-ARNt^{Sec} kinase (**PSTK**) et la conversion de la phosphoséryl-ARNt^{Sec} (Sep-ARNt^{Sec}) en Sec-ARNt^{Sec} par la sélénocystéine synthase (**SecS**) (Turanov et al, 2013).

La présence d'une activité kinase responsable de la conversion du Ser-ARNt^{Sec} en Sep-ARNt^{Sec} fut reportée en 1970 (Maenpaa & Bernfield, 1970), mais c'est seulement en 2004 que la **PSTK** fut identifiée par une approche de génomique comparative (Carlson et al, 2004). Cette enzyme est capable de discriminer le conjugué Ser-ARNt^{Sec} du conjugué Ser-tRNA^{Ser}. La longueur et la structure secondaire de la branche D de l'ARNt^{Sec} sont les facteurs déterminants de cette phosphorylation chez les eucaryotes (Wu & Gross, 1994). Par contre, chez les archéobactéries, l'enzyme reconnaît plutôt la branche réceptrice de l'ARNt^{Sec} (Sherrer et al, 2008). La figure 5 ci-dessous résume la synthèse de la sélénocystéine.

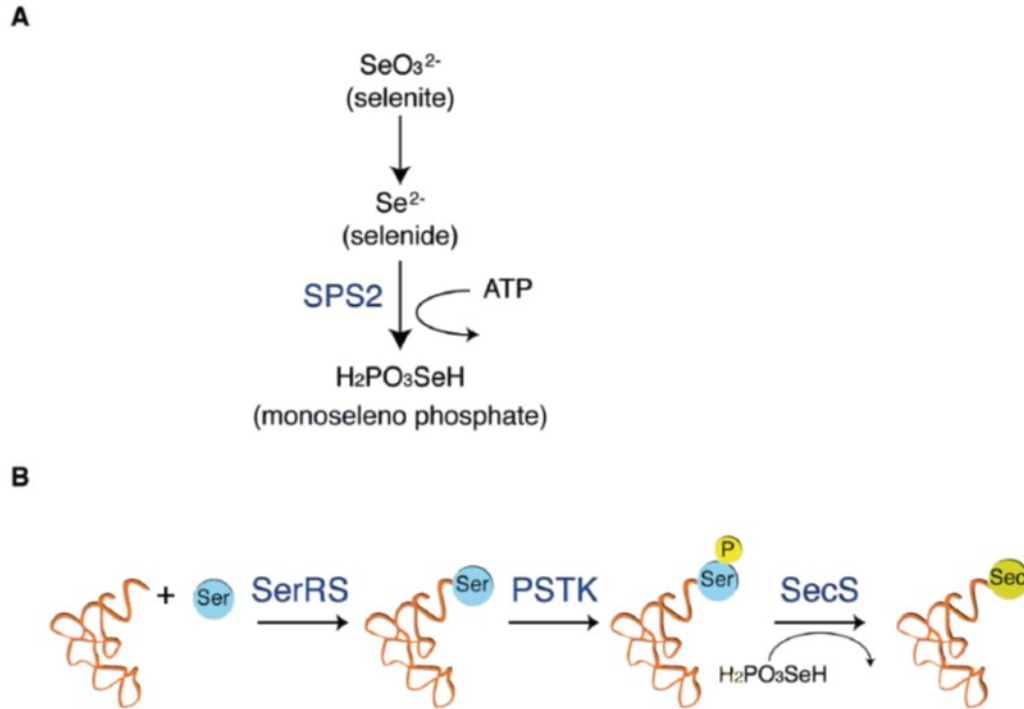


Figure 5. La voie de synthèse de la sélénocystéine. A. le donneur de sélénium, le monosélénophosphate ($\text{H}_2\text{PO}_3\text{SeH}$), est généré à partir de l'ion séléniure (Se^{2-}) par une réaction catalysée par la sélénophosphate synthétase 2 (SPS2). B. l'ARNt est chargé avec la sérine par la séryl-tRNA synthétase conventionnelle (SerRS). Chez les eucaryotes et les archéobactéries, le groupement séryl du conjugué Ser-ARNt^{Sec} est phosphorylé par la PSTK pour conduire au Sep-A lequel est finalement converti en Sec-ARNt^{Sec} par la sélénocystéine synthétase (SecS) en utilisant le monosélénophosphate. (Image produite par Takeuchi, 2009)

1.2.1.3. Conversion de la phosphosérine à la sélénocystéine

1.2.1.3.1. Génération du donneur de sélénium (SPS1/2)

La génération du donneur de sélénium, le monosélénophosphate est effectuée par la sélénophosphate synthétase (**SelD**) chez les bactéries. La sélénophosphate synthétase 1 (SPS1) et la sélénophosphate synthétase 2 (SPS2) ont tardivement été identifiés comme étant des homologues de SelD chez les eucaryotes (Guimaraes et al, 1996; Low et al, 1995). SPS2 est une sélénoprotéine dans la plupart des organismes eucaryotes. Des études ont montrées que contrairement à SPS1, seul SPS2 est capable de synthétiser du

monosélénophosphate *in vitro*. Par ailleurs, seul SPS2 est essentielle pour la synthèse des sélénocystéines *in vivo* (Xu et al, 2007a; Xu et al, 2007b). En outre, SPS1 est retrouvé chez les insectes qui ont perdu la machinerie de synthèse des sélénoprotéines indiquant ainsi que son rôle majeur n'est pas relié à la synthèse des sélénoprotéines (Chapple & Guigo, 2008.)

1.2.1.3.2. Conversion de Sep-ARNt^{Sec} à Sec-ARNt^{Sec} (SecS)

SecS est l'enzyme qui catalyse la dernière étape de synthèse de la sélénocystéine par un mécanisme ARNt-dépendant. Cette protéine de 48 kDa a initialement été nommée SLA (*Soluble liver antigen*) car des études ont montrées qu'elle co-immunoprécipite avec l'ARNt^{Sec} en utilisant des anticorps issus d'un sous-groupe de patients atteints d'une forme sévère d'hépatite chronique auto-immune (Costa et al, 2000). Deux études indépendantes ont tardivement identifié la SLA comme étant l'homologue de la selenocysteine synthétase (SecS) chez les eucaryotes et les archéobactéries (Xu et al, 2007a; Yuan et al, 2006). L'homologue humain de Sec est dénommé **SepSecS** et des études de complémentation *in vivo* ont montrées que d'une part **SepSecS** est capable de substituer SelA chez des souches d'*E. coli* déficientes en SelA, d'autre part, SepSecS assure la conversion du conjugué Sep-ARNt^{Sec} en Sec-ARNt^{Sec} en présence du sélénite de sodium et de la SelD (sélénophosphate synthetase d'*E. coli*) recombinante *in vitro* (Yuan et al, 2006). En outre, **SepSecS** possède une plus grande affinité vis-à-vis de Sep-ARNt^{Sec} comparativement à ARNt^{Sec} et Ser-ARNt^{Sec}. Cette observation confirme le fait selon lequel que contrairement à SelA d'*E. coli*, la synthèse de la sélénocystéine chez les eucaryotes et les archéobactéries procède d'une étape intermédiaire dans laquelle Sec-ARNt^{Sec} est généré en utilisant Sep-ARNt^{Sec} et le monosélénophosphate comme substrats (Xu et al, 2007b). Le mécanisme de cette réaction a été élucidé par le groupe du Dr. Palioura en 2009. La structure cristallographique du complexe **Sep-ARNt^{Sec}/SepSecS** et des essais enzymatiques suggèrent un mécanisme pyridoxal-phosphate-dépendant pour la formation de Sec-ARNt^{Sec}. **SepSecS** se lie à l'ARNt^{Sec} via son bras accepteur TΨC (où Ψ indique la pseudouridine) constitué de 13 paires de bases contrairement à 12 chez les

tRNAs canoniques. Cette liaison de l'ARNt^{Sec} induirait un changement de conformation au niveau du site actif de **SepSecS**. Ce changement permettrait à une phosphosérine liée de manière covalente à l'ARNt^{Sec} et non à une phosphosérine libre de s'orienter adéquatement dans le site actif pour que la réaction puisse avoir lieu (Palioura et al, 2009).

1.3. Incorporation de la sélénocystéine (Sec)

Le mécanisme général d'incorporation de Sec chez les bactéries diffère de celui rencontré chez les eucaryotes. Dans les bactéries, l'incorporation est essentiellement attribuée à un jeu d'interactions entre **bSECIS** (*bacterial Selenocysteine Insertion Sequence*, une structure tige-boucle située immédiatement en aval du codon UGA dans l'ARNm d'une sélénoprotéine donnée) et **SelB**, le facteur d'élongation spécialisé dans l'incorporation de **Sec**. La région N-terminale de SelB présente des similitudes de séquence et est fonctionnellement homologue à EF-Tu, le facteur d'élongation généralement utilisé au cours d'une traduction classique. La région C-terminale de SelB quant à elle se lie à **bSECIS**. En outre, **SelB** se lie spécifiquement Sec-ARNt^{Sec}. Sec-ARNt^{Sec} désormais lié à SelB est directement apporté au codon UGA via l'interaction **bSECIS-SelB**, ce qui permet l'incorporation de Sec au sein de la chaîne polypeptidique naissante.

En revanche, chez les eucaryotes, l'élément **SECIS** est situé dans le 3'-UTR (*untranslated region*) des ARNms encodant des sélénoprotéines. L'élément SECIS des eucaryotes est formé de motifs hélices-boucle conservés dont la structure diffère de celle du bSECIS. L'incorporation de **Sec** requiert dans ce cas la protéine **SBP2** (*SECIS Binding Protein 2*) et **EFSec**, un facteur d'élongation spécialisé (Allmang & Krol, 2006b). La protéine ribosomale **RPL30** joue également un rôle dans ce mécanisme (Chavatte et al, 2005)

1.3.1. Facteurs agissant en *Cis*

1.3.1.1. *Selenocysteine Incorporation Sequence (SECIS)*

SECIS est un élément d'ARN contenant un motif tige-boucle qui est indispensable à l'incorporation de la sélénocystéine en réponse au codon UGA. Sa séquence, sa structure et sa localisation sur l'ARNm varient en fonction des règnes ou trois grands domaines de la vie.

1.3.1.1. a. Localisation dans l'ARNm

Tel que mentionné précédemment, l'élément SECIS est situé immédiatement en aval du codon UGA de la région codante d'une sélénoprotéine bactérienne (Böck, 2006). Contrairement aux Bactéries, l'élément SECIS est retrouvé en 3'UTR de l'ARNm des sélénoprotéines chez les eucaryotes et les archéobactéries, suggérant ainsi des similitudes au niveau du mécanisme d'insertion de **Sec** entre ces deux règnes. L'intérêt d'avoir l'élément **SECIS** en **3'UTR** plutôt que dans la région codante réside en ce que la séquence d'ARN n'est pas contrainte de maintenir à la fois l'activité codante et l'appariement des bases de l'élément **SECIS**. La localisation de SECIS dans le 3'-UTR introduit ainsi un degré de flexibilité en bouclant la séquence intermédiaire qui la sépare du codon UGA. SECIS peut donc en raison de cette flexibilité interagir avec des codons **Sec** distants, UGA. En outre, sa présence dans le **3'-UTR** permet aux ARNms de sélénoprotéines d'arborer plus d'un codon UGA (**Sec**). En effet, la présence du SECIS en **3'-UTR** lève la nécessité des structures en tige-boucle dans la région codante, permettant une flexibilité complète du site d'emplacement du codon UGA (Berry et al, 1933). La présence du **SECIS** en **3'-UTR** est un des éléments qui confère aux eucaryotes un mécanisme d'insertion de **Sec** différent de celui des bactéries. Par exemple, il permet un échange ou une transition rapide et efficace entre **EFsec** (*Sec-specific elongation factor*) seul et le complexe **Sec-ARNt^{Sec}/EFSec**, qui semble essentiel dans le cas de multiples codons UGA (Tujebajeva et al, 2000). C'est le cas de la sélénoprotéine P (SelP) dont

l'ARNm possède 10 à 18 codons Sec (UGA) contrairement à la majorité des ARNms de sélénoprotéines qui contiennent un codon UGA unique (Ma et al, 2002).

1.3.1.1.b. Structure secondaire

Bien que les similitudes de séquence des éléments SECIS soient faibles, la structure secondaire (2D) de ces derniers est bien conservée à travers les grands règnes de la vie. L'élément SECIS des bactéries est essentiellement formé d'une structure en tige-boucle constitué de 40 nucléotides. La séquence primaire varie d'une espèce à l'autre mais celle secondaire demeure conservée dans plusieurs organismes. La boucle apicale est requise pour la liaison du facteur d'élongation spécifique **SelB** (Böck, 2006).

La structure secondaire des éléments SECIS chez les eucaryotes fut proposée selon des études de sondage structurel (Walczack et al, 1998). Il existe deux formes ou types de SECIS fonctionnels chez les eucaryotes dénommées respectivement forme **1** et forme **2**. Ces deux formes ont des structures conservées contenant des boucles internes, des hélices et quatre paires de bases non-Watson-Crick consécutives appelés *quartet*. Ces formes présentent également une série de résidus A dans l'apex et A/G en position 5' du quartet (figure 6). La forme **2** possède une hélice additionnelle mais une boucle apicale plus courte (Fagegaltier et al, 2000b). Des données d'alignement de séquences disponibles sur le sélénoprotéome des eucaryotes font ressortir une collection de 62 forme **1** et 224 forme **2**, suggérant ainsi que la forme 2 est prédominante (Chapple et al, 2009). Cependant, des ARNms encodant la même sélénoprotéine peuvent arborer différentes formes de SECIS en fonction de l'espèce considérée. L'ARNm de SelM contient par exemple, un SECIS de type **2** chez les mammifères tandis que c'est la forme **1** qui est retrouvée chez zebrafish (Korotkov et al, 2002). Par ailleurs, l'introduction des mutations dans l'apex des formes 1 et 2 a montré que ces deux formes fonctionnent équitablement sous des conditions expérimentales données (Grundner-Culemann et al, 1999). Ainsi l'existence de deux types d'élément SECIS demeure inexplicée.

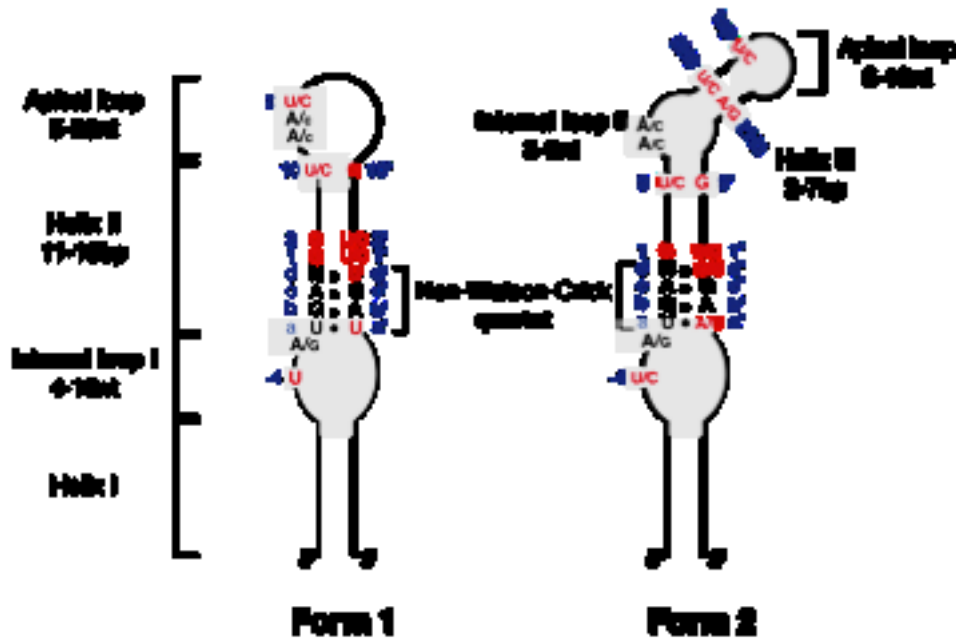


Figure 6. Modèles des structures secondaires des formes 1 et 2 de l'élément SECIS.

La séquence conservée et les motifs structuraux sont indiqués. Les nouveaux résidus conservés et retrouvés par alignement des SECIS sont indiqués en rouge (Chapple et al, 2009). La position des résidus conservés sont indiqués en bleu. abcd/a'b'c'd' indique les paires de bases formant le quartet non-Watson-Crick. Les numérotations (-4, 1, 2, 9, 10, 1' et 2') représentent les distances par rapport au quartet. La position 'z' est le premier nucléotide après le tandem A/Cs. Les positions 2H3/2'H3 représente la deuxième paire de base de l'Hélice III et 1ap est le premier nucléotide de la boucle apicale. Les structures sont tirées de (Fagegaltier et al, 2000b; Grundner-Culemann et al, 1999; Walczak et al, 1998; Walczak et al, 1996).

Le *quartet* formé de paires de bases non-Watson-Crick est essentiel pour l'incorporation de la sélénocystéine *in vivo*, il constitue le site de liaison de SBP2 (SECIS binding protein 2). C'est un motif qui contient en son centre un appariement de bases en cisaille et disposé en tandem, G.A/A.G (Fagegaltier et al, 2000b; Walczak et al, 1998). Ce motif est aussi retrouvé dans d'autres RNAs tel que l'ARN ribosomal, snRNAs et snoRNAs. Il constitue une structure conservée appelé kink-turn (*K-turn*) (Klein et al, 2001). Le motif *K-turn* étant une structure qui se lie aux protéines dans la plupart des cas et qui médie les interactions tertiaires au sein des ARNs. Ce motif d'environ 15 nucléotides est une structure double-brin qui se caractérise par un empilement de bases suivant la séquence 'hélice interne-boucle-hélice' (figure 7).

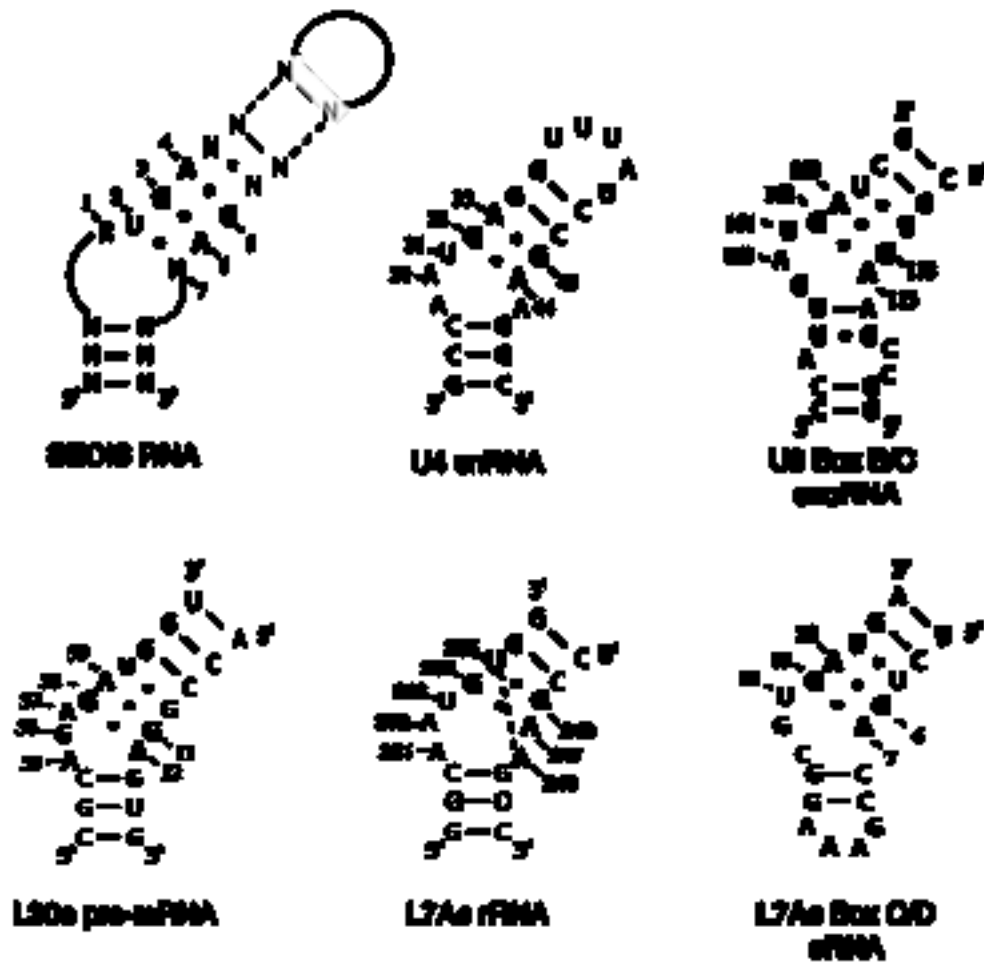


Figure 7. Structures secondaires du SECIS et d'autres RNAs contenant le motif *k-turn*. Les structures secondaires des ARNs : U4snRNA, L30e pré-mRNA, L7A2 rRNA, L7Ae box C/D sRNA proviennent des structures cristallographiques des complexes RNA-protéine correspondant (Chao & Williamson, 2004; Moore et al, 2004; Vidovic et al, 2000). Les structures de l'ARN SECIS et du U3Box B/C snoRNA ont été déterminées par des analyses de sondage structurel (Fagegaltier et al, 2000b; Marmier-Gourrier et al, 2003; Walczak et al, 1998; Walczak et al, 1996). La paire de base "sheared" G.A/A.G est indiquée en gras. La figure provient de (Allmang & Krol, 2006a).

1.3.1.2. *Selenocysteine codon Redefinition Element (SRE)*

Un autre facteur agissant en *Cis*, récemment reporté dans l'incorporation de la sélénocystéine est l'élément **SRE**. Il s'agit également d'une structure tige-boucle phylogénétiquement conservée et adjacent au codon UGA situé dans la région codante des ARNms de sélénoprotéines (figure 8B). Cet élément s'avère suffisant pour stimuler le recodage du codon UGA en sélénocystéine dans les ARNms synthétiques dépourvus de l'élément SECIS. Une meilleure activité de recodage est toutefois observée en présence de l'élément SECIS. Cette étude fut conduite en utilisant l'ARNm encodant la sélénoprotéine N (SelN), mais des approches bio-informatiques ont permis de déceler la présence de l'élément SRE dans d'autres sélénoprotéines tels que SPS2, SelH, SelO et SelT (Howard et al 2005). La structure bidimensionnelle de cet élément n'est cependant pas conservée d'un ARNm à l'autre. La présence sélective de l'élément SRE dans certaines et non pas dans toutes les ARNms sélénoprotéiques implique un rôle dans la régulation différentielle de l'expression des sélénoprotéines au niveau traductionnel. Quatre mutations ponctuelles entraînant la myopathie liée au SEPNI ont été retrouvées dans l'élément SRE de SelN. L'une d'entre elles déstabilise la structure secondaire du SRE via l'abolition d'une paire de base G-C, aboutissant conséquemment à la réduction de l'incorporation de Sec et parallèlement des niveaux de SelN (Maitiet et al, 2008). Ces données mettent en évidence l'importance de l'élément SRE pour la synthèse des sélénoprotéines.

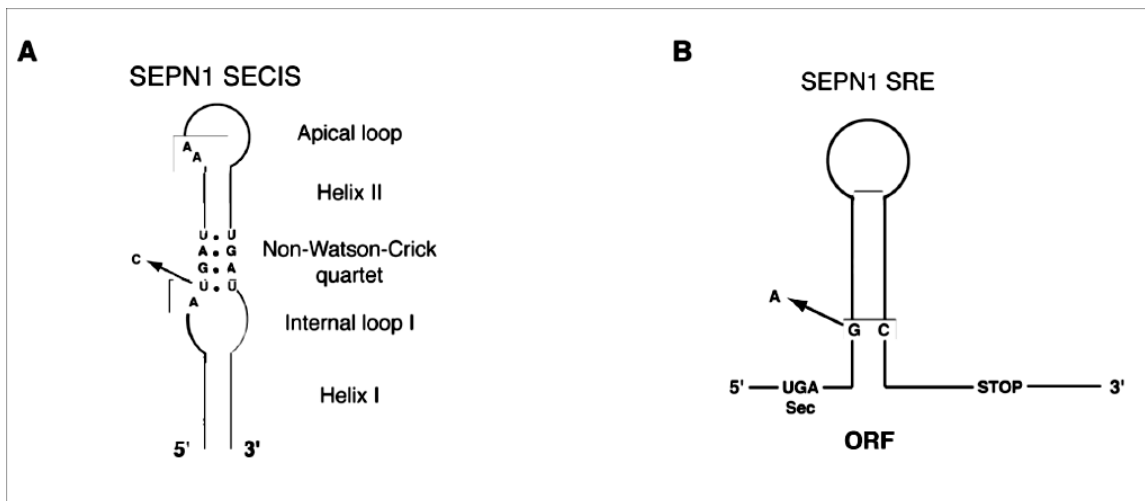


Figure 8. Les éléments SECIS et SRE des ARNms de SEP1 **A.** Structure secondaire de l'élément SECIS de SEP1. Le résidu conservé U dans le quartet non-Watson-Crick est essentiel pour la reconnaissance de SBP2 et la mutation de U en C abolit la liaison de SBP2 (flèche). Cette mutation fut initialement retrouvée dans l'élément SECIS de la sélénoprotéine N (SEPN) provenant d'un patient atteint de myopathie relié au SEP1 (Maiti et al, 2008). **B.** Structure secondaire de l'ARN SRE. Le motif tige-boucle est localisé au sein du cadre de lecture ouvert (ORF) de certains ARNms de sélénoprotéines en occurrence la sélénoprotéine N (SEP1) (Howard et al, 2005). La mutation du G en A fut retrouvée chez un patient atteint de myopathie SEP1-dépendante (Allamand et al, 2006).

1.3.2. Facteurs agissant en *Trans*

1.3.2.1. EFSec

Le recodage du codon UGA en sélénocystéine requiert un facteur d'élongation spécifique au conjugué sélénocystéyl-ARNt^{[Ser]^{Sec}}. Chez les procaryotes et les eucaryotes, il est respectivement connu sous les acronymes **SeIB** et **EFsec**. Le facteur procaryotique est une protéine dont la liaison avec l'ARN est GTP-dépendante. Il est constitué d'une part d'un domaine N-terminal spécifique à ARNt^{Sec} et d'autre part, d'un domaine C-terminal qui lie l'élément SECIS (figure 9). Son domaine N-terminal présente des similitudes de séquences et une homologie fonctionnelle avec le facteur d'élongation EFTu (facteur conventionnel chez les procaryotes) (Forchhammer, K et al, 1998). EFsec assure la livraison du Sec-ARNt^{Sec} au site A du ribosome pour l'incorporation de Sec, se dissocie

subséquemment du ribosome et de l'élément SECIS et se réassocie avec Sec-ARNt^{Sec} pour une nouvelle incorporation de sélénocystéine (Thanbichler, M. et al, 2000). Des analyses *in silico* des génomes eucaryotiques ont permis d'identifier le facteur d'élongation sélénocystéique chez les murins, **eEFSec** (Tujebajeva, R. M et al 2000). Contrairement à SelB, eEFsec interagit indirectement avec l'élément SECIS par le biais de son interaction avec SBP2 (SECIS Binding protein 2) en présence de Sec-ARNt^{Sec}, formant par conséquent la machinerie d'insertion de **Sec** (Tujebajeva, R. M et al 2000).

Tout comme SelB, le domaine N-terminal de eEFsec exhibe des similitudes de séquence avec le facteur d'élongation général EF1A et contient des domaines homologues aux domaines G1-G4 et GTP (Fagegaltier et al, 2000a). La longueur du domaine C-terminal varie selon les organismes.

EFSec contient un signal d'export et de localisation nucléaire dans les domaines N et C-terminal respectivement. La localisation subcellulaire d'EFSec dépend des lignées cellulaires et semble être influencé par les niveaux et la localisation de SBP2 (de Jesus et al, 2006). Le facteur EFSec des archéobactéries est également connu sous le nom de SelB et présente une séquence caractéristique du SelB bactérien et de l'eEFSec eucaryotique (Rother et al, 2000). Sa structure cristallographique révèle une forme générale de très près semble à celle observée pour le facteur d'initiation de la traduction IF2/eIF5B (Leibundgut et al, 2005). Une telle ressemblance amène à s'interroger sur les similitudes mécanistiques entre l'incorporation de Sec et l'initiation de la traduction.

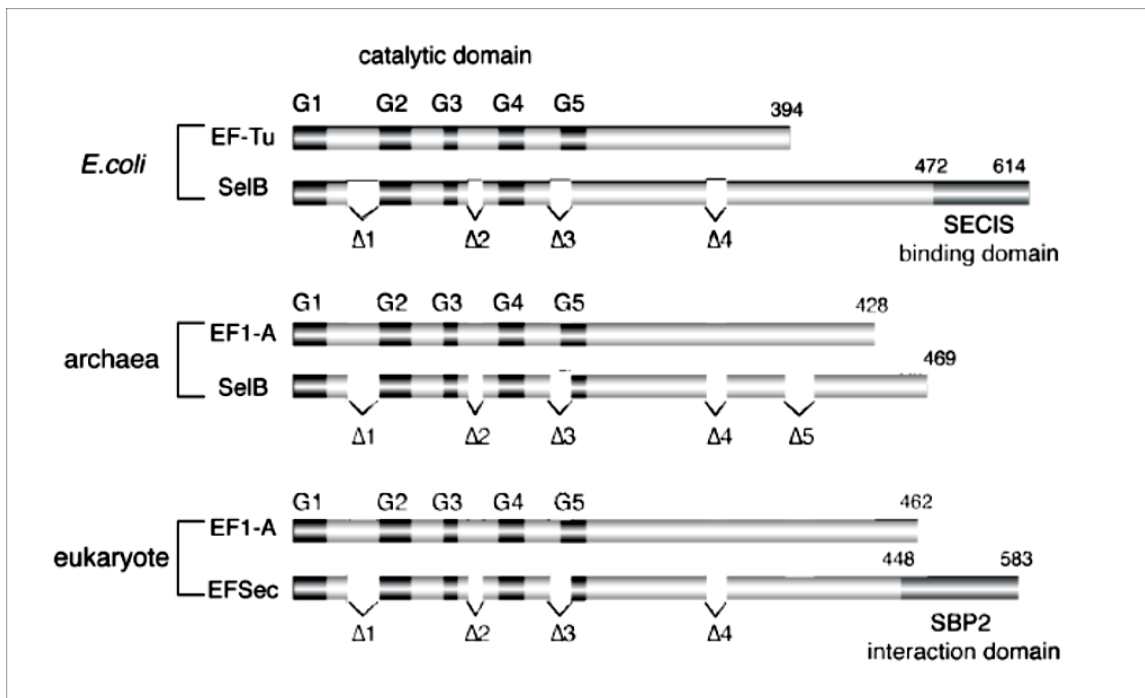


Figure 9. Représentation des facteurs d'élongation spécifiques à la sélénocystéine.

Les facteurs d'élongation spécialisés pour l'insertion de Sec chez *E. coli*, les archéobactéries et les eucaryotes (SelB ou EFSec) sont comparés avec les facteurs d'élongation classiques à savoir EF-Tu ou EF1-A. Les domaines de liaison du GTP sont indiqués par G1-G5. Δ1-Δ5 sont des régions de délétion relative aux facteurs EF-Tu ou EF1-A. Chez *E. coli* et chez les eucaryotes, l'extension C-terminal du facteur d'élongation contient respectivement le domaine de liaison à SECIS et le domaine d'interaction avec SBP2. (Allmang & Krol, 2006b).

1.3.2.2. SBP2

Parmi les facteurs agissant en *trans* lors de l'incorporation de la sélénocystéine, SBP2 (*SECIS binding protein 2*) est le facteur le mieux caractérisé à ce jour. SBP2 fut découvert par chromatographie d'affinité sur un extrait de testicule de rats en utilisant l'élément SECIS comme ligand (Copeland, P. R., et al, 1999). SBP2 joue un rôle majeur dans l'insertion de la sélénocystéine. Elle exerce sa fonction en s'associant aux éléments **SECIS** et en recrutant le complexe eEFSec-Sélénocysteinyl-ARNt^{Sec} au ribosome (Tujebajeva, R. M. et al, 2000). En effet, **SBP2** se lie spécifiquement d'une part à la grande sous-unité du ribosome et d'autre part à une boucle interne située au cœur de

l'élément SECIS (figure 10). **SBP2** possède un domaine de liaison à l'ARN similaire à celui de la protéine ribosomale **L30**, et des protéines eRF1 et SUP1, un suppresseur de l'arrêt de la traduction chez les levures (Copeland, P. R. et al 2000). La mutagenèse de ce domaine de liaison conservé abolit la liaison au **SECIS** mais pas au ribosome. Des études récentes rapportent des mutations survenant naturellement et affectant l'interaction **SBP2-SECIS** et l'efficacité du décodage de **Sec** (Copeland, P. R. et al 2001). Chez les patients atteints de dystrophie musculaire vertébrale, une mutation homozygote ponctuelle au sein du domaine de liaison à l'ARN de SBP2 a été identifiée. SBP2 se liant dans ce cas à l'élément SECIS de SelN. Cette mutation réduit l'affinité de liaison de SBP2 et les niveaux de SelN se trouvent régulés à la baisse. L'affinité de liaison de SBP2 varie selon les ARNms sélénoprotéiques et s'impose selon l'état actuel des connaissances sur le sujet comme étant le facteur déterminant dans la traduction différentielle des ARNms sélénoprotéiques et la sensibilité de ces derniers au mécanisme de désintégration des ARNms non-sens (ARNms ayant un codon stop prématuré) (Squires, J. E. et al 2007). Les études de localisation cellulaire suggèrent que SBP2 subit une translocation nucléocytoplasmique et forme un complexe avec la machinerie d'incorporation de Sec dans le noyau, permettant ainsi aux ARNms sélénoprotéiques de contourner ce mécanisme de dégradation des ARNms non-sens (de Jesus, L. A et al 2006). Par ailleurs, l'état redox de SBP2 régule la localisation subcellulaire de SBP2 et l'activité du décodage de la sélénocystéine. Le stress oxydatif induit la séquestration nucléaire de SBP2 et possiblement les ARNms auxquels il est associé, entraînant ainsi la régulation à la baisse de la synthèse sélénocystéine (Papp, L. V., et al 2006)

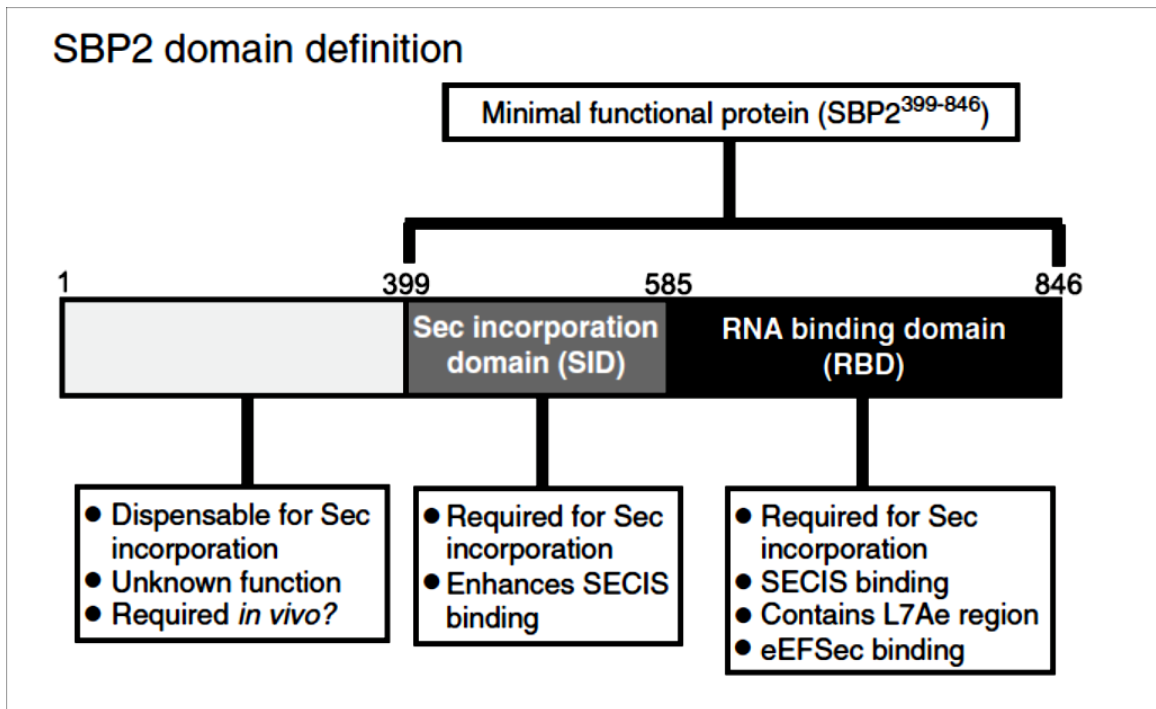


Figure 10. Domaines structuro-fonctionnels de SBP2. La figure est tirée (revue Sélénium 3^{ème} ed par Dolph L. Hatfield et al 2011)

1.3.2.3. RpL30 (*Ribosomal protein L30*)

La deuxième protéine de liaison à l'élément SECIS à avoir été fonctionnellement caractérisée est L30, une composante de la large sous-unité du ribosome chez les eucaryotes. Contrairement à SBP2, L30 est omniprésente et abondamment exprimée dans les tissus de mammifères. Bien qu'elle soit majoritairement associée au ribosome, des fractions de L30 sont retrouvées dans le noyau dans lequel il s'associe aux pré-ARNms de L30 afin de prévenir l'épissage de ces derniers. La fraction cytoplasmique de L30 quant à elle se lie aux ARNms matures de L30 afin d'inhiber leur traduction. Une telle boucle d'autorégulation suggère que l'expression de L30 est fortement contrôlée dans les cellules de mammifères (Dolph L. Hatfield et al, 2011).

Les sites de liaison de L30 et de SBP2 à l'élément SECIS se chevauchent. Des études de liaison et des essais d'incorporation de Sec ont montré d'une part que L30 promeut le

décodage de Sec et d'autre part que L30 et SBP2 sont en compétition pour le site de liaison à l'élément SECIS en présence d'une concentration élevée en magnésium. Le magnésium semble induire la formation du motif *k-turn* dans la structure de l'élément SECIS, favorisant par conséquent la liaison préférentielle de L30 au détriment de SBP2 (Chavatte, L. et al, 2005). Durant le recodage du codon UGA, L30 agirait en délogeant SBP2 d'un ARNm sélénoprotéique par le biais d'un changement conformationnel au sein de l'élément SECIS ce qui permet à la machinerie de décodage de Sec d'être recrutée à la grande sous-unité du ribosome. La dissociation de SBP2 de l'élément SECIS permet la translocation du Selenocysteyl-tRNA^{Sec} au site A du ribosome et stimule l'activité GTPase de eEFSec, qui par conséquent permet de recycler ces facteurs *trans*, lesquels seront réutilisés dans les cycles d'incorporation subséquents (Chavatte, L. et al, 2005).

1.4. SecP43

Identifié par une expérience de PCR dégénérée criblant des protéines de liaison d'ARN, **SecP43**, qui fait l'objet de ce mémoire est une protéine de 32 kDa hautement conservée et comprenant deux domaines de liaison ribonucléoprotéiques dans la région N-terminale et une extrémité C-terminale riche en acides aminés polaires et acides (Feng Ding et al, 1999). **SecP43** et son ARNm correspondant sont généralement exprimés tant dans les tissus de rat que dans des lignées cellulaires de mammifères. Afin d'élucider sa fonction biologique, des anticorps spécifiques à **SecP43** ont permis d'identifier ses partenaires d'interactions. Ainsi l'application d'un extrait cellulaire de la lignée native Hela sur une colonne d'anticorps anti-**SecP43** a permis de purifier un ARN de 90 nucléotides qui correspond à l'ARNt^{Sec} tel que suggéré par les données du séquençage direct et par buvardage de type *Northern* (Feng Ding et al, 1999). Par ailleurs, une protéine de 48 kDa correspondant à SecS, une composante de la machinerie de biosynthèse de Sec fut également identifiée par immuno-précipitation. Selon les données de cette étude, SecP43 forme un complexe ternaire avec Sec-ARNt^{Sec} et SecS. De plus, le retrait ciblé d'une de ces deux protéines affectait l'affinité de liaison de l'autre (Xu XM. et al, 2005). **SecP43** interagit également avec le complexe eEFsec/Sec-ARNt^{Sec} *in vitro* et avec le complexe SecS/SPS1 *in vivo* (Small-Howard et al, 2006).

En outre, des études d'interférence par l'ARN, ont montré que la suppression de **SecP43** dans les lignées NIH3T3 et TCMK-1 conduit à une réduction de la méthylation du groupement 2-hydroxyribosyl situé en position Wobble (Um34) (figure 11) sur l'ARNt^{Sec}, et par conséquent, à une réduction du niveau d'expression de la glutathione peroxydase 1 (GPX1), une sélénoprotéine impliquée dans le stress oxydatif. D'autres études de même nature ont montré que la double suppression de **SecP43** et de SecS entraîne la réduction de l'expression des sélénoprotéines de façon générale (Xu XM et al, 2005).

La localisation subcellulaire de **SecP43** a été déterminée par microscopie confocale à fluorescence. **SecP43** est principalement retrouvé dans le noyau tandis que SecS est abondamment retrouvé dans le cytoplasme. La co-transfection de ces deux protéines provoque une translocation nucléaire de SecS, suggérant ainsi que **SecP43** médierait la navette de SecS et de Sec-ARNt^{Sec} entre les différents compartiments cellulaires. Toutes ces données prises ensemble établissent une fonction putative de SecP43 dans la biosynthèse de la sélénocystéine via un jeu d'interactions avec des partenaires multiples (Xu XM. et al, 2005). Une activité enzymatique pourrait également lui être attribuée en raison de l'impact de son absence sur la méthylation du résidu Um34 de l'ARNt^{Sec}. Toutefois aucune caractérisation structuro-fonctionnelle n'a jusqu'à présent été menée dans le but de saisir les spécificités des interactions entre **SecP43** et ses multiples partenaires d'interactions.

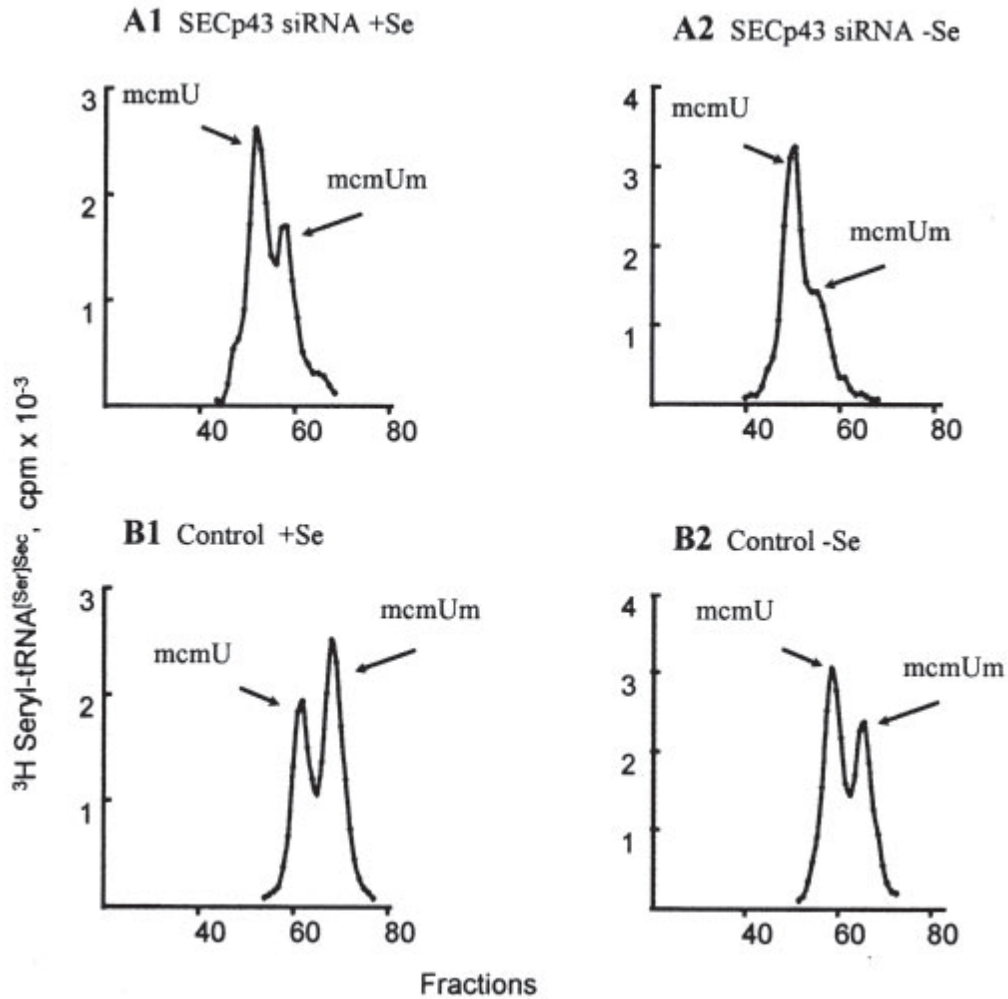


Figure 11. Effet de l'inhibition de SecP43 par siRNA sur la synthèse de Sec-ARNt^{Sec}Um34. L'ARNt total fut isolé des cellules NIH3T3 et aminoacylé avec la sérine [³H]. Le conjugué [³H]séryl-tRNA qui en résulta fut purifié par chromatographie sur une colonne RPC-5. [³H]Séryl-ARNtSec fut ensuite transfecté de concert avec la construction SecP43 siRNA (panneau A) ou avec le vecteur control (panneau B) et la croissance des cellules est faite en présence (+Se) ou en absence (-Se) du sélénium additionnel. Les variations des positions d'éluions des deux isoformes de tRNA^{[Ser]^{Sec} reflètent l'usage d'un gradient de NaCl légèrement différent pour chaque élution. Figure tirée de (Xu XM. et al, 2005).}

1.5. Autres protéines

En 2000, la **nucléoline**, une protéine bien connue pour son implication dans la synthèse du ribosome fut identifiée comme une protéine de liaison d'ARNm sélénocystéique dans une librairie d'ADNc dans laquelle, fut effectué un criblage utilisant comme sonde l'ARNm de la glutathione peroxydase radiomarquée. L'interaction nucléoline-SECIS fut subséquemment confirmée par des essais de retard sur gel d'électrophorèse (Wu, R. et al, 2000). Contrairement à SBP2 qui lie préférentiellement certains ARNms sélénoprotéiques par rapport à d'autres, la **nucléoline** affiche moins de discrimination vis-à-vis de divers ARNms. Ainsi, il est peu probable que la **nucléoline** contribue considérablement dans la hiérarchie des facteurs qui gouvernent la synthèse des sélénoprotéines mais elle pourrait néanmoins jouer un rôle dans la traduction et le transport de Sec (Squires, J. et al, 2007).

NSEP1 (*Nuclease Sensitive Element Binding Protein 1*) est également une protéine qui s'associe à l'élément SECIS et qui semble être fonctionnellement impliquée dans la synthèse sélénoprotéique dans les cellules de mammifères (Shen et al, 1998).

Tel qu'indiqué précédemment, le patron de synthèse des sélénoprotéines dépend de la biodisponibilité du sélénium dans la cellule. Une avancée majeure dans le domaine est la découverte que la protéine **eIF4a3**, qui établit un pont entre le niveau de sélénium disponible et l'expression différentielle des sélénoprotéines. eIF4a3 appartient à la famille des ATPases RNA-dépendantes. Cette protéine est prédominante dans le noyau cellulaire. Malgré ses similitudes avec les facteurs d'initiation canoniques, eIF4a3 n'a aucun rôle connu dans la synthèse protéique. La fonction normale des facteurs initiation canoniques est la liaison des ARNms épissés durant la formation du complexe de jonction entre les exons. Ils sont aussi impliqués dans le mécanisme de dégradation des ARNms non-sens, une réaction de surveillance qui élimine les ARNms contenant des codons stop prématurés. Par contre, eIF4a3 se lie sélectivement à un sous-ensemble d'éléments SECIS et régule l'expression des sélénoprotéines au niveau de la traduction des ARNms (Hatfield, 2011).

1.6. Modèles d'incorporation de Sec

Le mécanisme par lequel la machinerie d'insertion de **Sec** empêche l'arrêt du ribosome au codon UGA demeure très controversé. Cependant deux modèles distincts ont été proposés (figure 12).

Dans le premier modèle, SBP2 se déplace avec le ribosome. Le ribosome s'immobilisant au codon Sec, UGA permet à l'élément SECIS d'interagir avec SBP2, conduisant ainsi le complexe Sec-ARNt^{Sec}/EFSec au site A du ribosome (Donovan et al, 2008; Kinzy et al, 2005). Ce modèle fut proposé en tenant compte des découvertes selon lesquelles SBP2 interagit quantitativement avec le ribosome mais ne peut se lier simultanément à l'élément SECIS. Dans le deuxième modèle, SBP2 lié à l'élément SECIS avant la traduction, recrute le complexe EFSec/Sec-ARNt^{Sec} avant le décodage du codon UGA. Ensuite, approchant le ribosome, L30 déloge SBP2 par compétition pour l'élément SECIS. Le complexe SBP2/EFSec/Sec-ARNt^{Sec} est par conséquent livré au site A du ribosome (Chavatte et al, 2005).

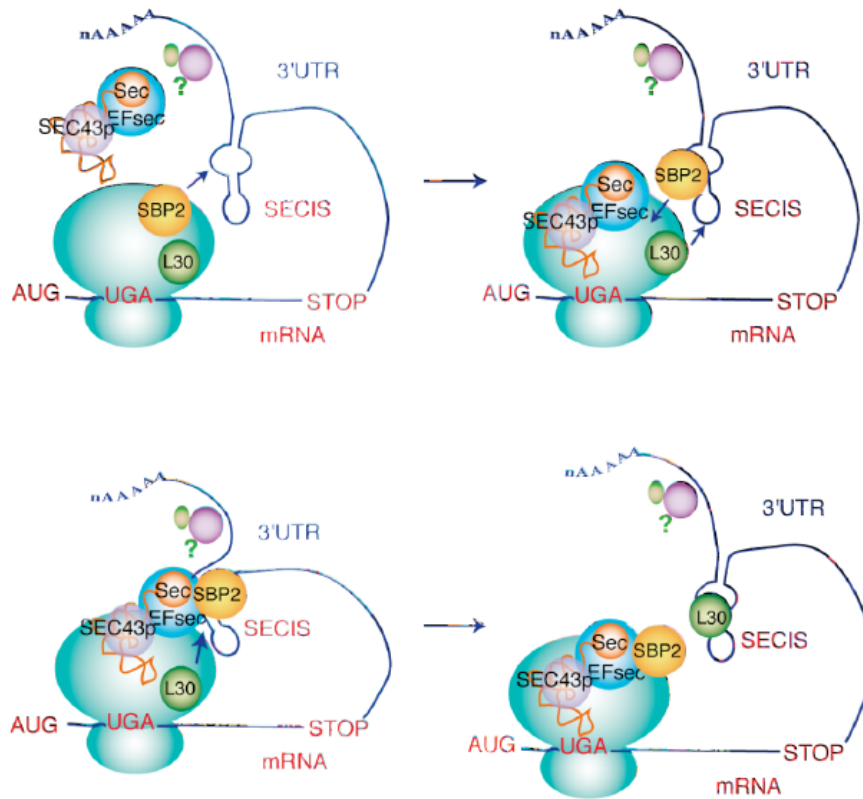
A**B**

Figure 12. Modèles d'incorporation de la sélénocystéine. **A.** Dans les bactéries, SelB s'associe à la fois à Sec-ARNt^{Sec} et à l'élément bSECIS situé directement en aval du codon Sec (UGA). Sec-ARNt^{Sec} porté par SelB est directement conduit au codon UGA via l'interaction bSECIS-SelB, permettant ainsi l'incorporation de Sec dans la chaîne polypeptidique naissante. **B.** Chez les eucaryotes, deux modèles sont proposés. Dans le premier, SBP2 attaché au ribosome, interagit avec l'élément SECIS et le complexe EFSec/Sec-ARNt^{Sec} pour la livraison de ce dernier au site A du ribosome. Par la suite, L30 déloge SBP2 de l'élément SECIS. Dans le deuxième modèle, SBP2 lié l'élément SECIS, recrute le complexe EFSec/Sec-ARNt^{Sec} à l'élément SECIS. Ensuite, L30 attaché au ribosome déloge SBP2 de tel enseigne que le complexe SBP2/EFSec/Sec-ARNt^{Sec} soit légué au ribosome. La figure est tirée de (Allmang & Krol, 2006b).

1.7. Objectifs et grandes lignes de ce mémoire

L'identification de SecP43 a été conduite en 1999 par le groupe du professeur Grabowski. Des données préliminaires sur sa fonction, sa localisation subcellulaire et ses partenaires d'interaction respectifs ont été fournies notamment par des techniques de *gene silencing* utilisant des ARNs d'interférence, de microscopie confocale et de coimmunoprécipitation. Toutefois s'agissant des partenaires d'interaction de SecP43, les observations faites *in vivo* et *in vitro* ne sont pas toujours concordantes. Par ailleurs SecP43 n'a jamais été exprimé et purifié afin d'être caractérisée structurellement. Des groupes de recherche ont tout de même effectué une synthèse *in vitro* des domaines ribonucléiques de SecP43. La structure de ces domaines chimiquement synthétisés fut déterminée par RMN. De telles méthodes ont été utilisées en lieu et place des méthodes d'expression classique car l'homologue humain de SecP43 forme facilement des corps d'inclusion lorsqu'exprimé dans les cellules bactériennes tel qu'*E. coli*. Compte tenu de la complexité du mécanisme d'incorporation de Sec, qui met notamment en jeu plusieurs facteurs dont la dynamique d'interaction demeure non précise, il nous semble que formuler une hypothèse précise sur le mécanisme d'action de SecP43 au sein de cette machinerie serait imprudent. Ainsi notre démarche dans ce mémoire consistera à apporter d'avantage d'informations structurelles sur SecP43 en utilisant une approche *in vitro*. Les objectifs ultimes de ce projet étaient donc premièrement de trouver les conditions optimales d'expression de SecP43, il était ensuite question de la purifier afin de la cristalliser et finalement, de caractériser qualitativement et quantitativement ses interactions avec SepSecS et ARN^{Sec}.

Chapitre 2. Matériels et Méthodes

2. Méthodes

2.1. Souches bactériennes décrites dans le mémoire

Plusieurs souches bactériennes ont été générées pour l'optimisation de l'expression des protéines SecP43 et SepSecS. Le tableau suivant résume les particularités de chacune de ces souches.

Tableau III : Souches bactériennes obtenues par transformation de la souche BL21(DE3)

Souches	Plasmides d'expression	Caractéristiques des plasmides
S ₁	pETDuet-1-SecP43	<ul style="list-style-type: none"> ● pETDuet-1 (Novagen, résistance à l'ampicilline, deux sites de clonage multiples précédés par le promoteur T7) ● pHT (insertion de la séquence du His-tag et du site de clivage par la protéase TEV en aval du promoteur T7 dans le plasmide pET24d (Novagen)) ● pRARE (extrait de la souche Rosetta(DE3)pLysS (Novagen), résistance au chloramphénicol, encode les ARNts humains) ● pBB535 (construit par Dr.Bernd Bukau, résistance à la Spectinomycine, encode les chaperones DnaK et DnaJ, disponible sur addgene.org)
S ₂	pETDuet-1-SecP43 + pRARE	
S ₃	pETDuet-1-SecP43 + pBB535	
S ₄	pETDuet-1-SecP43codonstop	
S ₅	pETDuet-1-SecP43codonstop + pRARE	
S ₆	pETDuet-1-SecP43codonstop + pBB535	
S ₇	pHT-SepSecS + pRARE	
S ₈		

Tableau IV : Séquences des amorces

Paires d'amorces	Séquences des amorces	Code des couleurs
NHis-Secp43-F	5'-AACTGCAGGAAAACCTGTATTTTCAGGGA GCGGCCAGCCTGTGGATGGGC-3'	● site de clivage par PstI
Secp43-R	5'-CCC AAGCTTCTACATCATGGCAGGGATC-3'	● site de clivage par la TEV protéase
Secp43codonstop-F	5'-AAAGCGAGCCGTGTATGAAGCCAGTGGAAATAGTCAGATG-3'	● site de clivage par HindIII
Secp43codonstop-R	5'-ATATTCCACTGGCTTTCATACACGGCTCGCTTAGGGATTGC-3'	● codon stop
SepsecS-F	5'-GGGGTACCAACCGCGAGAGCTTCGCGGC-3'	● site de clivage par KpnI
SepsecS-R	5'-CCGAAGCTTTCATGAAGAAGCATCCTG-3'	

2.2. Tampons divers

La composition des différents tampons et solutions utilisés lors des expériences décrites dans ce mémoire est indiquée ci-dessous.

- **Milieu de culture LB**

10 g NaCl, 10 g Tryptone, 5 g Yeast extract, 1L H₂O

- **Milieu d'auto-induction 4×LAC-8ZYM**

La préparation des solutions stock a été effectuée de la manière suivante :

8ZYM component	50 × LAC	20 × NPSC	Trace elements
4 g <i>Yeast extract</i>	25 g glycerol	5,35 g NH ₄ Cl	1 g FeSO ₄ ×7H ₂ O
8 g Bactotryptone	2,5 g glucose	3,22 g Na ₂ SO ₄	8,8 g ZnSO ₄ ×7H ₂ O
100 mL H ₂ O	10 g lactose	6,8 g KH ₂ PO ₄	0,4 g CuSO ₄ ×7H ₂ O
	100 mL H ₂ O	7,1 g Na ₂ HPO ₄	0,15 g MnSO ₄ ×4H ₂ O
		100 mL H ₂ O	0,1 g Na ₂ B ₄ O ₇ ×10H ₂ O
			0,05 g (NH ₄) ₆ MO ₇ O ₂ ×4H ₂ O
			1 L H ₂ O

Pour la préparation de 400 mL du milieu 4×Lac 8ZYM, les proportions suivantes ont été utilisées :

8ZYM component	372 mL
1M MgSO₄	800 µL
50 × LAC	32 mL
20 × NPSC	20 mL
Trace elements	400 µL
Total	400 mL

- **Laemmli sample buffer (2X)**

1 M Tris/Cl pH 6,8	5 mL
β-mercaptoethanol	4 mL
SDS	1,84 g
Glycerol	8 mL
Bromophenol blue	une pincée
H₂O	23,2 mL
Total	40 mL

- **Tampon CP**

0,5 Tris-HCl, 1 mM EDTA, pH 8.5

2.3. Culture cellulaire

Afin de tester l'expression de SecP43 dans les cellules humaines, la lignée cellulaire HEK 293-6E (*Human embryonic kidney cells from fibroblasts*) fut utilisée. Le milieu de culture F17 (Invitrogen) fut supplémenté avec 4 mM de glutamine, 0.1% pluronic F18 (agent tensio-actif) et 25 µg/mL de G418 (antibiotique). L'agent de transfection utilisé est le PEI (polyéthylènimine) qui forme des complexes avec le plasmide d'expression pTT5

dans lequel le gène de SecP43 a été cloné. Le plasmide pTT5-GFP fut employé comme contrôle positif.

2.4. Transfection des cellules HEK293-6E en suspension

À partir d'une culture de maintenance ayant une densité cellulaire de 10^6 cellules/mL, 1,8 mL de cellules fut ajouté à chacun des puits d'une plaque de six puits. Un mélange ADN plasmidique/PEI fut ensuite effectué dans 200 μ L de milieu F17 préalablement chauffé à 37°C. Après incubation du mélange à température pièce pendant 15 min, ce dernier est rajouté aux cellules contenues dans la plaque de six puits. La plaque est finalement incubée à 37°C avec 5% de CO₂. Afin d'optimiser l'efficacité de la transfection, plusieurs ratios stœchiométriques ADN/PEI ont été criblés à savoir 1/2, 1/3 et 1/4. Ces ratios indiquent que la quantité d'ADN plasmidique demeure constante tandis que la concentration en PEI est variée. Des échantillons de cellules prélevées après 24, 48 et 72 heures d'incubation ont été analysés sur gel SDS-PAGE pour déterminer le temps d'incubation optimal pour l'expression de la protéine d'intérêt à savoir SecP43. Les conditions d'expression optimales furent choisies pour la transfection à grande échelle.

2.5. Clonage de SecP43 et de SepSecS

Les séquences d'ADN qui codent les protéines SecP43 et SepSecS furent respectivement amplifiées à partir des plasmides pUC19-SecP43 et pUC19-SepSecS fournis par (Dan Su de Yale University, Laboratoire du Dr. Söll) en utilisant les amorces répertoriés dans le tableau IV. Un gradient de température a été effectué afin de déterminer la température d'hybridation optimale. La séquence codant pour l'étiquette histidine est situé en aval du promoteur T7 dans le plasmide pETDuet-1. L'amorce directe contient à son extrémité 5' la séquence codant le site de clivage par la protéase TEV (*Tobacco Etch Virus*). Les fragments d'ADN d'environ 1 kb et de 1,5 kb respectivement obtenus suite à l'amplification par PCR furent analysés sur un gel d'agarose de 1 %. Ces fragments, ainsi que les vecteurs d'expression pETDuet-1 (Novagen) et pHT (pET-24d modifié par B. Bessette et Dr. M. Smith), furent parallèlement digérés par les enzymes de restriction Pst

I et Hind III dont les sites de restriction sont respectivement situés en 5' et en 3' de l'insert encodant la protéine SecP43. L'insert encodant SepSecS quant à lui fut digéré par les enzymes HindIII et Kpn I. Les réactions de digestion et de ligation (par la T4 DNA ligase) se sont déroulées selon les conditions recommandées par le manufacturier. Le produit de la ligation fut finalement utilisé pour transformer les cellules compétentes XL1-Blue (souche de clonage) par *Heat Shock*, une technique de transformation qui repose sur la création des pores dans la membrane plasmique bactérienne suite à une augmentation soudaine de la température. Ainsi cette technique comporte une phase d'incubation du mélange (cellules + ADN) sur glace pendant 30 min. Cette phase est suivie d'une incubation à 42°C pendant 1 min et finalement le mélange est retransféré sur glace pour 2 min d'incubation.

La dernière construction a consisté en l'insertion d'un codon stop par mutagenèse dirigée dans la séquence codante de SecP43 après le 540^e codon. Les amorces utilisées pour la mutagenèse dirigée (*quick change*) se trouvent dans le tableau IV. Suite à l'ensemencement sur une boîte de pétri des bactéries nouvellement transformées, les clones sélectionnés au hasard furent séquencés afin d'en trouver ceux qui portaient l'insert. Les plasmides issus des clones positifs ont finalement été utilisés pour la transformation des bactéries compétentes BL21 DE3 (souche d'expression).

2.6. Expression des protéines SecP43 et SepSecS

Afin de trouver les conditions optimales d'expression de nos deux protéines en général, et de SecP43 en particulier, plusieurs conditions d'expressions ont été testées.

2.6.1. Induction à l'IPTG (Isopropyl β -D-1-thiogalactopyranoside)

750 μ L d'une pré-culture bactérienne fut utilisé pour l'inoculation de 50 ml de milieu LB (Luria-Bertani) contenant l'antibiotique approprié. Le mélange a été incubé à 37°C avec agitation (220 rpm) jusqu'à l'atteinte d'une densité cellulaire (D.O.600 nm) d'environ 0,6 à 0,7. L'expression est ensuite induite par l'ajout de 1 mM d'IPTG. Le mélange a finalement été incubé à 20°C, 25°C et 37°C, respectivement, et des prélèvements

ponctuels furent effectués aux intervalles de temps constants pendant 24h. Les échantillons prélevés ont été analysés sur gel SDS-PAGE pour identifier les conditions optimales d'expression et pour suivre la cinétique d'expression des deux protéines d'intérêt.

2.6.2. Auto induction

Le protocole d'auto-induction suit l'article de (Studier, F.W, 2005). Les températures d'incubation testées pour les deux souches bactériennes générées ont été 20°C, 25°C et 37°C. Des prélèvements de 1 mL de culture furent effectués pour évaluer la cinétique d'expression et pour déterminer les conditions thermo-temporelles optimales d'expression. Les temps de prélèvement ont été les suivants: 12, 24, 36, 48, 60 et 72h. Dans le but de collectionner les culots de bactéries, les échantillons prélevés après différents temps d'incubation ont été centrifugés à 13000 rpm pendant 1 min. Les culots ainsi obtenus ont été ré-suspendus dans un volume de Laemmli *sample buffer* calculé selon la formule qui suit : (Volume de tampon de lyse = 50 × D.O.). Les culots ré-suspendus ont finalement été chauffés à 100°C pendant 10 min pour lyser les cellules. Une centrifugation additionnelle a permis la séparation du lysat des débris cellulaires. Les fractions des protéines solubles et insolubles (culot) ont été analysées par gel SDS-PAGE. La condition d'expression optimale identifiée (milieu 4×Lac-8ZYM, 25°C, 72h) suite à cette analyse a été reproduite à grande échelle.

2.7. Purification des protéines SecP43 et SepSecS

Les protéines de fusion N-terminal-hexahistidine ont été surexprimées dans les cellules *E. coli* BL21 (DE3) en utilisant les constructions pETDuet-1-SecP43, pHT-SepSecS et le vecteur pRARE (extrait de la souche Rosetta) codant pour les ARNt humains des codons rares chez *E. coli*. Pour chaque protéine, 4 litres de culture induite ont été centrifugés et les culots ont été ré-suspendus dans du tampon de lyse (pour SecP43 : 20 mM sodium phosphate, 0,5 M NaCl, 40 mM imidazole, pH 7,4, DNase (1µg DNase/mg de culot) ; pour SepSecS : 20 mM Tris pH 8,0, 50 mM NaCl, 10% (v/v) glycerol, 1mM 2-mercaptoethanol, 10 µL PLP (*pyridoxal phosphate*)). Le volume en (mL) du tampon de lyse ajouté étant équivalent à 10 fois la masse en gramme (g) du culot correspondant. Les

bactéries ré-suspendues ont été lysées par le *cell disruptor* (Constant systems, 27 000 kpsi). Le lysat fut centrifugé deux fois (rotor SS34, 18 000 rpm, 30 min à 4°C), et le surnageant a été appliqué sur un système de chromatographie à performance rapide (ÄKTA Purifier, GE Healthcare) avec une colonne IMAC préalablement chargée au Ni²⁺ (HisTrap TM HP, 5 ml, GE Healthcare). L'élution des protéines étiquetées fut réalisée par un gradient linéaire d'imidazole en 50 min à un débit de 1 ml/min. Suite à l'application du lysat, la colonne fut d'abord lavée avec 50 ml du tampon **A** (SecP43 : 20 mM sodium phosphate, 40 mM imidazole, 0,5 M NaCl pH7,4 ; SepSecS : 20 mM Tris pH 8, 50 mM imidazole, 300 mM NaCl, 1 mM 2-mercaptoethanol, 10 µM PLP). L'élution a été faite avec les tampons **B** (20 mM sodium phosphate pH 7,4, 500 mM imidazole, 0,5 M NaCl) pour SecP43 et tampon **C** (20 mM Tris pH 8, 500 mM imidazole, 300 mM NaCl, 1 mM 2-mercaptoethanol, 10 µM PLP) pour SepSecS. Les fractions contenant les protéines ont été combinées, et concentrées en vue d'un changement de tampon requis pour la réaction de clivage de l'étiquette par la protéase TEV (25 mM NaPi, 125 mM NaCl, 5 mM DTT pH 7,4 à 20°C pendant 16 h ; le ratio stœchiométrique enzyme/substrat étant de 1/70). L'échantillon fut réappliqué sur la colonne IMAC afin de récupérer les protéines désormais dépourvues de leur étiquette. Les protéines ont été finalement purifiées davantage sur une colonne d'exclusion stérique (Superdex TM 75 10/300 GI GE Healthcare) avec un débit de 0,5 ml/ml dans les tampons **D** (20 mM Tris pH 7,5, 200 mM NaCl) pour SecP43 et tampon **E** (20 mM Tris pH 8, 200 mM NaCl 1 mM DTT, 10 µM PLP) pour SepSecS. En vue d'une utilisation éventuelle, les protéines purifiées ont été stockées à -80°C dans les tampons d'élution respectifs avec et sans glycérol 10%.

2.8. Analyse de SecP43 par spectrométrie de masse.

Des aliquots de SecP43 purifiés avant et après l'insertion du codon stop ont été soumis à la plateforme de protéomique de l'Institut de recherche en immunologie et oncologie (IRIC) de Montréal pour des analyses de spectrométrie de masse de type MS/MS. Les échantillons ont préalablement été digérés par la trypsine avant d'être ionisés par le spectromètre de masse. Les peptides générés ont été identifiés par le logiciel Mascot et d'autres logiciels développés par la dite plateforme. La reconstitution de la séquence a été faite manuellement.

2.9. Analyse structurale de SecP43

2.9.1. Essais de cristallisation

Des essais de cristallisation par diffusion de vapeur de la goutte assise ont été réalisés en utilisant des plaques de 96 puits. Les concentrations protéiques testées ont été de 5, 15, 25 et 40 mg/ml. Le pH de la solution protéique ([20 mM Tris, 50 mM NaCl] ou [20 mM phosphate, 50 mM NaCl]) a aussi été varié (6,5, 7, 7,5 et 8). Les différentes solutions de criblage proviennent de la compagnie *Molecular Dimensions* (CSS1&2, Morpheus, PACT premier, JSCG+, SS1&2, Morpheus Green, Midas). Le robot 'Mosquito' a été utilisé pour le mélange des gouttes de solutions protéiques et de celles de l'agent précipitant.

2.9.2. Mesure et analyse des données de BIOSAXS (Small Angle x-rays scattering)

Les données de diffraction ont été mesurées sur un générateur de rayons X à CHESS (*Cornell High Energy Synchrotron Source*, Ithaca, USA). Les caractéristiques du faisceau utilisé sont les suivantes : Station F2, E = 9.881 keV ; diamètre du faisceau : 250 μm \times 250 μm ; flux = 2×10^{10} photons/sec. Le détecteur utilisé a été le *dual Pilatus 100k-S SAXS/WAXS* pouvant détecter à la fois les rayons à petit et à grand angles. La distance échantillon-détecteur a été ajustée à tel enseigne que les angles minimale et maximale soient respectivement de $q_{\text{min}} = 0.008 \text{ \AA}^{-1}$ et de $q_{\text{max}} = 0.8 \text{ \AA}^{-1}$ où q représente le **momentum transfer** ou **la norme du vecteur de diffraction** qui est une fonction de l'angle de diffraction θ comme l'indique la formule : $q = \frac{4\pi}{\lambda} \sin(\theta/2)$. λ étant la longueur d'onde du faisceau de rayons X. 25 μL d'échantillon fut ajouté dans la cellule et une série de 10 expositions de 30 secondes chacune ont successivement été appliquées au tampon dans un premier temps et à l'échantillon protéique correspondant dans un deuxième temps. Afin de calculer l'intensité de diffraction moyenne, chaque exposition a été inspectée minutieusement afin de déceler des signes de dégradation protéique due à l'irradiation aux rayons x (aucune n'a été trouvée). L'intensité de diffraction du tampon a été soustraite de celle de l'échantillon protéique pour n'avoir que l'intensité de diffraction due à la protéine.

Les patrons de diffraction ont été enregistrés à différentes concentrations protéiques afin de confirmer l'effet de la concentration sur l'intensité de diffraction. Cette relation permet d'une part de valider le jeu de données en cas de linéarité. Le rayon de giration, R_g fut initialement évalué en utilisant l'approximation de Guinier (Guinier and Fournet, 1955). Dans cette approximation, le rayon de giration correspond à la pente du graphique du logarithme de $I(q)$ en fonction de q^2 . Toutefois, cette approximation ne s'applique pas efficacement aux protéines non-structurées ou partiellement non-structurées (K. Moncoq, et al 2004). Le rayon de giration fut aussi évalué en utilisant la suite GNOM (Svergun, 1992), un algorithme qui calcule la fonction de distribution des distances $p(r)$ d'une protéine, qui est un histogramme des distances inter-particulaires dans la protéine. Sa valeur est égale à zéro lorsque r est supérieur à D_{max} , distance maximale dans la protéine.

La conformation en solution de **SecP43** a été déterminée en utilisant une méthode *Ab initio*. Le programme DAMMIN représente la protéine comme un ensemble de sphères (*dummy atoms*) de rayon $r_0 \ll D_{max}$ compactées à l'intérieur d'une sphère de diamètre D_{max} (Svergun, 1999). La structure de DAM est définie par un vecteur de configuration X qui contient $N \approx (D_{max}/r_0)^3$ composantes ; $X(i) = 1$ si le *dummy atom* appartient à la protéine et $X(i) = 0$ dans le cas contraire. Par appariement simulé, le programme cherche une configuration compatible avec les données expérimentales tandis qu'une pénalité de décontraction assure la connectivité et la compacité de la solution. Aucune condition particulière sur la forme des particules n'a été imposée comme contrainte dans nos calculs.

2.10. Synthèse *in vitro* et purification de l'ARNt^{Sec}

L'orthologue humain de l'ARNt^{Sec} a été cloné dans le vecteur pUC19, exprimé dans *E. coli* DH5 α et synthétisé *in vitro* par la polymérase T7 selon le protocole développé par le laboratoire du Dr. Legault (Département de biochimie et médecine moléculaire, Université de Montréal). La réaction de synthèse a d'abord été optimisée à petite échelle. Le volume réactionnel étant de 100 μ L, les conditions de la réaction d'optimisation sont résumées dans le tableau ci-dessous.

Tableau V : Conditions des réactions de synthèse de l'ARNt^{Sec}. *

Solution	Volume (μL)	Final Conc.
400 mM Tris pH 8 (1)	10	40 mM
0.5 M MgCl ₂ (4)	5	25 mM
1 M DTT (7) (freshly made)	5	50 mM
25 mM Spermidine (8)	4	1 mM
50 mM each NTP pH 8.1 (200 mM total NTPs)	8	4 mM each NTP
1% Triton X-100 (9)	10	0.1%
2 μ M Template mix	10	0.2 μ M
T7 RNA Polymerase ~1.3 mg/mL	1	0.026 mg/mL
ddH ₂ O	47	
TOTAL VOLUME	100 μL	

(*) Ce tableau est tiré du protocole développé par le groupe du Dr. Legault.

Les paramètres ayant été variés à des fins d'optimisation ont été respectivement la concentration en MgCl₂, en polymérase T7, en matrice d'ADN, en GTP et la durée d'incubation de la réaction allant de 1 à 6 h à 37°C. La condition optimale déterminée par électrophorèse sur gel d'acrylamide fut retenue pour la synthèse à grande échelle se faisant notamment dans un volume réactionnel de 25 mL. L'ARNt^{Sec} ainsi synthétisé a été purifié par des techniques d'électrophorèse sur gel d'acrylamide. L'extraction du gel a été réalisée par la méthode *crush and soak* développée par le protocole du Dr. Legault. L'ARNt^{Sec} recueillie fut finalement quantifiée et incubé à 37°C avec la phosphatase

alcaline de veau (CIP) pendant 4 heures pour se débarrasser des groupements phosphate en 5'. Cette étape étant nécessaire pour éventuel marquage au ^{32}P . Les conditions typiques de cette réaction pour 10 mg d'ARN ont été les suivantes : 1 ml de 10 X Tampon CP (0,5 M TrisHCl, 1 mM EDTA, pH 8,5) + 2,8 mL H₂O + 200 µL CIP 20 U/mg RNA). La réaction a été arrêtée par l'ajout de 1/25 volume (400 µL) d'une solution de 0,5 M EDTA pH 8,0 et l'enzyme fut inactivé par l'incubation du mélange réactionnel à 65°C pendant 5 min. l'ARNt^{Sec} fut finalement purifiée par chromatographie d'échange ionique en utilisant la colonne DEAE Sephacel (GE Healthcare). Les tampons d'équilibration et d'éluion étant respectivement (10 mM Tris pH 7,6, 1 mM EDTA pH 8,0, 0,2 M NaCl) et (10 mM Tris pH 7,6, 1 mM EDTA pH 8,0, 1,5 M NaCl).

2.11. Interaction SecP43-SepSecS-ARNt^{Sec}

2.11.1. Essais de réticulation entre SecP43 et SepSecS

5 µl de chacune des solutions de protéines purifiées, étiquetées *Histag* et de concentration identique (10 pmol/µl) ont été mélangés pendant 5 min à 22°C. Le tampon F (20 mM phosphate pH 7,7, 200 mM NaCl, 10 µl PLP) a été utilisé comme contrôle négatif. 90 µl de tampon CLB (50 MES-KOH, 150 mM NaCl, 1 mM EDTA, pH 6,5) ont ensuite été ajouté et le mélange a été incubé à 22°C pendant 30 min. Plusieurs agents de réticulation ont été testés. Il s'agit de formaldéhyde, de DSS (disuccinimidyl suberate) et de DSP (Dithiobissuccinimidyl propionate). L'agent de réticulation a été ajouté à différentes concentrations (0,05, 0,1, 0,2, 0,4, 0,8, 1,6 %) et les échantillons furent incubés à 22°C, suivi de l'ajout d'un volume de Laemmli *sample buffer* et analysés par SDS-PAGE et immunobuvardage de type *Western*. La réaction de réticulation fut également effectuée en présence d'ARNt^{Sec} afin de vérifier si ce dernier stimule l'interaction SecP43/SepSecS.

2.11.2. Essais de retard de mobilité sur gel d'électrophorèse

Pour ces essais, la concentration en ARNt^{Sec} fut maintenue constante tandis que celle des protéines SecP43 et SepSecS fut variée. Afin d'uniformiser la conformation de l'ARNt^{Sec}, l'échantillon d'ARNt fut chauffé à 95°C pendant 2 min et immédiatement refroidi sur

glace pendant 5 min. Aux mélanges ARN-protéine sont ajoutées du tampon GLB (2 mM DTT + 10% glycérol) pour compléter le volume à 10 μ L. Les échantillons ont préalablement été incubés sur glace pendant 20 min avant d'être déposés sur un gel natif de 10% d'acrylamide. La migration a été faite à 250 V pendant 3 heures. Le gel a été ensuite coloré au *Sybr gold* et finalement révélé par le *ChemiDoc MP* (Bio-rad).

Chapitre 3. Résultats

3.1. Optimisation de l'expression de SecP43

SecP43 est une protéine de 32 kDa (calculé selon la séquence) qui demeure très peu caractérisée biochimiquement malgré son importance dans la synthèse et l'incorporation des sélénocystéines. SecP43 est par sa nature une protéine qui s'agrège rapidement dans les corps d'inclusion au cours d'une surexpression classique dans les souches bactériennes comme *E. coli* BL21. Ainsi, afin de trouver les conditions d'expression optimales qui garantissent d'une part un rendement substantiel et d'autre part une solubilité accrue de SecP43, nous avons testé les différentes méthodes, milieux, plasmides et cellules d'expression présentés ci-après.

3.1.1. Expression de SecP43 dans les cellules humaines HEK 293

Etant donné que nous travaillons avec l'homologue humain de SecP43, il a été jugé approprié de tester la surexpression de cette protéine dans les cellules HEK293-6E en suspension suivant le protocole décrit dans la section 2.3. (Matériels et méthodes). L'avantage de la surexpression dans cette lignée cellulaire humaine réside en ceci qu'elle permettrait la modification post-transcriptionnelle de SecP43 au cas échéant, ce qui est impossible dans le système bactérien. La cinétique d'expression de SecP43 dans cette lignée cellulaire (figure 13) permet d'apprécier l'effet du temps d'incubation et du ratio d'ADN/PEI sur l'expression de SecP43. Notons que le PEI (polyéthylèneimine) est un agent de transfection positivement chargé qui forme des complexes avec l'ADN. Ces complexes se lient aux résidus anioniques de la surface cellulaire et sont transférés à l'intérieur des cellules par endocytose (Longo PA et al, 2013). SecP43 semble migrer vers 50 kDa alors qu'on s'attend à un fragment d'environ 34 kDa (protéine + étiquette His). S'agirait-il d'une modification post-traductionnelle ? Les signaux de faible intensité qui apparaissent au dessus du signal de SecP43 correspondent probablement à des protéines qui interagissent non spécifiquement avec les anticorps.

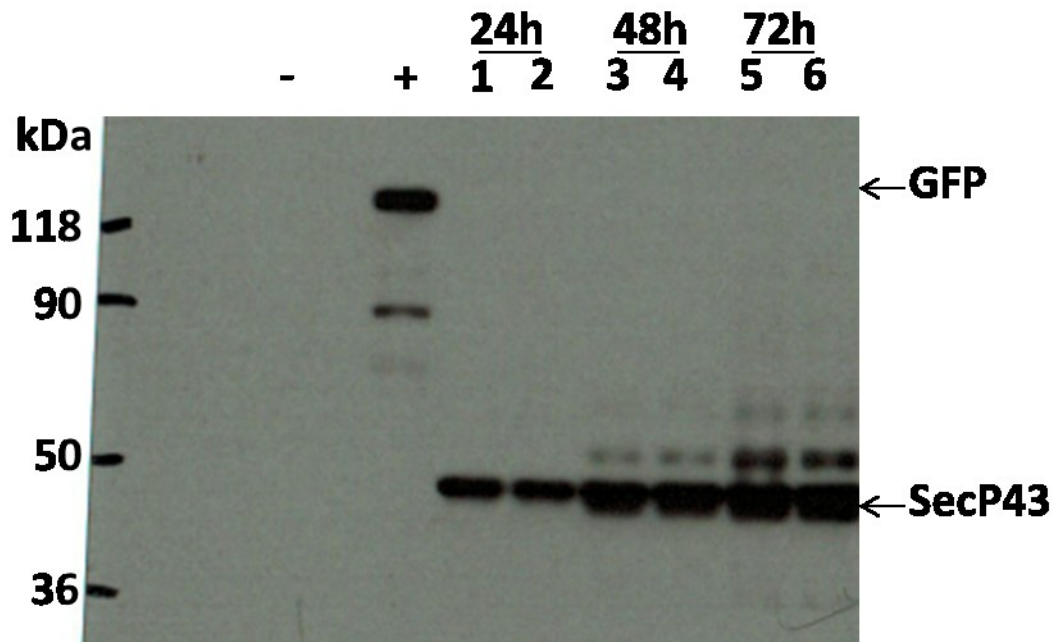


Figure 13. Test de surexpression transitoire de SecP43 dans les cellules HEK293-6E.

Analyse des échantillons prélevés après 24, 48 et 72h post transfection sur gel SDS-PAGE 10 %. La détection des protéines a été réalisée par immunobuvardage de type « western » utilisant des anticorps anti-Histag. (-) Contrôle négatif, plasmide pTT5 vide ; (+) contrôle positif, plasmide pTT5-GFP ; le ratio pTT5-SecP43/PEI des pistes (1-3-5) est de 1/3 et celui des pistes (2-4-6) est de 1/4. Le ratio $\frac{1}{4}$ donne le meilleur rendement.

La figure ci-dessus montre que l'expression de SecP43 dans les cellules HEK293 a bien fonctionnée, mais que le rendement n'était pas suffisant pour une purification à grande échelle. En effet, SecP43 n'a pas pu être par coloration au bleu de coomassie dont la limite de détection est d'environ 20 ng. Par ailleurs, les tentatives d'augmentation de l'expression se sont soldées par des échecs en raison des contaminations bactériennes du milieu de culture. Il a donc été nécessaire d'utiliser d'autres approches pour améliorer notre rendement.

3.2. Expression de SecP43 dans la souche BL21 DE3 d'*E. coli*

3.2.1. Expression dans le milieu LB

Le gène de SecP43 fut initialement cloné dans le vecteur d'expression pETDuet1 qui contient un promoteur T7 inducible par l'ajout de l'IPTG. Ce vecteur fut utilisé pour la transformation des cellules compétentes *E. coli* de souche BL21 DE3. Nous avons dans un premier temps testé l'expression de SecP43 dans le milieu LB (Luria-Bertani) en modifiant la température et la durée de l'incubation suite à l'ajout de l'IPTG. Des prélèvements successifs ont ensuite été analysés sur gel SDS-PAGE (figure 14). On remarque qu'en dépit des variations de températures et de temps d'induction, SecP43 (34 kDA, protéine+ étiquette) n'est pas efficacement exprimé dans le milieu LB. En effet, comparativement au contrôle négatif, on n'observe aucun signal qui apparait exclusivement après 1, 2, 3 ou 16h d'induction. Les signaux de fortes intensités observées ne représentent que l'expression des protéines endogènes retrouvées dans le contrôle négatif.

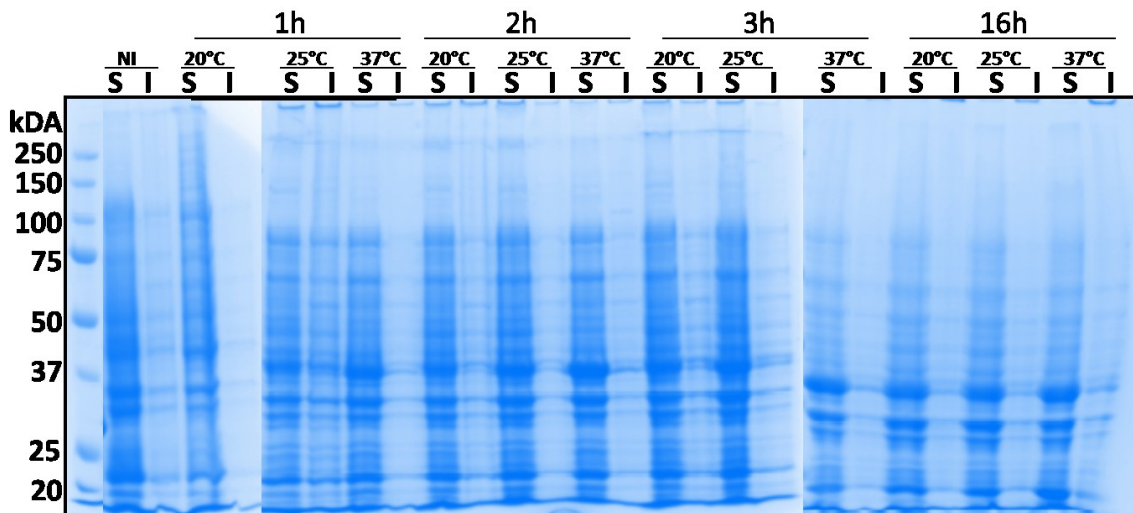


Figure 14. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d'*E.coli*. Analyse SDS-PAGE des échantillons prélevés après 1, 2, 3, et 16h d'incubation suite à l'induction par l'IPTG. La détection des protéines sur le gel SDS-PAGE 10% a été effectuée par coloration au bleu de Coomassie. Les lettres S et I désignent respectivement les fractions des protéines solubles et insolubles. NI est l'acronyme de non-induit par l'IPTG. Il s'agit du contrôle négatif.

3.2.2. Milieu LB ; effet du vecteur pBB535

L'échec de la surexpression de SecP43 dans les conditions décrites plus haut pourrait d'une part être attribué au fait que nous tentions d'exprimer une protéine humaine dans une souche bactérienne qui n'est pas nécessairement adaptée pour cette protéine. De ce fait, c'est possible de postuler qu'en cas de repliement défectueux, SecP43 soit une cible du mécanisme de dégradation intracellulaire des protéines inadéquatement repliées. Ainsi nous avons jugé nécessaire d'ajouter une deuxième composante dans notre système d'expression à savoir le vecteur pBB535. Ce vecteur encode la chaperonne DnaJ et sa co-chaperonne DnaK qui appartiennent aux familles des protéines Hsp40 et Hsp70 respectivement. Ces protéines sont connues pour faciliter le repliement des protéines en réponse à une dénaturation due au choc thermique. La souche utilisée dans le test d'expression précédent a donc été transformée avec le vecteur pBB535 et un nouveau test de surpression similaire fut effectué (figure 15). Ainsi, tel qu'on peut le constater, l'ajout de pBB535 n'a pas promu la surexpression de SecP43 tel que nous l'avions pensé. Les signaux mis en évidence par la flèche représentent probablement la chaperonne DnaK qui fait environ 68 kDa. Face à ce nouveau résultat, nous avons par conséquent eu recours à une approche alternative, l'auto-induction.

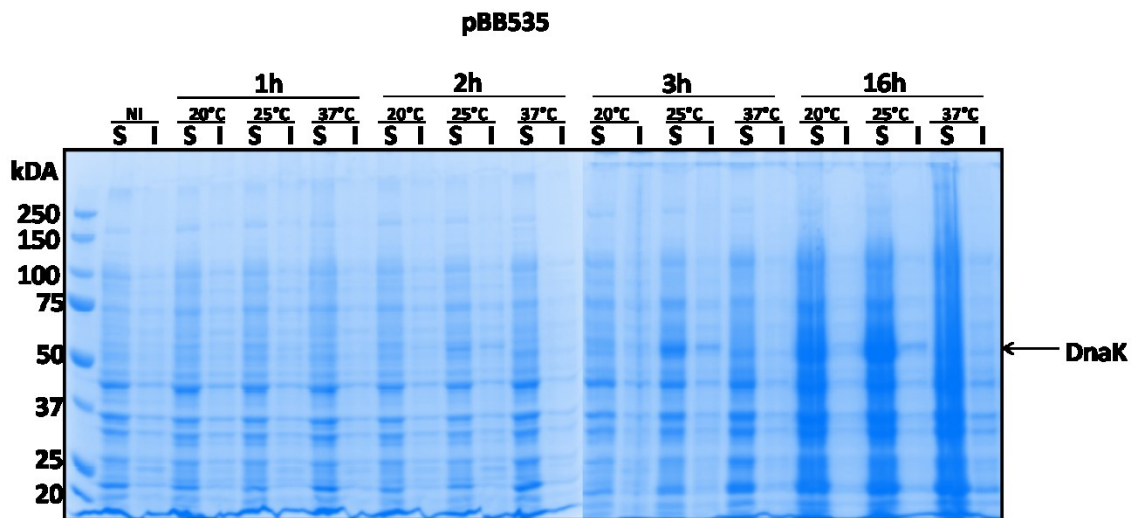


Figure 15. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d'*E. coli*. Effet du vecteur additionnel pBB535. Des échantillons sont prélevés après 1, 2, 3 et 16h d'incubation ont été ensuite analysés sur gel SDS-PAGE 10% coloré au bleu de coomassie. Les lettres

S et I désignent respectivement les fractions des protéines solubles et insolubles. NI est l'acronyme de Non-induit par l'IPTG. Il s'agit du contrôle négatif. La flèche indique l'expression de la chaperonne DnaK encodée par le vecteur pBB535.

3.2.3 Expression dans le milieu d'auto-induction

Le système d'induction T7 est hautement effectif et largement utilisé pour la production d'ARNs et de protéines à partir de séquences codantes dans la bactérie *E. coli* (F.W. Studier, et al 1986). Dans les cellules hôtes telles que BL21 (DE3), la séquence codante de la l'ARN polymérase T7 est présente dans le chromosome sous le contrôle du promoteur LacUV5. La séquence codante de la protéine d'intérêt est quant à elle intégrée dans un plasmide sous le contrôle du promoteur T7, et n'est donc reconnaissable que par l'ARN polymérase T7. Lorsque le promoteur LacUV5 n'est pas induit, les niveaux de la polymérase T7 et des protéines cibles, devraient être faibles et la croissance cellulaire devrait être optimale. Cependant, suite à l'ajout d'un inducteur tel que l'IPTG (isopropyl- β -D-thiogalactoside), la polymérase T7 sera produite et transcrita toute séquence d'ADN placée sous le contrôle du promoteur T7. La polymérase T7 est si spécifique et active que la quantité d'ARN cible produite lors de la transcription peut être comparable à la quantité d'ARN ribosomale dans la cellule. Si l'ARN cible possède un signal d'initiation de la traduction approprié, il en résultera une accumulation substantielle de la protéine cible.

Un des problèmes inhérents à l'utilisation du système d'expression T7 est qu'en raison de la très haute activité de la polymérase T7, un niveau basal de cette polymérase peut entraîner l'expression substantielle de la protéine cible et ce, sans l'ajout de l'IPTG. Si la protéine cible est suffisamment toxique pour la cellule hôte, l'établissement du plasmide cible dans la souche d'expression peut être compromise ou impossible et peut dans certains cas entraîner l'instabilité et l'accumulation des mutations dans la souche d'expression (F.W. Studier et al, 1991). Un moyen efficace de réduire l'expression basale de la polymérase T7 consiste à placer la séquence de l'opérateur Lac (site de liaison du répresseur Lac) en aval du site d'initiation du promoteur T7, créant ainsi un promoteur T7lac (F.W. Studier et al, 1990). Lié à la séquence de l'opérateur, le répresseur Lac

interfère avec l'établissement du complexe d'élongation par la polymérase T7 au niveau du promoteur T7lac et réduit considérablement le niveau d'ARNm cibles produits (F.W. Studier et al, 1990). Donc si le répresseur Lac est suffisamment concentré pour occuper tous les sites de liaison dans la cellule, l'expression de la protéine cible est réprimée mais l'induction avec l'IPTG débloque à la fois les promoteurs lacUV5 et T7lac, ce qui conduit à de hauts niveaux d'expression.

Des études portant sur la stabilité, la croissance et l'induction des souches du système d'expression T7 ont permis d'établir les effets de certains constituants du milieu de culture sur les paramètres suscités. Par ailleurs, d'autres études ont montré que de faibles quantités de lactose peuvent stimuler l'expression d'une protéine cible dans certains milieux complexes de culture. Finalement, des mécanismes bien étudiés ont montré que le glucose est capable de prévenir l'induction par le lactose. Sur la base de ces observations, un milieu d'auto-induction fut mis au point dans le but d'effectuer une induction plus douce des protéines. Ce dernier contient à la fois du glycérol, du glucose et du lactose. Nous avons testés plusieurs milieux d'auto-induction caractérisés par des concentrations en glycérol, glucose et lactose variables (2.2.). Pendant la phase initiale de croissance, les cellules utilisent principalement du glucose comme source de carbone car ce dernier est plus facilement métabolisable. Dans la phase de croissance plus tardive, le glycérol est ensuite utilisé. Finalement lorsque la saturation est atteinte, l'induction de l'expression est déclenchée par l'absorption du lactose (figure 16). Comparativement à l'induction à l'IPTG, l'auto-induction favorise une expression lente et contrôlée par le métabolisme de la cellule hôte, et de cette manière le niveau d'expression d'une protéine cible potentiellement toxique est désormais contrôlé. Ainsi, la protéine cible échappe aux mécanismes de dégradation et peut bénéficier de l'énergie métabolique disponible pour se replier correctement. Tout ceci favorise donc une meilleure solubilité et un rendement amplifié.

Préalablement aux tests de surexpression, puisqu'il s'agit d'une protéine humaine, nous avons incorporés le vecteur pRARE à la souche BL21 (DE3) transformée par le vecteur pET28a-SecP43. Ce vecteur encode des ARNt humains pour la traduction des codons rares dans la souche BL21 (DE3). Comme alternative, nous avons aussi testé l'effet de

l'expression de la chaperonne DnaK encodé par le vecteur pBB535. La figure 17 met en évidence l'analyse SDS-PAGE des échantillons prélevés lors du test de surexpression de SecP43 dans le milieu d'auto-induction 4xLac-8ZYM (2.2.). Cette figure montre que la souche contenant le pRARE promeut l'expression de SecP43 et ce, indépendamment de la température d'incubation quoique l'expression ne survient pas au même moment, étant entendu qu'elle est régie par la densité cellulaire qui elle à son tour dépend de la température d'incubation. On note que SecP43 est visiblement exprimé dans la souche qui contient le vecteur pRARE contrairement à la souche pBB535 pour laquelle on observe uniquement la chaperonne DnaK. On observe également que SecP43 subit probablement une protéolyse dans la région C-terminale, tel que l'indique le signal d'environ 25 kDa située juste en-dessous des protéines de 34 kDa. Cette observation a été confirmée par immunobuvardage de type *Western* utilisant des anticorps *anti-histag* qui interagissent avec l'étiquette histidine présente dans la partie N-terminale de SecP43 (figure 18). La condition qui a été retenue pour l'expression à grande échelle et la purification fut la suivante : 25°C, 72 h, milieu d'auto-induction : 4x Lac 8ZYM (2.2).

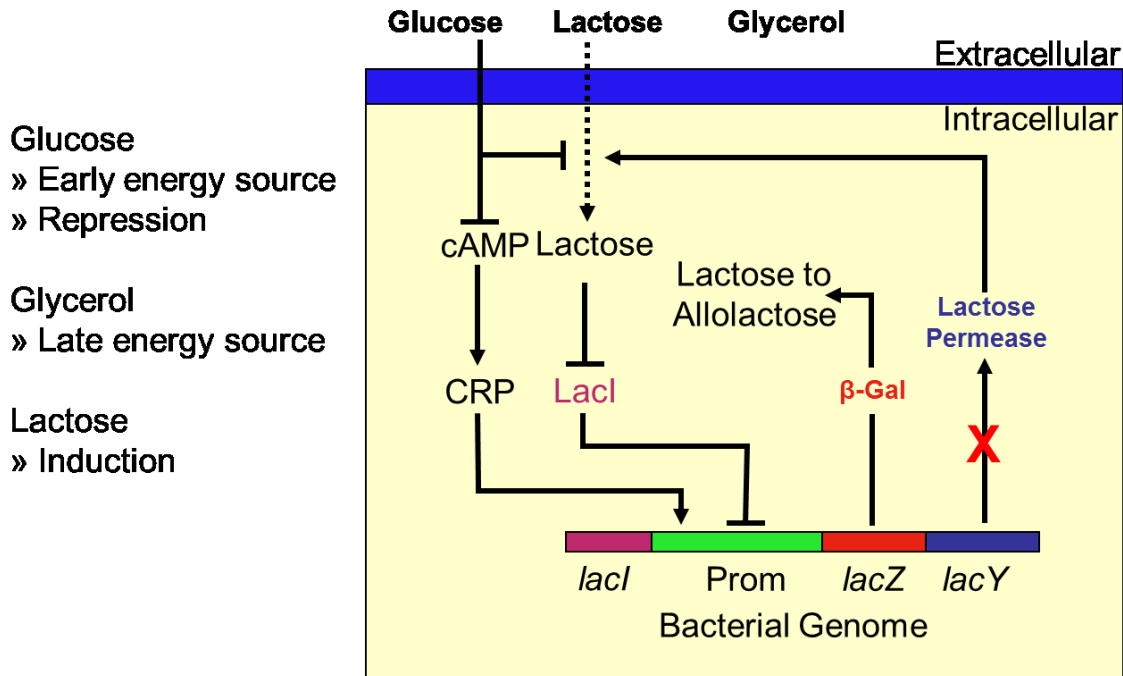


Figure 16. Principe du fonctionnement de l'auto-induction. La figure est issue du protocole de Trevor Sweeney (Imperial college, London).

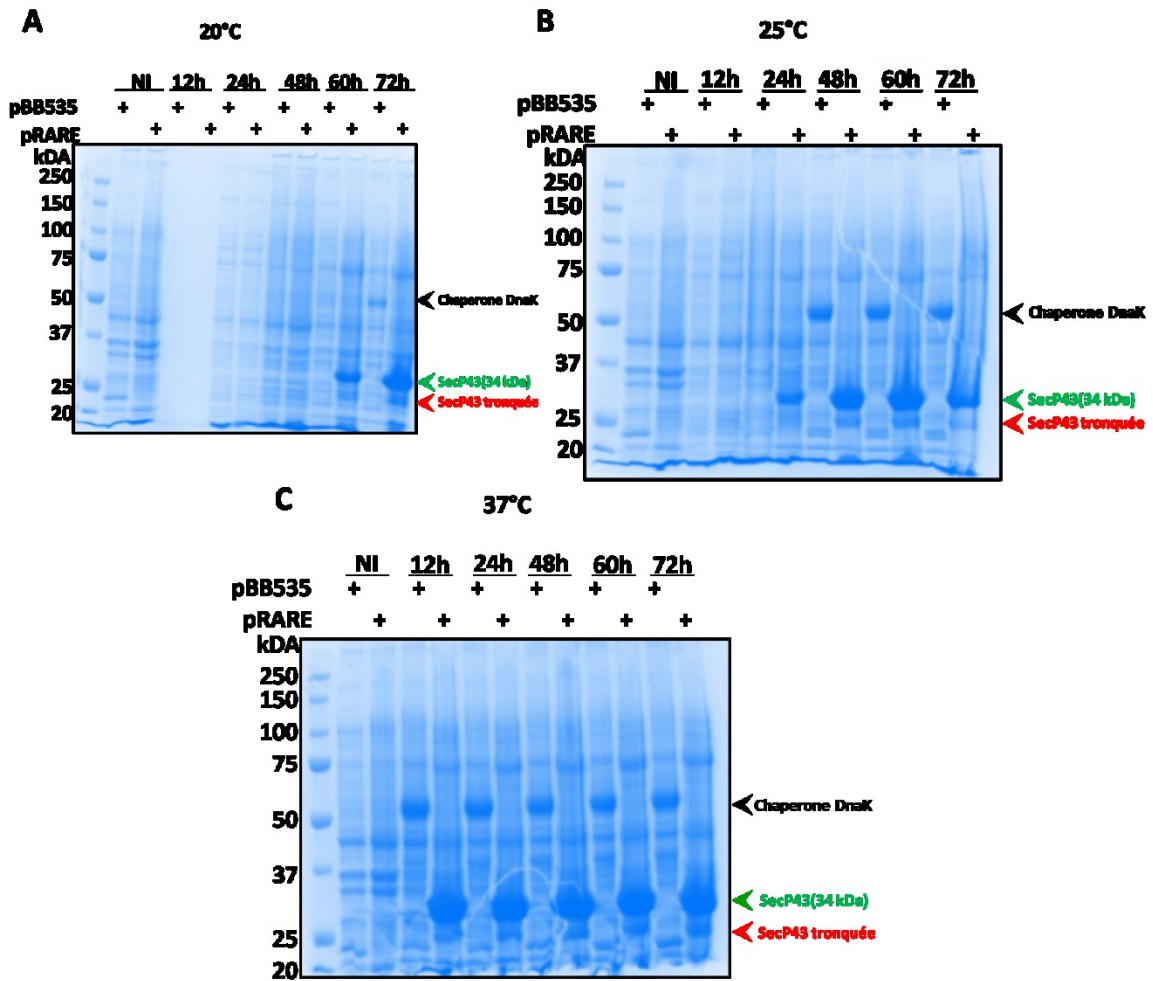


Figure 17. Test de surexpression de SecP43 dans les cellules de souche BL21 (DE3) d’*E. coli*. Effet des vecteurs pRARE et pBB535 dans le milieu d’auto-induction 4xLac-8ZYM. Les échantillons ont été prélevés toutes les 12h et analysés sur gel SDS-PAGE 10 %. Seules les fractions solubles ont été mises sur gel. NI est l’acronyme de non-induit. Il s’agit du contrôle négatif. Les signaux représentant la chaperone DnaK, la protéine SecP43 et un produit probablement tronqué de SecP43 sont indiqués par les flèches.

3.3. Purification de SecP43

SecP43 fut purifié en deux étapes incluant premièrement une chromatographie d'affinité (IMAC-Histrap) qui sert à capter la protéine d'intérêt tout en se débarrassant de la plupart des protéines indésirables du lysat cellulaire, et deuxièmement d'une chromatographie d'exclusion stérique (étape de polissage).

3.3.1. Chromatographie d'affinité IMAC-Histrap

La figure 18 décrit les étapes suivies lors de la phase de captation de SecP43 sur une colonne IMAC chargée au nickel. Rappelons que la protéine de fusion contient une étiquette (6×His) qui interagit avec le nickel. Les fractions collectées suite au gradient d'imidazole ont été analysées sur gel SDS-PAGE 10% afin de confirmer l'identité de la protéine par son poids moléculaire. La détection des protéines par immunobuvardage de type *western* indique que SecP43 subit probablement une protéolyse dans la région C-terminale car l'étiquette Histidine est située en N-terminale.

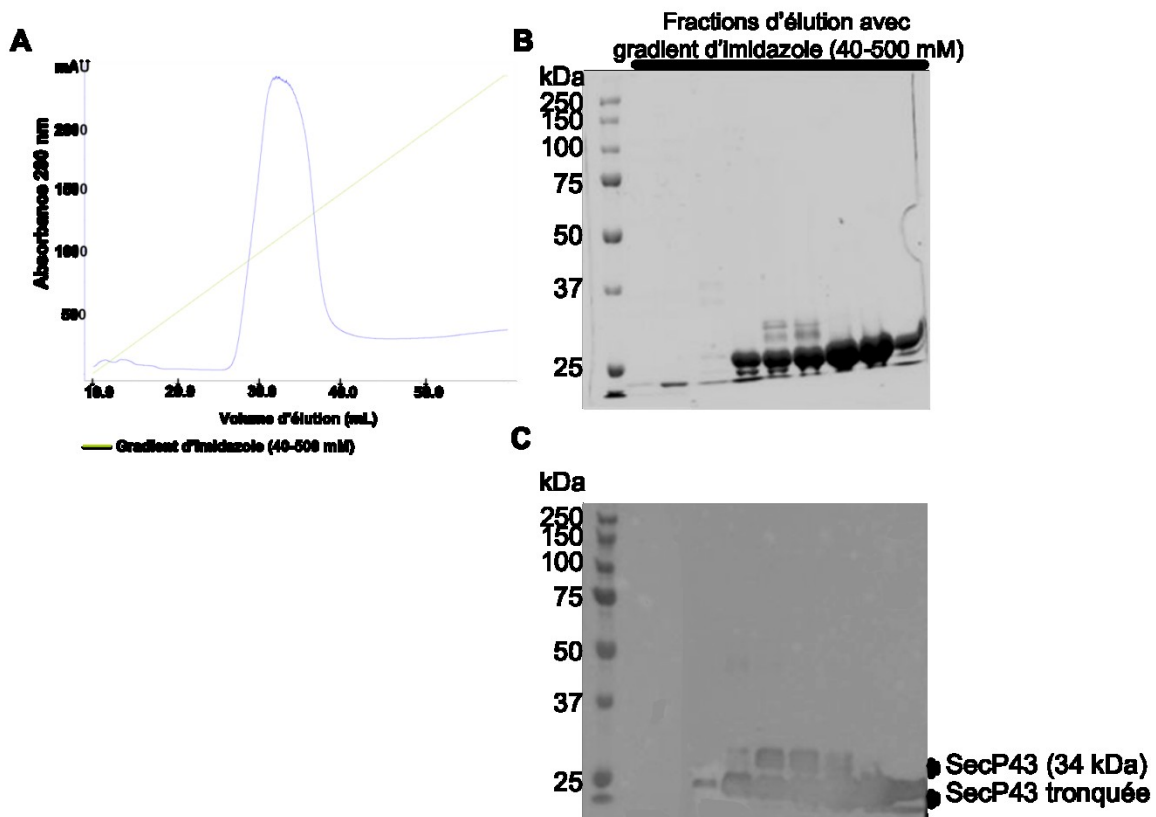


Figure 18. Purification de SecP43 par chromatographie d'affinité (IMAC). A. Profil d'élution de SecP43 sur une colonne IMAC chargé au nickel. B. Les fractions d'élutions ont été analysées sur gel SDS-PAGE 10 % et la détection des protéines a été effectuée par coloration au bleu de Coomassie et (C) par immunobuvardage de type *Western* en utilisant des anti-corps *anti-Histag*. SecP43 et SecP43 tronquée sont indiqués par les flèches.

3.3.1.2. Étape de polissage par filtration sur gel

Les fractions collectées à l'étape précédente ont été rassemblées et injectées sur une colonne d'exclusion stérique (Superdex 75) (2.2). La figure 19 évalue le degré de pureté de SecP43 dans différentes fractions d'élution suite à la chromatographie d'exclusion stérique. Tel que l'indique le gel SDS-PAGE, la fraction F2 semble être la plus pure. C'est également dans cette fraction que la dimérisation de SecP43 (signal d'environ 50 kDa) semble se produire en raison d'une concentration plus élevée.

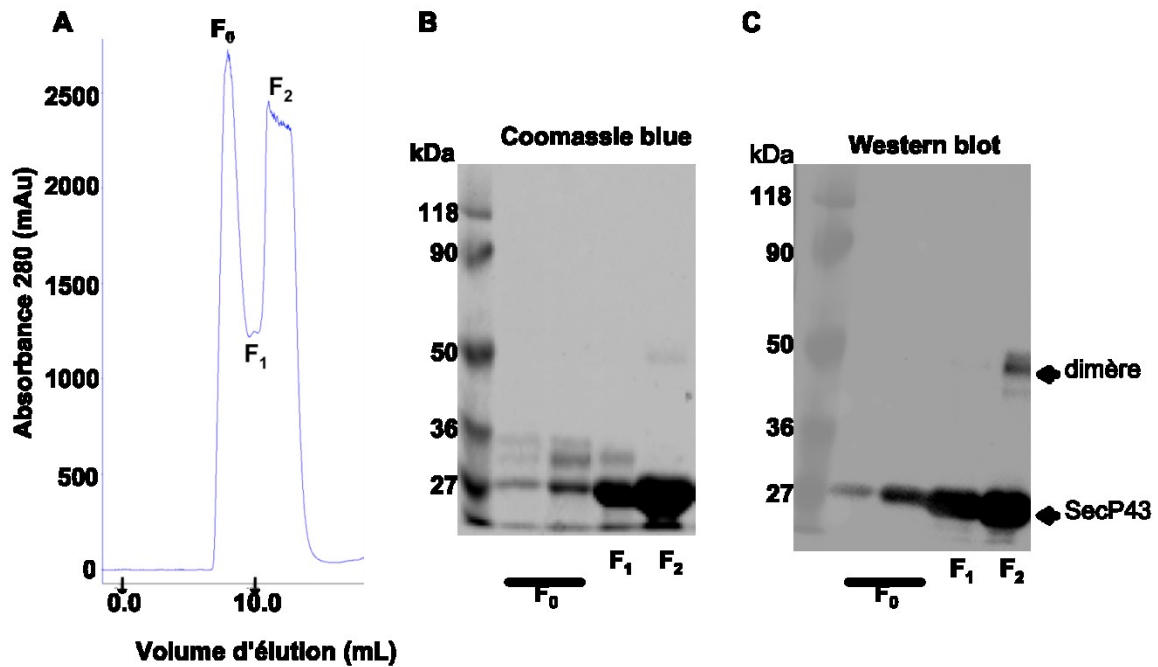
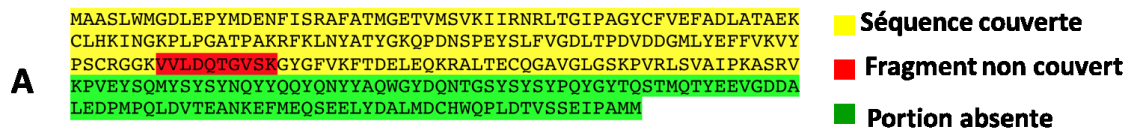


Figure 19. Purification de SecP43 et évaluation de la pureté sur gel SDS-PAGE 10%. (A) Les fractions F₀, F₁, et F₂ du chromatogramme ont été analysées sur gel SDS-PAGE 10%. La détection des protéines a été effectuée par coloration au bleu de Coomassie (B) et par *Western blot* (C). Les flèches indiquent la protéine SecP43 tronquée (environ 25 kDa) et probablement un dimère d'environ 50 kDa.

3.3.1.3. Étude la protéine purifiée par spectrométrie de masse

Afin de localiser le site qui fait l'objet d'une protéolyse dans la séquence de SecP43, nous avons en plus de déterminer la masse exacte de la protéine par spectrométrie de masse, fait un séquençage de type MS-MS. Pour ce faire, nous avons d'abord clivé l'étiquette histidine grâce au traitement par la protéase TEV (2.2). Dans le séquençage MS-MS, la protéine d'intérêt est d'abord digérée par la trypsine qui coupe après une lysine ou une arginine. Les peptides ainsi générés (Annexe) sont ensuite ionisés. Le premier analyseur (MS) sélectionne un ion (ion parent), à l'exclusion des autres ions. L'ion parent est ensuite fragmenté en "ions fils" par collision avec un gaz inerte. Les ions fils sont ensuite séparés par le second analyseur (MS) puis détectés (Morgan et al, 2005). La figure 20 combine les résultats de ces deux analyses. Notons que la reconstitution de la séquence à partir des fragments générés a été faite manuellement. Ainsi, tandis que les 180 premiers acides aminés ont été détectés, on remarque que la portion verte de la séquence n'a pas été détectée dans notre échantillon (figure 20A). Cela s'explique probablement par la dégradation de cette région par les protéases. La protéolyse est bien effective puisque la masse détectée lors du *ESI scan (Electron spray ionization)* est d'environ 20 kDa (figure 20 B) et non de 32 kDa (taille de la protéine complète).



Le pourcentage de couverture de la séquence est d'environ 40,76 %

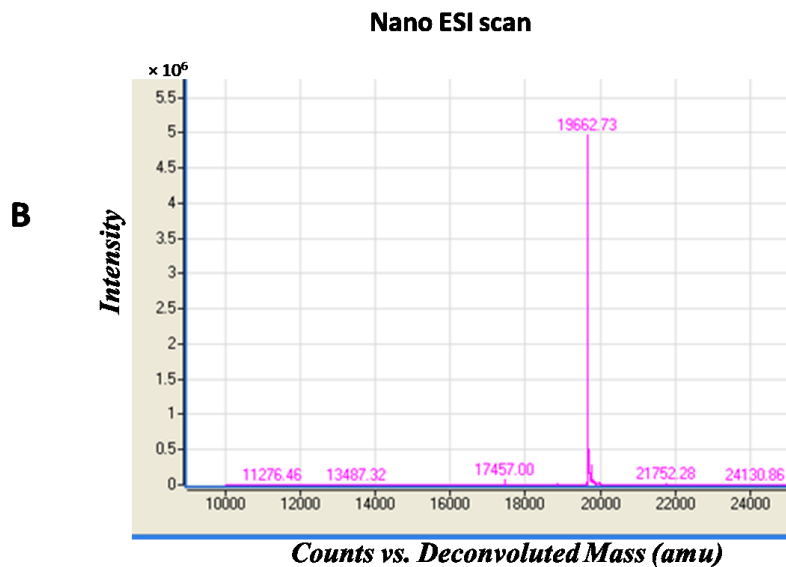


Figure 20. Analyse de la séquence de SecP43 par spectrométrie de masse. Un échantillon purifié de SecP43 a été utilisé. **A** séquençage MS-MS de l'échantillon purifié. La séquence complète de SecP43 est ici représentée. La portion verte n'a pas été détectée dans l'échantillon. **B** détermination de la masse de SecP43 contenue dans l'échantillon analysé. ESI (Electron spray ionization)

3.4. Purification de SecP43 tronquée

Conformément aux données de spectrométrie de masse, un codon stop a été inséré après la valine 180 dans la séquence du gène de SecP43 par mutagenèse dirigée (2.2). La purification de la protéine tronquée a été effectuée suivant la même procédure que celle décrite pour la protéine complète. Toutefois, l'étiquette histidine a été clivée par la protéase TEV avant l'étape de polissage par filtration sur gel. Notons que la protéase TEV contient elle aussi l'étiquette et sera par conséquent retenue lors de la réinjection de l'échantillon sur la colonne IMAC. La figure 21 résume toutes les étapes de purification. La figure 22 quant à elle montre la variation du chromatogramme de SecP43 en fonction de la concentration. On note qu'à une concentration de 13 μM le profil du pic est symétrique. Cependant, le pic devient asymétrique à mesure que la concentration

augmente. Ceci s'explique probablement par le fait qu'à très haute concentration, SecP43 se dimérise d'où l'inclinaison du pic vers la gauche.

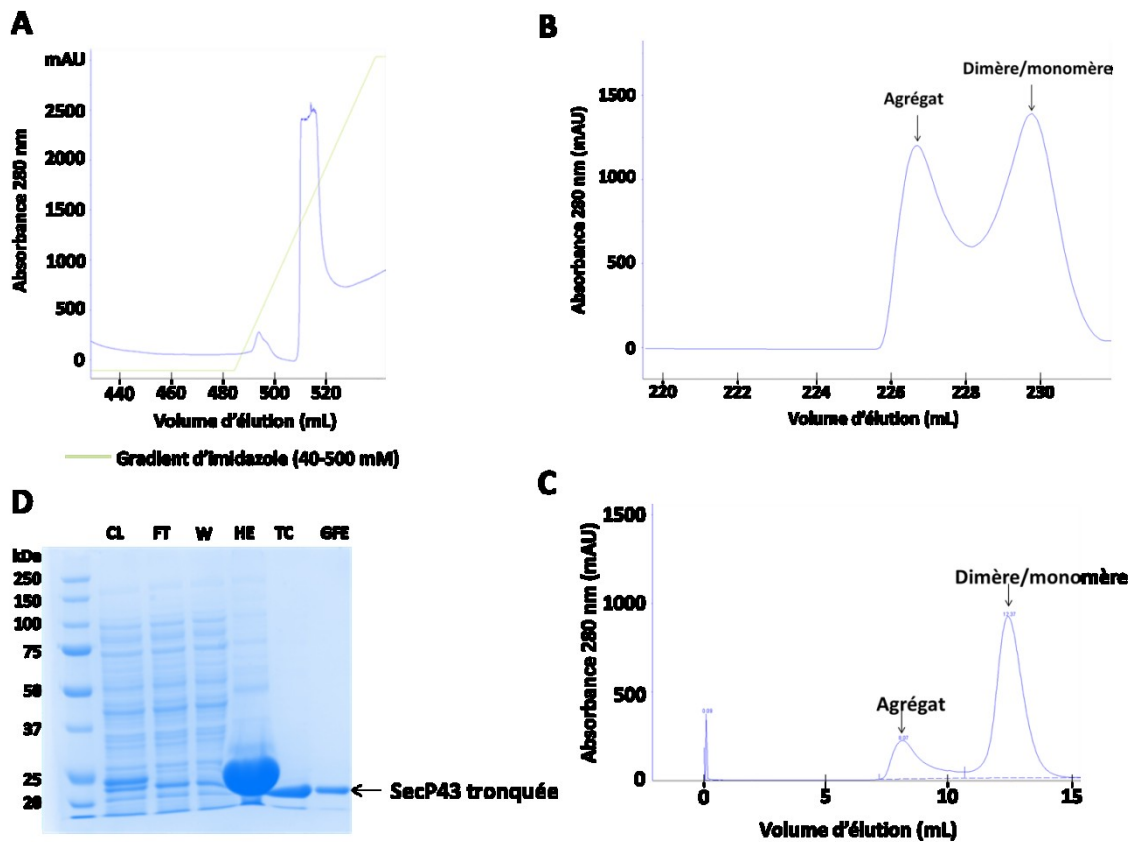


Figure 21. Purification de SecP43 tronquée. **A.** Étape de captation de SecP43 suite à la lyse cellulaire. **B.** Éluions des fractions collectées à l'étape A préalablement traitées à la protéase TEV pour cliver le *Histag* avant d'être réinjectées sur la colonne IMAC. **C.** Elution des fractions collectées sur filtration sur gel (colonne Superdex 75). **D.** Analyse des fractions d'éluion de chaque étape sur gel SDS-PAGE 10 %. La détection des protéines a été effectuée par coloration au bleu de Coomassie. CL (*cell lysate*), FT (*flow through*), W (*wash*), HE (*histrap elution*), TC (*TEV cleavage*), GFE (*gel filtration elution*).

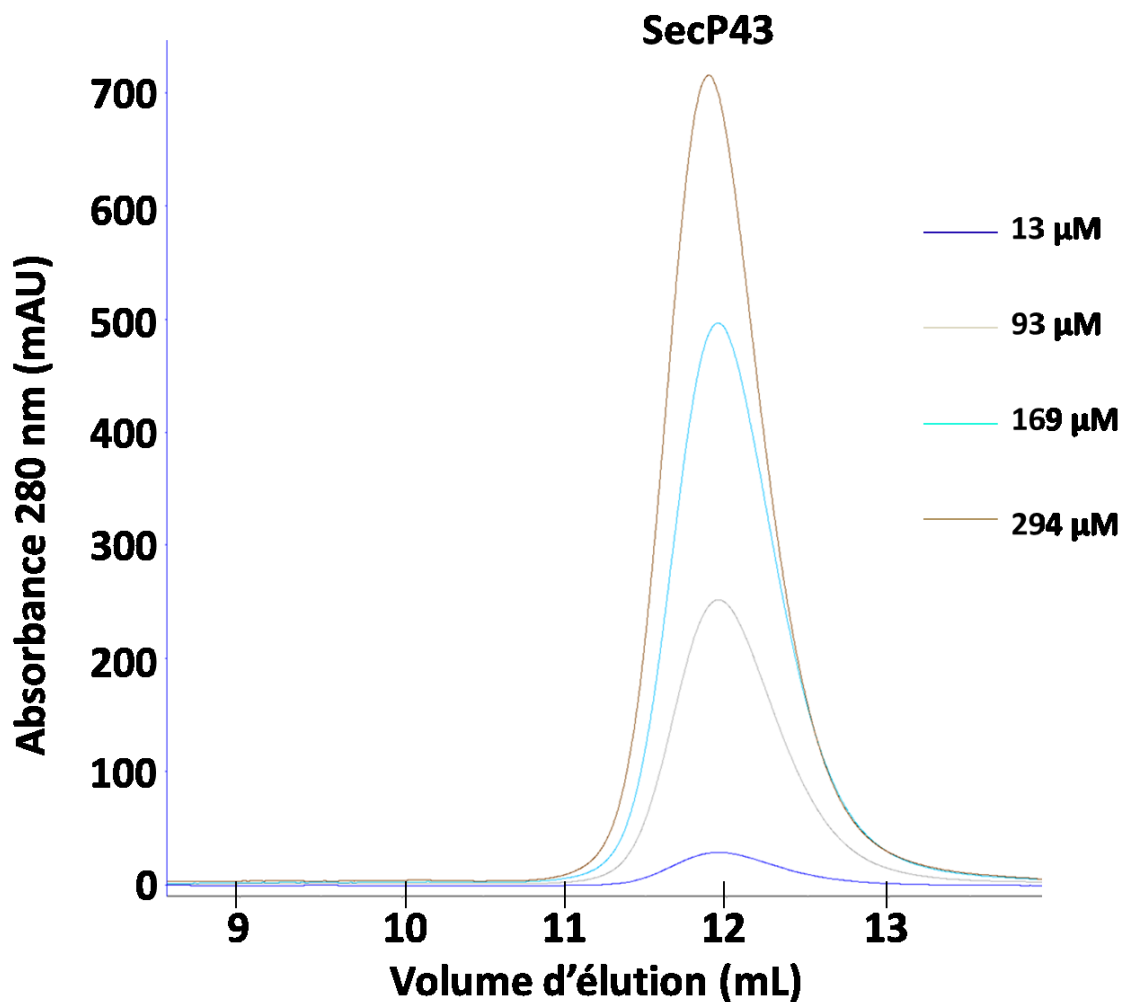


Figure 22. Analyse de la dimérisation de SecP43 par chromatographie d'exclusion stérique. Les profils correspondant aux différentes concentrations ont été superposés afin de les comparer.

3.5. Vérification de la masse et de la séquence par spectrométrie de masse

Un aliquot de SecP43 tronquée et purifiée a été soumis aux mêmes analyses que lors de la purification de la protéine complète. Les données primaires de cette analyse sont compilées dans le tableau IX (Annexe). Les résultats de ces analyses sont indiqués dans la figure suivante. On note que 93,3 % de la séquence escomptée a pu être détectée

(figure 23 A). Par ailleurs l'échantillon contient à la fois le monomère et le dimère (figure 23 B). La masse obtenue est bien conforme à la masse de la protéine tronquée, calculée à partir de la séquence, soit en environ 19736.7 Da.

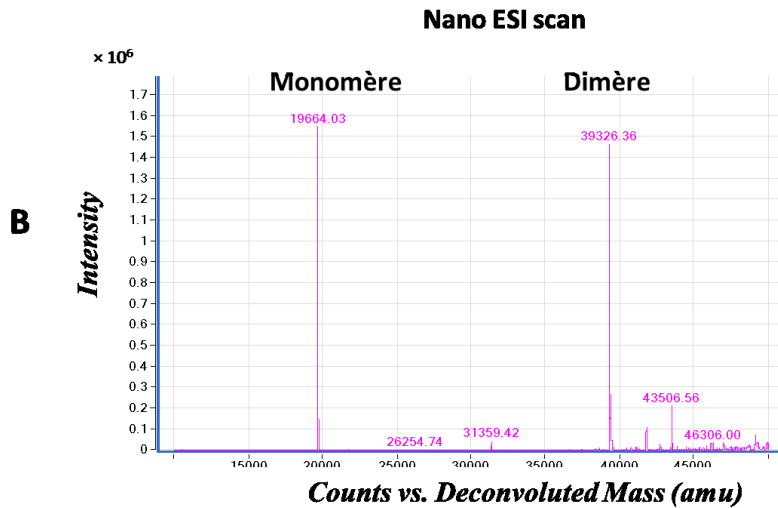
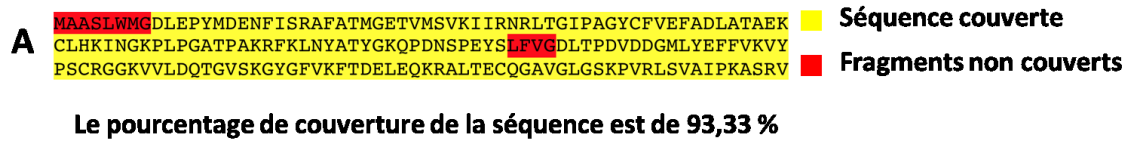


Figure 23. Analyse de la séquence de SecP43 tronquée par spectrométrie de masse. Un échantillon purifié de SecP43 tronquée a été analysé. **A.** Séquençage MS-MS de l'échantillon purifié. La séquence tronquée de SecP43 est ici représentée. **B.** Détermination de la masse de SecP43 contenue dans l'échantillon analysé.

3.6. Caractérisation structurelle de SecP43

3.6.1. Essais de cristallisation

L'objectif initial était de cristalliser la protéine complète seule, et en complexe avec ses partenaires éventuels tels que l'ARNt^{Sec} et SepSecS. Afin d'augmenter la probabilité de succès lors de la cristallisation, nous avons effectué quelques modifications au sein de la séquence de SecP43 en nous basant sur les prédictions des algorithmes RONN et Xtalpred (figure 24 et 25). Le logiciel RONN prédit à partir de la séquence primaire les régions naturellement non-structurées dans une protéine. Le logiciel Xtalpred calcule la probabilité de cristallisation d'une protéine et la classe selon une échelle qualitative. Cette classification est basée sur les statistiques des protéines microbiennes non secrétées et elle est optimisée pour l'identification des cibles cristallisables au sein de larges familles de protéines. Ainsi, tenant compte de cette classification d'une part et du fait que SecP43 était assujéti à une protéolyse en C-terminal (3.3.1.3), un codon stop a notamment été inséré au gène de SecP43 en aval de la valine 180, résultant ainsi en une version tronquée qui ne contient que les deux domaines de reconnaissance de l'ARN (RRM) en N-terminal. Malgré cette troncature, aucune condition des 11 *screens* (Molecular dimension, contenant chacun 96 agents précipitants différents) testés n'a permis de cristalliser SecP43 car la flexibilité induite par la boucle reliant les deux domaines RRM suggère l'existence probable de plusieurs conformations en solution. Les cristaux que nous avons obtenus n'étaient en fait que des cristaux de sel (non montré). C'est la raison pour laquelle nous avons eu recours à une technique de diffraction en solution telle que le BIOSAXS (*Biological Small Angle X-rays Scattering*) pour caractériser la structure de SecP43.

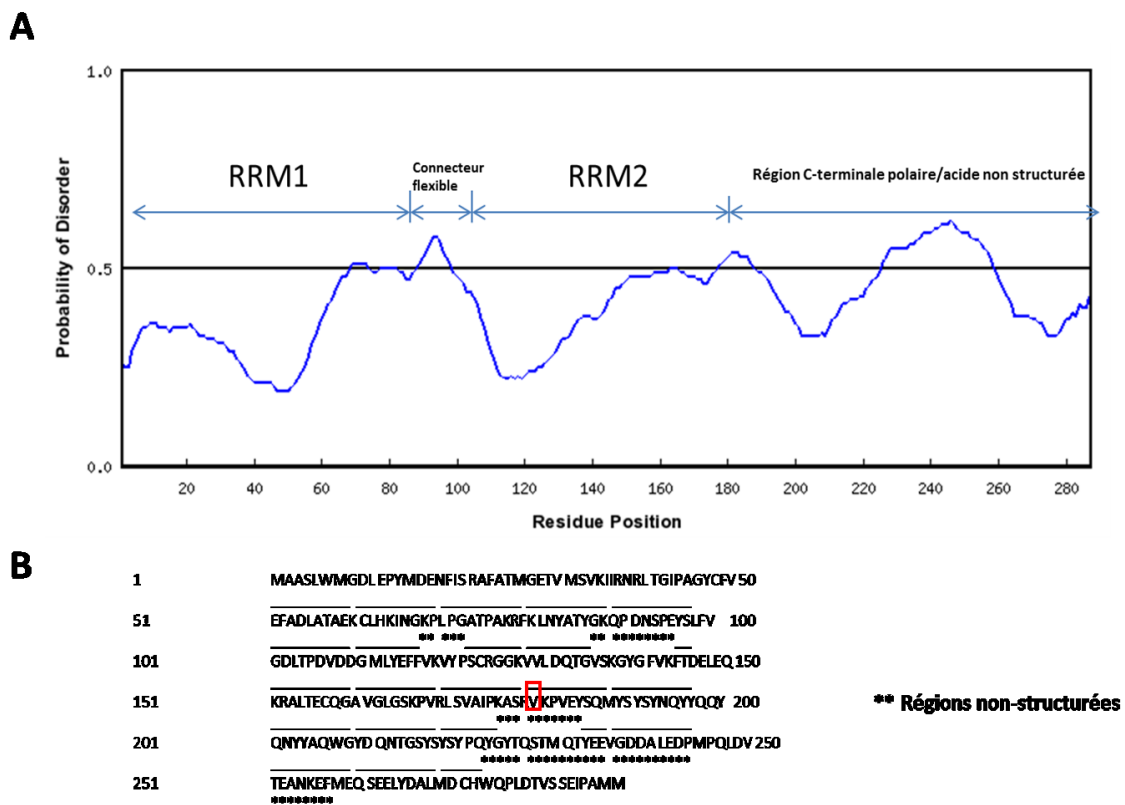


Figure 24. Domaines topologiques de SecP43 selon l’algorithme RONN. A. SecP43 est formé de domaines de liaisons de l’ARN connectées par une boucle probablement flexible. La région C-terminale qui comprend les résidus 181 à 287 est non structurée car elle présente beaucoup plus de régions désordonnées. B. Mise en évidence de la valine 180 (en rouge), résidu en aval duquel un codon stop fut inséré (www.app.strubi.ox.ac.uk/RONN).

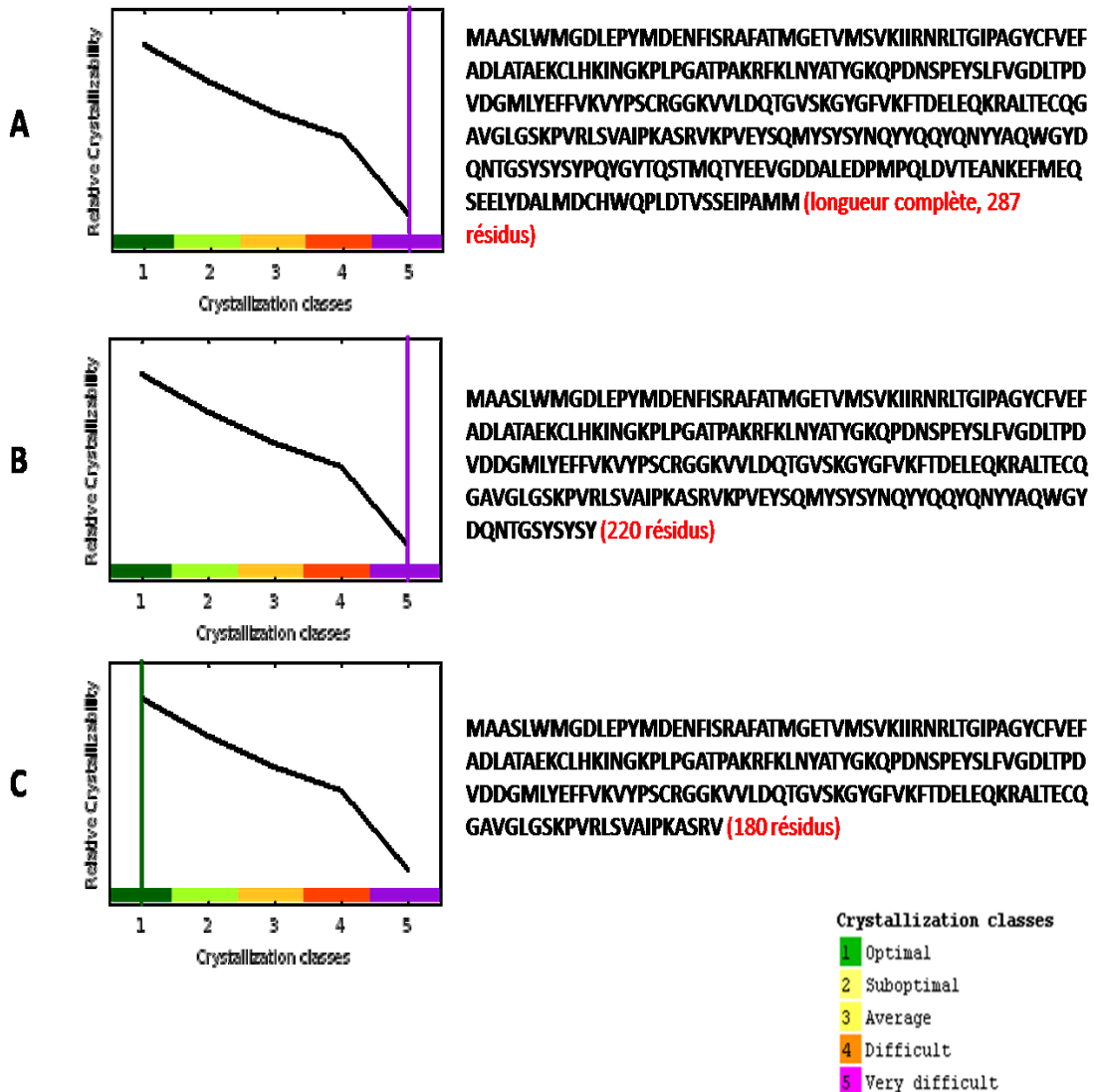


Figure 25. Probabilité de cristallisation selon la séquence de SecP43. Le calcul des probabilités et la classification ont été effectués par le logiciel Xtalpred (www.ffas.burnham.org/XtalPred.cgi/xtal.pl). **A.** La probabilité de cristallisation relative de la séquence complète de SecP43 est très faible en raison la région C-terminale. **B.** En absence d'une partie de la région C-terminale la cristallisation demeure peu probable. **C.** L'élimination complète de cette région non structurée augmente significativement les chances de réussite

3.6.2. Caractérisation structurale de SecP43 par BIOSAXS (*Biological-Small Angle X-rays Scattering*)

3.6.2.1. Description de la technique de BIOSAXS

La diffraction des rayons X à des petits angles est une technique fondamentale de faible résolution en plein essor dans l'étude des macromolécules biologiques. L'avantage majeur de cette méthode réside en ceci qu'elle permet de donner des informations structurales sur des systèmes partiellement ou totalement désordonnés. Le SAXS permet d'étudier la structure native des particules dans un environnement physiologique et d'analyser les changements conformationnels en réponse aux variations des conditions externes (Svergun DI et al, 2003).

Dans une expérience de diffraction, une solution de protéine est exposée à un faisceau de rayons X dont la longueur d'onde (λ) est typiquement de 0,15 nm. L'intensité de dispersion $I(s)$ est enregistrée comme une fonction du vecteur de dispersion dont l'intensité communément appelée *moment de transfert*, s ou q est donnée par la relation $s = q = \frac{4\pi}{\lambda} \sin\left(\frac{\theta}{2}\right)$ où θ représente l'angle entre le rayon incident et le rayon dispersé. L'intensité de dispersion du tampon est soustraite de l'intensité de la solution pour avoir l'intensité spécifique du soluté. La position et l'orientation aléatoire des particules en solution entraînent une distribution d'intensité isotrope qui, pour un échantillon monodisperse composé de particules n'interagissant pas les unes avec les autres, est proportionnel à la moyenne de la dispersion due à une particule unique dans toutes les orientations (moyenne sphérique) (Svergun DI et al, 2002). Ceci donne donc contrairement à la diffraction par un cristal, un patron de diffraction à une dimension (1D) (figure 26). C'est à partir de cette courbe unidimensionnelle qu'on extrait grâce à des logiciels spécialisés, les paramètres structurels tels que la forme, la taille, l'état d'oligomérisation, et l'état de repliement de la protéine étudiée. Le SAXS est pour ainsi dire, une technique complémentaire aux techniques de plus haute résolution telles que la cristallographie aux rayons X ou la RMN. Cependant elle s'applique à toutes les macromolécules sans discrimination de taille contrairement à la RMN.

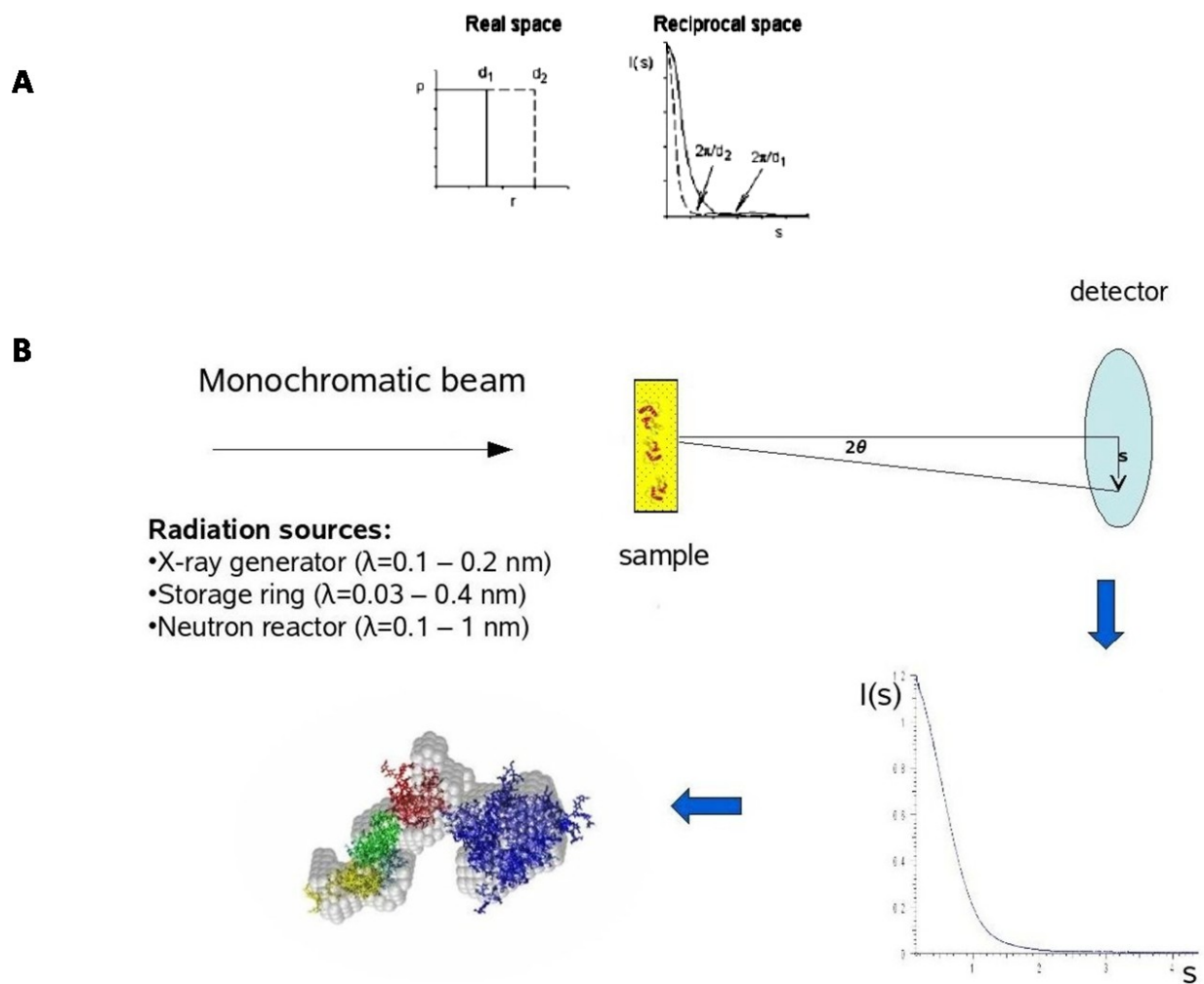


Figure 26. Représentation schématique d’une expérience de diffraction à des petits angles (SAXS). A. Transformation de Fourier permettant de passer de l’espace réel à l’espace réciproque. B. Représentation d’une expérience classique de SAXS. (Svergun. et al, 2003)

3.6.2.2. Collecte des données de diffraction de SecP43 tronquée par BIOSAXS

Concernant SecP43, nous avons mesuré la diffraction de cette protéine dans deux tampons différents dans lesquelles cette protéine est stable. Il s’agit notamment des tampons à base de tris et de phosphate. Les données et les paramètres de l’expérience de

diffraction sont résumés dans le tableau VI. Les valeurs inscrites correspondent à l'échantillon de concentration 0,55 mg/mL dans le tampon à base de phosphate.

Tableau VI: Collection de données et paramètres dérivés des données de dispersion

Paramètres	Données
<i>Beam Diameter</i>	250 μm \times 250 μm
<i>Wavelength</i> (\AA)	1.5418
<i>q range</i> (\AA^{-1})	0.008 \AA^{-1} -0.8 \AA^{-1}
<i>Exposure time</i> (s)	300
<i>Concentration range</i>	0.55-3.73
<i>Temperature</i> (K)	283
Structural parameters#	
<i>I(0)</i> (from Guinier)	0.139 +/- 0.002
<i>Rg</i> (\AA) (from Guinier)	22.673 +/- 10.329
<i>Dmax</i> (\AA)	77.65
<i>Porod volume estimate</i> (\AA^3)	25421.70
Molecular-mass determination#	
<i>Molecular mass Mr</i> [from <i>I(0)</i>]	23850.51 +/- 2.38
<i>Calculated monomeric Mr from sequence</i> (Da)	19736.7
Software employed	
<i>Data processing</i>	RAW
<i>Ab initio analysis</i>	DAMMIN
<i>Validation and averaging</i>	DAMMAVER
<i>Three-dimensional graphics representations</i>	PymoL

(#) Les calculs concernent les données de l'échantillon SecP43-Tris-0,55 mg/mL.

3.6.3. Validation de la collecte des données de diffraction de SecP43 tronquée

Les courbes de diffractions à une dimension obtenues sont représentées sur la figure 27 A. les signaux des tampons Tris et phosphate qui y sont également montrés et servent de contrôle négatif. La figure 27 B montre les signaux de diffraction dus à la protéine SecP43 dans les échantillons de concentration différente. Ces courbes ont été obtenues en soustrayant le signal dû aux tampons des signaux de la figure 27 A. Les informations obtenues à partir des données de SAXS ne peuvent être pris en compte que si les échantillons sont monodisperses. Ainsi, afin de vérifier la validité de notre collecte de données, nous avons dans un premier temps fait varier la concentration de nos échantillons pour analyser l'impact de ces variations sur le comportement des facteurs tels que le rayon de gyration (R_g) de la protéine et l'intensité de dispersion $I(s)$ des rayons X (Tableau VII). Le rayon de gyration devrait pour une protéine repliée (globulaire), uniconformationnelle (peu dynamique) et non agrégée demeurer constant tandis que l'intensité de dispersion $I(s)$ devrait varier proportionnellement à la concentration en protéine des échantillons irradiés. Par contre, dans le cas d'une protéine flexible et dynamique qui adopte plusieurs conformations en solution, le rayon de gyration pourrait varier d'une mesure à l'autre. C'est cette tendance qui est observée dans la figure 29 A. On note que le rayon de gyration de SecP43, qui est une mesure de la taille d'une particule est variable quoique on puisse estimer une moyenne d'environ 26 Å (figure 29 A).

Par ailleurs, l'approximation de Guinier varie d'un échantillon à l'autre. On note toutefois que les échantillons SecP43-Phosphate-0,55 mg/mL, SecP43-Tris-1,5 mg/mL et SecP43-Tris 3 mg/mL sont plus ou moins monodisperse (figure 28) car les points sont plus ou moins alignés suivant la droite rouge qui délimite la région de Guinier. En revanche, les échantillons SecP43-Tris-0,5 mg/mL, SecP43-tris-0,7 mg/mL et SecP43-Tris-0,9 mg/mL présente une légère agrégation compte tenu de la déviation des points vers le dessus de la droite rouge (figure 28). Cette Observation paraît contre-intuitive puisqu'elle suggère que des échantillons de plus faibles concentrations s'agrègent plus facilement que ceux de plus haute concentration. En effet, l'approximation de Guinier peut être biaisée dans le cas des protéines flexibles (Véronique B et al, 2012). L'approximation de Guinier stipule

qu'à de très petits angles de dispersion, il existe une région dite de **Guinier** dans laquelle la courbe $\ln(I)$ en fonction de s^2 devient linéaire pour un échantillon monodisperse. Dans cette région, le produit entre le rayon de gyration (R_g) et le moment de transfert (q) est typiquement de 1,3. En présence d'agrégats de protéines, on observe une déviation des points vers le dessus de la droite rouge tandis qu'une déviation vers le bas de cette droite indique qu'il y a répulsion inter particulaire.

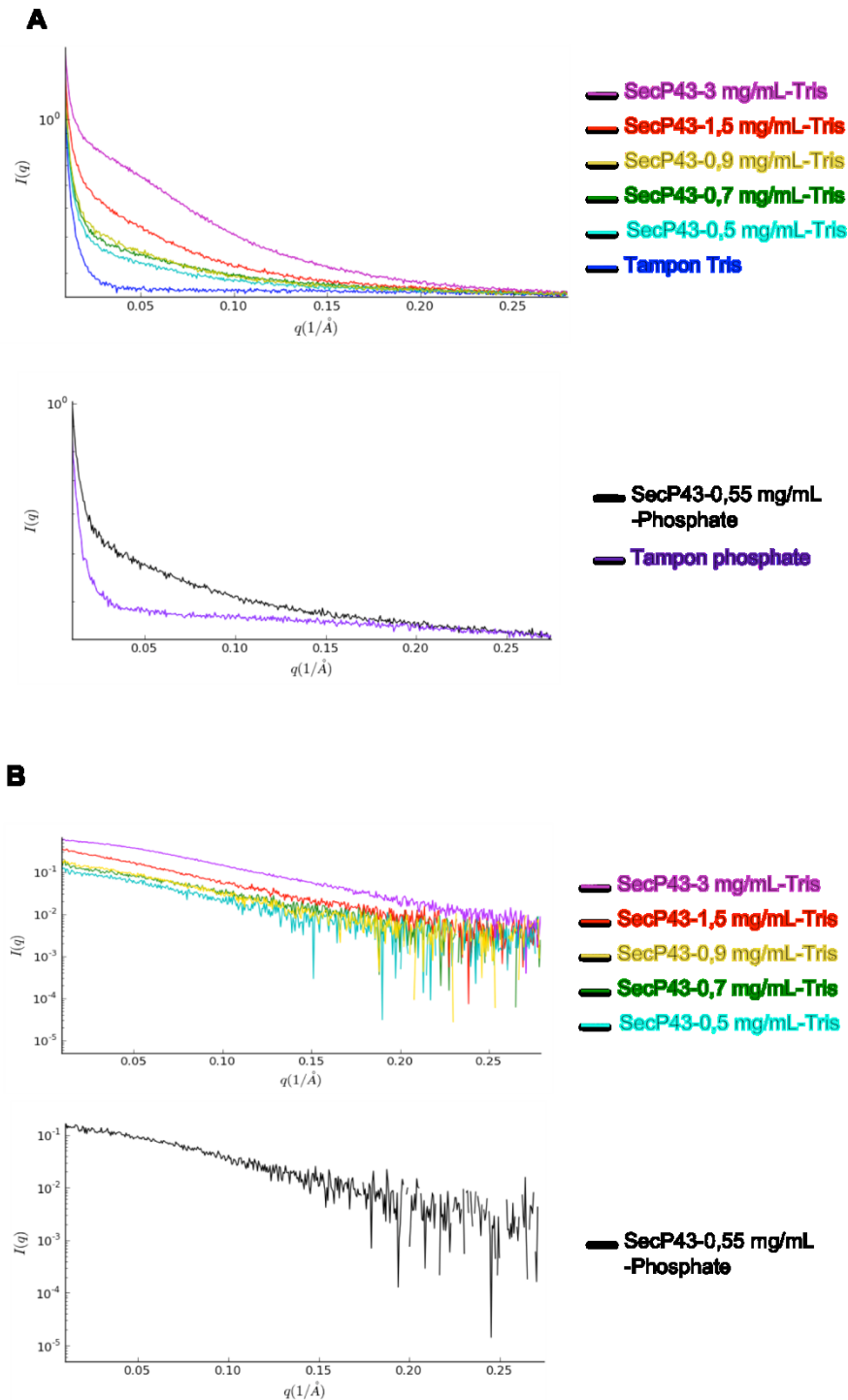


Figure 27. Validation de la collecte des données de diffraction par SAXS. Analyse de la version tronquée de SecP43. La monodispersité des échantillons fut préalablement vérifiée par chromatographie d'exclusion stérique. **A.** Courbes de diffractions unidimensionnelles obtenues pour différentes concentrations de SecP43 dans les tampons Tris et Phosphate. La différence entre les signaux des tampons et les échantillons protéiques corrobore la diffraction due à la protéine. **B.** Courbes de

diffraction unidimensionnelles obtenues par soustraction du signal du aux tampons Tris et Phosphate respectivement. La soustraction a été effectuée par le logiciel RAW.

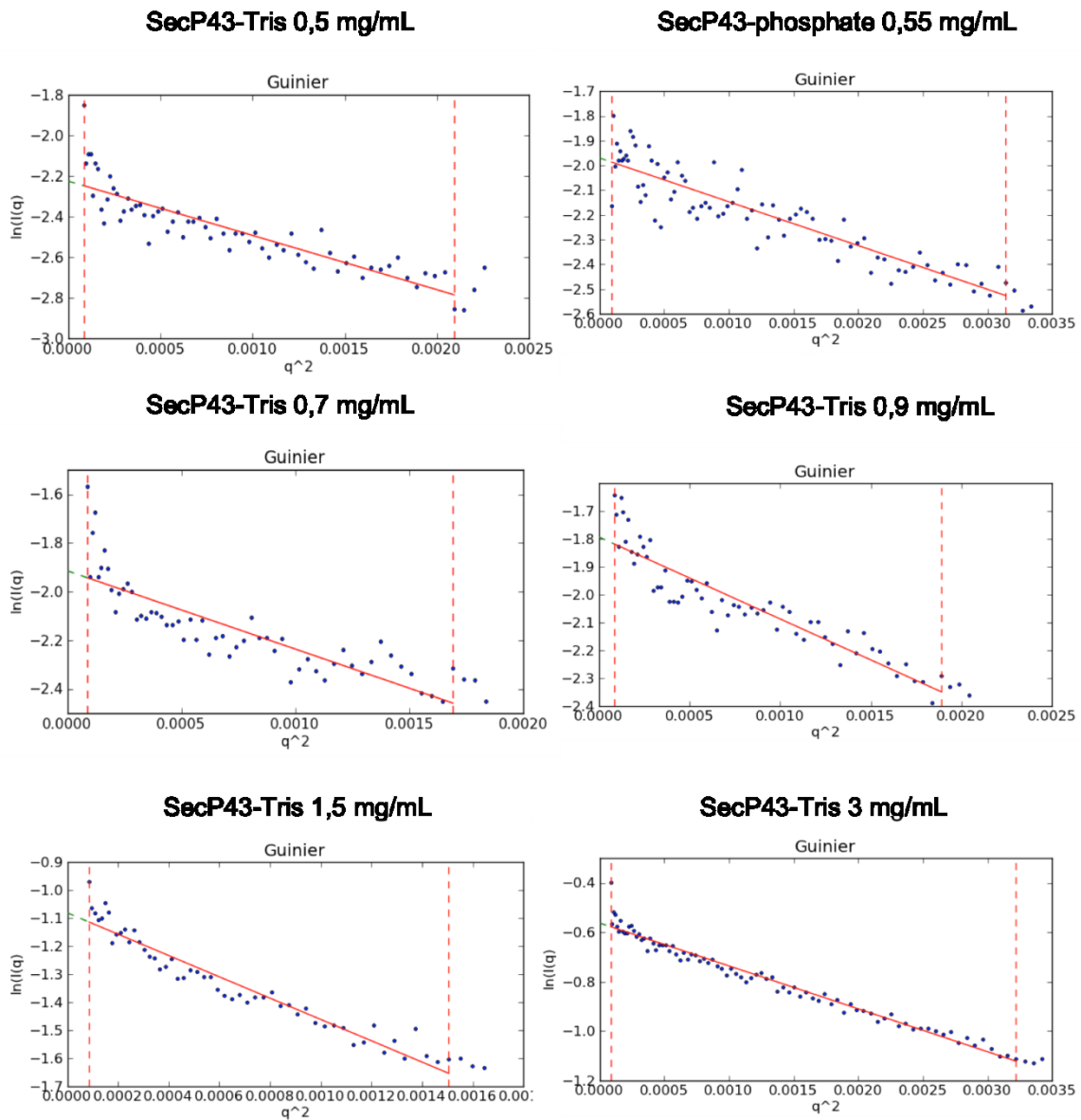


Figure 28 : Approximation de Guinier. L'approximation de Guinier donne un indice sur l'homogénéité d'un échantillon. Cette approximation fut calculée par le logiciel RAW (Svergun et al 2011). La valeur du produit qR_g étant en moyenne 1,3. Les échantillons SecP43-Phosphate-0,55 mg/mL, SecP43-Tris-1,5 mg/mL et SecP43-Tris 3 mg/mL sont plus ou moins monodisperses tandis que les échantillons SecP43-Tris-0,5 mg/mL, SecP43-tris-0,7 mg/mL Et SecP43-Tris-0,9 mg/mL présentent une légère agrégation en raison de la déviation des point vers le dessus de la droite rouge.

Tableau VII : Paramètres des courbes de diffraction unidimensionnelles*

Échantillons\Paramètres	I₀	Rg	MM (kDa)
SecP43-Tris-0,5 mg/mL	0,10814	28,34907	20,34002
SecP43-Phosphate-0,55 mg/mL	0,13949	23,05256	23,85091
SecP43-Tris-0,7 mg/mL	0,14532	31,06603	19,80326
SecP43-Tris-1,5 mg/mL	0,34002	33,77001	21,36977
SecP43-Tris-3 mg/mL	0,56989	22,82238	17,8647

(*) Les paramètres ont été calculés par le logiciel RAW

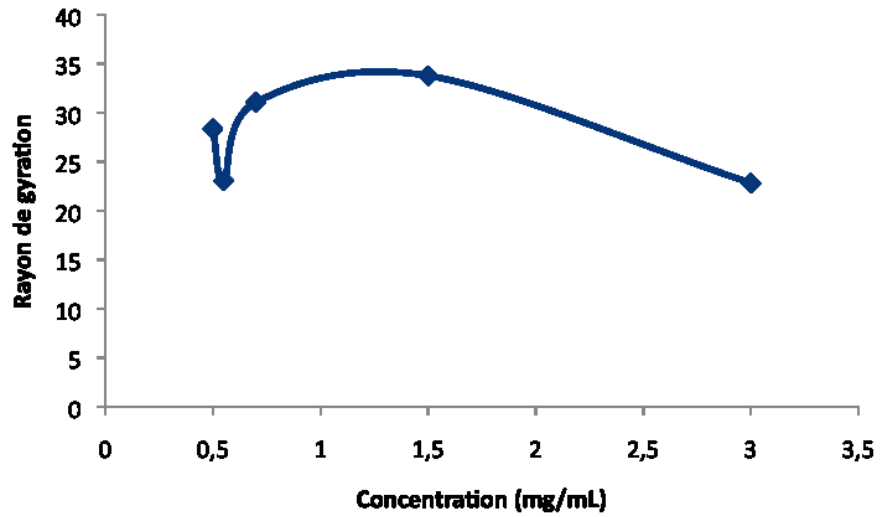
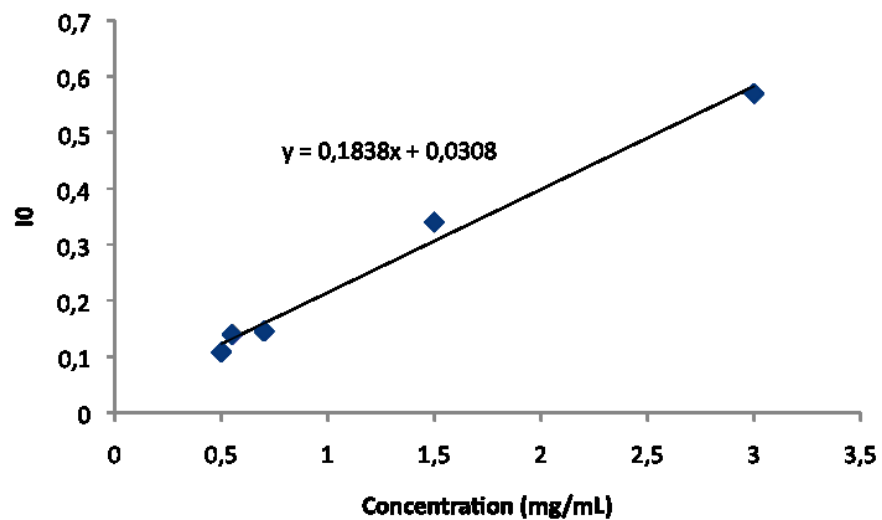
A**B**

Figure 29. Comportement des paramètres Rg et I_0 . A. Rayon de gyration en fonction de la concentration de la protéine SecP43 tronquée. On note que le rayon de gyration (Rg) est variable. D. L'intensité I_0 est proportionnelle à la concentration de protéine tel qu'escompté. Les valeurs des I_0 et Rg ont également été calculées par le logiciel RAW.

3.6.4. Fonction de distribution des distances $p(r)$

Intéressons nous maintenant à la fonction de distribution des distances $p(r)$ des échantillons analysés. La fonction $p(r)$ est un histogramme représentant la distribution des distances interparticulaires au sein d'une protéine. La plus grande distance mesurée équivaut au D_{\max} qui est la valeur retenue comme diamètre de la sphère permettant de calculer la forme de la protéine d'intérêt. Cette fonction a été calculée automatiquement par le logiciel GNOM (Svergun 1999) et elle permet également de détecter la présence éventuelle d'une protéine à plusieurs domaines et la dimérisation selon le profil obtenu (figure 30). Nous avons calculé la fonction $p(r)$ des échantillons distincts afin de voir l'effet de la concentration sur la conformation de SecP43. On remarque en effet que le profil de l'échantillon de SecP43-Tris-0,9 mg/mL est caractéristique d'une protéine à plusieurs domaines flexibles l'un par rapport à l'autre. Il s'agit en occurrence des domaines RRM1 et RRM2 de SecP43 tronquée. Le D_{\max} moyen pour les échantillons SecP43-Tris-0,5 mg/mL, SecP43-Phosphate-0,55 mg/mL, SecP43-Tris-0,7 mg/mL et SecP43-Tris-3mg/mL échantillon est d'environ 77 Å. Par contre, les échantillons SecP43-Tris-0,9 mg/mL et SecP43-Tris-1,5 mg/mL s'écartent légèrement de cette moyenne avec des D_{\max} de 89,76 Å et de 92,34 Å respectivement (figure 30). Cette différence s'explique probablement par l'adoption d'une conformation plus étendue de SecP43 dans ces échantillons.

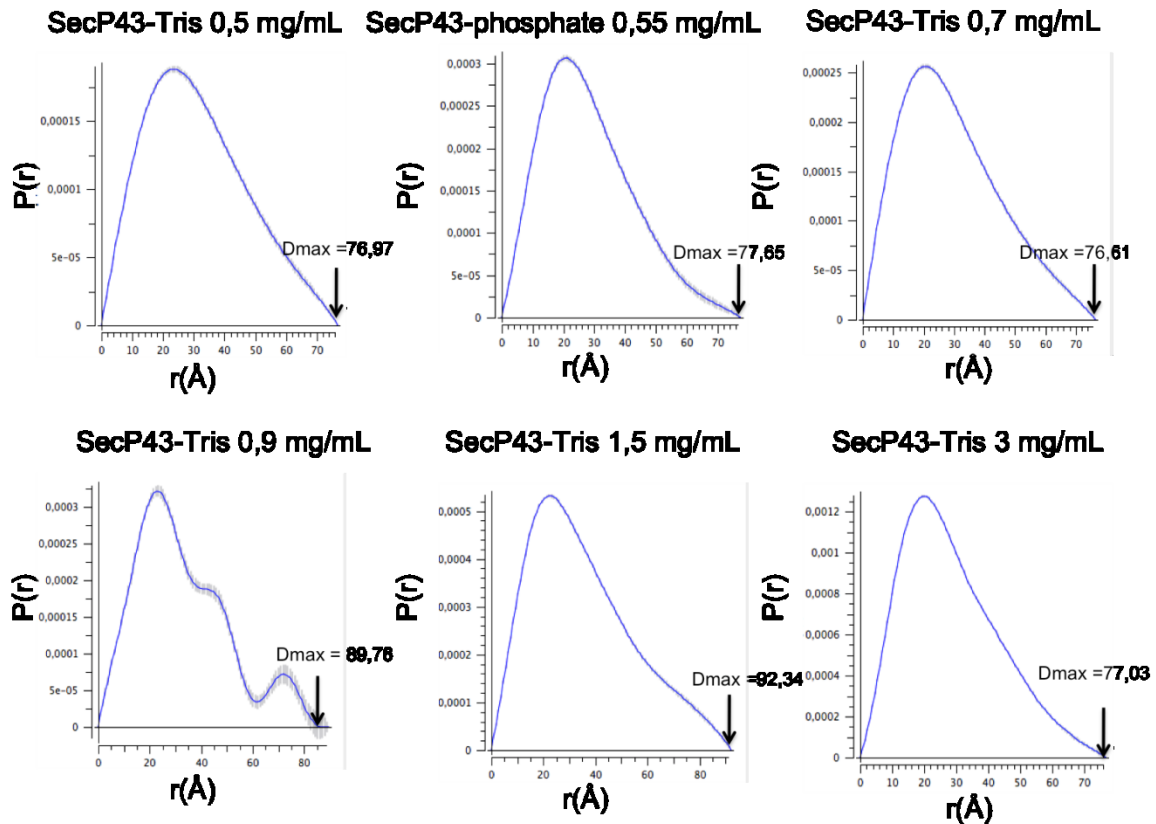


Figure 30. Fonctions de distribution de distances $p(r)$ et effet de la concentration en SecP43. Les courbes de $p(r)$ ont été déterminées de manière automatique par le logiciel GNOM de la suite ATAS (Svergun). Les flèches indiquent les valeurs des D_{\max} pour chaque échantillon.

3.6.5. Fusion des courbes de diffraction des échantillons SecP43-Tris

Étant donné la variabilité du profil de diffraction de SecP43 notamment le rayon de gyration, nous avons fait fusionner les courbes d'une même série de données par le logiciel Almerge de la suite ATSAS afin de d'étudier le repliement (kratky plot) et de calculer une enveloppe qui représenterait la conformation moyenne de SecP43 en solution. La courbe résultant de cette fusion et les paramètres y étant associés sont représentés sur la figure 31. On note que l'agrégation naguère observée par sur l'approximation de Guinier de certains échantillons primaires semble inexistante pour la courbe fusionnée (figure 31 B). Le D_{\max} quant à lui est réduit à 74, 48 \AA (figure 31 C)

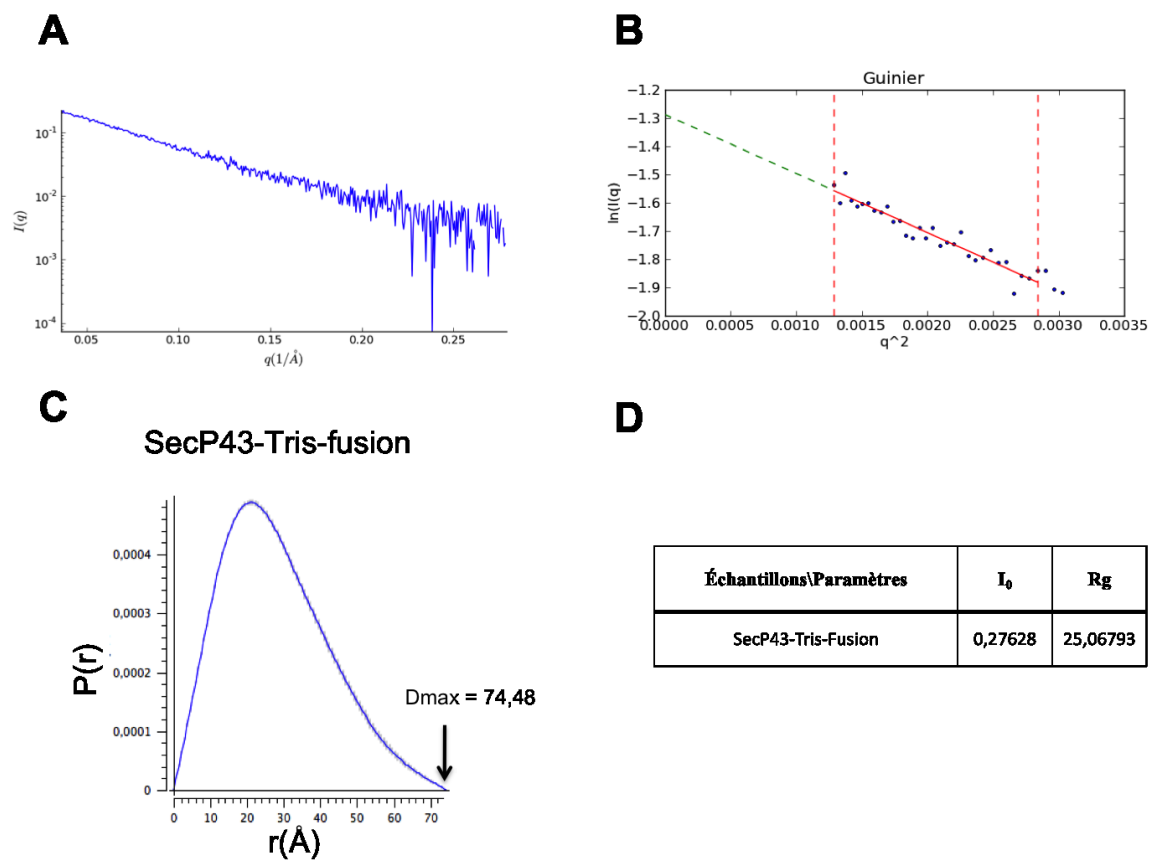


Figure 31 : Courbe résultant de la fusion des Courbes de diffraction des échantillons SecP43-Tris. **A.** Courbe de diffraction unidimensionnelle de la résultante de la fusion des courbes SecP43-Tris-0,5-0,7-0,9-1,5-3 mg/mL par le logiciel Almerge. **B.** Approximation de Guinier de la courbe résultante. La valeur du produit qR_g étant d'environ 1,31. **C.** fonction de distribution des distances. **D.** Paramètres I_0 et R_g issue de l'approximation de Guinier calculé par le logiciel RAW.

3.6.6. Repliement de SecP43 (*Kratky plot normalisé*)

Le repliement de SecP43 est une des informations que nous pouvons extraire de l'analyse des données de diffraction. Le repliement est ici évalué par l'allure du graphique dénommé *Kratky plot* qui est la courbe représentative du produit $I(q) \times q^2$ en fonction de q . toutefois, Afin de calculer une courbe *Kratky* indépendante de la taille des protéines étudiées, il convient de normaliser le calcul en exprimant cette fois-ci $(q/R_g)^2 \times I(q)/I_0$ en fonction de q/R_g (Véronique B et al, 2012). On note en effet que la version tronquée de SecP43 n'est pas complètement repliée, car la courbe correspondant au *kratky* effectué

sur la courbe résultant de la fusion des courbes de concentration différentes n'indiquent pas une distribution gaussienne parfaite tel que représenté à la figure 32 A. Par ailleurs, l'allure de la courbe observée s'apparente à celle de la protéine P67 (figure 32 A et B) et permet de la classer dans la famille des protéines à domaines multiples pouvant adopter plusieurs conformations selon les partenaires d'interactions et les conditions de solubilisation.

Remarquons que la courbe de la figure 32 B ne rencontre pas l'origine du repère car une partie infime des données aux petits angles n'apparaissait pas dans le fichier .text obtenu suite à la fusion par Almerge. L'analyse de la courbe n'en est pas pour autant affectée puisque le début d'une courbe de kratky normalisée est similaire pour tous les types de protéines (figure 32 A)

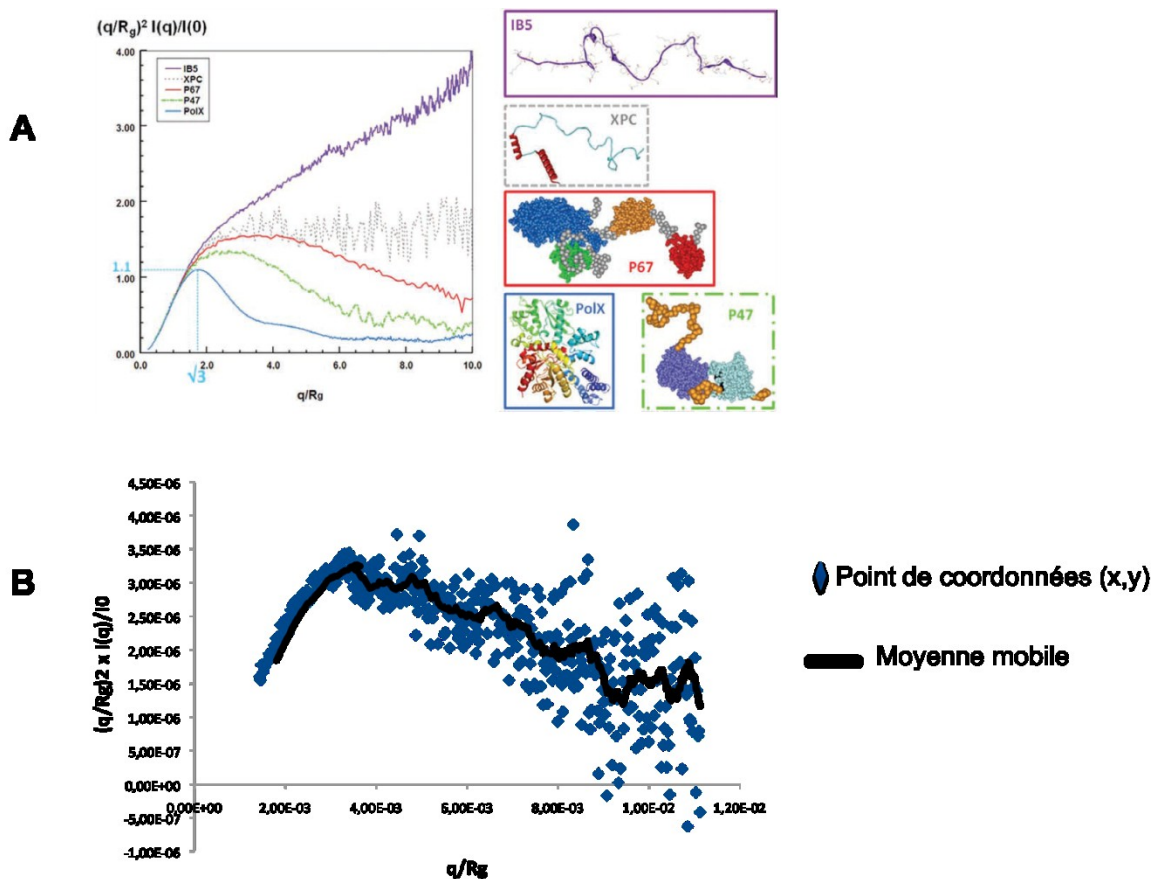


Figure 32. Repliement de SecP43 selon le profil de la courbe Kratky normalisée. A. Exemple d'allures de la courbe de Kratky selon l'état de repliement de différentes protéines (Véronique

B et al, 2012). **B.** La modification des paramètres q et $I(q)$ par les constantes I_0 et R_g a été effectuée par le logiciel Excel. Les constantes I_0 et R_g utilisées sont celles issues de l'approximation de Guinier de la courbe de fusion.

3.6.7. Calcul des enveloppes

Les données extraites des fichiers des distributions de distances $p(r)$ ont ensuite été utilisées pour le calcul des enveloppes de SecP43. Le logiciel DAMMIF (*online*, <http://www.embl-hamburg.de/biosaxs/atsas-online>) fut premièrement utilisé pour générer des modèles ayant des profils de diffraction unidimensionnels compatibles à la courbe de diffraction expérimentale. Les modèles obtenus ont ensuite été moyennés par le logiciel DAMAVER et le modèle résultant fut affiné par le logiciel DAMFILT (figure 33). Ces calculs ont été effectués en ligne grâce à la suite ATSAS (Servgun, <http://www.embl-hamburg.de/biosaxs/atsas-online/>). Selon les modèles de SecP43 obtenus, il s'agit une protéine contenant des régions de densité importante (volumineuse) et des régions de densité plus étroite (figure 33). On note que la différence entre les enveloppes des échantillons SecP43-Tris-fusion et SecP43-Phosphate-0,55 mg/mL est minime. Les régions volumineuses abriteraient probablement les domaines de reconnaissance de l'ARN (RRM) typiquement formés de deux hélices alpha et d'un feuillet beta. La région moins dense quant à elle contiendrait en son sein des boucles reliant les deux domaines RRM ou les séquences adjacentes à ces deux domaines. Pour mieux les localiser, nous avons superposé (grâce au logiciel SUPCOMB) ces enveloppes avec les structures atomiques résolues par RMN, des domaines RRM synthétiques de SecP43 (figure 34). Il se trouve que chacun de ces modèles accommode parfaitement les motifs RRM pris individuellement (figure 34 A et B). Par contre, lorsqu'on essaye de faire une co-superposition, l'un des motifs RRM s'incruste inadéquatement dans l'enveloppe. Ceci s'explique probablement par le fait que, lors de la première superposition, le logiciel SUPCOMB réduit la différence de densité entre la structure cristalline du domaine RRM1(1-90^e résidus) et l'enveloppe. Ainsi, l'insertion subséquente du domaine RRM2 (94-184^e résidus) pourrait être décalée de sa position réelle.

SecP43-Phosphate-0,55 mg/mL

SecP43-Tris-fusion

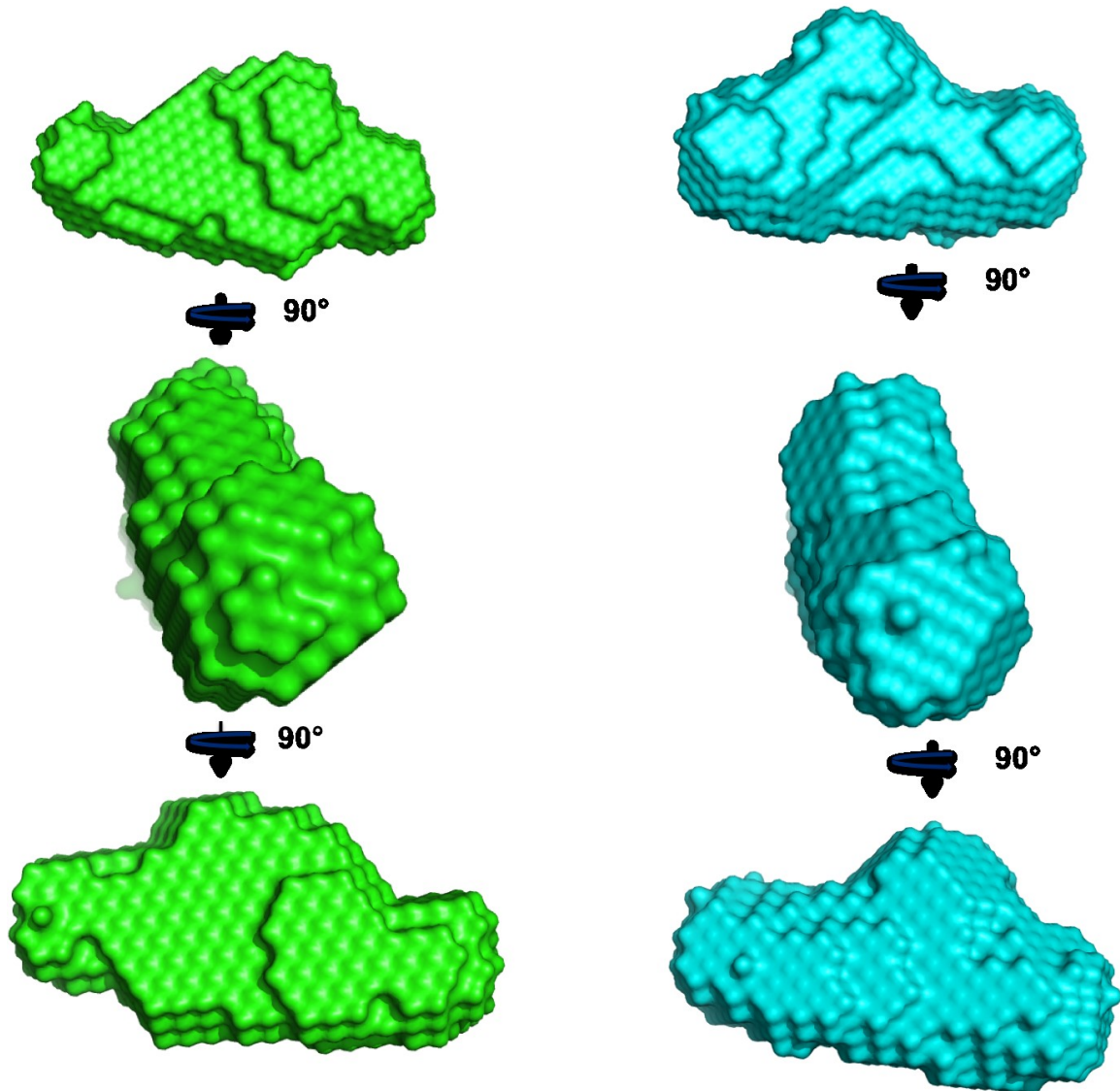


Figure 33. Modèles de basse résolution de SecP43 tronquée. Les enveloppes ont été calculées par le logiciel DAMMIF en ligne (<http://www.embl-hamburg.de/biosaxs/atsas-online>) de manière à être compatible avec les courbes de diffraction expérimentale. Les coefficients de corrélation par rapport aux données brutes étant respectivement de $\chi^2(1,070$ et $1,059)$. 10 modèles distincts furent générés pour chaque les échantillons SecP43-Phosphate-0,55 mg/mL et SecP43-Tris-fusion. Ces modèles furent ensuite moyennés par DAMAVER. Les valeurs de NSD respective étant de $0,588 \pm 0,029$ et $0,588 \pm 0,013$. Le modèle moyen de chaque échantillon est ensuite affiné par DAMFILT.

SecP43-Phosphate-0,55 mg/mL

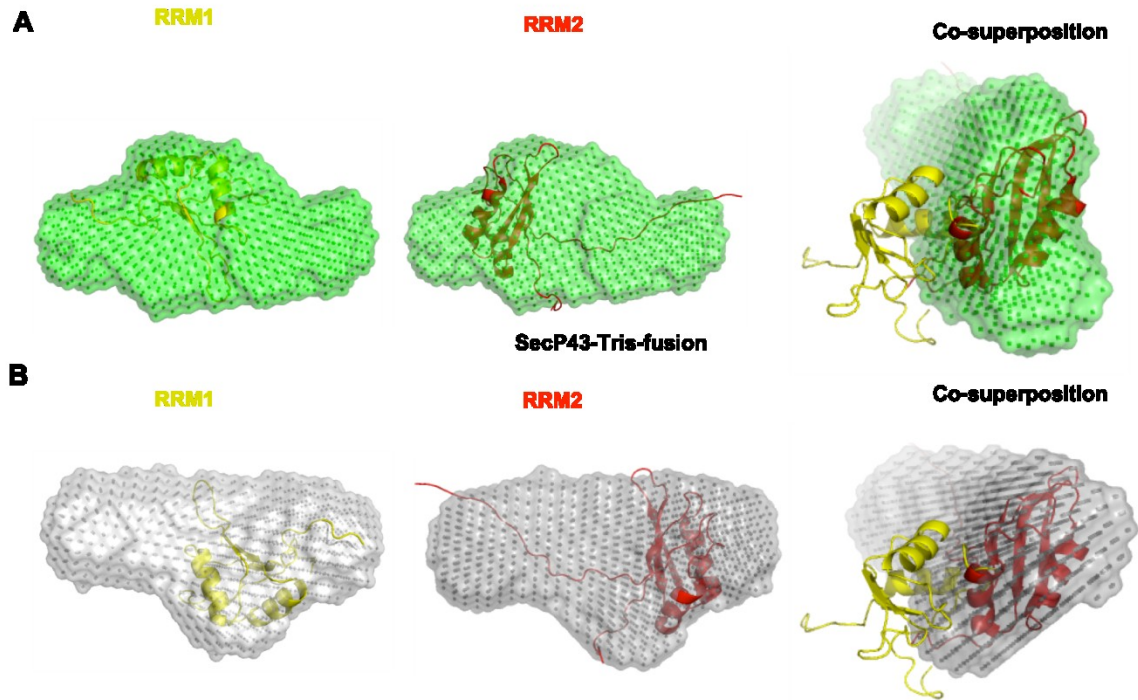


Figure 34. Superposition des structures atomiques des RRM de SecP43 dans les enveloppes calculées. Les structures des domaines synthétiques de reconnaissance de l'ARN, RRM1 et 2 ont été résolues par RMN (identification PDB : 2DIV et 2DHG) (Imai T. et al, 2006). La superposition des structures dans les enveloppes a été effectuée par le logiciel SUPCOM de la suite ATSAS (Svergun). **A.** Superposition des motifs RRM1 et RRM2 dans l'enveloppe de l'échantillon SecP43-Phosphate-0,55 mg/mL. **B.** Superposition des motifs RRM1 et RRM2 dans l'enveloppe de l'échantillon SecP43-Tris-fusion

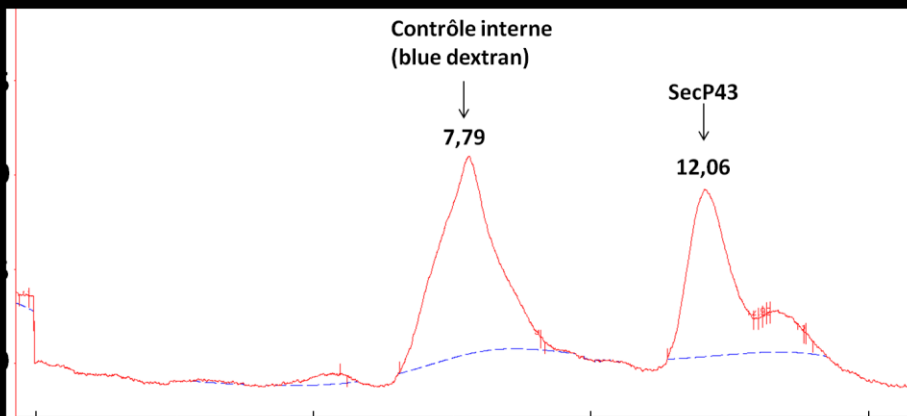
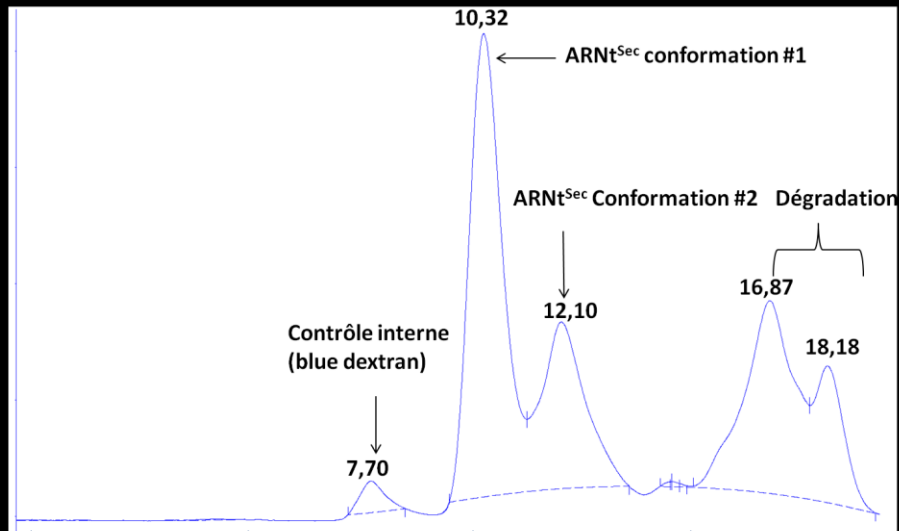
3.8. Études des interactions SecP43/ARNt^{Sec}/SepSecS

3.8.1. Analyse des interactions SecP43/ARNt^{Sec} par chromatographie d'exclusion stérique

L'un des objectifs que nous nous sommes fixés dans ce projet était de caractériser les interactions entre SecP43 et certains de ses partenaires d'interactions présumés tels que SepSecS et l'ARNt^{Sec} *in vitro*. Notons à cet effet que SecP43 est également désigné *eukaryotic tRNA associated protein* dans la littérature (Small Howard et al 2006). Selon les études menées par d'autres groupes, SecP43 ne s'associerait à l'ARNt^{Sec} qu'en présence d'autres facteurs du sélénosome tels que le facteur d'élongation EFsec (Small Howard et al 2006). De la même manière, sa liaison avec SepSecS ne serait possible qu'en présence la SPS1 impliquée dans la synthèse de Sec (Xu XM. et al, 2005). Ainsi, afin de confirmer ou d'infirmer ces observations, nous avons étudié les interactions entre SecP43/ARNt^{Sec} /SepSecS par diverses techniques comprenant entre autres la chromatographie par filtration sur gel, des essais de réticulation analysés par immunobuvardage de type *western* et finalement par des essais de retard sur gel (EMSA).

Les études d'interactions par filtration sur gel n'ont pas permis de montrer une interaction stable entre SepSecS et SecP43 d'une part, et entre SepSecS et ARNt^{Sec} d'autre part (données non montrées). Par contre, une observation différente a été faite dans le cas de SecP43 et ARNt^{Sec}. Pour cette expérience, SecP43 et l'ARNt^{Sec} ont d'abord été injectés séparément afin de bien distinguer leurs volumes d'élutions qui sont respectivement d'environ 12,06 mL et 10,32 mL (conformation #1 de l'ARNt^{Sec}) et 12,10 mL (conformation #2 de l'ARNt^{Sec}) lorsque le débit est fixé à 0.5 mL/min (figure 35 A). La figure 35 A montre également la dégradation de l'ARNt^{Sec} probablement due à la présence des nucléases. Lorsque nous avons co-injecté SecP43 et ARNt^{Sec}, nous observé un léger déplacement du pic d'élution correspondant probablement à un complexe instable entre SecP43 et ARNt^{Sec} (figure 35 B). On note en effet que ce déplacement semble se renforcer pour des concentrations en SecP43 plus élevées. Par ailleurs, notons que la conformation #1 de l'ARNt^{Sec} semble moins favorisée que la conformation #2

suite à l'ajout de SecP43. Cela se traduit notamment par la disparition du pic d'élution de 10,32 mL correspondant au contrôle (figure 35 A). Ainsi, puisque les volumes d'élution respectifs de SecP43 et de la conformation #2 de l'ARNt^{Sec} sont très proches (12,06 mL et 12,10 mL), nous sommes en droit de suggérer qu'il s'agit probablement d'une co-élution de ces deux molécules se caractérisant par le profil de la figure 35 B. En principe, un complexe stable entre ces deux partenaires devrait avoir un volume d'élution d'au plus 11 mL. Or, le pic le plus décalé vers la gauche (ratio 1:4) a un volume d'élution d'environ 11,76 mL. Ceci pourrait s'expliquer par le fait que la méthode de calibration utilisée ne soit pas nécessairement adaptée pour les protéines non globulaires, en occurrence SecP43. Par ailleurs, la dégradation progressive de l'ARNt^{Sec} tel qu'observée sur la figure 35 B lui conférerait probablement une taille inférieure à sa taille réelle (29700 Da).



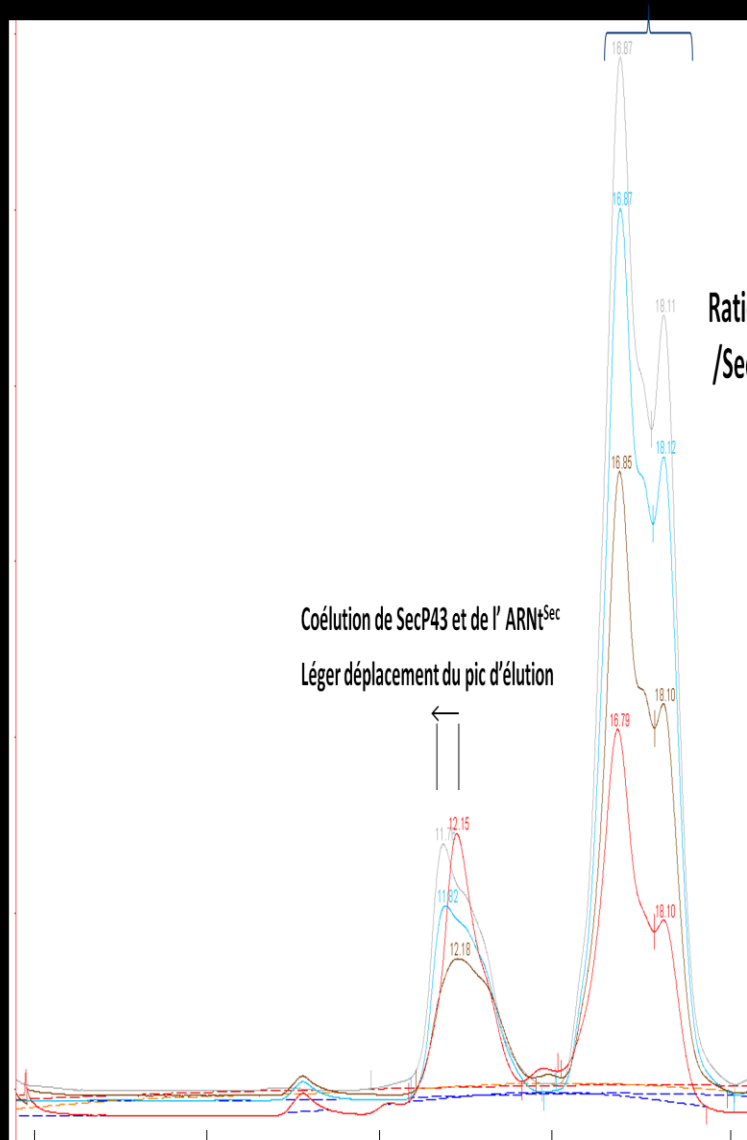


Figure 35. Analyse de l'interaction SecP43/ARNt^{Sec} par chromatographie d'exclusion stérique. A. Volumes d'éluion respectifs de SecP43 et ARNt^{Sec} injectés séparément sur la colonne Superdex 75. B. Des mélanges de 100 μ L contenant chacun 8 μ M d'ARNt^{Sec} et des concentrations croissantes de SecP43 à 8, 16, 24 et 32 μ M ont été préparées. Les mélanges ont été injectés sur la colonne

Superdex 75 (GE Healthcare). Les profils d'élution ont été superposés afin d'en étudier les variations. Seule l'absorbance à 260 nm est indiquée ici afin de simplifier le graphique. Un contrôle interne (bleu de dextran) a été ajouté.

3.8.2. Analyse des interactions SecP43/ARNt^{Sec}/SepSecS par réticulation au formaldéhyde

Etant donné que l'interaction SecP43/SepSecS qui a clairement été établie *in vivo* (Xu XM. et al, 2005) n'a pas pu être montrée par chromatographie d'exclusion stérique, nous avons postulé qu'il s'agirait d'une interaction faible ou transitoire et donc facilement déstabilisable par un effet de dilution. Pour détecter une interaction faible, nous avons eu recours à la technique de réticulation utilisant le formaldéhyde comme agent de réticulation. Comparativement à d'autres agents de réticulation, le formaldéhyde a l'avantage de seulement faire le pontage entre deux protéines très rapprochées l'une de l'autre en raison de sa petite taille. Par ailleurs, le formaldéhyde est connu pour sa rapidité d'action de formation de pontages et la stabilisation des interactions protéine-protéine transitoires (B. W. Sutherland et al, 2008). Le mécanisme de pontage procède par une attaque nucléophile du groupement C=O du formaldéhyde par le groupe NH₂ d'une lysine ou d'une arginine tel que représenté sur la figure 36.

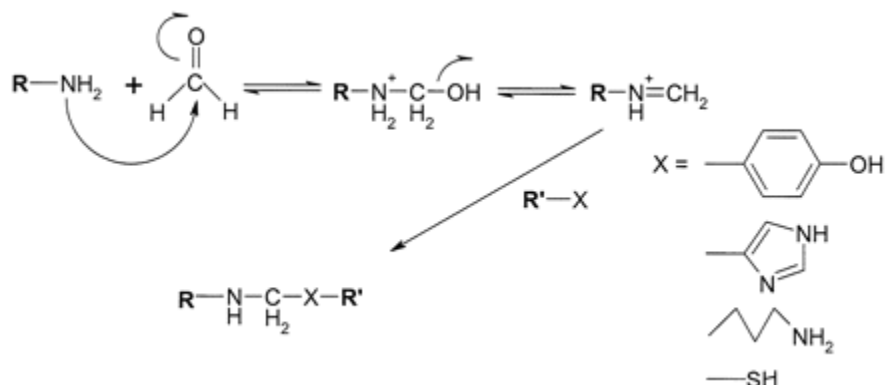


Figure 36. Mécanisme de réticulation des protéines par le formaldéhyde via la formation d'un cation immonium. (John A. Kiernan, 2000)

Nous avons dans un premier temps vérifié si une interaction directe entre SecP43 et SepSecS était détectable (figure 37). On note que SecP43, contrairement à SepSecS ne semble pas former des oligomères malgré l'augmentation de la concentration de l'agent

de réticulation. SepSecS se dimérise même en absence d'agent de réticulation et l'ajout du formaldéhyde promeut l'amplification du signal correspondant au trimère (figure 37). Aucun signal correspondant au complexe SecP43/SepSecS (environ 74 kDa) n'est identifiable. Nous avons dans un deuxième temps ajouté l'ARNt^{Sec} afin de voir si ce dernier était capable de promouvoir l'interaction entre les deux protéines d'intérêt (figure 38). Il s'avère dans ce cas qu'aucun signal correspondant au complexe SepSecS/SecP43/ARNt^{Sec} n'était perceptible malgré l'ajout de l'ARNt^{Sec} non chargé au système. Notons également qu'à partir d'une certaine concentration de formaldéhyde, l'intensité du signal de SecP43 s'affaiblit. Cela s'explique probablement par la formation d'entités multimériques dont le signal est inférieur à la limite de détection.

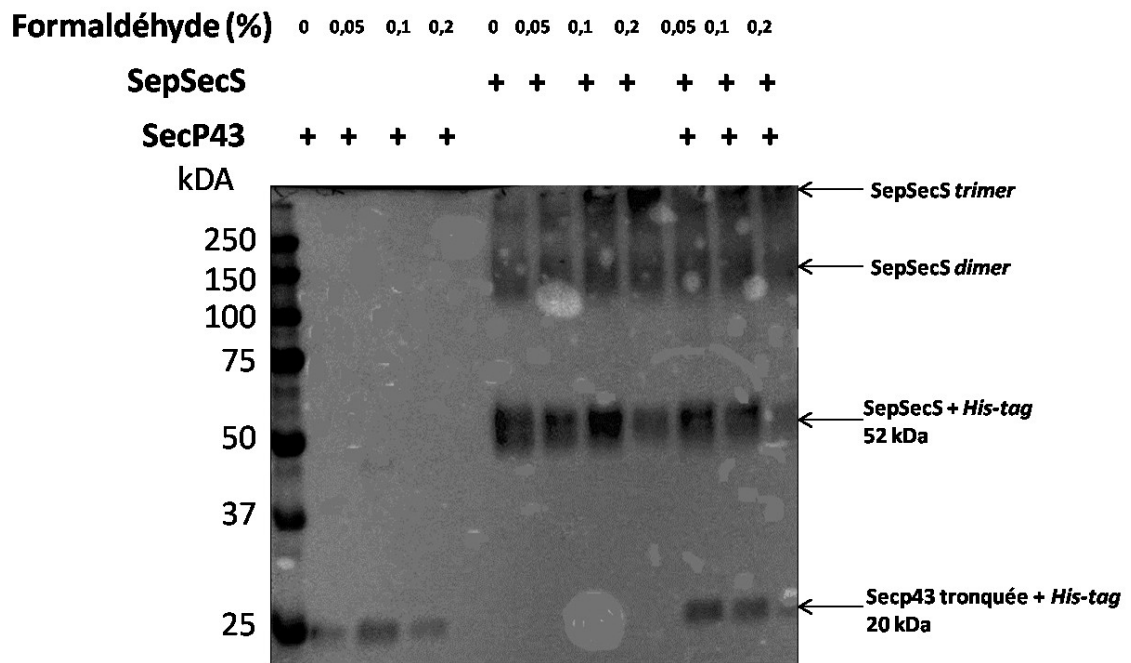


Figure 37. Essai de réticulation entre SecP43 et SepSecS. Les quatre premières pistes contiennent uniquement la protéine SecP43 avec un gradient de concentration en formaldéhyde. Les quatre pistes suivantes contiennent exclusivement la protéine SepSecS et le même gradient de concentration en agent de réticulation est appliqué. Ces huit premiers puits servent de contrôle négatif. Les trois dernières pistes contiennent les mélanges des deux protéines en présence du formaldéhyde. Les signaux représentant les espèces monomériques, dimériques et trimériques de chacun de ces protéines sont indiqués par les flèches. Les échantillons ont été analysés par SDS-PAGE 10 % et révélés par immunobuvardage de type *Western* utilisant les anticorps anti-His. Les échantillons des huit premières pistes représentent les contrôles négatifs.

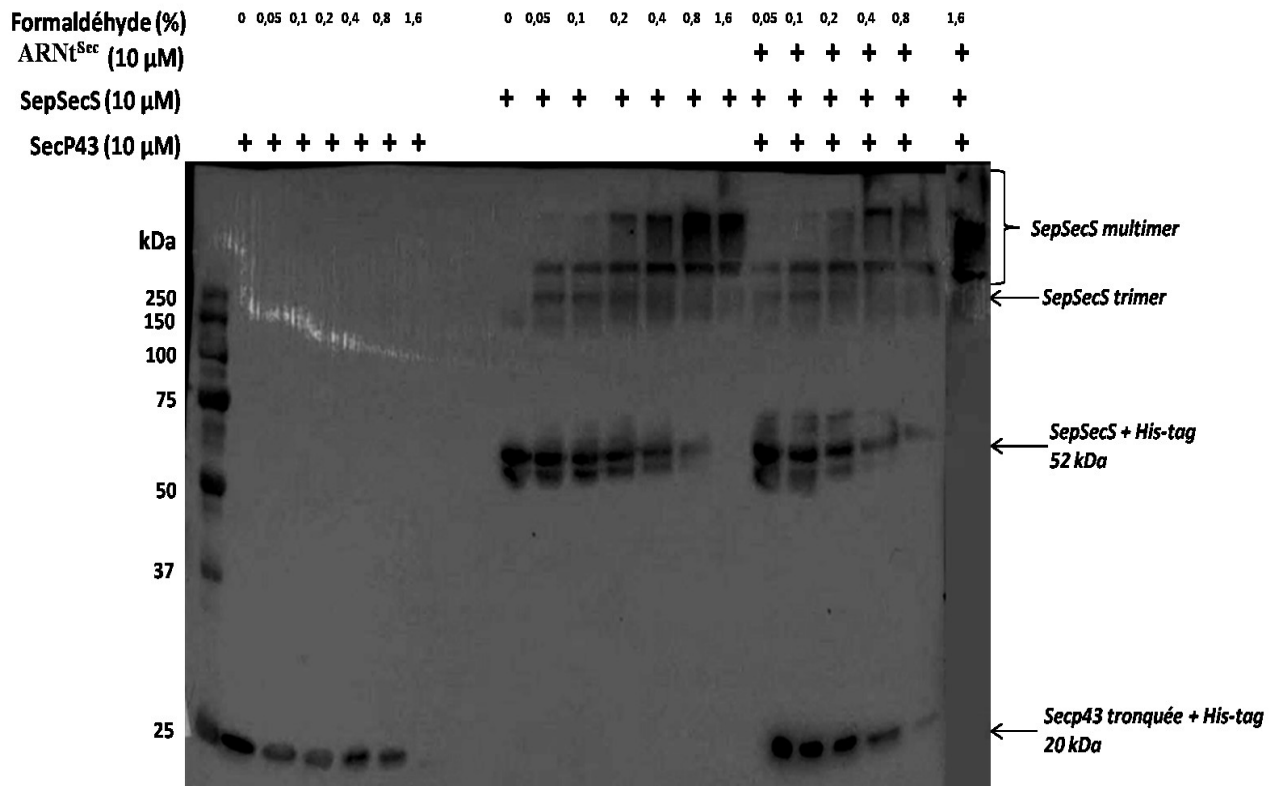


Figure 38. Essai de réticulation entre SecP43, SepSecS et effet de l'ARNt^{Sec}. Les sept premières pistes contiennent uniquement la protéine SecP43 avec un gradient de concentration en formaldéhyde. Les sept pistes suivantes contiennent exclusivement la protéine SepSecS et le même gradient de concentration en agent de réticulation est appliqué. Ces quatorze premiers puits servent de contrôle négatif. Les sept dernières pistes contiennent les mélanges des deux protéines et de l'ARNt^{Sec} en présence du formaldéhyde. Les signaux représentant les espèces monomériques, dimériques et trimériques de chacun de ces protéines sont indiqués par les flèches. Les échantillons ont été analysés par SDS-PAGE 10 % et les protéines ont été révélées par immunobuvardage de type *Western* utilisant les anticorps anti-His. Les échantillons des quatorze premières pistes représentent les contrôles négatifs.

3.8.3. Analyse des interactions SecP43/ARNt^{Sec}/SepSecS par retard sur gel

La détection d'interactions étant très peu évidente avec la technique de réticulation, nous avons en dernier ressort eu recours à la technique de retard sur gel pour analyser la formation potentielle de complexes entre nos trois composantes d'intérêt. Les échantillons ont été préparés dans un minimum de volume afin de prévenir la déstabilisation des complexes potentiels par dilution. Ainsi il s'agissait de voir premièrement si l'interaction SecP43/ARNt^{Sec} était détectable sans l'intermédiaire d'un tiers partenaire. La figure 39 A montre à cet effet que l'interaction SecP43/ARNt^{Sec} déjà appréhendée par chromatographie d'exclusion semble faible étant donné que le déplacement de l'ARNt^{Sec} ne survient qu'à partir d'un ratio stœchiométrique SecP43/ARNt^{Sec} de 104 (figure 39 A). Nous sommes même en droit de supposer qu'il s'agit plus d'un encombre stérique que d'une interaction spécifique. Par contre dans le cas de SepSecS, le déplacement de l'ARNt^{Sec} survient pour un ratio stœchiométrique équimolaire (figure 39 B). Comparativement à SecP43, SepSecS présente donc plus d'affinité pour l'ARNt^{Sec} non chargé. Finalement, nous avons intégré SepSecS au mélange SecP43/ARNt^{Sec} afin de tester si SepSecS stimulerait l'interaction SecP43/ARNt^{Sec}. Contre toute attente, les résultats suggèrent que SecP43 réduit l'interaction SepSecS/ARNt^{Sec} car le déplacement de l'ARNt^{Sec} qui jadis survenait pour un ratio stœchiométrique équimolaire, survient pour un ratio d'environ 3,3 en présence de SecP43 (Figure 40). Ces résultats suggèrent que SecP43 induit probablement un changement conformationnel au sein de SepSecS.

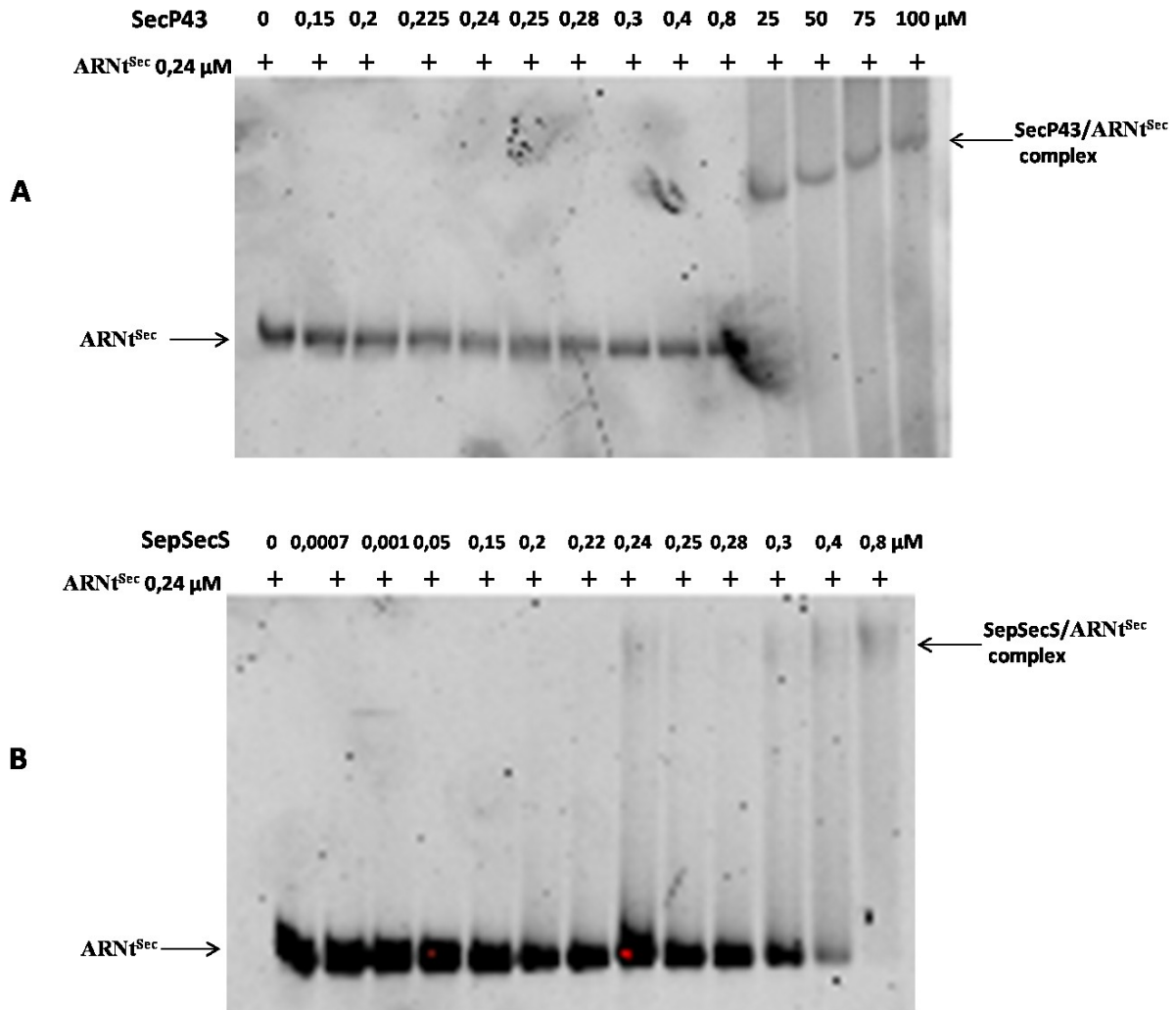


Figure 39. Analyse des interactions SecP43/ARNt^{Sec} et SepSecS/ARNt^{Sec} par essai de retard sur gel. Des mélanges ARN/protéines ont été effectués en maintenant la concentration d'ARNt^{Sec} constante et en variant la concentration des protéines (0-100 μM). Les échantillons ont été analysés sur gel natif d'acrylamide 10 %. La détection de l'ARNt^{Sec} a été faite par coloration au Sybr gold. **A.** Analyse de l'interaction SecP43/ARNt^{Sec}. **B.** Analyse de l'interaction SepSecS/ARNt^{Sec}.

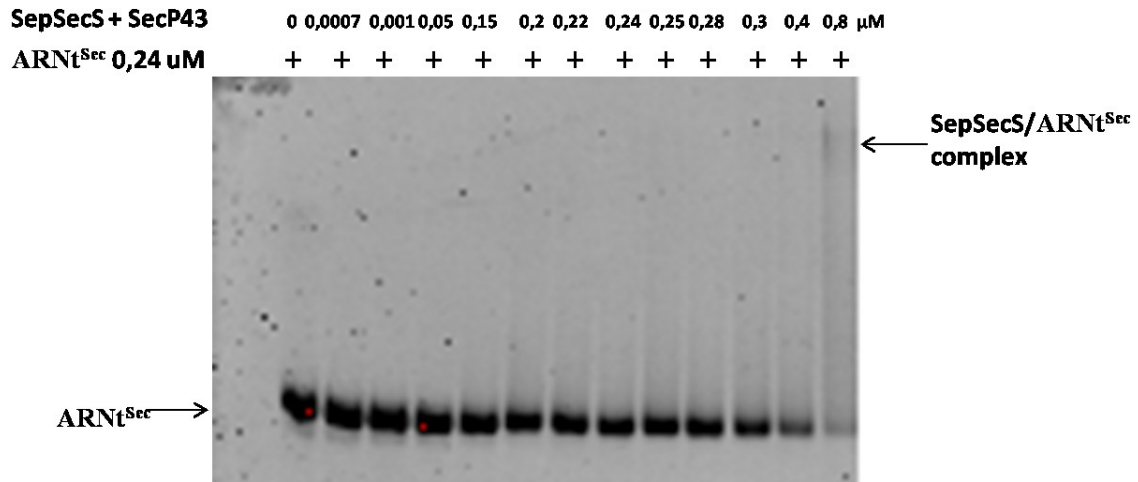


Figure 40. Analyse des interactions SecP43/ARNt^{Sec} et SepSecS/ARNt^{Sec} par essai de retard sur gel. Des mélanges ARN/protéines ont été effectués en maintenant la concentration d'ARNt^{Sec} constante et en variant la concentration des protéines (SecP43 + SepSecS) (0-0,8 μM). Les échantillons ont été analysés sur gel natif d'acrylamide 10 %. La détection de l'ARNt^{Sec} a été faite par coloration au Sybr gold.

Chapitre 4. Discussion

4.1 Mise en contexte

En dépit de sa présence à l'état de traces dans l'organisme humain, le sélénium joue un rôle important et unique parmi les autres microéléments car c'est le seul dont l'insertion dans les protéines est génétiquement encodée sous la forme du 21^{ème} acide aminé, la sélénocystéine. Selon l'état actuel des connaissances sur le sujet, 25 sélénoprotéines ont été identifiées dans le protéome humain. Bien que la fonction biologique des certaines de ces sélénoprotéines demeure un mystère, des fonctions diverses ont été attribuées à certaines d'entre elles. Des études ont montrées qu'elles interviennent entre autre dans le système de défense contre les radicaux libres, dans la régulation de l'état d'oxydoréduction de la cellule et une dans large variété de voies métaboliques spécifiques.

Une homéostasie anormale du sélénium dans l'organisme ainsi que des mutations dans les gènes des sélénoprotéines ont été causalement liés à une variété de pathologies humaines. Par contre, l'importance des composantes du **sélénosome** qui désigne l'ensemble des facteurs ribonucléoprotéiques appartenant à la machinerie de biosynthèse des sélénoprotéines, est moins bien connue (Rachel L et al 2012). Cependant des études plus poussées suggèrent que les enzymes responsables de la formation de la sélénocystéine et du décodage du codon UGA sont essentielles pour la santé et le développement chez plusieurs espèces en général et chez l'humain en particulier. Par exemple, il est connu que l'interaction SECIS/SBP2 est critique pour la santé humaine étant donné que des mutations dans le site de liaison de SBP2 où dans son domaine de liaison à l'ARN entraînent d'une part la réduction de la synthèse des sélénoprotéines et d'autre part l'établissement d'une variété de phénotypes pathologiques (Azevedo MF et al, 2010).

À la lumière de leurs multiples fonctions, les sélénoprotéines ont au cours des années récentes émergées comme étant des biomarqueurs potentiels pour plusieurs maladies telles que le diabète et certains types de cancer (Roman M et al, 2013).

Le mécanisme de synthèse des sélénoprotéines est de plus en plus cerné. Il procède notamment par deux étapes incluant séquentiellement une phase de synthèse de la sélénocystéine directement sur son ARNt apparenté dans laquelle les facteurs SerRS, PSTK, SPS2, SepSecS et ARNt^{Sec} sont impliqués et une phase d'incorporation dans la chaîne polypeptidique mettant principalement en jeu les facteurs SECIS, SBP2, EFSec, L30, et Sec-ARNt^{Sec}. La plupart des facteurs cités ci haut ont été plus ou moins caractérisés moléculairement par certains groupes. Par contre, pour la plupart d'entre eux des études structurales font encore défaut à l'exception des protéines SepSecS, PSTK et le domaine IV d'EFSec. Les protéines SepSecS et PSTK ont été cristallisées en complexe avec l'ARNt^{Sec} (Palioura et al 2009 ; Shiho Chiba et al 2010). Les protéines SBP2 et L30 ont notamment été étudiées extensivement et le mécanisme et la spécificité de leurs interactions respectives avec l'élément SECIS ont été élucidés (Abby L Bifano et al 2013). Par ailleurs, des essais d'incorporation *in vitro* et *in vivo* ont établi l'interaction SBP2 /SECIS comme étant le facteur limitant de l'incorporation de Sec dans les sélénoprotéines.

Toutefois, la protéine SecP43 découverte en 1999 semble elle également occuper une place primordiale dans la synthèse des sélénoprotéines. En effet, les travaux du groupe du Dr. Hatfield ont montré que la suppression de SecP43 dans les cellules NITH3T3 et TCMK-1 entraînait la réduction de la méthylation du groupe 2'-hydroxylribosyl en position (UM34) de Sec-ARNt^{Sec} et donc de l'expression de la glutathion peroxydase 1 (GPX1). Selon la même étude, SecP43 interagit avec SepSecS en présence de Sec-ARNt^{Sec} et l'absence d'une des protéines affaiblit la liaison de l'autre à Sec-ARNt^{Sec}. Finalement, SecP43 principalement retrouvé dans le noyau semble promouvoir la translocation de SepSecS entre divers compartiments cellulaires (Xue-Ming Xu et al 2005).

Malgré cette avancée majeure sur les fonctions de SecP43, de nombreuses zones d'ombres demeurent non éclairées notamment sur son mécanisme d'action moléculaire, mais aussi sur la phase principale dans laquelle elle rentre en jeu. L'un de nos buts était situer SecP43 dans l'une des phases clé du mécanisme de biosynthèse de Sec qui inclue la phase de synthèse de Sec suivie de celle d'incorporation dans la chaîne polypeptidique

en élongation. Pour ce faire nous avons utilisé des approches *in vitro* dans le but de confirmer ou d'infirmer les hypothèses et les observations découlant des études *in vivo* présentées dans la littérature. Tout naturellement, notre démarche s'est élaborée suivant le schéma qui suit : surexpression et purification de SecP43, tentative de caractérisation structurale et essais d'interactions avec SepSecS, et l'ARNt^{Sec}.

4.2. Surexpression et purification de SecP43

Nous avons initialement tenté d'exprimer SecP43 dans les cellules HEK2936E étant entendu qu'il s'agit de l'homologue humain de SecP43. L'expression à petite échelle bien qu'ayant fonctionné n'a pas été transférée à des plus grands volumes à cause du risque de contamination bactérienne très élevé en absence des instruments conçus pour contrôler plus efficacement ce type de culture tel que les bioréacteurs. On remarque néanmoins que le signal de 50 kDa visualisée (figure 13) par *western blot* ne reflète pas la taille de la protéine d'intérêt (34 kDa, incluant le *his-tag*). Nous expliquons difficilement cet écart car nous pensions à première vue qu'il s'agissait d'une modification post-traductionnelle qui augmenterait de ce fait le poids moléculaire de la protéine. Cependant, aucune modification n'est prédite par les logiciels de la plateforme *Expasy* (www.expasy.org/).

L'analyse par spectrométrie de masse aurait pu trancher cette problématique avec plus de fidélité, mais nous ne disposons pas d'une quantité suffisante d'échantillon pour effectuer cette analyse. C'est donc la raison pour laquelle la surexpression de SecP43 fut concrétisée et le rendement amélioré dans les cellules d'*E. coli* BL21(DE3) en utilisant la procédure d'auto-induction.

Contrairement à l'induction classique à l'IPTG, la technique d'auto-induction favorise, une expression lente et douce, induite par le lactose à l'atteinte d'une densité cellulaire élevée. Ceci mène souvent à un meilleur rendement qui s'explique par le fait que la protéine est exprimée à la transition entre la phase logarithmique et celle de stagnation. Ainsi, sa production n'affecte pas la croissance bactérienne et les mécanismes de dégradation sont moins actifs en cas d'une protéine toxique. Finalement, une expression lente favorise le repliement tout en limitant l'agrégation dans les corps d'inclusion, processus qui s'est avéré récurrent dans le cas de l'induction à l'IPTG. SecP43 fut

surexprimée et purifiée sous sa forme soluble avec un rendement d'environ 20 mg/litres de culture. Il s'agit d'une avancée majeure dans le domaine car c'est bien la première fois que cette protéine fut purifiée avec succès.

Cependant, comme on peut le constater, SecP43 a été assujettie à la protéolyse intra et extra cellulaire (figure 17, 18 et 19). En effet, la région C-terminale très peu structurée selon les prédictions du logiciel RONN (figure 24) explique cette observation. La flexibilité de ce segment rendrait SecP43 susceptible à l'action de diverses protéases. Cette protéolyse persistait malgré l'ajout, d'un cocktail d'inhibiteurs de protéase. Pourrait-on à la lumière de cette observation suggérer que SecP43 subit une auto-protéolyse ? Quoiqu'il en soit, nous avons tenu compte de la protéolyse et des probabilités de cristallisation calculés par le logiciel Xtalpred pour insérer un codon stop en aval du 180ème acide aminé pour exprimer une protéine tronquée.

4.3. Caractérisation structurelle de SecP43

Les essais de cristallisation de SecP43 n'étaient pas productifs en dépit des différentes conditions et tampons de solubilisations utilisées. Plusieurs facteurs pourraient expliquer ce résultat. Premièrement, SecP43 semble adopter plusieurs conformations en solution en raison de sa flexibilité. Ceci se traduit par une variation du rayon de gyration d'une mesure à l'autre d'une part et de la variation du Dmax d'autre part (figure 29 et 30). Deuxièmement, l'analyse de la diffraction en solution de SecP43 (SAXS), montre clairement que même en absence, de la région C-terminale, le repliement de SecP43 demeure partiel quelque soit le tampon utilisé (*kratky plot normalisé* et fonctions de distribution des distances $p(r)$) (figure 30 et 32). Ainsi, SecP43 adopterait une multitude de conformations en solution. Ce dynamisme présumé est incompatible avec la formation des cristaux étant entendu que celle-ci requiert l'homogénéité de l'échantillon (état multimérique et conformations). Dans le même ordre d'idées, nous pouvons à partir des enveloppes calculées par DAMMIF conjecturer sur la forme et le comportement de SecP43 en solution. La partie N-terminale de SecP43 est constituée de deux domaines relativement bien repliés qui appartiennent selon l'analyse de leur séquence à la famille

des protéines de reconnaissance de l'ARN. Ces deux motifs (RRM1 et RRM2) seraient reliés par une région coudée et hautement flexible constituée des acides aminés **GKQPDNSPEY** (prédiction par RONN). Remarquons que la présence d'une Glycine et de deux Prolines dans cette séquence d'acides aminés explique assez bien la flexibilité de cette région. Contrairement aux autres acides aminés, les valeurs des angles dièdres ψ et ϕ de la proline ne sont pas circonscrites par les limites du diagramme de Ramachandran (Ramachandran et al, 1963). Ainsi, bien qu'étant structurés, les deux motifs RRM de SecP43 oscilleraient l'un par rapport à l'autre autour de cette région potentiellement coudée. C'est donc probablement pour cette raison que les enveloppes calculées (DAMMIF) ont toutes une forme non globulaire (figure 34). Les deux motifs RRM s'insèrent aisément des les enveloppes correspondant aux échantillons SecP43-Tris-fusion et SecP43-Phosphate-0,55 mg/mL. Cependant la co-superposition des deux motifs RRM dans chacune de ses enveloppes paraît impossible compte tenu des limitations du logiciel SUPCOMB. Toutefois, nous sommes en droit de suggérer que ces enveloppes puissent accommoder les deux motifs RRM simultanément car ces derniers occupent des positions distinctes lorsque traité individuellement.

4.4. Interaction SecP43/SepSecS/ARNt^{Sec}

Notre principale interrogation dans la deuxième partie du mémoire était d'analyser les interactions entre SecP43 et ARNt^{Sec} non chargé. La littérature fait état d'une interaction entre ces deux composantes seulement en présence d'EFsec ou de SepSecS. Précisons que le ARNt^{Sec} doit dans ce cas être au préalable chargé par la Sélénocystéine (Sec) qui y est attaché par une liaison covalente au niveau de l'extrémité CCA de l'ARNt (Xu XM. et al 2005 ; Small-Howard, A et al 2006).

L'interaction SecP43/ARNt^{Sec} fut dans un premier temps évaluée par chromatographie d'exclusion stérique. On constate que le déplacement du pic d'éluion observé (figure 35) n'illustre pas la formation d'un complexe stable étant donné qu'un tel complexe aurait une taille d'environ 50 kDa et donc un volume d'éluion inférieur à 10,5 mL selon la courbe d'étalonnage. Cependant, il est important de considérer que le volume d'éluion dépend également de la forme de l'analyte. Par conséquent une conclusion hâtive sur la stabilité de l'interaction pourrait être biaisée d'autant plus que l'effet de la diffusion est

important dans le cas des interactions faibles. Alternativement, nous avons fait usage de la technique de retard sur gel qui permet de contrôler la dilution des échantillons lors de la préparation. Nous constatons que l'interaction SecP43/ARNt^{Sec} est très faible car le retard de l'ARNt^{Sec} n'est perceptible qu'à des concentrations élevées de SecP43 (ratio protéine/ARNt^{Sec} d'environ 100). SepSecS semble avoir plus d'affinité pour l'ARNt^{Sec} quoique ce dernier soit non chargé.

Cette observation est en adéquation avec les études menées par d'autres groupes qui stipulaient qu'une interaction directe entre SecP43 et ARNt^{Sec} n'était pas démontrable *in vitro* (Ding, F. et al 1999). Notons toutefois que ces études ont été réalisées avec l'ARNt^{Sec} enzymatiquement chargé isolé des réticulocytes de lapin. Il se trouve donc que cette observation est confirmée par notre approche utilisant l'ARNt^{Sec} produit par la polymérase T7 et non chargé. L'absence d'interaction peut également s'expliquer par le fait que la partie C-terminale de SecP43 est absente. En effet, les deux motifs RRM ne seraient certainement pas suffisants pour promouvoir une interaction forte. Nous pouvons donc sur la base de ce résultat émettre l'hypothèse selon laquelle SecP43 ne serait impliqué dans la phase de synthèse de Sec qu'après le chargement de l'ARNt^{Sec} et l'association d'autres facteurs.

Afin d'analyser l'effet de SecP43 sur le complexe SepSecS/ARNt^{Sec}, nous l'avons ajouté au mélange SepSecS/ARNt^{Sec} et nous nous attendions à l'apparition d'un complexe de plus haut poids moléculaire. Contre toute attente, nous avons remarqué que l'ajout de SecP43 semble réduire l'affinité de SepSecS pour l'ARNt^{Sec} non chargé. Ceci mène à l'hypothèse que l'interaction SecP43/SepSecS qui n'a pas pu être démontrée par chromatographie ou par essai de réticulation, ne se produit qu'en présence d'ARNt^{Sec}. En d'autres termes, SecP43 provoquerait un changement de conformation de SepSecS qui lui à son tour perdrait son affinité avec l'ARNt^{Sec}.

4.5. Fonction de SecP43

À la lumière des nos résultats, nous pouvons affirmer en conformité avec les travaux décrits par d'autres groupes que l'interaction de SecP43 avec l'ARNt^{Sec} est dépendante d'autres facteurs tels que SepSecS, EFsec, SBP2. Tout comme SBP2 qui comporte 70% de régions désordonnées, et qui possède plusieurs partenaires d'interactions comme (Roman M et al 2013), SecP43 une protéine intrinsèquement désordonnée et donc joue probablement un rôle de protéine d'échafaudage. Doté d'un signal de localisation nucléaire et d'un signal d'export nucléaire, SecP43 serait aussi bien impliqué dans la phase de synthèse cytoplasmique de Sec que dans la phase de l'incorporation (cytoplasmique et nucléaire) de ce dernier dans la chaîne polypeptidique (Small-Howard, A et al 2006).

La problématique qui reste à résoudre est de savoir comment une protéine aussi flexible peut avoir une activité enzymatique comme le suggèrent des observations faites sur l'impact de l'inhibition de SecP43 sur la méthylation du groupement 2'-OH de l'uridine 5-methoxycarbonylmethyl situé sur la base *wobble* du Sec-ARNt^{Sec}? Le lien de causalité entre l'inhibition de SecP43 et la réduction de cette méthylation serait-il direct ou indirect? Pour répondre objectivement à cette question, nous avons essayé d'aligner les séquences de quelques méthyltransférases humaines (*Accession numbers* : P21964.2, P22061.4 et O43463.1) avec celle de SecP43. Aucune similarité n'a été trouvée. Le pourcentage de similitude de séquence le plus élevé n'était que de 14% ce qui est insignifiant. Par ailleurs, d'un point de vue structurel, les motifs structuraux caractéristiques de la classe I des méthyltransférases ne sont pas retrouvés chez SecP43. Il s'agit notamment d'un motif structurel commun qui inclue un feuillet distorsionné composé de sept brins β (Djordjevic S et al 1997).

Au regard de cette disparité entre SecP43 et la grande famille des méthyltransférases, et compte tenu de sa très forte flexibilité, nous sommes en droit de supposer que SecP43 ne catalyse pas directement la méthylation de l'ARNt^{Sec}. Elle pourrait faciliter la mise en place d'un complexe ou d'une enzyme non connue pour l'instant qui se chargerait de la méthylation de l'ARNt^{Sec}. Notons à juste titre que l'inhibition de SecP43 réduit la méthylation sans pour autant l'abolir complètement (figure 11) (Xu XM. Et al, 2005).

Notons néanmoins qu'il est aussi possible d'envisager un scénario dans lequel SecP43 acquerrait une activité enzymatique par dimérisation démontré *in vitro* par spectrométrie de masse et par SAXS et *in vivo* par la technique du PCA (*Protein-fragment Complementation Assay*) développée par le Dr. Michnick du Département de biochimie et médecine moléculaire de l'Université de Montréal.

Saisissons cette parenthèse pour faire un bref détour sur les conclusions tirées de l'étude PCA réalisée par Étienne Pageau-Crevier dans le contexte de ses travaux au laboratoire du Dr. Baron. Après transfection des cellules HEK293T par des plasmides contenant les différents gènes codant les protéines du sélénosome, il a été démontré qu'en absence de l'ARNtSec, SecP43 interagit faiblement avec SBP2 tel qu'établi par coimmunoprécipitation (Small-Howard et al, 2006). En outre SecP43 en plus de se dimériser, interagit faiblement avec SPS1. C'est précisément via cette interaction, que l'interaction avec SepSecS serait possible ce qui consolide les modèles d'interactome proposés jusqu'à présent.

4.6. Conclusion et perspective

Bien que les facteurs ribonucléoprotéiques impliqués dans la machinerie de synthèse et d'insertion de la sélénocystéine soient pour la plupart cernés, le mécanisme d'action de certaines protéines comme SBP2, L30, SecP43 et bien plus récemment de la nucléoline n'a pas encore été élucidé faute d'études structurales. SecP43 apparaît comme l'un des éléments clé du sélénosome en raison de ses interactions multiples avec les autres facteurs de cette machinerie. Nous avons notamment vu qu'il est difficile de démontrer l'interaction entre SecP43, ARNt^{Sec} et SepSecS *in vitro*. SecP43 semble être une protéine d'échafaudage qui mérite d'être étudiée extensivement. C'est la raison pour laquelle nos travaux, bien incomplets, se sont articulés autour de cette protéine. Nous avons certes réussi à la purifier, mais sa caractérisation structurale par SAXS demeure partielle étant entendu que le SAXS est une technique complémentaire à la cristallographie et à la RMN. Par ailleurs, la stabilisation de la protéine complète SecP43 demeure un défi puisque sa région C-terminale est très sensible aux protéases. Ainsi les observations faites sur ses interactions avec SepSecS et ARNt^{Sec} peuvent être biaisés en raison de l'absence de cette région.

Le complexe ribonucléoprotéique qui régit la synthèse et l'insertion de Sec n'est sûrement pas un complexe statique. Sa formation procéderait d'une séquence d'étapes bien précises qu'il convient de cerner. Plusieurs modèles ont été proposés en se référant à l'intensité des interactions protéine-protéine au sein de ce complexe. Nous avons tentés dans le même d'ordre d'idées de quantifier les interactions SecP43/ARNt^{Sec} et SepSecS/ARNt^{Sec} par ITC (*isothermal titration calorimetry*) en vain. Il faudrait donc dans le futur optimiser cette réaction afin de quantifier ces interactions. En outre, il serait intéressant d'étudier, les changements de conformation de SecP43 dus à une titration avec l'ARNt^{Sec} par SAXS. Ceci nous permettra de voir si l'ARNt^{Sec} stimule le repliement de SecP43 en se référant aux courbes de *kratky*. Alternativement, la mise au point d'un milieu d'auto-induction radio-marquée permettrait de mieux cerner la dynamique de SecP43 par RMN. En outre, la présence ou non d'une activité de méthyltransférase demeure difficilement démontrable dans la mesure où un test enzymatique nécessite une interaction enzyme/substrat forte ce qui ne semble pas être le cas pour SecP43 et

l'ARNt^{Sec} non chargé. Pour pallier à ce problème les études d'interaction futures devront se baser sur l'ARNt^{Sec} isolé du lapin car un tel ARNt^{Sec} contiendrait assurément les modifications nécessaires pour sa liaison avec SecP43 et bien d'autres facteurs. Finalement, il serait intéressant d'optimiser l'expression de SBP2 et d'EFsec afin de voir si un complexe stable peut être formé avec SecP43, SepSecS et ARNt^{Sec} *in vitro*. L'analyse d'un tel complexe permettrait d'élucider les acides aminés impliqués dans l'interaction entre ces facteurs.

Références

Atkins J. F. and Gesteland R. F. (2000) « The twenty-first amino acid », *Nature* **407**: 463-465.

Anton A. Turanov, Alexei V. Lobanov, Dolph L. Hatfield and Vadim N. Gladyshev. (2013) « UGA codon position-dependent incorporation of selenocysteine into mammalian selenoproteins », *Nucleic Acids Res* **41**(14): 6952–6959.

Azevedo M. F., Barra G. B., Naves L. A., Ribeiro Velasco L. F., Godoy Garcia Castro P., de Castro L. C., Amato A. A., Miniard A., Driscoll D. and Schomburg L. (2010) « Selenoprotein-related disease in a young girl caused by nonsense mutations in the SBP2 gene », *JCEM* **95**(8): 4066–4071.

Alexey V. Lobanov, Dolph L. Hatfield, Vadim N. and Gladyshev. (2009) « Eukaryotic selenoproteins and selenoproteomes », *Biochim Biophys Acta* **1790**: 1424–1428.

Arner E. S. (2009) « Focus on mammalian thioredoxin reductases-Important selenoproteins with versatile functions », *Biochim Biophys Acta* **1790**(6): 495-526

Allmang C., Wurth L. and Krol A. (2009) « The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated », *Biochim Biophys Acta* **1790**(11):1415-23

Allmang C. and Krol A. (2006a) « SECIS RNAs and K-turn binding proteins. A survey of evolutionary conserved RNA and protein motifs. Selenium, its molecular Biology and role in human Health 2nd edition DL Hatfield », *Kluwer Academic Publishers* **5**: 51-61.

Allmang C. and Krol A. (2006b) « Selenoprotein synthesis: UGA does not end the story », *Biochimie* **88**(11):1561-1571.

Amberg R., Urban C., Reuner B., Scharff P., Pomerantz SC., McCloskey J.A., and Gross H. J., (1993) « Editing does not exist for mammalian selenocysteine tRNAs », *Nucleic Acids Res* **21**(24): 5583-5588.

Berry M. J., Martin G. W., and Low S. C., (1997) « RNA and protein requirements for eukaryotic selenoprotein synthesis », *Biomed Environ Sci* **10**(2-3):182-9.

Besse D. and Moroder L. (1997) « Synthesis of selenocysteine peptides and their oxidation to diselenide-bridged compounds », *J Pept Sci* **3**: 442-453.

Biteau B., Labarre J., and Toledano M. B., (2003) « ATP-dependent reduction of cysteinesulphinic acid by *S. cerevisiae* sulphiredoxin », *Nature* **425**: 980-984.

Sutherland B. W., Toews J., and Kast J. (2008) « Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions », *JMS* **43**(6): 699–715.

Baron C., Westhof E., Böck A. and Giegé R. (1993) « Solution structure of selenocysteine-inserting tRNA(Sec) from *Escherichia coli*. Comparison with canonical tRNA(Ser) », *J Mol Biol* **231**(2): 274-92.

Baron C., Sturchler C., Wu X. Q., Gross H. J., Krol A. and Bock A. (1994) « Eukaryotic selenocysteine inserting tRNA species support selenoprotein synthesis in *Escherichia coli* », *Nucleic Acids Res* **22**(12): 2228-2233.

Bosl M. R., Takaku K., Oshima M., Nishimura S. and Taketo M. M. (1997) «Early embryonic lethality caused by targeted disruption of the mouse selenocysteine tRNA gene (Trsp) », *Proc Natl Acad Sci USA* **94**(11): 5531-5534.

Böck A. and Hatfield D. L. (2006) « A Selenium metabolism in prokaryotes. Selenium, its molecular biology and role in human health », *Second edition Springer Science + Business Media*: 9-28.

Baron C. and Böck A. (1991) « The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNA(Sec) of *Escherichia coli* is the determinant for binding to elongation factors SELB or Tu », *J Biol Chem* **266**(30): 20375-20379.

Bréchet V. R. and Dominique D. (2012) « How Random are intrinsically Disordered Proteins? A Small Angle Scattering Perspective », *Current Protein and Peptide Science* (13): 55-75 55.

Bellinger F. P., Arjun V. R., Maricclair A. R. and Berry M. J. (2009) « Regulation and function of selenoproteins in human disease », *Biochem J* **422**: 11–22.

Cardey, B., and Enescu, M. (2007) « Selenocysteine versus cysteine reactivity: a theoretical study of their oxidation by hydrogen peroxide », *J Phys Chem A* **111**: 673-638.

Carlson B. A., Xu X. M., Gladyshev V. N. and Hatfield D.L. (2005) « Selective rescue of selenoprotein expression in mice lacking a highly specialized methyl group in selenocysteine tRNA », *J Biol Chem* **280**(7): 5542-5548.

Carlson B. A., Xu X. M., Kryukov G. V., Rao M., Berry M. J., Gladyshev V. N. and Hatfield D. L. (2004) « Identification and characterization of phosphoseryl-tRNA[Ser]Sec kinase », *PNAS* **101**(35): 12848-12853.

Chapple C. E. and Guigo R. (2008) « Relaxation of selective constraints causes independent selenoprotein extinction in insect genomes », *PLoS ONE* **3**(8): e2968

Chapple C. E., Guigo R. and Krol A. (2009) « SECISaln, a web-based tool for the creation of structure-based alignments of eukaryotic SECIS elements », *Bioinformatics* **25**(5): 674-675.

Costa M., Rodríguez-Sánchez J., Czaja A. and Gelpí C. (2000) « Isolation and characterization of cDNA encoding the antigenic protein of the human tRNA(Ser)Sec complex recognized by autoantibodies from patients with type-1 autoimmune hepatitis », *Clin Exp Immunol* **121**(2): 364-374.

Chavatte L., Brown B. A., and Driscoll D. M. (2005) « Ribosomal protein L30 is a component of the UGAselenocysteine recoding machinery in eukaryotes », *Nat Struct Mol Biol* **12**(5): 408-16.

Copeland P. R., and Driscoll D. M. (1999) « Purification, redox sensitivity, and RNA binding properties of SECIS- binding protein 2, a protein involved in selenoprotein biosynthesis », *J Biol Chem* **274**(36): 25447-25454.

Copeland P. R., Fletcher J. E., Carlson B. A., Hatfield D. L., and Driscoll D. M. (2000) « A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs », *EMBO J* **19**(2):306-314.

Copeland P. R., Stepanik V. A., and Driscoll D. M. (2001) « Insight into mammalian selenocysteine insertion: domain structure and ribosome binding properties of Sec insertion sequence binding protein 2 », *Mol Cell Biol* **21**(5): 1491-1498.

Dikiy A., Novoselov S.V., Fomenko D. E., Sengupta A., Carlson B. A., Cerny R. L., Ginalski K., Grishin N.V., Hatfield D. L. and Gladyshev V. N. (2007) « SelT, SelW, SelH, and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family », *Biochemistry* **46**: 6871–6882.

Diamond A. M., Choi I. S., Crain P. F., Hashizume T., Pomerantz S. C., Cruz R., Steer C. J., Hill K. E., Burk R. F., McCloskey J. A., et al. (1993) « Dietary selenium affects methylation of the wobble nucleoside in the anticodon of selenocysteine tRNA([Ser]Sec) », *J Biol Chem* **268**(19): 14215-14223.

Ding F. and Grabowski P. J. (1999) « Identification of a protein component of a mammalian tRNA (Sec) complex implicated in the decoding of UGA as selenocysteine », *RNA* **5**(12): 1561-1569.

Donovan J., Caban K., Ranaweera R., Gonzalez-Flores J. N. and Copeland P. R. (2008) « A novel protein domain induces high affinity selenocysteine insertion sequence binding and elongation factor recruitment», *J Biol Chem* **283**(50): 35129-35139.

De Jesus L. A., Hoffmann P. R., Michaud T., Forry E. P., Small-Howard A., Stillwell R. J., Morozova N., Harney J. W., Berry M. J. (2006) « Nuclear Assembly of UGA

Decoding Complexes on Selenoprotein mRNAs: a Mechanism for Eluding Nonsense-Mediated Decay? », *Mol Cell Biol* **26**(5): 1795-1805.

Ding F., and Grabowski P. J. (1999) « Identification of a protein component of a mammalian tRNA(Sec) complex implicated in the decoding of UGA as selenocysteine », *RNA* **5**:1561-1569.

Djordjevic S. and Stock A. M. (1997) « Crystal structure of the chemotaxis receptor methyltransferase CheR suggests a conserved structural motif for binding S-adenosylmethionine », *Structure* **5**:545– 558.

Ferguson A. D., Labunskyy V. M., Fomenko D. E., Arac D., Chelliah Y., Amezcua C. A., Rizo J., Gladyshev V. N. and Deisenhofer J. (2006) « NMR Structures of the Selenoproteins Sep15 and SelM Reveal Redox Activity of a New Thioredoxin-like Family », *J Biol Chem* **281**(6): 3536-3543.

Flohe L. (2009) « The labour pains of biochemical selenology: The history of selenoprotein biosynthesis », *Biochim Biophys Acta* **1790**(11):1389-403.

Fagegaltier D., Lescure A., Walczak R., Carbon P., Krol A. (2000b) « Structural analysis of new local features in SECIS RNA hairpins », *Nucleic Acids Res* **28**(14): 2679-2689.

Studier F. W. and B.A. Moffatt. (1986) « Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes », *J Mol Biol* **189**: 113–130.

Studier F. W. (1991) « Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system », *J Mol Biol* **219**(1): 37-44.

Studier F. W., Rosenberg A. H., Dunn J. J., Dubendorff. (1990) « Use of T7 RNA polymerase to direct expression of cloned genes », *Methods Enzymol.* **185**: 60–89.

Finlayson A. J., MacKenzie S. L., and Finley J. W. (1979) « Reaction of alanine-3-sulfinic acid with 2-mercaptoethanol », *Can J Chem* **57**: 2073-2077.

Gromer S., Urig S., and Becker K. (2004) « The thioredoxin system – from science to clinic », *Med Res Rev* **24**, 40-89.

Gromer S., Eubel J. K., Lee B. L., Jacob J. (2005) « Human selenoproteins at a glance ». *Cell Mol Life Sci* **62**(21): 2414-2437.

Guimaraes M. J., Peterson D., Vicari A., Cocks B. G., Copeland N. G., Gilbert D. J., Jenkins N. A., Ferrick D. A., Kastelein R. A., Bazan J. F. and Zlotnik A. (1996) « Identification of a novel selD homolog from eukaryotes, bacteria, and archaea: is there an autoregulatory mechanism in selenocysteine metabolism? » *Proc Natl Acad Sci USA* **93**(26): 15086-15091.

Grundner-Culemann E., Martin G. W., 3rd, Harney J. W. and Berry M. J. (1999) « Two distinct SECIS structures capable of directing selenocysteine incorporation in eukaryotes », *RNA* **5**(5): 625-635.

Hondal R. J., and Ruggles E. L. (2011) « Differing views of the role of selenium in thioredoxin reductase », *Amino Acids* **41**:73-89.

Herbette S., Roeckel-Drevet P. and Drevet J. R. (2007) « Seleno-independent glutathione peroxidases. More than simple antioxidant scavengers », *FEBS J* **274** (9): 2163-2180.

Hatfield D., Portugal F. H. (1970) « Seryl-tRNA in mammalian tissues: chromatographic differences in brain and liver and a specific response to the codon, UGA », *Proc Natl Acad Sci USA* **67** (3): 1200-1206.

Hondal R. J., Marino S. M., and Gladyshev V. N. (2013) « Selenocysteine in thiol/disulfidexlike exchange reactions », *Antioxid Redox Signal* **18**:1675-1689.

Howard M. T., Aggarwal G., Anderson C. B., Khatri S., Flanigan K. M., Atkins J. F. (2005) « Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons », *EMBO J* **24**(8): 1596-1607.

Ishii T. M., Kotlova N., Tapsoba F., Steinberg S. V. (2013) « The long D-stem of the selenocysteine tRNA provides resilience at the expense of maximal function », *J Biol Chem* **288**(19):13337-44.

Imai T., Tsuda K., Muto Y., Inoue M., Kigawa T., Terada T., Shirouzu M., Yokoyama S. (2006) « Solution structure of the c-terminal RNA recognition motif in tRNA selenocysteine associated protein », *Protein data bank* (PDB).

Jukes T. H. (1990) « Genetic code 1990 Outlook », *Experientia* **46**: 1149-1157.

John Kiernan. (2000) « Formaldehyde Glutaraldehyde and Fixation » *Microscopy Today* **00**(1): 8-12.

Kryukov G. V., Castellano S., Novoselov S. V., Lobanov A. V. , Zehtab O., Guigo R. and Gladyshev V. N. (2003) « Characterization of mammalian selenoproteomes », *Science* **300** (5624): 1439-1443.

Kim L. K., Matsufuji T., Matsufuji S., Carlson B. A., Kim S. S., Hatfield D. L. and Lee B. J. (2000) « Methylation of the ribosyl moiety at position 34 of selenocysteine tRNA[Ser]Sec is governed by both primary and tertiary structure », *RNA* **6**(9): 1306-1315.

Korotkov K. V., Novoselov S. V., Hatfield D. L., Gladyshev V. N. (2002) « Mammalian selenoprotein in which selenocysteine (Sec) incorporation is supported by a new form of Sec insertion sequence element », *Mol Cell Biol* **22**(5): 1402-1411.

Klein D. J., Schmeing T. M., Moore P.B. and Steitz T. A. (2001) « The kink-turn: a new RNA secondary structure motif », *EMBO J* **20**: 4214–4221.

Kinzy S. A., Caban K. and Copeland P. R. (2005) « Characterization of the SECIS binding protein 2 complex required for the co-translational insertion of selenocysteine in mammals », *Nucleic Acids Res* **33**(16):5172-5180.

Lobanov A. V., Fomenko D. E., Zhang Y., Sengupta A., Hatfield D. L., Gladyshev V. N. (2007) « Evolutionary dynamics of eukaryotic selenoproteomes: large selenoproteomes may associate with aquatic life and small with terrestrial life », *Genome Biol* **8**(9): R198

Lee B. J., Worland P. J., Davis J. N., Stadtman T. C. and Hatfield D. L. (1989) « Identification of a selenocysteyltRNA(Ser) in mammalian cells that recognizes the nonsense codon, UGA », *J Biol Chem* **264**(17): 9724-9727.

Longo P. A., Kavran J. M., Kim M. S. and Leahy D. J. (2013) « Transient mammalian cell transfection with polyethylenimine (PEI) », *Methods Enzymol* **529**: 227-40.

Low S. C., Harney J. W. and Berry M. J. (1995) « Cloning and Functional Characterization of Human Selenophosphate Synthetase, an Essential Component of Selenoprotein Synthesis », *J Biol Chem* **270**(37): 21659-21664.

Leibundgut M., Frick C., Thanbichler M., Böck A. and Ban N. (2005) « Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors », *EMBO J* **24**(1): 11-22.

Leinfelder W., Zehelein E., Mandrand-Berthelot, M. A., and Böck A. (1988) « Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine », *Nature* **331**:723-725.

Lu J Holmgren A. (2009) « Selenoproteins », *J Biol Chem* **284**(2): 723-727.

Muttenthaler M. and Alewood P. F. (2008) « Selenopeptide chemistry » *J Pept Sci* **14**(12): 1223-1239.

Maenpaa P. H., and Bernfield M.R. (1970) « A specific hepatic transfer RNA for phosphoserine », *Proc Natl Acad Sci USA* **67**(2): 688-695.

Ma S., Hill K. E., Caprioli R. M. and Burk R. F. (2002) « Mass spectrometric characterization of full-length rat selenoprotein P and three isoforms shortened at the C terminus. Evidence that three UGA codons in the mRNA open reading frame have alternative functions of specifying selenocysteine insertion or translation termination », *J Biol Chem* **277**(15): 12749-12754.

Maiti B., Arbogast S., Allamand V., Moyle M. W., Anderson C. B., Richard P., Guicheney P., Ferreiro A., Flanigan K. M. and Howard M. T. (2008) « A mutation in the SEPNI1 selenocysteine redefinition element (SRE) reduces selenocysteine incorporation and leads to SEPNI1 related myopathy », *Hum Mutat* **30**(3): 411-6.

Morgan J. W., Hettick J. M., Russell D. H. (2005) « Peptide sequencing by MALDI 193-nm photodissociation TOF MS », *Meth Enzymol* **402**: 186–209.

Novoselov S. V., Lobanov A.V., Hua D., Kasaikina M. V., Hatfield D. L. and Gladyshev V. N. (2007) « A highly efficient form of the selenocysteine insertion sequence element in protozoan parasites and its use in mammalian cells », *Proc Natl Acad Sci USA* **104**(19): 7857-7862.

Papp L.V., Lu J., Striebel F., Kennedy D., Holmgren A. and Khanna K. K. (2006) « The redox state of SECIS binding protein 2 controls its localization and selenocysteine incorporation function », *Mol Cell Biol* **26**: 4895–4910.

Papp L. V., Lu J., Holmgren A. and Khanna K. K. (2007) « From selenium to selenoproteins: synthesis, identity, and their role in human health », *Antioxid Redox Signal* **9**(7): 775-806.

Palioura S., Sherrer R. L., Steitz T. A., Söll D. and Simonovic M. (2009) « The human SepSecS-ARNtSec complex reveals the mechanism of selenocysteine formation », *Science* **325**(5938): 321-5

Rederstorff M., Krol A. and Lescure A. (2006) « Understanding the importance of selenium and selenoproteins in muscle function », *Cell Mol Life Sci* **63**(1): 52-59.

Ruggles E. L., Snider, G. W., Hatfield D. L., Berry M. J., Gladyshev V.N. and Hondal R. J. (2012) « Chemical basis for the use of selenocysteine, In Selenium: Its Molecular Biology and Role in Human Health », *Review Selenium* 73-83.

Rachel L. Schmidt and Miljan Simonović. (2012) « Synthesis and decoding of selenocysteine and human health », *Croat Med J* **53**(6): 535-50.

Ramachandran G. N., Sasisekharan V. & Ramakrishnan C. (1963) « Stereochemistry of polypeptide chain configurations », *J Mol Biol* **7**: 95–99.

Rother M., Wilting R., Commans S. and Böck A. (2000) « Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon *Methanococcus jannaschii* », *J Mol Biol* **299**(2): 351-358.

Svergun D. I. and Koch M. H. J. (2003) « Small-angle scattering studies of biological macromolecules in solution », *Rep Prog Phys* **66** (10): 1735–82.

Svergun D. I. and Koch MHJ (2002) « Advances in structure analysis using small-angle scattering in », *Curr Opin Struct Biol* **12** (5): 654–660.

Small-Howard A., Morozova N., Stoytcheva Z., Forry E. P., Mansell J. B., Harney J. W., Carlson B. A., Xu X. M., Hatfield D. L., and Berry M. J. (2006) « Supramolecular complexes mediate selenocysteine incorporation *in vivo* », *Mol Cell Biol* **26**: 2337–2346
4.

Squires J. E., Stoytchev I., Forry E. P. and Berry M. J. (2007) « SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay », *Mol Cell Biol* **27**(22): 7848-7855.

Squires J. E. and Berry M. J. (2008) « Eukaryotic selenoprotein synthesis: mechanistic insight incorporating new factors and new functions for old factors », *IUBMB Life* **60**(4): 232-235.

Squires J. E., Stoytchev I., Forry E. P. and Berry M. J. (2007) « SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay », *Mol Cell Biol* **27**(22): 7848-7855.

Sturchler C., Westhof E., Carbon P. and Krol A. (1993) « Unique secondary and tertiary structural features of the eucaryotic selenocysteine tRNA(Sec) », *Nucleic Acids Res* **21**(5): 1073-1079.

Sturchler C., Lescure A., Keith G., Carbon P., Krol A., Ibba M. and Söll D. (1994) « Base modification pattern at the wobble position of *Xenopus* selenocysteine tRNA(Sec) », *Nucleic Acids Res* **22**(8): 1354-1358.

Sherrer R. L., Ho J. M. and Söll D. (2008) « Divergence of selenocysteine tRNA recognition by archaeal and eukaryotic O-phosphoseryl-ARNtSec kinase », *Nucleic Acids Res* **36**(6):1871-1880.

Shen Q., Wu R., Leonard J. L. and Newburger P. E. (1998) « Identification and molecular cloning of a human selenocysteine insertion sequence-binding protein. A bifunctional role for DNA-binding protein B », *J Biol Chem* **273**(10): 5443-5446.

Tujebajeva R. M., Copeland P. R., Xu X. M., Carlson B. A., Harney J. W., Driscoll D. M., Hatfield D. L. and Berry M. J. (2000) « Decoding apparatus for eukaryotic selenocysteine insertion », *EMBO Rep* **1**(2): 158-163.

Toppo S., Flohe L., Ursini F., Vanin S., and Maiorino M. (2009) « Catalytic mechanisms and specificities of glutathione peroxydases: variations of a basic scheme », *Biochim Biophys Acta* **1790**: 1486-14500.

Wu X. Q. and Gross H. J. (1993) « The long extra arms of human tRNA((Ser)Sec) and tRNA(Ser) function as major identify elements for serylation in an orientation-dependent, but not sequence-specific manner », *Nucleic Acids Res* **21**(24): 5589-5594.

Wu X. Q. and Gross H. J. (1994) « The length and the secondary structure of the D-stem of human selenocysteine tRNA are the major identity determinants for serine phosphorylation », *EMBO J* **13**(1):241-248.

Walczak R., Carbon P. and Krol A. (1998) « An essential non-Watson-Crick base pair motif in 3'UTR to mediate selenoprotein translation », *RNA* **4**(1): 74-84.

Walczak R., Westhof E., Carbon P., Krol A. (1996) « A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs », *RNA* **2**(4): 367-379.

Wu R., Shen Q., Newburger P. E. (2000) « Recognition and binding of the human selenocysteine insertion sequence by nucleolin », *J Cell Biochem* **77**(3): 507-516.

Carlson X. B., Irons R., Mix H., Zhong N., Gladyshev V., Hatfield D. L. (2007) « Selenophosphate synthetase2 is essential for selenoprotein biosynthesis », *Biochem J* **404**(1): 115-120.

Xu X. M., Carlson B. A., Mix H., Zhang Y., Saira K., Glass R. S., Berry M. J., Gladyshev V. N., Hatfield D. L. (2007) « Biosynthesis of Selenocysteine on Its tRNA in Eukaryotes », *PLoS Biology* **5**(1): e4

Xu X. M., Mix H., Carlson B. A., Grabowski P. J., Gladyshev V. N., Berry M. J. and Hatfield D. L. (2005) « Evidence for direct roles of two additional factors, SECp43 and soluble liver antigen, in the selenoprotein synthesis machinery », *J Biol Chem* **280**(50): 41568-75.

Yuan J., Palioura S., Salazar J. C., Su D., O'Donoghue P., Hohn M. J., Cardoso A. M., Whitman W. B., Soll D. (2006) « RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea », *Proc Natl Acad Sci USA* **103**(50): 18923-18927.

Annexe

Tableau VIII : Peptides générés et détectés par le spectromètre de masse avant l'insertion du codon stop.

Peptides générés	Modifications*	Abondance (%)
VYPSCR	Carbamidomethyl (C)-C5	3,63
LNATYQK	Deamidated (NQ)-N2	3,63
FTDELEQK	Aucune	1,81
VYPSCRGGK	Carbamidomethyl (C)-C5	1,81
LSVAIPKASR	Phospho (ST)-S9	3,63
FTDELEQKR	Deamidated (NQ)-Q7	3,63
FKLNATYQK	Deamidated (NQ)-N4, Phospho (Y)-Y5	5,45
INGKPLPGATPAK	Deamidated (NQ)-N2	3,63
RFKLNATYQK	aucune	1,81
AFATMGETVMSVK	Oxidation (M)-M5, Oxidation (M)-M10, Oxidation (M)-M5 Oxidation (M)-M10	7,27
INGKPLPGATPAKR	Deamidated (NQ)-N2	3,63
GYGFVKFTDELEQK	Aucune	1,81
INGKPLPGATPAKRFK	aucune	1,81
CLHKINGKPLPGATPAK	Deamidated (NQ)-N6, Carbamidomethyl (C)-C1, Carbamidomethyl (C)-C1 Deamidated (NQ)-N6	5,45
AFATMGETVMSVKIIR	Oxidation (M)-M5, Oxidation (M)-M10, Oxidation (M)-M5 Oxidation (M)-M10	7,27
ALTECQGAVGLGCKPVR	Carbamidomethyl (C)-C5	1,81
GYGFVKFTDELEQKR	aucune	1,81
CLHKINGKPLPGATPAKR	Carbamidomethyl (C)-C1, Carbamidomethyl (C)-C1 Deamidated (NQ)-N6	3,63
AFATMGETVMSVKIIRNR	Aucune	1,81
LTGIPAGYCFVEFADLATAEK	Carbamidomethyl (C)-C9, Phospho (Y)-Y8,	5,45

AASLWMGDLEPYMDENFISR	Oxidation (M)-M13,Oxidation (M)-M6 Oxidation (M)-M13	3,63
ALTECQGAVGLGCKPVRLSVAIPK	Carbamidomethyl (C)-C13	1,81
NRLTGIPAGYCFVEFADLATAEK	Carbamidomethyl (C)-C11,Deamidated (NQ)-N1 Carbamidomethyl (C)-C11,Phospho (ST)-T4	5,45
LTGIPAGYCFVEFADLATAEKCLHK	Carbamidomethyl (C)-C9 Carbamidomethyl (C)-C22	1,81
IIRNRLTGIPAGYCFVEFADLATAEK	Carbamidomethyl (C)-C14	1,81
QPDNSPEYSLFVGDLPDVEDDGMLEYEFFVK	Oxidation (M)-M23	3,63
LTGIPAGYCFVEFADLATAEKCLHKINGKPLPGATPAKR	Deamidated (NQ)-N27	1,81
LNATYTGKQPDNSPEYSLFVGDLPDVEDDGMLEYEFFVK	Deamidated (NQ)-Q9 Deamidated (NQ)-N12,Oxidation (M)-M31	5,45
FKLNATYTGKQPDNSPEYSLFVGDLPDVEDDGMLEYEFFVK	Oxidation (M)-M33	3,63

(*) Les modifications spécifiques et les acides aminés du fragment modifié sont indiqués. Ex : C5 (Cystéine 5)

Tableaux IX : Peptides générés par la digestion de SecP43 tronquée et détectés par le spectromètre de masse

Peptides générés	Modifications*	Abondance (%)
SVKIIR	aucune	0,61
VYPSCR	Carbamidomethyl (C)-C5	1,23
LSVAIPK	aucune	0,61
LGSKPVR	aucune	0,61
PGATPAKR	aucune	0,61
KGYGFVK	aucune	0,61
ELEQKR	aucune	0,61
GLGSKPVR	aucune	0,61
NYATYGK	aucune	0,61
QGAVGLGSK	aucune	0,61
ADLATAEK	aucune	0,61
LDQTGVSK	aucune	0,61
GETVMSVK	Oxidation (M)-M5	1,23

PLPGATPAK	aucune	0,61
KVYPSKR	Carbamidomethyl (C)- C6	1,23
DENFISR	aucune	0,61
SKGYGFVK	aucune	0,61
LPGATPAKR	aucune	0,61
VGLGSKPVR	aucune	0,61
LNYATYGK	aucune	0,61
LYEFFVK	aucune	0,61
FADLATAEK	aucune	0,61
VYPSCRGGK	Carbamidomethyl (C)- C5	1,23
VRLSVAIPK	aucune	0,61
AVGLGSKPVR	aucune	0,61
VSKGYGFVK	aucune	0,61
PLPGATPAKR	aucune	0,61
FTDELEQK	aucune	0,61
MDENFISR	aucune	0,61
TDELEQKR	aucune	0,61
SVAIPKASRV	aucune	0,61
GKPLPGATPAK	aucune	0,61
GAVGLGSKPVR	aucune	0,61
GVSKGYGFVK	aucune	0,61
LSVAIPKASR	aucune	0,61
VVLDQTGVS	aucune	0,61
KLNYATYGK	aucune	0,61
PVRLSVAIPK	aucune	0,61
TMGETVMSVK	aucune	0,61
EFADLATAEK	aucune	0,61
KPLPGATPAKR	aucune	0,61
KFTDELEQK	aucune	0,61
NGKPLPGATPAK	Deamidated (NQ)-N1	1,23
ATMGETVMSVK	Oxidation (M)- M3 Oxidation (M)-M8	1,23
FTDELEQKR	aucune	0,61
QGAVGLGSKPVR	aucune	0,61
KVVLDQTGVS	aucune	0,61
YMDENFISR	aucune	0,61
GKPLPGATPAKR	aucune	0,61
VEFADLATAEK	aucune	0,61
FKLNYATYGK	aucune	0,61

GKVVLDQTGVS	aucune	0,61
INGKPLPGATPAK	Deamidated (NQ)-N2	1,23
RALTECQGAVGL	Carbamidomethyl (C)-C6	0,61
QGAVGLGSKPVRL	aucune	0,61
GGKVVLDQTGVS	aucune	0,61
KFTDELEQKR	aucune	0,61
FATMGETVMSVK	Oxidation (M)-M9	1,23
NGKPLPGATPAKR	Deamidated (NQ)-N1	1,23
CQGAVGLGSKPVR	Carbamidomethyl (C)-C1	0,61
ALTECQGAVGLGSK	Carbamidomethyl (C)-C5	1,23
FVEFADLATAEK	aucune	0,61
RFKLNATYTGK	aucune	0,61
AFATMGETVMSVK	Oxidation (M)-M5, Oxidation (M)-M10, Oxidation (M)-M5 Oxidation (M)-M10	2,46
DQTGVSKGYGFVK	aucune	0,61
VKFTDELEQKR	aucune	0,61
INGKPLPGATPAKR	Deamidated (NQ)-N2	1,23
ECQGAVGLGSKPVR	Carbamidomethyl (C)-C2	0,61
VDDGMLYEFFVK	Oxidation (M)-M5	0,61
LDQTGVSKGYGFVK	aucune	0,61
CFVEFADLATAEK	Carbamidomethyl (C)-C1	0,61
TECQGAVGLGSKPVR	Carbamidomethyl (C)-C3	1,23
FADLATAEKCLHK	Carbamidomethyl (C)-C10	0,61
GGKVVLDQTGVS	aucune	0,61
LEPYMDENFISR	aucune	0,61
FVKFTDELEQKR	aucune	0,61
KINGKPLPGATPAKR	aucune	0,61
ALTECQGAVGLGSKPV	Carbamidomethyl (C)-C5	0,61
GFVKFTDELEQKR	aucune	0,61
DLEPYMDENFISR	aucune	0,61
LHKINGKPLPGATPAK	aucune	0,61

GYGFVKFTDELEQK	aucune	0,61
LTECQGAVGLGSKPVR	Carbamidomethyl (C)-C4	0,61
HKINGKPLPGATPAKR	Deamidated (NQ)-N4	1,23
ALTECQGAVGLGSKPVR	Carbamidomethyl (C)-C5, Carbamidomethyl (C)-C5 Deamidated (NQ)-Q6, Carbamidomethyl (C)-C5 Phospho (ST)-S13, Carbamidomethyl (C)-C5	2,46
VVLDQGTGVSKGYGFVK	aucune	0,61
VEFADLATAEKCLHK	Carbamidomethyl (C)-C12	0,61
CLHKINGKPLPGATPAK	Carbamidomethyl (C)-C1	1,23
YGFVKFTDELEQKR	aucune	0,61
AFATMGETVMSVKIIR	Oxidation (M)-M10	0,61
LHKINGKPLPGATPAKR	Deamidated (NQ)-N5	1,23
GYGFVKFTDELEQKR	aucune	0,61
RALTECQGAVGLGSKPVR	Carbamidomethyl (C)-C6, Carbamidomethyl (C)-C6 Deamidated (NQ)-Q7	1,85
ALTECQGAVGLGSKPVRL	Carbamidomethyl (C)-C5	0,61
FVEFADLATAEKCLHK	Carbamidomethyl (C)-C13	0,61
PAGYCFVEFADLATAEK	Carbamidomethyl (C)-C5	0,61
CLHKINGKPLPGATPAKR	Deamidated (NQ)-N6, Carbamidomethyl (C)-C1, Carbamidomethyl (C)-C1 Deamidated (NQ)-N6	1,85
GGKVLDQGTGVSKGYGFVK	Deamidated (NQ)-Q8	1,23
ALTECQGAVGLGSKPVRLS	Carbamidomethyl (C)-C5	0,61
VYPSCRGGKVLDQGTGVSK	Carbamidomethyl (C)-C5	1,23
DLTPDVDDGMLYEFFVK	aucune	0,61

RALTECQGAVGLGSKPVRL	Carbamidomethyl (C)-C6	0,61
SKGYGFVKFTDELEQKR	aucune	0,61
RGGKVVLDTQGVSKGYGFVK	aucune	0,61
FKLNYATYGKQPDNSPEYS	aucune	0,61
LTGIPAGYCFVEFADLATAEK	Carbamidomethyl (C)-C9	0,61
EQKRALTECQGAVGLGSKPVR	Carbamidomethyl (C)-C9	0,61
ALTECQGAVGLGSKPVRLSVAIPK	Carbamidomethyl (C)-C5	1,23
ATAEKCLHKINGKPLPGATPAKR	Deamidated (NQ)-N11, Carbamidomethyl (C)-C6 Deamidated (NQ)-N11	1,23
DQTGVSKGYGFVKFTDELEQKR	Deamidated (NQ)-Q2	0,61
RALTECQGAVGLGSKPVRLSVAIPK	Carbamidomethyl (C)-C6	1,23
ADLATAEKCLHKINGKPLPGATPAKR	Deamidated (NQ)-N14, Carbamidomethyl (C)-C9, Carbamidomethyl (C)-C9 Deamidated (NQ)-N14	1,85
ALTECQGAVGLGSKPVRLSVAIPKASRV	Carbamidomethyl (C)-C5	1,23
FTDELEQKRALTECQGAVGLGSKPVR	Carbamidomethyl (C)-C14, Deamidated (NQ)-Q7 Carbamidomethyl (C)-C14, Carbamidomethyl (C)-C14 Deamidated (NQ)-Q15	2,46
VVLDTQGVSKGYGFVKFTDELEQKR	aucune	0,61
FTDELEQKRALTECQGAVGLGSKPVRL	Carbamidomethyl (C)-C14	0,61
RALTECQGAVGLGSKPVRLSVAIPKASRV	Carbamidomethyl (C)-C6	0,61

NRLTGIPAGYCFVEFADLATAEKCLHK	Carbamidomethyl (C)-C11 Carbamidomethyl (C)-C24	0,61
GGKVLDQGTGVSQGYGVKFTDELEQKR	Deamidated (NQ)-Q8	1,23
GYGVKFTDELEQKRALTECQGAVGLGSKPVR	Carbamidomethyl (C)-C20	1,23
YGVKFTDELEQKRALTECQGAVGLGSKPVR	Deamidated (NQ)-Q12 Carbamidomethyl (C)-C19	0,61
LTGIPAGYCFVEFADLATAEKCLHKINGKPLPGATPAKR	Carbamidomethyl (C)-C9 Carbamidomethyl (C)-C22	0,61
NRLTGIPAGYCFVEFADLATAEKCLHKINGKPLPGATPAKR	Carbamidomethyl (C)-C24, Carbamidomethyl (C)-C11 Carbamidomethyl (C)-C24	1,23

(*) Les modifications spécifiques et les acides aminés du fragment modifié sont indiqués. Ex : C5 (Cystéine 5)