



HAL
open science

La free R Méthionine sulfoxyde réductase (fRMsR) de Neisseria meningitidis : Mécanisme, catalyse et spécificité structurale

Marouane Libiad

► **To cite this version:**

Marouane Libiad. La free R Méthionine sulfoxyde réductase (fRMsR) de Neisseria meningitidis : Mécanisme, catalyse et spécificité structurale. Biochimie, Biologie Moléculaire. Université de Lorraine, 2012. Français. NNT : 2012LORR0335 . tel-01749661

HAL Id: tel-01749661

<https://hal.univ-lorraine.fr/tel-01749661v1>

Submitted on 29 Mar 2018

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

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



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par **Marouane LIBIAD**

**La « free » R Méthionine sulfoxyde réductase (fRMsR) de
Neisseria meningitidis : Mécanisme, catalyse
et spécificité structurale**

Le 12 Octobre 2012

Membres du jury :

| | | |
|---------------|--------------------------|------------------------------------------------------------------|
| Rapporteurs : | Corinne LIONNE | Chargé de recherche, CNRS/CPBS, Montpellier |
| | Jean-François COLLET | Professeur, Université catholique de Louvain/WELBIO, Brussels |
| Examineurs : | Sarah SANGLIER-Cianféran | Chargé de recherche, CNRS/ IPHC, Strasbourg |
| | Jean Marie DUPRET | Professeur, Université Paris Diderot-Paris 7, Paris |
| | Pierre LEROY | Professeur, Université de Lorraine, Nancy |
| | Sandrine BOSCHI-MULLER | Professeur, Université de Lorraine, Nancy, Directeur de thèse |
| | Guy BRANLANT | Professeur, Université de Lorraine, Nancy, Co-directeur de thèse |

**Laboratoire ARN-RNP, structure-fonction-maturation, Enzymologie Moléculaire et Structurale
UMR CNRS-UL 7214. Biopôle, Faculté de Médecine, B.P. 184
54506 Vandoeuvre-Les-Nancy - FRANCE**

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par **Marouane LIBIAD**

**La « free » R Méthionine sulfoxyde réductase (fRMsR) de
Neisseria meningitidis : Mécanisme, catalyse
et spécificité structurale**

Le 12 Octobre 2012

Membres du jury :

| | | |
|---------------|--------------------------|------------------------------------------------------------------|
| Rapporteurs : | Corinne LIONNE | Chargé de recherche, CNRS/CPBS, Montpellier |
| | Jean-François COLLET | Professeur, Université catholique de Louvain/WELBIO, Brussels |
| Examineurs : | Sarah SANGLIER-Cianféran | Chargé de recherche, CNRS/ IPHC, Strasbourg |
| | Jean Marie DUPRET | Professeur, Université Paris Diderot-Paris 7, Paris |
| | Pierre LEROY | Professeur, Université de Lorraine, Nancy |
| | Sandrine BOSCHI-MULLER | Professeur, Université de Lorraine, Nancy, Directeur de thèse |
| | Guy BRANLANT | Professeur, Université de Lorraine, Nancy, Co-directeur de thèse |

**Laboratoire ARN-RNP, structure-fonction-maturation, Enzymologie Moléculaire et Structurale
UMR CNRS-UL 7214. Biopôle, Faculté de Médecine, B.P. 184
54506 Vandoeuvre-Les-Nancy - FRANCE**

REMERCIEMENTS

Je tiens dans un premier temps à remercier Madame le Docteur Christiane BRANLANT, directrice de l'UMR 7214 CNRS-UL, pour m'avoir accueilli dans son laboratoire.

Je tiens à remercier sincèrement Madame le Professeur Sandrine BOSCHI-MULLER, qui, en tant que Directeur de thèse, s'est toujours montré disponible tout au long de ce travail. Je lui suis sincèrement reconnaissant pour sa générosité, sa patience, la rigueur scientifique qu'elle m'a transmise, la formation, les discussions scientifiques, les conseils, le suivi et la très forte implication dans ce travail et sans qui ce travail n'aurait pas abouti. Merci encore pour son aide tout au long de ma thèse jusqu'à la rédaction de ce manuscrit.

Je tiens à exprimer également ma gratitude à Monsieur le Professeur Guy BRANLANT, responsable de l'équipe Enzymologie Moléculaire au moment où j'ai rejoint l'équipe et co-encadrant de ce travail de thèse, pour ses enseignements, la formation et la rigueur scientifique qu'il m'a transmise. Je le remercie également pour les discussions scientifiques que nous avons eues au cours de ces quatre années de thèse, ainsi que pour ses idées et conseils.

J'aimerais adresser un remerciement particulier à Monsieur Christophe GAUTHIER et à Monsieur le Docteur Christophe JACOB pour m'avoir co-encadré avec Madame le Professeur Sandrine BOSCHI-MULLER respectivement en stage de Master-1 et de Master-2.

Je souhaite remercier Madame le Docteur Corinne LIONNE et Monsieur le Professeur Jean-François COLLET qui me font l'honneur d'accepter de juger ce travail en tant que rapporteurs. Je remercie également Madame le Docteur Sarah SANGLIER-CIANFERANI et Messieurs le Professeur Pierre LEROY et le Professeur Jean Marie DUPRET d'avoir accepté de juger ce travail.

Mes remerciements s'adressent à Monsieur le Docteur Guillaume Béchade, Madame Stéphanie PETIOT et Madame le Docteur Sarah SANGLIER-CIANFERANI du laboratoire LSMBO à Strasbourg (UMR 7509 CNRS-ULP) pour les nombreuses analyses de spectrométrie de masse. Je vous remercie également pour m'avoir si bien accueilli et initié au monde de la spectrométrie de masse. Je tiens à présenter mes profonds remerciements à Monsieur le Professeur Kai TITTMANN pour m'avoir si bien accueilli dans son laboratoire et pour m'avoir initié à la technique de Quench-flow.

Merci aux docteurs François TALFOURNIER, Sophie RAHUEL-CLERMONT et Hortense MAZON pour les enseignements, les discussions, l'aide et la bonne humeur.

Je remercie également les stagiaires Master-1 que j'ai eu la chance d'encadrer et tout particulièrement Arek et Abbas, avec qui je me suis lié d'amitié.

Je remercie grandement tous les membres du Laboratoire AREMS, anciens et nouveaux, et en particulier Carine, Sabrina et Séverine pour leur précieuse assistance technique. Je n'oublierai pas évidemment de remercier les thésards que j'ai pu côtoyer durant ma thèse, Xavier, Baptiste, Raphael, Samia et Antoine. Je remercie particulièrement Samia pour les pauses café et les discussions partagées.

Un grand merci à mes « potes » de la musculation, Antoine et Thomas, pour les bons moments passés ensemble. Un grand merci à Benjamin et Régis pour les repas partagés et les discussions passionnantes ou pas qu'on a pu avoir.

J'adresse un remerciement très particulier à Monsieur Alexandre KRIZNIK, avec qui je me suis lié d'une très forte amitié, tout d'abord pour sa disponibilité, ses compétences, ses conseils précieux, les discussions scientifiques et autres, et son aide indispensable. Je remercie également sa femme Kahina pour ses délicieux plats. Merci encore à Alex et à toute sa famille pour tous les moments partagés.

Merci à ma chérie de m'avoir supporté, écouté, soutenu et encouragé, et surtout pour sa patience inépuisable et surtout pendant cette dernière année de thèse. Sa présence et les moments passés ensemble m'ont donné la force morale pour arriver au bout de ce travail. Je termine par lui présenter toute mon affection.

Merci à toute ma famille et tout particulièrement à mon papa et à ma maman pour leurs soutien moral, leurs sens de l'écoute, leurs conseils et encouragements, et à ma petite sœur Fatima Zohra, et à mon petit frère Youssef. Je vous aime beaucoup et j'espère que vous serez là pour me soutenir le jour de ma soutenance.

TABLE DES MATIERES

ABREVIATION

INTRODUCTION

| | |
|-----------------------------------------------------------------------------------------------------------|-----------|
| Préambule | 1 |
| PARTIE I : OXYDATION DE LA Met PAR LES ESPECES RÉACTIVES DE L'OXYGÈNE ET SES CONSEQUENCES | 2 |
| I. 1 Oxydation de la Met par les espèces réactives de l'oxygène | 2 |
| I.1.1 Le radical hydroxyle | 3 |
| I.1.2 Le radical superoxyde | 4 |
| I.1.3 Le peroxyde d'hydrogène | 5 |
| I.1.4 Le peroxydinitrite | 6 |
| I. 2 Conséquences de l'oxydation des Met | 7 |
| I.2.1 Modulation de l'activité des protéines : perte en partie ou totale de l'activité de protéines | 9 |
| I.2.2 Modulation de l'activité des protéines : augmentation de l'activité de protéines.. | 9 |
| I.2.3 Dégradation des protéines | 10 |
| I.2.4 Implication dans des pathologies | 10 |
| PARTIE 2 : LES MÉTHIONINE SULFOXYDE RÉDUCTASES | 11 |
| II. 1 Rôles physiologiques des Msr | 11 |
| II.1.1 Rôle dans la résistance au stress oxydant | 11 |
| II.1.2 Rôle dans la réparation et la régulation | 12 |
| II.1.3 Rôle dans la virulence de bactéries pathogènes | 12 |
| II. 2 Propriétés enzymatiques et structurales des Msr | 13 |
| II.2.1 Structure tridimensionnelle des Msr | 13 |
| II.2.2 Mécanisme et catalyse | 16 |
| II. 3 Les sous-classes de Msr | 18 |
| PARTIE 3 : LES fRMsrS | 23 |

| | |
|-----------------------------------------------------------------------------------|-----------|
| III. 1 Les protéines à domaine GAF | 23 |
| III.1.1 Organisation des domaines GAF..... | 23 |
| III.1.2 Les domaines GAF dans les phosphodiesterases | 24 |
| III.1.3 Les domaines GAF dans les adénylate cyclases..... | 25 |
| III.1.4 Les domaines GAF dans le facteur de transcription NifA | 26 |
| III.1.5 Structure tridimensionnelle du domaine GAF | 27 |
| III.1.6 Le mode de fixation des nucléotides cycliques | 29 |
| III.1.7 La signalisation vers le domaine effecteur | 32 |
| | |
| III. 2 Les fRMsr | 35 |
| III.2.1 Distribution dans les trois domaines du vivant et rôle physiologique..... | 35 |
| III.2.2 Mécanisme..... | 36 |
| III.2.3 Structure tridimensionnelle..... | 36 |
| | |
| Objectifs | 38 |

RESULTATS

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| PARTIE I : Caractérisation du mécanisme de la fRMsr de <i>N. meningitidis</i> | 40 |
| I. 1 Stratégie utilisée | 40 |
| I. 2 Production et purification des fRMsr sauvage et mutées de <i>N. meningitidis</i> | 42 |
| I. 3 Dosage des groupements thiols libres avant et après ajout de Met-O, détermination de la stoechiométrie de formation de Met en absence de réducteur et détermination des p<i>K</i> des Cys..... | 43 |
| I. 4 Détermination des paramètres cinétiques des fRMsr sauvage et mutées à l'état stationnaire | 47 |
| I. 5 Caractérisation des intermédiaires réactionnels..... | 52 |
| I.5.1 Etudes du mutant monocystéinique C84A/C94A/C136A | 52 |
| I.5.1.1 Détermination de la masse des formes oxydées du mutant monocystéinique en conditions dénaturantes | 52 |
| I.5.1.2 Détermination de l'état oligomérique du mutant C84A/C94A/C136A sous forme oxydée en solution | 55 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| I.5.1.3 Conclusion concernant les formes oxydées du mutant monocystéinique | 57 |
| I.5.2 Etude de la fRMs _r C118A/C136A | 58 |
| I.6 Conclusion | 60 |
| | |
| PARTIE II : Détermination de la vitesse des différentes étapes du mécanisme | 62 |
| II. 1 Mise en évidence de l'étape cinétiquement limitante..... | 62 |
| II.1.1 Etudes en absence de réducteur | 62 |
| II.1.1.1 Stratégie utilisée | 62 |
| II.1.1.2 Détermination des vitesses de réduction de la Met-O, étape réductase, et de la formation du pont disulfure..... | 64 |
| II.1.1.3 Paramètres cinétiques des deux premières étapes du mécanisme..... | 65 |
| II.1.2 Etude en présence de réducteur : Détermination de la vitesse de réduction du pont disulfure Cys84-Cys118 de la fRMs _r par la Trx réduite | 69 |
| II.1.3 Nature de l'étape cinétiquement limitante du mécanisme global | 71 |
| II. 2 Etude en absence de réducteur des fRMs_r C84A/C136A et C84A/C94A/C136A ... | 72 |
| | |
| PARTIE III : Caractérisation des acides aminés impliqués dans la catalyse de l'étape réductase et dans la spécificité de substrat | 76 |
| III. 1 Stratégie utilisée..... | 77 |
| III. 2 Dosage des groupements thiols libres et détermination de la stœchiométrie de formation de Met en absence de réducteur..... | 78 |
| III. 3 Détermination des paramètres cinétiques à l'état stationnaire..... | 79 |
| III. 4 Détermination des propriétés de fluorescence | 81 |
| III. 5 Détermination des paramètres cinétiques de l'étape réductase | 83 |
| III. 6 Conclusion | 85 |
| | |
| DISCUSSION ET PERSPECTIVES | |
| I. Mécanisme des fRMs_r | 89 |
| I.1 Etape réductase | 90 |
| I.2 Formation du(des) pont(s) disulfure(s) | 92 |
| I.3 Le devenir de l'intermédiaire acide sulfénique en l'absence de la Cys84 | 93 |
| II. Catalyse et spécificités structurales des fRMs_r | 94 |
| II.1 Catalyse de l'étape réductase | 94 |

II.2 Spécificités structurales 97

MATERIELS ET METHODES

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| I. Matériel biologique et milieux de culture | 101 |
| I. 1 Souches bactériennes | 101 |
| I. 2 Vecteurs plasmidiques | 101 |
| I. 3 Milieux de culture bactérienne | 102 |
| I.3.1 Milieu riche Luria Bertani (LB) | 102 |
| I.3.2 Milieu riche SOB | 102 |
| II. Techniques de génie génétique | 102 |
| II. 1 Minipréparation et analyse d'ADN plasmidique | 102 |
| II. 2 Clonage de la fRMSr de <i>N.meningitidis</i> | 102 |
| II. 3 Mutagenèse dirigée du matériel plasmidique | 103 |
| III. Production et purification de la fRMSr sauvage et des mutants | 104 |
| III. 1 Production | 104 |
| III. 2 Purification | 104 |
| III.2.1 Sonication | 104 |
| III.2.2 Filtration sur gel ACA 54 | 104 |
| III.2.3 Chromatographie échangeuse d'anions | 105 |
| III.2.4 Chromatographie par interactions hydrophobes | 105 |
| III.2.5 Concentration et précipitation au sulfate d'ammonium | 105 |
| III.2.6 Gels Bilans de purification des protéines | 106 |
| III.2.7 Détermination de la masse moléculaire par spectrométrie de masse | 106 |
| IV. Etudes enzymatiques de la fRMSr sauvage et des différents mutants | 106 |
| IV. 1 Dosage des groupements thiols | 107 |
| IV.1.1. Dosage des groupements thiols libres par le 5,5'-dithiobis (2-nitro) benzoate (DTNB) | 107 |
| IV.1.2. Titration des résidus Cys par le 2,2'-dipyridyl disulfure (2PDS) | 107 |
| IV.2. Détermination des paramètres cinétiques KM et kcat en présence du système réducteur Trx /Trx réductase/NADPH à l'état stationnaire | 107 |
| IV. 3 Etude en absence de réducteur | 108 |
| IV.3.1 . Détermination de la stoechiométrie de formation de Met..... | 108 |

| | |
|------------------------------------------------------------------------------------------------------|------------|
| IV.3.2. Détermination de la vitesse d'apparition de la Met et de disparition des thiols libres | 108 |
| IV.3.3. Techniques de fluorescence | 109 |
| IV.3.3.1. Spectres de fluorescence des enzymes | 109 |
| IV.3.3.2. Etudes cinétiques | 109 |
| IV.3.4. Caractérisation des formes oxydées par spectrométrie de masse..... | 110 |
| IV.3.5. Détermination de l'état oligomérique | 110 |
| IV. 4 Réduction des ponts disulfure par la Trx | 111 |
| IV.4.1. Approche chromatographique | 111 |
| IV.4.2. Cinétique rapide | 111 |

ANNEXES

| | |
|------------------------------------------|------------|
| Annexes | 113 |
| Annexe 1 | 114 |
| Annexe 2 | 155 |
| REFERENCES BIBLIOGRAPHIQUES | 156 |

ABBREVIATIONS

ABREVIATIONS

| | |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2PDS | 2,2'-dithiodipyridine |
| AC | Adénylate Cyclase |
| ADN | Acide désoxyribonucléique |
| AMPc | Adénosine monophosphate cyclique |
| CaMKII | Protéine kinase II |
| DMSO | Diméthylsulfoxyde |
| DTNB | 5,5'-dithiobis (2-nitro) benzoate |
| DTT | Dithiothréitol |
| ERO | Espèces réactives de l'oxygène |
| EDTA | Acide éthylène diamine tétraacétique |
| ϵ | Coefficient d'extinction molaire |
| FADH ₂ | Flavine adénine dinucléotide (forme réduite) |
| FMNH ₂ | Flavine mononucléotide (forme réduite) |
| FPLC | "Fast protein liquid chromatography" |
| fRMs _r | "free <i>R</i> methionine sulfoxide reductase" |
| GAF | "mammalian cGMP-binding phosphodiesterases, cyanobacterial Adenylyl cyclases, and the <i>Escherichia coli</i> Formate-hydrogen lyase transcription factor" |
| GMPc | Guanosine monophosphate cyclique |
| k_2 | Constante de vitesse de second ordre |
| k_{cat} | Constante catalytique |
| K_M | Constante de Michaelis |
| k_{obs} | Constante de vitesse observée |
| K_S | Constante apparente d'affinité |
| <i>L,D</i> -Met- <i>R,S</i> -O | Mélange des des 4 diastéréoisomères (<i>D</i> -Met- <i>R</i> -O, <i>D</i> -Met- <i>S</i> -O, <i>L</i> -Met- <i>R</i> -O et <i>L</i> -Met- <i>S</i> -O) |
| MES | Acide 2-(<i>N</i> -morpholino)-ethanesulfonique |
| Met-O | Méthionine sulfoxyde |
| Msr | Méthionine sulfoxyde réductase |

Abréviations

| | |
|------------------------|-------------------------------------------------------------|
| NADH | Nicotinamide adénine dinucléotide (forme réduite) |
| NADPH | Nicotinamide adénine dinucléotide phosphate (forme réduite) |
| NO | Monoxyde d'azote |
| PAGE | Gel d'électrophorèse en polyacrylamide |
| PAS | "Per ARNt and Sim" |
| PCR | Réaction de polymérisation en chaîne |
| PDB | Banque de données protéiques |
| PDE | Phosphodiesterase |
| <i>Pfu</i> | ADN polymérase de <i>Pyrococcus furiosus</i> |
| REC | « Receiver domain » ou phosphoaccepteur cheY-like |
| SDS | Sodium dodecyl sulfate |
| SOD | Superoxyde dismutase |
| <i>Taq</i> | ADN polymérase de <i>Thermus aquaticus</i> |
| TFA | Acide trifluoroacétique |
| TNB ⁻ | Thionitrobenzoate |
| Tris | N-tris(hydroxyméthyle) éthyl-2-aminométhanesulfonate |
| Trx | Thiorédoxine |
| TR | Thiorédoxine réductase |
| <i>A. thaliana</i> | <i>Arabidopsis thaliana</i> |
| <i>B. subtilis</i> | <i>Bacillus subtilis</i> |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| <i>M. smegmatis</i> | <i>Mycobacterium smegmatis</i> |
| <i>M. tuberculosis</i> | <i>Mycobacterium tuberculosis</i> |
| <i>M. pulmonis</i> | <i>Mycoplasma pulmonis</i> |
| <i>N. meningitidis</i> | <i>Neisseria meningitidis</i> |
| <i>N. gonorrhoeae</i> | <i>Neisseria gonorrhoeae</i> |
| <i>S. cerevisiae</i> | <i>Saccharomyces cerevisiae</i> |
| <i>S. aureus</i> | <i>Staphylococcus aureus</i> |
| <i>X. campestris</i> | <i>Xanthomonas campestris</i> |

L'appellation des acides amines suit le code à une ou trois lettres.

INTRODUCTION

Préambule.

Les méthionine sulfoxyde réductases (Msr) catalysent la réduction des méthionines oxydées sous forme de méthionine sulfoxyde (Met-O) en Met en présence d'un réducteur qui est généralement la thiorédoxine (Trx) *in vivo*. Jusqu'en 2007, deux classes de Msr étaient connues, la MsrA réduisant spécifiquement l'isomère *S* de la Met-O préférentiellement incluse dans une chaîne polypeptidique et la MsrB spécifique de l'isomère *R* de la Met-O également de préférence incluse dans une chaîne polypeptidique. Récemment, une nouvelle classe de Msr a été identifiée chez *E. coli* (Lin *et al.*, 2007). Cette Msr, appelée fRMsr, est spécifique des Met-*R*-O libres. Elle présente un repliement tridimensionnel caractéristique des domaines GAF, distinct de ceux des MsrA et des MsrB qui eux-mêmes ont des repliements différents.

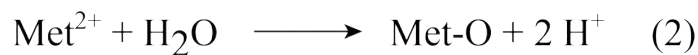
Les domaines GAF représentent une des plus grandes familles de domaines fixant des petites molécules, associés à des fonctions nombreuses, variées, essentielles au niveau cellulaire (Martinez *et al.*, 2002). C'est ainsi que, les protéines à domaine GAF jouent un rôle clé dans la signalisation des nucléotides cycliques. Les domaines GAF constituent des domaines senseurs typiquement retrouvés en tandem avec différents domaines effecteurs, généralement impliqués dans la dimérisation des protéines et dans la régulation allostérique des enzymes (Linder *et al.*, 2007). Cependant, aucune activité catalytique n'avait été décrite pour ces domaines GAF jusqu'à la mise en évidence d'une activité Msr portée par un domaine GAF, les fRMsr (Lin *et al.*, 2007).

Mon projet de thèse concernait d'une part la caractérisation du mécanisme catalytique de la fRMsr de *Neisseria meningitidis*, et d'autre part la caractérisation des facteurs structuraux et moléculaires impliqués dans la catalyse et la reconnaissance du substrat. Les deux premières parties de l'Introduction seront consacrées successivement aux mécanismes d'oxydation des Met et aux conséquences de cette oxydation ainsi qu'aux propriétés biochimiques, enzymatiques et structurales des Msr, alors que la troisième partie concernera les protéines à domaine GAF dont font partie les fRMsr.

PARTIE I: OXYDATION DE LA Met PAR LES ESPECES RÉACTIVES DE L'OXYGÈNE ET SES CONSEQUENCES

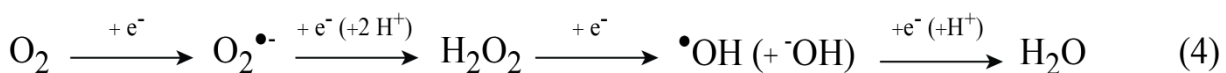
I. 1 Oxydation de la Met par les espèces réactives de l'oxygène

La méthionine (Met) peut être oxydée par deux voies majeures, impliquant une oxydation à un ou deux électrons. L'oxydation à deux électrons entraîne la formation d'un dication (réaction 1) qui en présence d'eau, va se transformer en Met-O (réaction 2), alors que l'oxydation à un électron entraîne la formation d'un radical cation très instable (réaction 3) qui peut se transformer en une série d'intermédiaires et de produits. L'oxydation à un électron de ce radical cation conduit au dication (Schoneich, 2005)



Plusieurs oxydants biologiques sont capables d'oxyder la Met selon des mécanismes à 1 ou 2 électrons. La plupart des études cinétiques s'est focalisée sur l'étude des mécanismes d'oxydation à deux électrons réversibles, qui sont les plus répandus, les mécanismes d'oxydation à un électron pourraient causer quant à eux des dommages oxydatifs irréversibles. Ces réactions d'oxydation dépendent de plusieurs paramètres dont la nature de l'espèce oxydante, le solvant ou l'accessibilité de la Met.

Les oxydants biologiques sont majoritairement les espèces réactives de l'oxygène (ERO), encore appelées espèces activées de l'oxygène, principalement le peroxyde d'hydrogène (H_2O_2), l'anion superoxyde ($\text{O}_2^{\bullet-}$) et le radical hydroxyle (OH^\bullet), qui sont des produits de la réduction de l'oxygène (réactions 4). Ces ERO sont produits en faibles quantités par tous les organismes aérobies, et ils peuvent s'accumuler transitoirement dans des conditions dites de stress oxydant, lorsqu'il existe un déséquilibre de la « balance » entre leurs vitesses de production et d'élimination par les systèmes antioxydants.

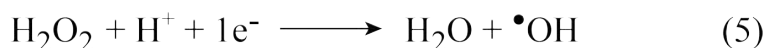


Dans la suite de l'Introduction, seront traités les mécanismes d'oxydation de la Met par le radical hydroxyle, le radical superoxyde, le peroxyde d'hydrogène et le peroxydinitrite, ainsi que l'oxydation catalysée par le Cu (II⁺).

I.1.1 Le radical hydroxyle

Le radical hydroxyle (OH[•]) est le plus réactif des ERO. Sa forte réactivité est limitée *in vivo* par sa faible diffusion à travers les membranes cellulaires. En conséquence sa durée de vie étant extrêmement faible (inférieure à la microseconde), il réagit presque uniquement sur son site de production avec toutes les biomolécules présentes en ce site (ADN, protéines, lipides...).

Il peut être généré dans la cellule par clivage réducteur d'H₂O₂ (réaction 5), par la dissociation du peroxydinitrite (réaction 6), par coupure homolytique d'H₂O₂ (réaction 7), ou par la réaction de Fenton en présence de Fe²⁺ ou de Cu⁺ (réaction 8).



L'oxydation de la Met par OH[•] procède selon un mécanisme d'oxydation à 1 électron, *via* dans un premier temps l'addition de OH[•] sur l'atome du soufre de la Met conduisant à la formation d'un radical hydroxysulfuranyl (réaction 9).



Ce radical hydroxysulfuranyl est très instable et au sein d'un peptide ou d'une protéine. C'est ainsi qu'il peut subir une déshydratation (Schoneich et Bobrowski, 1994). Cette dernière peut engendrer la formation d'intermédiaires radical cation à liaison à trois électrons entre le soufre et un atome d'azote de type S[•]:N (Figure 1, schéma A) ou entre le soufre et un atome d'oxygène de type S[•]:O (Figure 2, schéma B) (la notation « :. » représente une liaison à trois électrons où deux électrons sont localisés dans une orbitale σ -liante et un électron est localisé dans une orbitale σ^* -antiliante). Ces intermédiaires

Introduction

radicalaires vont pouvoir évoluer vers différentes formes (Asmus, 1979; Schöneich *et al.*, 1993 ; Schöneich, 2005).

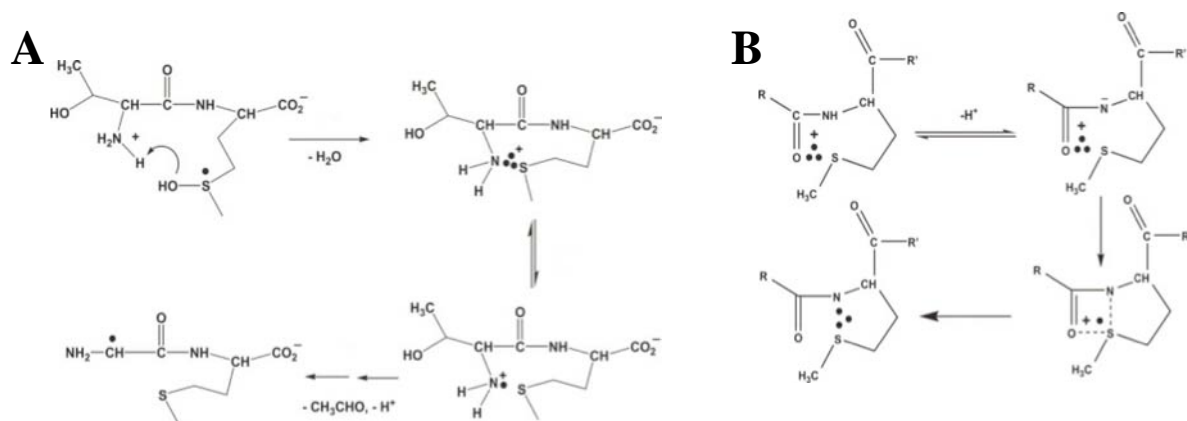
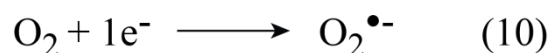


Figure 1 : Transfert de proton intramoléculaire de l'acide aminé N-terminal au niveau du radical hydroxysulfuranyl du dipeptide Thr-Met (Schéma A) et réarrangement de la liaison S.:O à la liaison S.:N (Schéma B). Schéma (A) : le radical hydroxysulfuranyl subit une déshydratation qui conduit à la formation de l'intermédiaire de type S.:N. Un transfert d'électron de l'atome de soufre vers l'atome d'azote permet un clivage homolytique de la liaison Cα-Cβ de la Thr, et la formation de l'acétaldéhyde. Schéma (B) : la déprotonation de l'azote permet un réarrangement des électrons et la formation de l'intermédiaire de type S.:N (Schoneich, 2005).

Le radical hydroxysulfuranyl peut réagir avec l'oxygène pour former un azasulfonium, qui pourra s'hydrolyser pour former un sulfoxyde (Miller *et al.*, 1996, Miller *et al.*, 1998), ou encore réagir directement avec une molécule d'oxygène pour former un intermédiaire radical peroxyde, qui pourra ensuite réagir *via* une oxydation à deux électrons pour former un sulfoxyde (Schoneich and Bobrowski, 1994). La fonction sulfoxyde ne pourra se former qu'en présence d'oxygène et de manière pH-dépendante (Schoneich *et al.*, 1993).

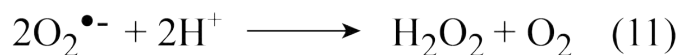
I.1.2 Le radical superoxyde

Le radical superoxyde $O_2^{\bullet-}$, est formé majoritairement par réduction monoélectronique de l'oxygène au niveau de la chaîne respiratoire (réaction 10) (Cadenas and Davies, 2000).



Il peut également être produit par l'auto-oxydation de certains neuromédiateurs (adrénaline, noradrénaline, dopamine...), de thiols (cystéine), de coenzymes réduits ($FMNH_2$, $FADH_2$), la détoxification des xénobiotiques par le système des cytochromes P450 (Halliwell et Gutteridge, 1999 ; Halliwell, 2006), ou encore dans certains cas par la NADH-

déshydrogénase. Sa durée de vie est relativement longue (jusqu'à quelques dizaines de secondes) et de ce fait, peut diffuser bien au-delà de son lieu de production. Sa dismutation conduit à la formation de H_2O_2 qui est catalysée *in vivo* par une superoxyde dismutase (SOD) (réaction 11).

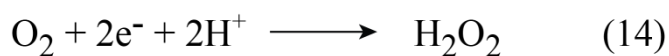


La toxicité des radicaux superoxydes s'exerce plutôt d'une manière indirecte. En effet, en réagissant avec H_2O_2 , les radicaux $\text{O}_2^{\bullet-}$ peuvent donner naissance à OH^{\bullet} (réaction 12, dite réaction de Haber Weiss, catalysée par Fe^{3+}) ou bien encore, en réagissant avec des radicaux NO^{\bullet} , conduire à l'apparition de peroxy-nitrites (réaction 13).



I.1.3 Le peroxyde d'hydrogène

Le peroxyde d'hydrogène H_2O_2 , communément appelé eau oxygénée, est une petite molécule non chargée, dont la production provient principalement de la dismutation d' $\text{O}_2^{\bullet-}$, spontanée ou catalysée par les SOD. Il peut également résulter de la réduction biélectronique de l'oxygène (réaction 14) en présence d'oxydases (aminoacide oxydases, glycolate oxydase, urate oxydase...) qui se trouvent principalement dans des organites cellulaires bien individualisés comme les peroxysomes.



Du fait de sa nature non chargée, H_2O_2 diffuse aisément à travers les membranes des cellules. Il peut également facilement générer OH^{\bullet} en présence de cations métalliques tels que Fe^{2+} (réaction 8, dite réaction de Fenton) ou Cu^{2+} (Wardman and Candeias, 1996).

La réaction d' H_2O_2 sur la Met est un processus d'oxydation à deux électrons résultant d'un transfert d'oxygène (Chu and Trout, 2004, Chu *et al.*, 2004). Initialement, le mécanisme était considéré comme une réaction de type $\text{S}_{\text{N}}2$ (ou substitution nucléophile bimoléculaire en une seule étape) avec une scission hétérolytique de la liaison O-O du peroxyde. Un autre

Introduction

mécanisme a été proposé impliquant la formation d'un intermédiaire molécule d'eau oxydée ($\text{H}_2\text{O}-\text{O}$) (Figure 2). Au cours de ce mécanisme, deux à trois molécules d'eau vont stabiliser les charges de l'état de transition grâce à la formation de liaisons hydrogène avec H_2O_2 , et il y aura un transfert d'un atome d'hydrogène à l'oxygène distal de H_2O_2 .

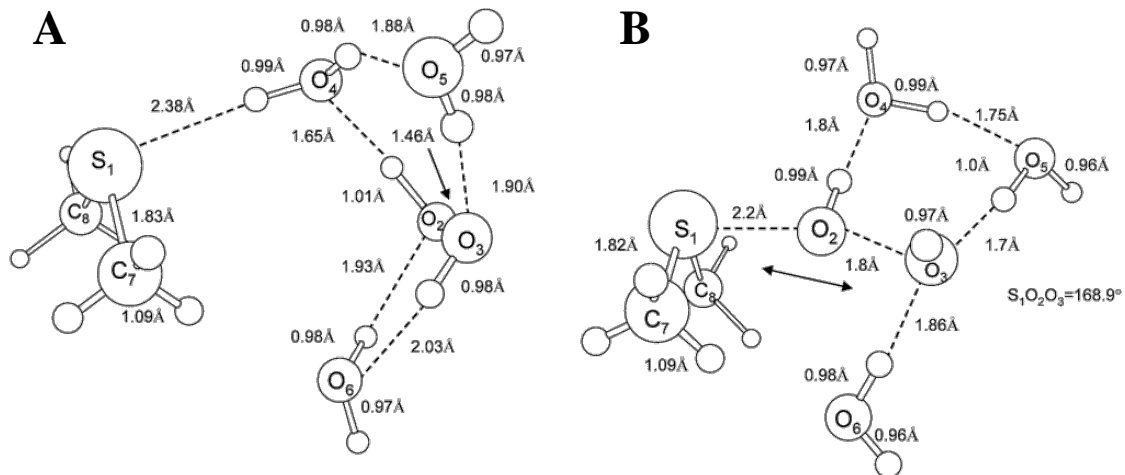


Figure 2 : Représentation du groupement réactif (A) et de l'état de transition (B) de l'oxydation d'une molécule de diméthyl sulfure (DMS) par H_2O_2 (d'après Chu et Trout, 2004). Dans ce modèle, trois molécules d'eau (O_4 , O_5 et O_6) stabilisent la distribution de charge induite par la séparation entre les deux oxygènes de H_2O_2 (O_2 et O_3). C_7 - S_1 - C_8 représente la molécule de DMS. La réaction d'oxydation du DMS par H_2O_2 se produit *via* l'allongement de la liaison O-O de H_2O_2 et la réduction de la distance de la liaison S-O, entraînant la rupture de la liaison O-O et la formation de la liaison S-O. Deux à trois molécules d'eau stabilisent la séparation de charge de l'état de transition par l'intermédiaire d'interactions spécifiques, y compris des liaisons hydrogène

I.1.4 Le peroxydinitrite

Le peroxydinitrite provient de la réaction entre NO^\bullet (produit par les macrophages, les neutrophiles...) et $\text{O}_2^{\bullet-}$ (réaction 13). Il présente une faible durée de vie. Les sites de production du peroxydinitrite sont associés aux sources d' $\text{O}_2^{\bullet-}$, puisque NO^\bullet , contrairement au $\text{O}_2^{\bullet-}$, présente une demi-vie très courte, et ne peut pas diffuser à travers les membranes.

Le peroxydinitrite va oxyder les Met suivant une réaction d'oxydation à deux électrons (Perrin et Koppenol, 2000). Un mécanisme catalytique de cette oxydation a été proposé (Figure 3). La première étape de ce mécanisme consiste en l'attaque nucléophile de l'atome de soufre de la Met sur l'atome d'azote de HOONO entraînant la formation d'un adduit entre ces deux molécules. La perte d'un groupement hydroxyle au niveau de cet adduit ainsi que l'addition d'une molécule d'eau au niveau du soufre de la Met entraîne la formation de

Met-O et de nitrite. Si l'addition de la molécule d'eau a lieu sur l'atome d'azote, il y a alors formation de Met et de nitrate. Le nitrite formé va lui pouvoir réagir avec une autre Met, selon un mécanisme d'oxydation à un électron, entraînant entre autre la formation d'éthylène.

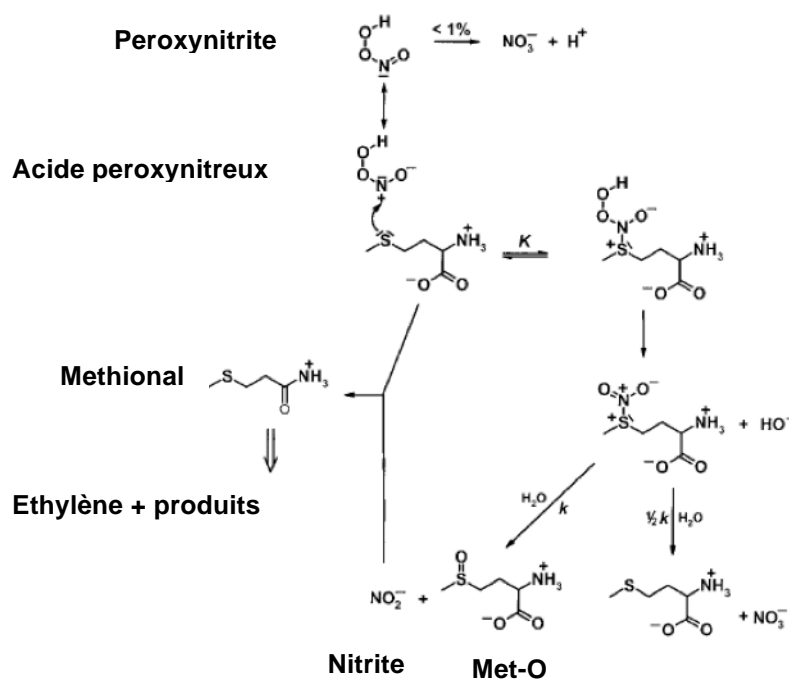


Figure 3 : Mécanisme d'oxydation de la Met par le peroxy-nitrite (d'après Perrin et Koppenol, 2000).

L'attaque nucléophile de l'atome de soufre de la Met sur l'atome d'azote de l'acide peroxy-nitrique conduit à la formation d'un adduit, qui après la perte d'un groupement hydroxyle et l'addition d'une molécule d'eau, libère la Met-O et le nitrite.

I.2 Conséquences de l'oxydation des Met

L'oxydation des Met en Met-O constitue un mécanisme de modification post-traductionnelle non enzymatique des protéines, qui entraîne une modification des propriétés de la chaîne latérale de la Met qui est plutôt hydrophobe et flexible, devient moins hydrophobe et plus polaire dans la Met-O (Gellman, 1991). D'autre part, le soufre de la fonction thioéther forme un centre prochiral dont l'oxydation conduit à la formation des deux stéréoisomères de configuration *R* et *S* (Figure 4). L'ensemble des études réalisées sur l'oxydation de la Met suggère qu'il n'y a pas de préférence pour la formation d'un stéréoisomère par rapport à un autre. Par contre, il existe une enzyme du système de

Introduction

détoxication, la monooxygénase 3 à flavine, qui catalyse l'oxydation stéréosélective des Met et qui est décrite présenter une préférence pour la formation du stéréoisomère *S* de la Met-O (Ripp *et al.*, 1999 ; Duescher *et al.*, 1994).

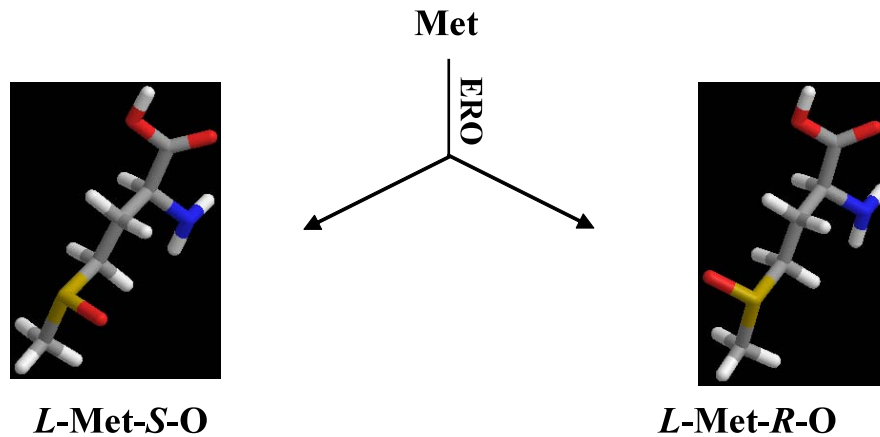


Figure 4 : Représentation des isomères *R* et *S* de la *L*-Met-O.

Les molécules sont représentées en mode bâton. Les atomes de carbone, d'oxygène, d'azote, d'hydrogène et de soufre sont représentés respectivement en gris, rouge, bleu, blanc et jaune. La figure a été réalisée avec le logiciel ChemBiodraw (CambridgeSoft).

En fonction de la localisation des résidus Met oxydés, les conséquences peuvent être différentes. Ainsi les résidus Met accessibles au solvant sont beaucoup plus facilement oxydés que ceux enfouis au sein d'une protéine. L'oxydation des Met en surface ne semble pas avoir un effet marqué sur la conformation d'une protéine (Levine *et al.*, 1999), sauf si ce résidu est situé dans un site actif ou dans une poche de liaison (c'est le cas par exemple de l' α 1-protéinase (Swaim et Pizzo, 1988). Par contre l'oxydation des Met enfouies au sein de régions hydrophobes des protéines peut engendrer des changements conformationnels locaux marqués, pouvant aller dans certains cas jusqu'à la perte de la fonction de certaines protéines. Ainsi, l'oxydation de résidus Met est décrite comme étant impliquée dans l'étape de reconnaissance des protéines non fonctionnelles à dégrader, dans des mécanismes de modulation de la fonction de certaines protéines comme par exemple la calmoduline, et comme un mécanisme de détoxication qui piègerait les ERO. L'oxydation de résidus Met pourrait également jouer un rôle dans certaines voies de signalisation en réponse à un excès d'oxydants dans la cellule (Schoneich, 2005). Des exemples illustrant des conséquences de l'oxydation des Met sont décrits dans les paragraphes suivants.

I.2.1 Modulation de l'activité des protéines : perte en partie ou totale de l'activité de protéines

L'oxydation du résidu Met146 ou Met147 de la calmoduline, résidus situés dans le domaine C-terminal, entraîne un changement conformationnel de la calmoduline et une diminution de 30 % de son affinité pour le calcium (Yao *et al.*, 1996). Ces Met jouent également un rôle important dans la fixation de la calmoduline à ses cibles. C'est le cas des ATPase membranaires dépendantes de la concentration en Ca^{2+} , dont l'activation est diminuée de 50-60 % lorsque les Met144 et Met145 de la calmoduline sont oxydées (Bartlett *et al.*, 2003). De plus, la calmoduline oxydée se fixe sur l'ATPase sur un site distinct du site de liaison de la calmoduline non oxydée, conduisant à la formation d'un complexe non productif (Gao *et al.*, 2001).

L'oxydation des résidus Met présents en positions 412, 432 et 446 de la fibronectine plasmatique, protéine responsable de l'adhésion des cellules à la matrice extracellulaire, entraîne la diminution de l'interaction de celle-ci avec la gélatine. La présence de cette dernière empêche l'oxydation des résidus Met de la fibronectine, ce qui implique leur localisation au sein du site de liaison (Miles et Smith, 1993).

Il a également été montré que l'oxydation des résidus Met de l'actine, présents en position 176, 190 et 269, cause une inhibition complète de sa polymérisation et affecte sévèrement la stabilité des filaments d'actine, qui rapidement dépolymérisent. Ces résultats suggèrent que l'oxydation de certains résidus Met perturberait les interactions non-covalentes spécifiques qui stabilisent normalement la structure des filaments d'actine (Dalle-Donne *et al.*, 2002).

Enfin, l'oxydation de résidus Met conservés de petites protéines du choc thermique, localisés dans une hélice α amphipatique à la surface de la protéine, entraîne des changements conformationnels et une perte complète de son activité chaperone (Härndahl *et al.*, 2001).

I.2.2 Modulation de l'activité des protéines : augmentation de l'activité de protéines

Dans certains cas, l'oxydation de résidus Met au sein des protéines n'est pas délétère mais peut entraîner une augmentation de l'activité biologique des protéines. C'est le cas de l'ubiquitine pour laquelle l'oxydation du résidu Met1 en Met-O entraîne une augmentation de l'activité de dégradation des protéines de 50% (Bamezai *et al.*, 1990). De la même manière, l'activation du composant 5 du complément humain (C5) peut être réalisée non pas

par clivage de la chaîne peptidique mais par oxydation de certains résidus Met en Met-O (Vogt et Hesse, 1994).

I.2.3 Dégradation des protéines

L'oxydation des résidus Met a été proposée être un marqueur de la dégradation des protéines par la machinerie cellulaire. Dans ce cas, la dégradation des protéines est indépendante de l'ATP et elle ne nécessite pas d'ajout d'ubiquitine. C'est par exemple le cas pour la calmoduline oxydée qui est dégradée par le protéasome 20S, avec un taux de dégradation qui dépend du nombre de résidus Met oxydés. La dégradation est corrélée au taux de structures secondaires mais pas à l'hydrophobicité de surface de la protéine modifiée, ce qui suggère que des changements dans les structures secondaires doivent permettre d'exposer des sites spécifiques dans la calmoduline oxydée qui sont protéolysés par le protéasome de manière nonprocessive (Ferrington *et al.*, 2001).

I.2.4 Implication dans des pathologies

Plusieurs études ont reporté que le taux des Met-O dans les protéines augmente au cours du vieillissement et dans certaines maladies (Wells-knecht *et al.*, 1997 ; Stadtman *et al.*, 2005)

Dans la maladie d'Alzheimer, l'oxydation d'une Met en position 35 du peptide β -amyloïde ($A\beta$) ralentit sa fibrillation et favorise son agrégation sous forme de plaques amyloïdes structurées en feuillets- β toxiques pour les cellules corticales (Hou *et al.*, 2002). De même, dans le cas de la maladie de Parkinson, la fibrillation de l' α -synucléine est inhibée suite à l'oxydation de résidus Met, ce qui a pour conséquence de favoriser son agrégation. Ces agrégats d' α -synucléine sont cytotoxiques pour les cellules dopaminergiques d'une zone du cerveau, la substance noire, qui participe au contrôle des mouvements, et conduisent à la destruction de ces cellules (Glaser *et al.*, 2005).

PARTIE 2 : LES MÉTHIONINE SULFOXYDE RÉDUCTASES

Dans cette partie, seules les propriétés des MsrA et MsrB seront décrites. Les MsrA et MsrB sont retrouvées dans presque tous les organismes vivants. Chez les bactéries, il n'existe en général qu'une seule forme de MsrA et de MsrB, alors que chez les mammifères il existe une seule forme de MsrA mais quatre formes de MsrB appelées MsrB1, MsrB2, MsrB3A et MsrB3B. La forme MsrB1 est une sélénoenzyme cytoplasmique, MsrB2 et MsrB3B sont mitochondriales alors que MsrB3A est localisée dans le réticulum endoplasmique (Bar-Noy et Moskovitz, 2002 ; Moskovitz *et al.*, 2002, Kim et Gladishev, 2004a).

II.1. Rôles physiologiques des Msr

Les Msr sont décrites être impliquées dans la résistance au stress oxydant, dans la réparation des protéines oxydées, dans des mécanismes de régulation et dans la virulence de certaines bactéries pathogènes (Ezraty *et al.*, 2005; Moskovitz, 2005; Weissbach *et al.*, 2005). Il est à noter que jusqu'à ce jour peu d'études ont montré un rôle physiologique avéré des MsrB.

II.1.1 Rôle dans la résistance au stress oxydant

L'importance des MsrA dans la résistance au stress oxydant a été particulièrement bien étudiée au cours de ces dix dernières années. Par exemple, l'interruption du gène *msrA* chez *E. coli* et *Saccharomyces cerevisiae* entraîne une augmentation de la sensibilité des cellules au stress oxydant, et la surproduction de MsrA chez *S. cerevisiae* entraîne une augmentation de la résistance des cellules au stress oxydant (Moskovitz *et al.*, 1995, Moskovitz *et al.*, 1997). D'autres études menées chez les eucaryotes ont également permis de mettre en évidence l'implication des Msr dans la lutte contre le stress oxydant. En effet, des souris ne possédant plus le gène *msrA* sont plus sensibles au stress oxydant (Salmon *et al.*, 2009). A l'inverse, la surproduction de la MsrA bovine chez la drosophile entraîne une augmentation de la résistance des drosophiles face au stress oxydant généré par le paraquat, et de leur durée de vie (Ruan *et al.*, 2002). Récemment, un rôle important dans la protection de la drosophile vis-à-vis du stress oxydant a également été montré pour la MsrB3A exprimée dans le réticulum endoplasmique et la MsrB1 (Lim *et al.*, 2012 ; Kwak *et al.*, 2012 ; Shchedrina *et al.*, 2011). Un autre exemple concerne l'implication de la MsrA (Kantorow *et*

al., 2004 ; Brennan *et al.*, 2009) et la MsrB (Marchetti *et al.*, 2005) dans la protection des cellules épithéliales de la lentille humaine vis-à-vis du stress oxydant, responsable de la formation de cataracte. De plus, des études récentes ont montré que la MsrB2 est largement exprimée au niveau de la rétine et protège les cellules épithéliales pigmentées de la rétine contre les entités oxydantes (Pascual *et al.*, 2010).

II.1.2 Rôle dans la réparation et la régulation

Plusieurs études *in vitro* ont permis de mettre en évidence la réparation par MsrA et/ou MsrB de différentes protéines oxydées sur des résidus Met. Ces oxydations peuvent ne pas modifier l'activité des protéines lorsque les résidus Met sont situés à la surface. C'est le cas par exemple de l'alpha-2-macroglobuline, un inhibiteur de protéase impliqué notamment dans la cascade de coagulation, ou de la glutamine synthétase (Levine *et al.*, 2000 ; Levine *et al.*, 1996). Ces cycles d'oxydation/réduction par les Msr des résidus Met de surface peuvent être considérés comme des mécanismes anti-oxydants permettant de diminuer la concentration intracellulaire en oxydants (Levine *et al.*, 1996).

Cependant, l'oxydation de certains résidus Met en Met-O peut aussi entraîner une modulation de l'activité biologique d'une protéine. La possibilité de réduire ces Met-O grâce aux Msr a donc conduit à proposer que ce mécanisme d'oxydation/réduction puisse aussi constituer une voie de régulation (Ciorba *et al.*, 1997). Un des exemples le mieux caractérisé est celui des canaux à potassium. L'oxydation des résidus Met en surface de ces canaux et la réaction inverse de réduction de la Met-O par les Msr contribuent à la modulation de la fonction de ces canaux. En effet, l'oxydation des Met provoque une modification de la conformation des canaux, ce qui empêche la fixation de la protéine kinase C (Ciorba *et al.*, 1997). Plus récemment, une régulation de la protéine kinase II (CaMKII) par oxydation réversible de résidus Met a été mise en évidence. Ainsi, l'oxydation d'une paire de résidus Met des domaines régulateurs permet l'activation de la CaMKII en absence de calcium et de calmoduline, cette activation étant réversée par action de la MsrA (Erickson *et al.*, 2008).

II.1.3 Rôle dans la virulence de bactéries pathogènes

Les Msr font partie des facteurs de virulence nécessaires à plusieurs microorganismes pathogènes pour résister aux ERO générées par les défenses de l'hôte. Ainsi, chez les bactéries comme *Mycobacterium smegmatis* et *Staphylococcus aureus*, la délétion du gène *msrA* entraîne une augmentation de la sensibilité aux ERO (Douglas *et al.*, 2004 ; Singh et

Moskovitz, 2003) et chez *Salmonella enterica* serotype Typhimurium, elle entraîne une augmentation de la sensibilité aux ERO mais également une diminution de la réplication du matériel génétique de la bactérie dans les macrophages activés et chez la souris (Denkel *et al.*, 2011). Les *Neisseriae* présentent elles la particularité de posséder une protéine multidomaine appelée PilB, comprenant notamment un domaine MsrA et un domaine MsrB. Chez *N. gonorrhoeae*, des études *in vivo* ont montré que PilB, présente au niveau périplasmique, confère à la bactérie une protection contre le H₂O₂ (Skaar *et al.*, 2002).

II.2. Propriétés enzymatiques et structurales des Msr

Les MsrA sont spécifiques des Met-S-O incluses dans des protéines ou libres alors que les MsrB réduisent spécifiquement les Met-R-O incluses dans des protéines ou libres. Les études menées sur les MsrA et MsrB ont montré que leur activité nécessitait la présence d'un réducteur, qui peut être le DTT *in vitro*, ou pour les Msr à au moins deux cystéines, l'une dite catalytique et l'autre de recyclage, la Trx *in vivo* (Boschi-Muller *et al.*, 2000 ; Olry *et al.*, 2004).

II.2.1 Structure tridimensionnelle des Msr

Les MsrA et MsrB ne présentent aucune identité de séquence et leurs structures tridimensionnelles sont également totalement différentes. Les Msr sont toutes deux des protéines monomériques de 170 à 211 résidus pour les MsrA et d'environ 150 résidus pour les MsrB. Plusieurs structures 3D de MsrA et MsrB ont été résolues par diffraction des rayons X (Lowther *et al.*, 2002; Ranaivoson *et al.*, 2008; Rouhier *et al.*, 2007; Taylor *et al.*, 2003; Tete-Favier *et al.*, 2000) et par RMN (Sal *et al.*, 2007 ; Breivik *et al.* 2008 ; Carella *et al.*, 2010 ; Aachmann *et al.*, 2011) .

Le repliement des MsrA est composé d'une partie centrale structurée d'environ 140 acides aminés entourée des extrémités N- et C-terminales, qui ne contiennent pas de structures secondaires régulières de type hélice α ou brin β . La partie centrale adopte un repliement mixte α/β , constitué d'un feuillet β antiparallèle à 5 brins, flanqué de 4 hélices α et d'un petit feuillet β antiparallèle à 2 brins surplombant le site actif (Figure 5). Le site actif est organisé autour de la Cys51 catalytique et il est constitué 1) d'une poche hydrophile, constituée des résidus invariants Glu94, Tyr82 et Tyr134, dans laquelle vient se fixer

Introduction

l'oxygène de la fonction sulfoxyde du substrat, 2) une poche hydrophobe constituée des résidus invariants Phe52 et Trp53 dans laquelle vient se fixer le groupement ϵ -méthyle du substrat, et 3) un site de reconnaissance de la chaîne principale du substrat constitué des résidus invariants Asp129 et Tyr197 (Gand *et al.*, 2007).

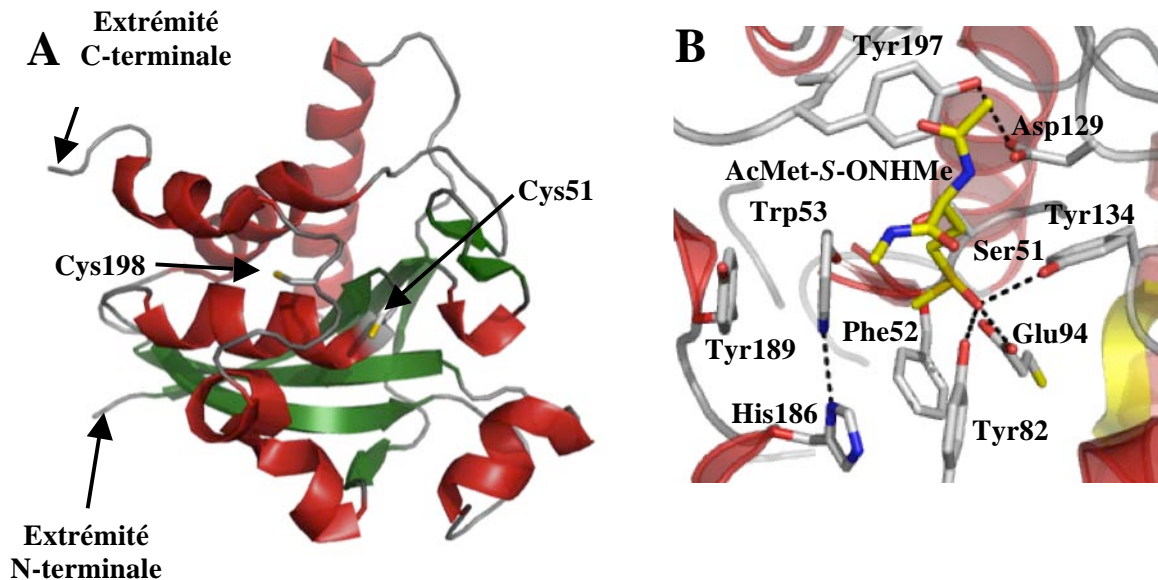


Figure 5 : Représentation schématique de la structure du monomère de la MsrA (A) et du site actif du mutant inactif C51S de la MsrA de *N. meningitidis* en complexe avec une molécule d'AcMet-S-ONHMe (B).

A) Architecture globale de la MsrA. Les chaînes latérales des deux cystéines (51 et 198) impliquées dans le mécanisme sont représentées. Les structures secondaires en hélices α sont représentées en rouge et celles en brins β sont représentées en vert.

B) Structure du site actif représentant les chaînes latérales des résidus conservés de la poche hydrophile (Tyr82, Glu94 et Tyr134), de la poche hydrophobe (Phe52, Trp53, His186 et Tyr189) et du site de reconnaissance de la chaîne principale du substrat (Asp129 et Tyr197). Le substrat AcMet-S-ONHMe est représenté en mode bâton et coloré en jaune.

La figure a été réalisée avec le logiciel Pymol d'après les fichiers PDB 3BQE et 3BQF.

Les MsrB présentent elles un repliement essentiellement de type β , avec le cœur de l'enzyme composé de deux feuillets β antiparallèles se faisant face pour former une structure proche du tonneau β , et flanquée de trois hélices α et cinq hélices 3_{10} (Figure 6). Le site actif des MsrB, organisé autour de la Cys117 catalytique, est formé 1) d'une poche hydrophile, constituée du résidu His103 et d'une molécule d'eau, dans laquelle vient se fixer l'oxygène de la fonction sulfoxyde du substrat, et 2) d'une poche hydrophobe, formée par la chaîne

latérale du résidu Trp65, impliquée dans la reconnaissance du groupement ϵ -méthyle du substrat (Neiers *et al.*, 2007).

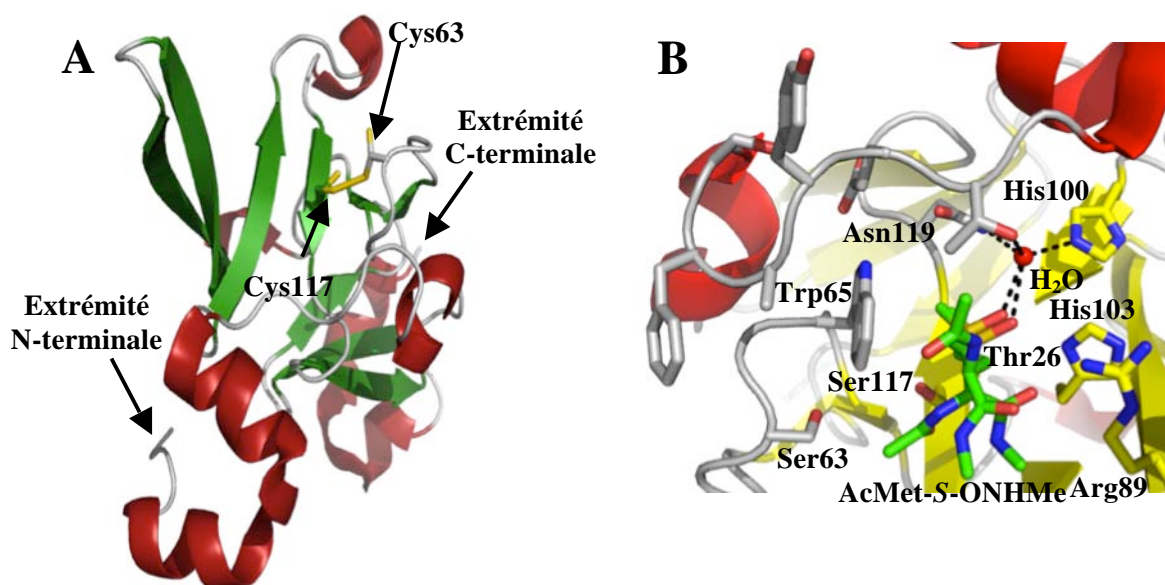


Figure 6 : Représentation schématique de la structure du monomère (A) et du site actif (B) de la MsrB de *N. meningitidis* en complexe avec une molécule d'AcMet-R-ONHMe.

A) Architecture globale de la MsrB de *N. meningitidis*. Les chaînes latérales de la Cys117 catalytique et de la Cys63 sont représentées, avec un mélange de forme réduite et de forme pont disulfure Cys63-Cys117. Les structures secondaires en hélices α sont représentées en rouge et celles en brins β sont représentées en vert.

B) Structure du site actif du mutant C63S/C117S de la MsrB de *N. meningitidis* en complexe avec le substrat AcMet-R-ONHMe (3 conformations), représentant les chaînes latérales des résidus conservés de la poche hydrophile (His103, His100, Thr26 et Asn119), de la poche hydrophobe (Trp65). Le substrat AcMet-S-ONHMe est représenté en mode bâton et coloré en vert. La figure a été réalisée avec le logiciel Pymol d'après les fichiers PDB 3HCG et 3HCH.

La différence de stéréosélectivité de substrat des MsrA et MsrB s'explique par la symétrie « inversée » de leur site actif, due à l'orientation du site hydrophile par rapport au site hydrophobe (Lowther *et al.*, 2002). D'autre part, la reconnaissance du substrat sulfoxyde reposant majoritairement sur des interactions de type liaisons de van der Waals entre la poche hydrophobe et les groupements méthyle ϵ et méthylène γ du substrat Met-O, explique que les MsrA et MsrB puissent réduire les Met-O incluses dans une chaîne peptidique mais aussi les Met-O libres (Gand *et al.*, 2007; Neiers *et al.*, 2007; Olry *et al.*, 2002).

II.2.2 Mécanisme et catalyse

Bien que présentant des structures tridimensionnelles différentes et des spécificités de substrat opposées, les MsrA et MsrB présentent un même mécanisme catalytique, qui repose sur la chimie de l'acide sulfénique, et la formation de ponts disulfures réductibles par la Trx (Boschi-Muller *et al.*, 2000; Boschi-Muller *et al.*, 2005; Olry *et al.*, 2002) (Figure 7).

Lors de la première étape, appelée étape réductase, la Cys catalytique de la Msr réalise une attaque nucléophile sur l'atome de soufre de la fonction sulfoxyde du substrat conduisant à un état de transition de type sulfurane. Le réarrangement de ce sulfurane, *via* la formation d'un sulfonium (Figure 8), permet de générer l'intermédiaire acide sulfénique sur la Cys catalytique et de libérer une mole de Met par mole d'enzyme. Cette étape est très rapide pour les MsrA comme les MsrB grâce à une catalyse assistée par le substrat et à la présence d'un catalyseur acide (Figure 9) (Antoine *et al.*, 2006; Boschi-Muller *et al.*, 2008; Neiers *et al.*, 2007).

Dans la deuxième étape, une deuxième Cys, appelée Cys de recyclage, vient attaquer l'acide sulfénique pour former un pont disulfure et libérer une mole d'eau par mole d'enzyme. Cette étape est cinétiquement limitée par la première étape, qui est l'étape cinétiquement déterminante dans le mécanisme à deux étapes qui conduit à la formation du pont disulfure intramoléculaire. Enfin, lors de la troisième étape, le pont disulfure de la Msr oxydée est réduit par la Trx. Cette étape est l'étape cinétiquement limitante du mécanisme global. Ce mécanisme catalytique est en accord avec le mécanisme cinétique des Msr qui est de type ping-pong en présence de Trx ou de DTT comme réducteur (Boschi-Muller *et al.*, 2005).

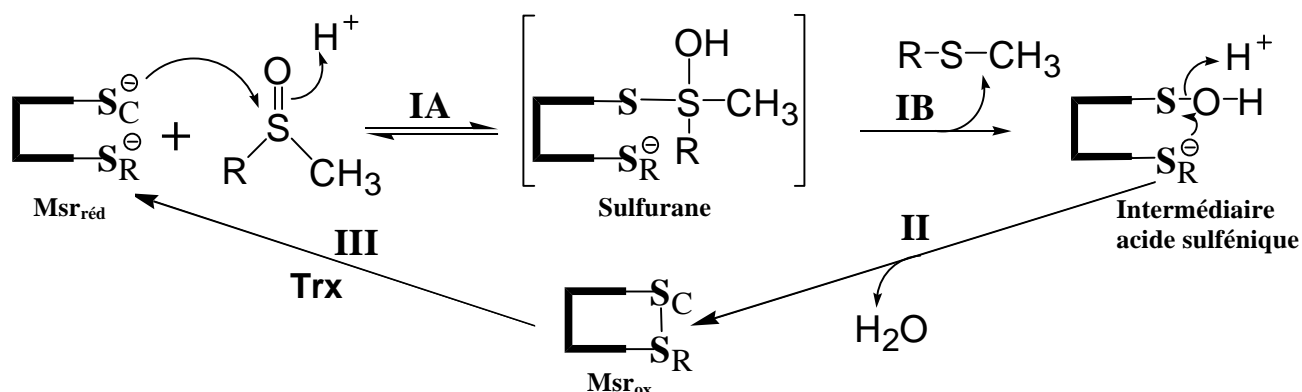


Figure 7 : Mécanisme catalytique des Msr à deux Cys.

La formation du complexe $\text{Msr}_{\text{red}}/\text{Met-O}$ n'est pas explicitée. L'étape réductase peut être divisée en deux sous-étapes. Le substrat sulfoxyde est attaqué par la Cys catalytique (S_C) lors de l'étape IA, conduisant à la formation de l'intermédiaire/état de transition sulfurane. Ce sulfurane doit se réarranger dans l'étape IB pour aboutir à l'intermédiaire acide sulfénique. L'attaque de la Cys de recyclage (S_R) sur cet intermédiaire permet la formation d'un pont disulfure intramoléculaire et la libération concomitante d'une molécule d'eau, lors de l'étape II. Enfin, le pont disulfure est réduit par la Trx au cours de l'étape III (d'après Olry *et al.*, 2002).

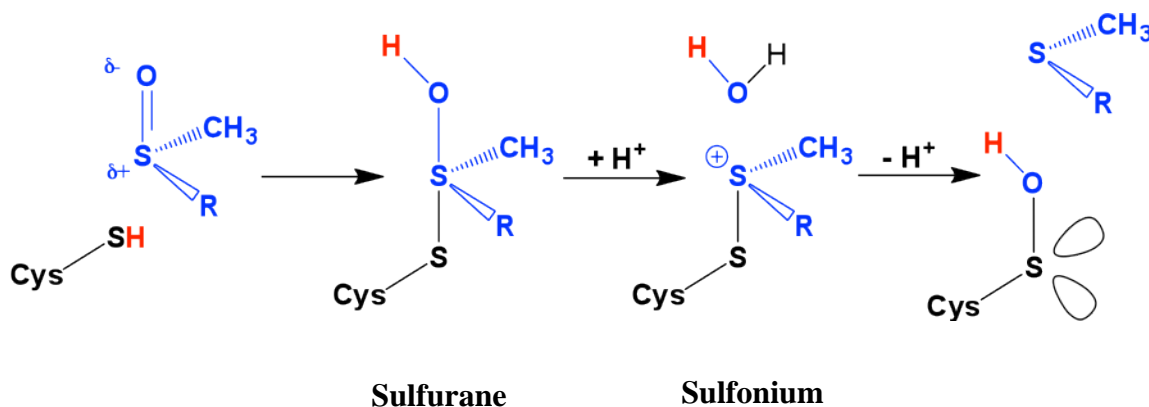


Figure 8 : Schéma du mécanisme de la formation de l'intermédiaire acide sulfénique.

Dans ce mécanisme, le réarrangement du sulfurane conduit à la formation d'un sulfonium. Une molécule d'eau attaque le soufre de la Cys catalytique, ce qui conduit à la libération de la Met et la formation de l'intermédiaire acide sulfénique. La molécule d'eau devrait être activée par un catalyseur basique (non représenté) (Robinet *et al.*, 2011).

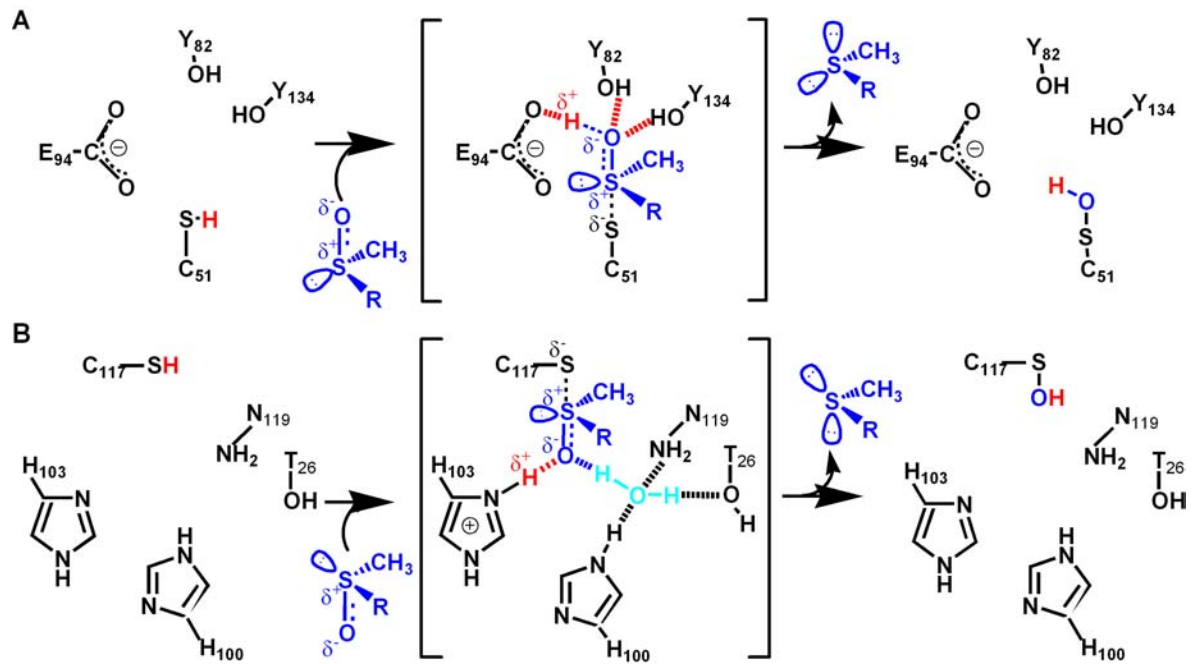


Figure 9 : Mécanisme catalytique proposé pour l'étape réductase des MsrA (A) et MsrB (B) de *N. meningitidis*.

Seuls les groupements fonctionnels des résidus catalytiques sont représentés : Cys51, Glu94, Tyr82, et Tyr134 pour les MsrA ; Cys117, His100, His103, Asn119 et Thr26 ainsi que la molécule d'eau formant un réseau de liaisons hydrogène pour les MsrB. RSOCH₃ et RSCH₃ représentent respectivement la Met-O et la Met.

Pour les MsrA et MsrB, le mécanisme de catalyse est semblable : le substrat se fixe au niveau du site actif sous une forme fortement polarisée, *via* des liaisons hydrogène avec l'oxygène de la fonction sulfoxyde. Un état de transition de type sulfurane est formé et est stabilisé *via* un réseau de liaisons hydrogène. Le réarrangement de ce sulfurane permet la libération d'une mole de Met par mole d'enzyme et la formation de l'intermédiaire acide sulfénique. L'hydrogène provenant de la Cys 51 est représenté en rouge, le substrat Met-O en bleu.

II.3. Les sous-classes de Msr

A l'exception de quelques MsrB « eucaryotes » qui ont une sélélocystéine catalytique, toutes les Msr possèdent une Cys catalytique conservée, en position 51 pour les MsrA et 117 pour les MsrB. Cependant, différentes sous-classes peuvent être distinguées en fonction du nombre et de la position des Cys impliquées dans le mécanisme de régénération (Boschi-Muller *et al.*, 2005) (Figure 10 et 11).

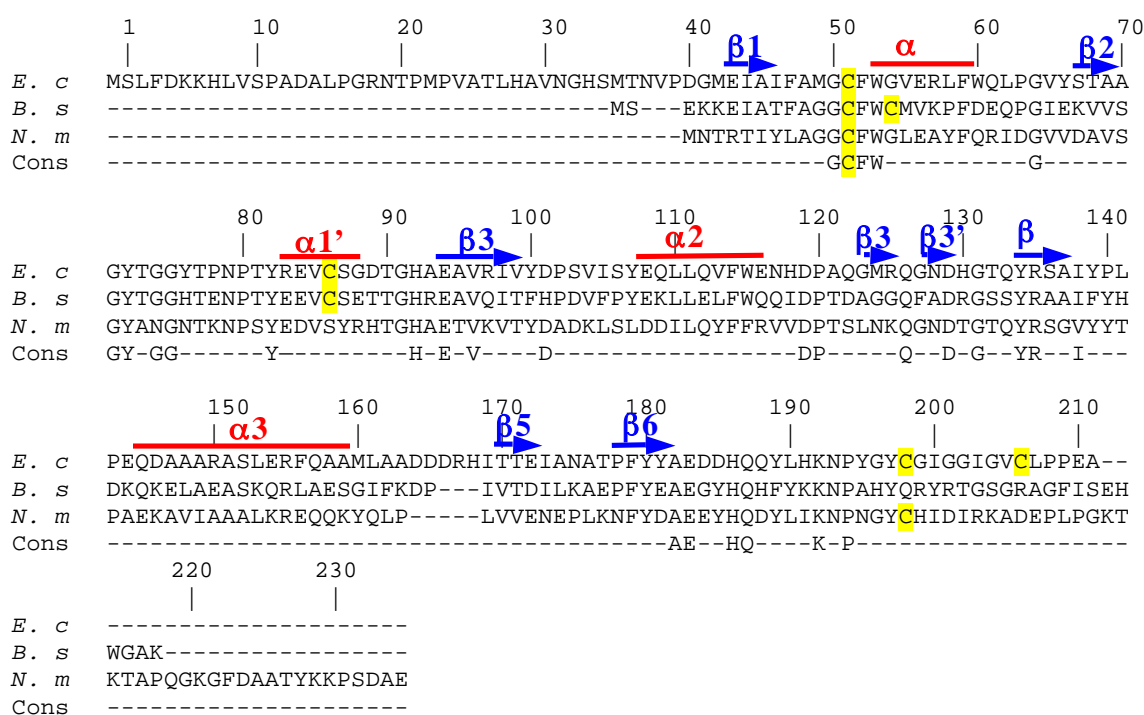


Figure 10 : Alignement des séquences des MsrA d'*E. coli* (*E. c*), de *B. subtilis* (*B. s*) et de *N. meningitidis* (*N. m*).

Chacune de ces enzymes est représentative d'une des trois sous-classes de MsrA, qui diffèrent par la position des résidus Cys de recyclage impliqués dans le mécanisme de régénération (*E. coli* : C51/C198/C206 ; *B. subtilis* : C51/C54 ; *N. meningitidis* : C51/C198). La numérotation utilisée est basée sur celle de l'enzyme d'*E. coli*. Les structures secondaires indiquées sont celles correspondant à l'enzyme d'*E. coli* (Tete-Favier *et al.*, 2000) : les hélices α sont indiquées par des traits rouges, les brins β par des flèches bleues. Les résidus Cys sont indiqués sur fond jaune. La séquence consensus (cons) présente les résidus conservés à plus de 90 % dans les séquences des MsrA putatives extraites des banques génomiques (génomomes entièrement séquencés).

La majorité des Msr possède deux Cys impliquées dans le mécanisme, conduisant à la formation d'un intermédiaire pont disulfure qui est réduit par la Trx. C'est le cas des MsrA de *N. meningitidis* et *Bacillus subtilis*, et des MsrB d'*E. coli*, de *N. meningitidis* et de *Xanthomonas campestris*. Certaines Msr ne possèdent elles que la Cys catalytique. C'est le cas en particulier de la MsrA de *Clostridium* sp. OhILAs, de la MsrB de *Mycoplasma pulmonis*, de la MsrB2 de souris, de la MsrB3 humaine, et de la MsrB1 d'*Arabidopsis thaliana* pour lesquelles la réduction de la Met-O conduit à la formation d'un intermédiaire acide sulfénique sur la Cys catalytique qui serait réduit soit par la Trx (cas des MsrB2 de souris et MsrB3 humaine), soit par le système glutathion/glutarédoxine (cas de la MsrA de *Clostridium* sp. OhILAs, de la MsrB de *M. pulmonis* et de la MsrB1 d'*A. thaliana*) (Kim et Kim, 2008 ; Lee et Kim, 2008 ; Neiers, 2007 ; Tarrago *et al.*, 2009 ; Kim *et al.*, 2011).

Introduction

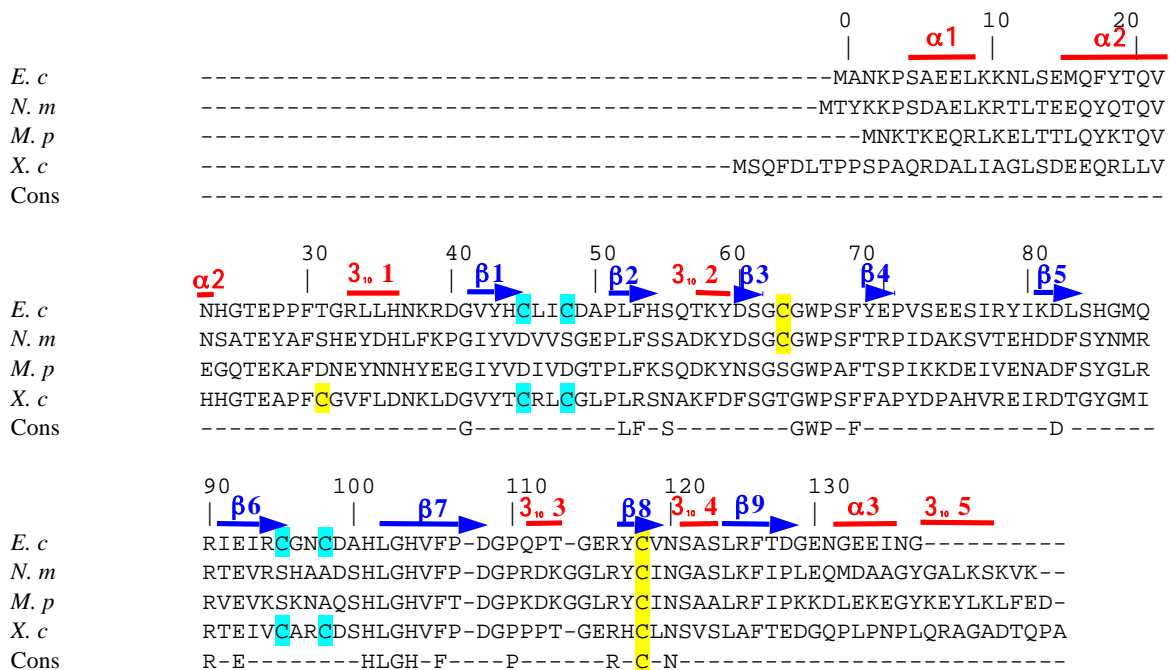


Figure 11 : Alignement des séquences des MsrB d'*E. coli* (*E. c*), de *N. meningitidis* (*N. m*), de *M. pulmonis* (*M. p*), et de *X. campestris* (*X. c*).

Les MsrB peuvent être divisées en trois sous-classes selon le nombre et la position des Cys impliquées dans le mécanisme de recyclage : deux Cys (cas de *N. meningitidis* et d'*E. coli* avec la Cys de recyclage en position 63, et de *X. campestris* avec une Cys en position 31) ou une seule Cys (cas de *M. pulmonis*). Certaines MsrB possèdent quatre Cys supplémentaires, impliquées dans la fixation d'un métal (cas d'*E. coli* et de *X. campestris*). La numérotation utilisée est celle de l'enzyme d'*E. coli*. Les structures secondaires indiquées sont celles correspondant à l'enzyme de *N. gonorrhoeae* (Lowther *et al.*, 2002) : les hélices α sont indiquées par des traits rouges, les brins β par des flèches bleues. Les résidus Cys impliqués dans la chélation d'un ion métallique sont indiqués sur fond bleu, les autres Cys sont indiqués sur fond jaune. La séquence consensus (cons) présente les résidus conservés à plus de 90 % dans les séquences de MsrB putatives extraites des banques génomiques (génomés entièrement séquencés).

Enfin, il existe une sous-classe de MsrA qui possède trois Cys impliquées dans le mécanisme. C'est le cas par exemple des MsrA d'*E. coli* et de souris pour lesquelles les Cys198 et 206 sont impliquées dans le mécanisme de régénération *via* la formation successive de deux ponts disulfures : le pont 51-198, puis 198-206, ce dernier étant celui préférentiellement réduit par la Trx. En absence de réducteur, la forme oxydée pont disulfure 198-206 peut réduire une seconde molécule de Met-O avec la formation d'un intermédiaire oxydé acide sulfénique sur la Cys51 et du pont disulfure 198-206. Une étude récente, réalisée sur la MsrA de souris, a montré que cette forme oxydée possède une activité oxydase *in vitro vis-à-vis* de résidus Met, en absence de Trx (Lim *et al.* 2011) (Figure 12).

Cette enzyme possède donc une double activité réductase et oxydase. Un mécanisme de régulation de ces deux activités a été proposé, dépendant de l'accessibilité du l'extrémité C-terminale, appelée domaine par les auteurs, contenant les deux Cys de régénération. Ainsi, quand ce domaine est accessible à la Trx, l'activité réductase serait prédominante, dans le cas contraire, probablement *via* la fixation d'une protéine régulatrice sur le domaine C-terminal, l'activité oxydase serait prédominante (Figure 13).

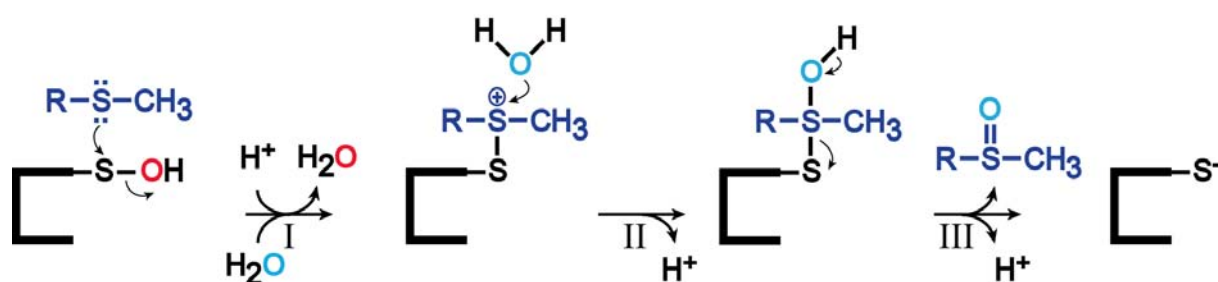


Figure 12 : Mécanisme proposé pour l'activité oxydase de la MsrA de souris (adapté d'après Lim *et al.*, 2011).

Étape (I) : le doublet libre de la Met réalise une attaque sur l'atome du soufre de l'acide sulfénique, ce qui conduit à la formation de l'intermédiaire sulfonium et à la libération d'une molécule d'eau *via* une catalyse acide assistée. Étape (II) : une molécule d'eau activée attaque le soufre du sulfonium pour former un état de transition de type sulfurane. Étape (III) : La cassure de la liaison S-S de l'intermédiaire sulfurane, favorisée par une catalyse basique, conduit à la libération de la Met-O et la restauration de l'enzyme sous sa forme réduite.

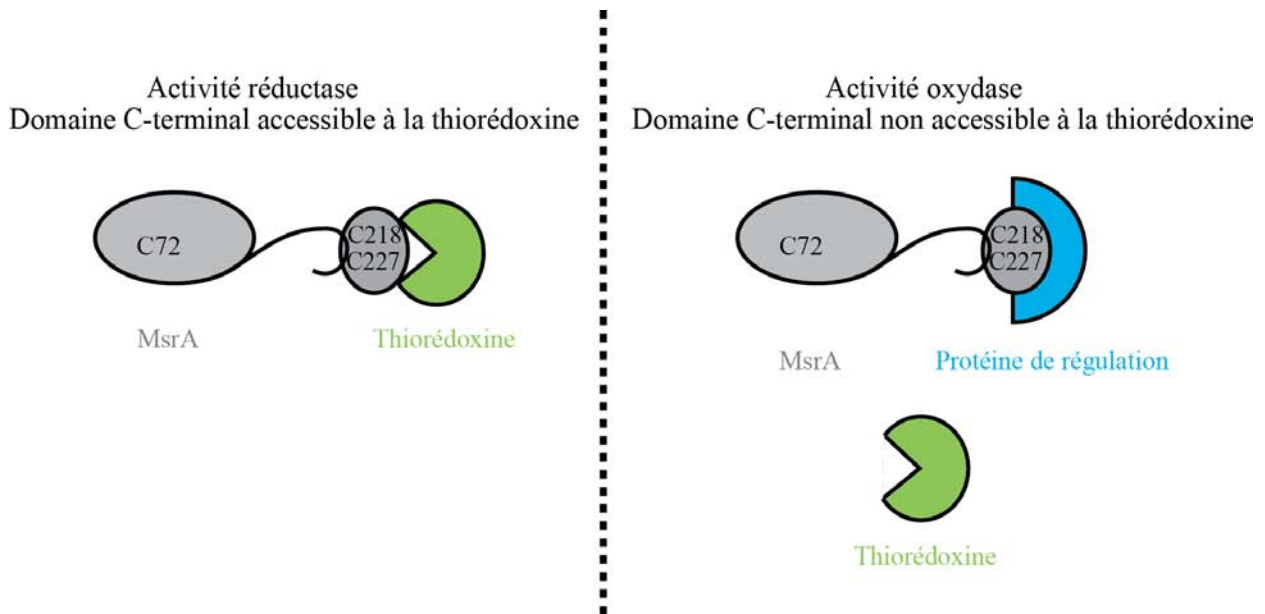


Figure 13 : Schéma du mécanisme potentiel pour la régulation des activités réductase et oxydase de la MsrA de souris (adapté de Lim *et al.*, 2011).

Lorsque le domaine C-terminal, contenant les Cys de recyclage, est accessible, la MsrA possède une activité réductase (A). Si ce site est occupé par une autre protéine appelée ici « protéine de régulation », la MsrA possède une activité oxydase (B). La C72 correspond à la C51 de la MsrA d'*E. coli*.

Indépendamment de ces différentes classes, plusieurs MsrB possèdent également deux motifs $C_{45}XXC_{48}$ et $C_{94}XXC_{97}$ permettant la fixation d'un cation métallique Zn^{2+} . C'est notamment le cas pour les MsrB d'*E. coli* et de *X. campestris*.

PARTIE 3 : Les fRMsr

Les fRMsr representent la troisième classe de Msr, qui a été mise en évidence en 2007 chez *E. coli* (Lin *et al.*, 2007) alors que leur structure tridimensionnelle était connue depuis 2000 (Ho *et al.*, 2000). Elles constituent à l'heure actuelle la seule famille de protéines à domaine GAF qui présente une activité catalytique. C'est pourquoi dans la suite de cette partie seront présentées dans une première section les protéines à domaines GAF autres que les fRMsr, qui elles feront l'objet de la section suivante.

III.1 Les protéines à domaine GAF

En 1997, l'étude bioinformatique réalisée par Aravind et Ponting qui a consisté à rassembler des domaines protéiques possédant des motifs structuraux similaires, a conduit à introduire et définir le terme GAF. L'acronyme GAF provient des initiales des noms des trois premières classes de protéines caractérisées contenant ce domaine : les cGMP-binding phosphodiésterases de mammifères, les Adénylyl cyclases Cyb1 et Cyb2 d'*Anabaena sp.*, et le « Formate-hydrogen lyase transcription activator » d'*Escherichia coli* (Aravind et Ponting, 1997).

Depuis, les domaines GAF ont été identifiés dans plus de 7400 protéines issues de la base de données non-redondante (<http://smart.embl-heidelberg.de>), mais dans la plupart des cas les ligands sont inconnus. Ce sont des domaines qui fixent des petites molécules, appelés « Small Molecule Binding Domain » (ou SMBD) et qui sont retrouvés chez les *Archaea*, les bactéries et les eucaryotes. Ils sont impliqués principalement dans différents processus de signalisation cellulaire.

III.1.1 Organisation des domaines GAF

Les domaines GAF sont retrouvés dans la plupart des cas dans des protéines multi-domaines qui participent d'une manière ou d'une autre à la transduction du signal. Ces domaines sont habituellement associés avec d'autres domaines régulateurs comme par exemple les domaines PAS (« Per ARNt and Sim »), des domaines récepteur REC (ou phosphoaccepteur cheY-like), ou des domaines histidine kinase (Figure 14) (Martinez *et al.* 2002 ; Galperin, 2006).

Les domaines GAF sont organisés dans la plupart des cas en tandem de deux domaines GAF appelé GAF-A et GAF-B, mais ils peuvent également exister sous forme d'un seul domaine ou même de six domaines GAF consécutifs. C'est le cas par exemple de l'histidine

Introduction

kinase *A7HHY3* de *Synechococcus* ou d'*Anaeromyxobacter dehalogenans* qui possède six domaines GAF dont le rôle est inconnu (Schultz, 2009) (Figure 14).

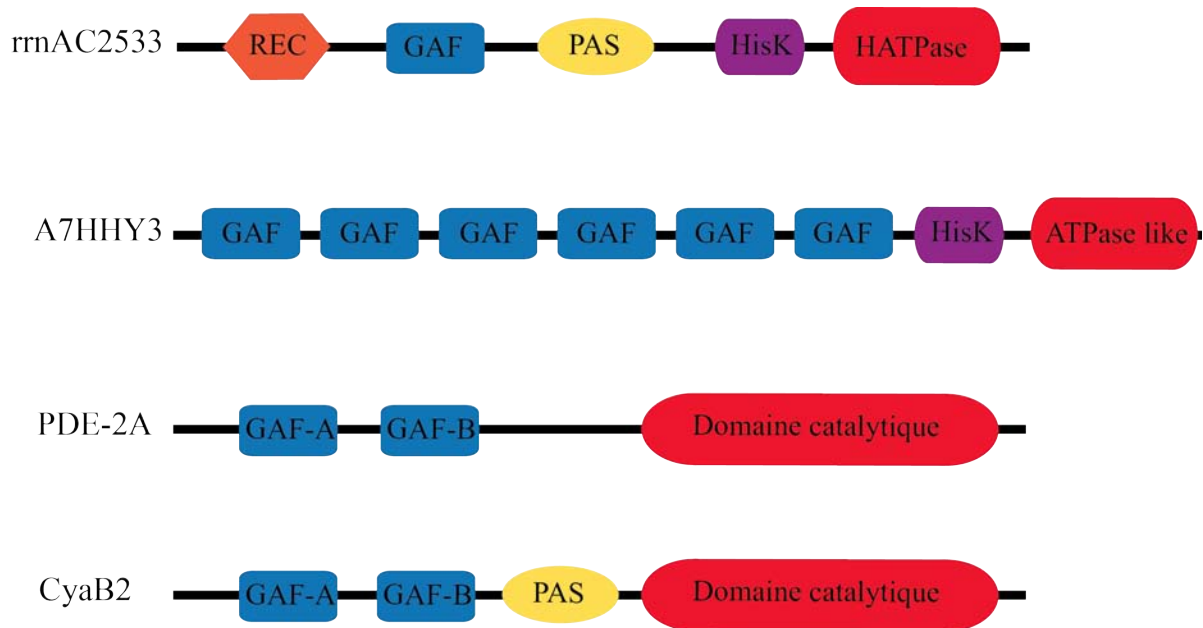


Figure 14 : Représentation schématique de différentes organisations génomiques des domaines GAF.

Les protéines représentées sont : rrnAC2533 : Adaptive-response sensory-kinase [*Haloarcula marismortui* ATCC 43049] ; A7HHY3 : Signal de transduction multi-senseur histidine kinase de *Synechococcus* ou d'*Anaeromyxobacter dehalogenans* ; PDE-2A : Phosphodiesterase 2A humaine ; CyaB2 : Adénylate cyclase B2 d'*Anabaena*.

REC, GAF, PAS, HisK, HATPase, et ATPase-like représentent respectivement les domaines « Receiver domain » ou phosphoaccepteur cheY-like représenté en orange, GAF représenté en bleu, « Per ARNt and Sim » représenté en jaune, « Histidine kinase » représenté en violet, « Histidine kinase-like ATPase » représenté en rouge, et « Adenosine triphosphatase like » représenté en rouge. Les domaines catalytiques sont représentés en rouge.

Dans la suite de cette introduction, seuls les exemples de protéines contenant des domaines GAF les mieux caractérisés à ce jour seront traités, à savoir les phosphodiesterases, les adénylate cyclases et le facteur de transcription NifA.

III.1.2 Les domaines GAF dans les phosphodiesterases

Chez les mammifères, les domaines GAF sont essentiellement retrouvés associés aux phosphodiesterases (PDE), enzymes importantes puisqu'elles contrôlent les taux des seconds messagers AMPc et GMPc. Seulement cinq des onze familles de PDE connues (PDE-2, -5, -6, -10, et -11) contiennent deux domaines GAF, appelés GAF-A et GAF-B, en tandem à l'extrémité N-terminale du domaine catalytique de la PDE. Deux fonctions différentes ont été décrites pour les domaines GAF des PDE (Figure 15): la fixation des

nucléotides cycliques et la dimérisation. Dans le cas de la PDE-2, le domaine GAF-A est impliqué dans la dimérisation alors que le domaine GAF-B est responsable de la fixation du ligand AMPc. En ce qui concerne les PDE-5 et -6, seul le domaine GAF-A possède les 2 fonctions, le domaine GAF-B étant incapable de fixer le GMPc (Martinez *et al.* 2002 ; Rybalkin *et al.* 2003 ; Muradov *et al.* 2004 ; Huang *et al.* 2004 ; Zoraghi *et al.* 2005).

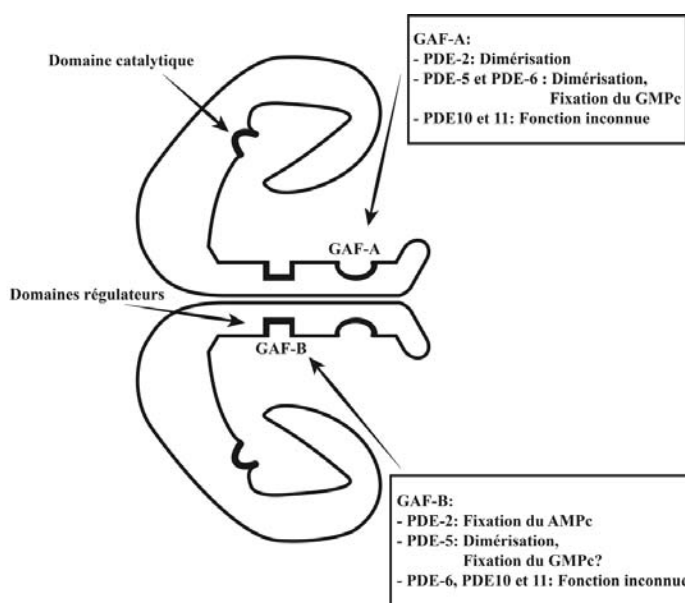


Figure 15 : Représentation schématique des rôles connus des domaines GAF des PDEs (d'après Zoraghi *et al.*, 2002).

Les PDE sont des protéines dimériques, mais il semblerait que la dimérisation de cette enzyme ne soit pas obligatoire pour qu'elles aient une activité catalytique efficace. Il a été montré que le domaine catalytique sous forme monomérique de la PDE-5 présente une activité catalytique similaire à celle observée sur la protéine entière (Fink *et al.*, 1999 ; Heikaus *et al.*, 2008). En 2009, une étude a permis de déterminer par diffraction des rayons X la structure 3D de la PDE-2A sous forme de dimère avec les domaines GAF-A et -B et le domaine catalytique (voir paragraphe III.1.7 page 32). Cette étude a permis de proposer pour cette dimérisation un rôle dans la régulation du domaine catalytique. En effet, en présence du substrat, le site actif adopte une conformation « open » (site accessible au solvant), alors qu'en son absence, il adopte une conformation « close » (site inaccessible au solvant) (Pandit *et al.*, 2009).

III.1.3 Les domaines GAF dans les adénylate cyclases

Chez la cyanobactérie *Anabaena*, les adénylate cyclases (AC) de classe III contiennent également un tandem de domaines GAF à leur extrémité N-terminale (Katayama *et al.*, 1997). Les AC CyaB1 et CyaB2 sont des enzymes membranaires qui produisent de l'AMPc à partir de l'ATP, un second messager impliqué dans plusieurs voies de signalisation

Introduction

cellulaires. Les AC présentent une organisation en domaines similaire à celle observée dans les PDE, à savoir un domaine N-terminal contenant les deux domaines GAF-A et GAF-B associés en tandem et un domaine catalytique C-terminal. En plus des deux domaines GAF, les AC contiennent aussi un domaine PAS, localisé entre les domaines GAF et le domaine catalytique, dont la fonction reste encore inconnue. Au contraire des PDE, la dimérisation des AC est obligatoire pour avoir une activité catalytique efficace (Kanacher *et al.* 2002 ; Martinez *et al.* 2005).

Les enzymes CyaB1 et CyaB2 sont activées par la fixation de l'AMPc au niveau du domaine GAF-B. Une AC chimérique composée des domaines GAF de l'enzyme PDE-2 et du domaine catalytique de CyaB1 peut être activée par le GMPc (Kanacher *et al.* 2002). Ceci suggère que les domaines GAF de ces deux enzymes possèdent des mécanismes communs de régulation des activités enzymatiques des domaines catalytiques voisins.

III.1.4 Les domaines GAF dans le facteur de transcription NifA

Les bactéries de type *Azotobacter vinelandii* utilisent l'azote de l'atmosphère en aérobie. Elles utilisent la nitrogénase pour fixer l'azote, et synthétisent trois types de nitrogénase : la nitrogénase-1 contenant du fer et du molybdène, la nitrogénase-2 contenant du fer et du vanadium, et la nitrogénase-3 contenant uniquement du fer (Premakumar *et al.* 1992 ; Waugh *et al.* 1995). NifA est un activateur transcriptionnel appartenant à la famille des protéines EBP (« Enhancer binding protein »), qui est requis pour l'expression des gènes codant la nitrogénase-1 chez *A. vinelandii*. Il est constitué de trois domaines: un domaine GAF à l'extrémité N-terminale, un domaine central de la famille des protéines AAA+ (« ATPases Associated with diverse cellular Activities »), et un domaine C-terminal de liaison à l'ADN.

L'activité de NifA est régulée en réponse à la fixation de l'azote *via* une interaction avec la protéine régulatrice NifL. La protéine NifL est une flavoprotéine qui joue le rôle de senseur de l'état redox ainsi que de la quantité d'azote présente dans la cellule. La fixation de la protéine NifL à la protéine NifA entraîne une inhibition de l'activité ATPase de la protéine NifA et donc de son rôle d'activation de la transcription par l'ARN polymérase dépendant du facteur σ_{54} . L'interaction entre les protéines NifA et NifL est contrôlée par le domaine GAF de NifA qui fixe le 2-oxoglutarate (Little et Dixon, 2003). Il a été montré que la fixation du 2-oxoglutarate sur le domaine GAF ne régule pas l'activité de NifA mais plutôt l'interaction de type protéine-protéine entre NifA et NifL (Martinez-Argudo *et al.*,

2005). En effet, il semblerait que la fixation du 2-oxoglutarate sur son domaine GAF conduirait à un changement conformationnel de NifA qui diminuerait son affinité pour la protéine NifL.

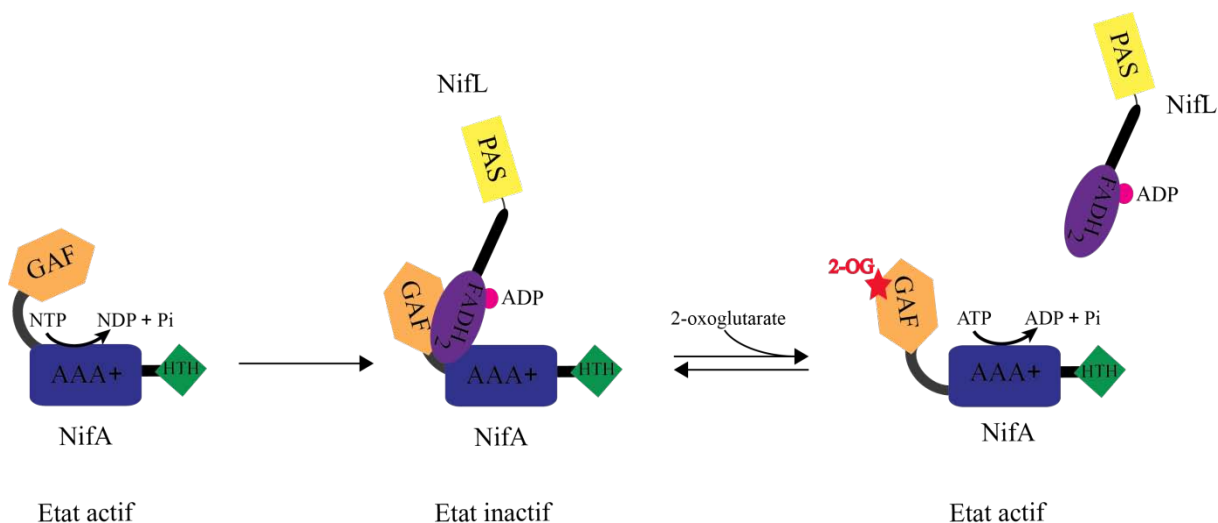


Figure 16 : Modèle schématique du rôle du domaine GAF dans la régulation de l'activité de NifA (adapté de Martinez-Argudo *et al.*, 2005).

En absence de NifL, et en présence ou absence de 2-oxoglutarate, la protéine NifA est active et le domaine AAA+ hydrolyse les nucléosides triphosphate. En présence de l'ADP et en absence de 2-oxoglutarate, la forme réduite de NifL (FADH₂) se lie à la protéine NifA et inhibe son activité. La forme réduite de NifL complexée à l'ADP n'est plus capable de se fixer et d'inhiber NifA, quand le 2-oxoglutarate (représenté par une étoile rouge) se fixe sur le domaine GAF de NifA.

FADH₂, HTH et AAA+ représentent respectivement les domaines de fixation de l'ADP, de fixation à l'ADN « Helix-Turn-Helix » et ATPase Associées à diverses Activités cellulaires. Le 2-OG représente la molécule de 2-oxoglutarate.

III.1.5 Structure tridimensionnelle du domaine GAF

La première structure tridimensionnelle décrite d'un domaine GAF a été la structure de la protéine ykg9 de *S. cerevisiae*, déterminée par diffraction des rayons X avec une résolution de 1,9 Å, que l'on sait aujourd'hui être une fRMsR (Ho *et al.*, 2000) (Figure 17). La protéine ykg9 de *S. cerevisiae* est un domaine GAF isolé, codé par le cadre ouvert de lecture ykg9, et qui ne fixe pas les nucléotides cycliques comme c'est le cas des domaines GAF des AC et des PDE. Le cœur du domaine GAF est constitué d'un feuillet central de 6 brins β antiparallèles, entouré de 4 hélices α. La protéine GAF existe sous forme dimérique en solution et dans le cristal (Figure 17). L'interface du dimère est constituée des résidus de la partie N-terminale, α3, β2, β3 et les boucles β2- β3, pour une surface d'interactions de 2200 Å². Les deux monomères du dimère de la protéine ykg9 sont positionnés tête-bêche.

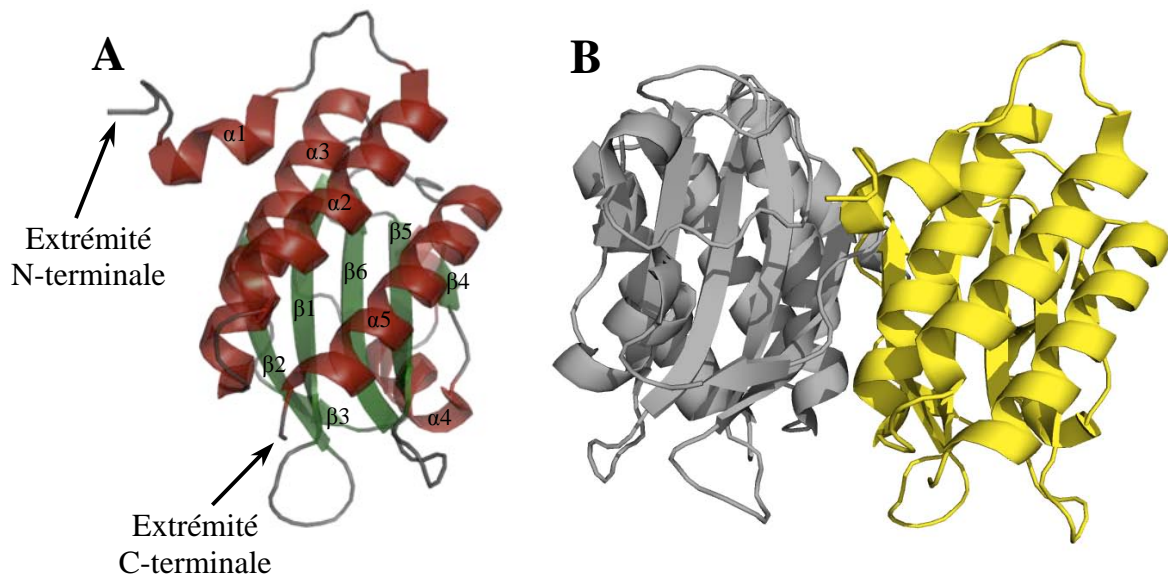


Figure 17 : Représentation schématique de la structure du monomère du domaine GAF de la protéine ykg9 de *S. cerevisiae* (A) et de son dimère (B).

(A) Architecture globale du domaine GAF. Les structures secondaires en hélices α sont représentées en rouge et celles en brins β sont représentées en vert. (B) Organisation du dimère. Les deux monomères sont représentés respectivement en gris et en jaune.

La figure a été réalisée avec le logiciel Pymol d'après le fichier PDB 1F5M.

La résolution, en 2002 et 2005, des structures tridimensionnelles des domaines GAF en tandem respectivement de la PDE-2A de souris à 2,9 Å et de l'AC CyaB2 d'*Anabaena* à 1,9 Å a montré l'existence, tout au moins dans le cristal, de deux modes de dimérisation des domaines GAF, c'est-à-dire en dimère parallèle ou en dimère anti-parallèle (Martinez *et al.*, 2002) (Figure 18). Ainsi, les domaines GAF de la PDE-2 forment un dimère parallèle dans lequel les domaines GAF-B contiennent chacun une molécule de GMPC et sont éloignés de 65 Å, alors que les domaines GAF-A dimérisent et ne contiennent pas de ligand. Par contre, les deux domaines GAF de CyaB2 forment un dimère anti-parallèle, i.e. en position tête-bêche, dans lequel le domaine GAF-A d'un monomère dimérise avec le domaine GAF-B de l'autre monomère. De plus, les domaines GAF-A et GAF-B fixent tous les deux l'AMPc ce qui conduit à la fixation de quatre molécules d'AMPc par dimère qui sont complètement enfouis dans la protéine. Dans les deux structures, les domaines GAF-A et GAF-B sont séparés par une longue hélice α qui est essentielle à la dimérisation. La question du rôle fonctionnel de ces deux modes de dimérisation et de l'influence de la fixation des nucléotidiques cycliques sur le mode de dimérisation reste encore sujet à débat (Martinez *et*

al., 2002 et 2005 ; Steegborn *et al.*, 2005 ; Tews *et al.*, 2005 ; Zhang *et al.*, 1997 ; Sunahara *et al.*, 1997).

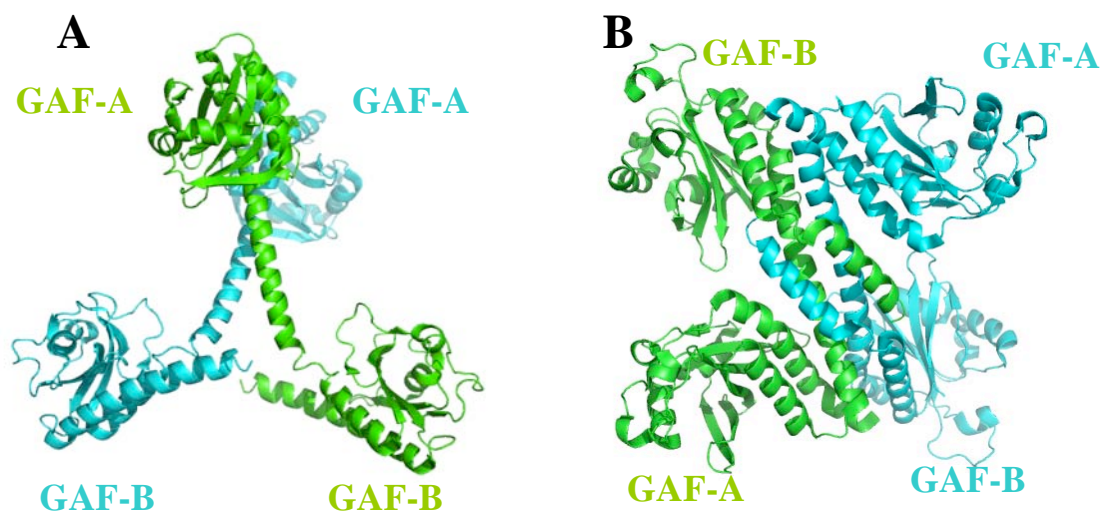


Figure 18 : Représentation schématique de la structure du dimère des domaines GAF en tandem de la PDE-2A de souris (A) et de l'AC CyaB2 d'*Anabaena* (B).

Organisation du dimère de tandem GAF. Les deux monomères sont représentés respectivement en bleu et en vert. Deux modes de dimérisation des domaines GAF, (A) le dimère des domaines GAF A et B de la PDE-2A de souris est de type parallèle, les hélices α du côté N-terminal pointe vers l'extérieur du dimère, et la surface d'interaction des hélices α « linker » qui lie le domaine GAF-A au domaine GAF-B, apparaît restreinte. (B) le dimère des domaines GAF A et B de l'AC CyaB2 est de type tête-bêche, la surface de dimérisation est formée d'une part par les hélices α du côté N-terminal qui se replient vers le centre du dimère et d'autre part par l'hélice α « linker ».

La figure a été réalisée avec le logiciel Pymol d'après les fichiers PDB 1MC0 et 1YKD.

III.1.6 Le mode de fixation des nucléotides cycliques

Les domaines GAF présentent une affinité très élevée pour leurs nucléotides spécifiques, de l'ordre du nanomolaire, ce qui pourrait s'expliquer par l'existence d'interactions conservées entre le domaine GAF et la partie invariante du nucléotide cyclique. En effet, le mode de fixation de l'AMPC dans les domaines GAF-A et GAF-B de l'AC est similaire à celui du GMPC au niveau du domaine GAF-B de la PDE-2A. Ce site de fixation est formé d'un côté par le feuillet β principal et de l'autre côté par les hélices α_3 et α_4 (Figure 19). Les nucléotides cycliques sont stabilisés *via* d'une part l'interaction entre le moment dipolaire de l'hélice α_3 et le groupement phosphate, et d'autre part un ensemble de liaisons hydrogène établies entre la chaîne principale du domaine GAF et l'oxygène du groupement phosphate.

Introduction

D'autre part, la spécificité pour l'AMPc ou le GMPc pourrait être liée à la présence ou à l'absence du résidu Arg291. En effet, l'Arg291 de l'AC CyaB2 établit une liaison hydrogène avec l'azote N1 de l'AMPc, liaison qui ne peut pas se former avec le N1 protoné du GMPc. A noter que dans le domaine GAF-B de la PDE-2A, l'équivalent du résidu Arg291 est le résidu Ile422 qui est en contact avec le noyau guanine du GMPc (Martinez *et al.*, 2002 ; Martinez *et al.*, 2005).

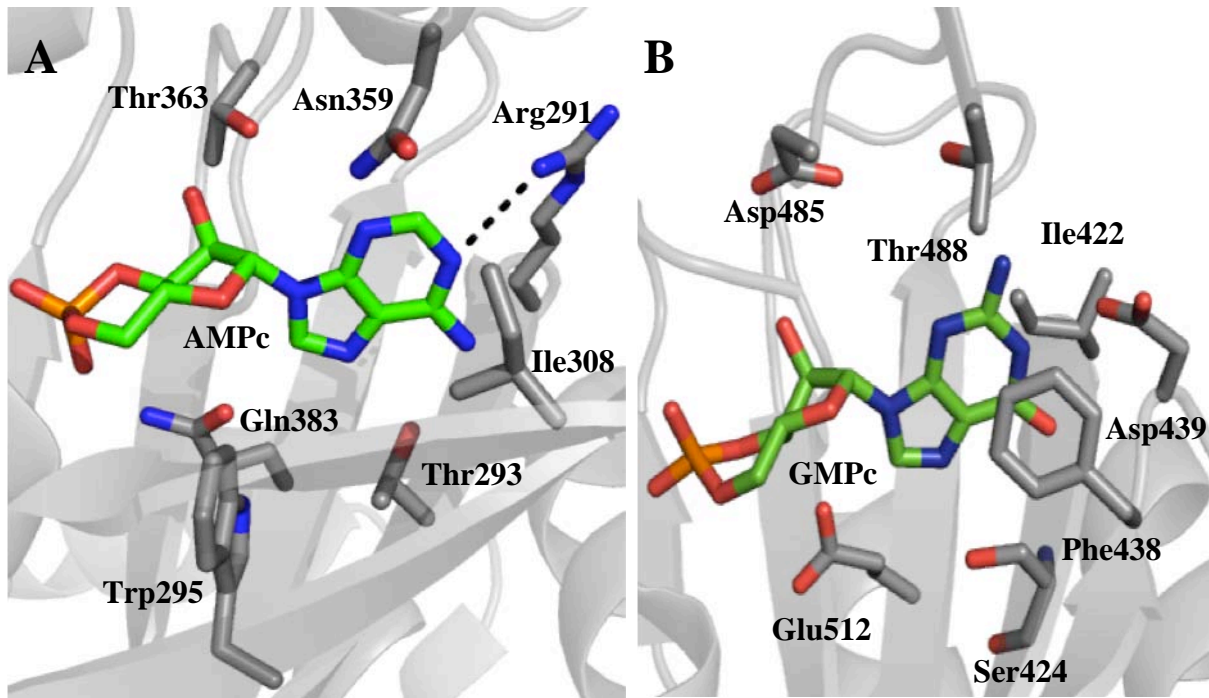


Figure 19 : Représentation schématique de la structure du site de fixation de l'AMPc de GAF-B de l'AC CyaB2 d'*Anabaena* (A), et de du site de fixation du GMPc de GAF-B de la PDE-2A de souris (B) (adapté de Martinez *et al.*, 2002 ; Martinez *et al.*, 2005).

Organisation du site de fixation des nucléotides cycliques AMPc (A) et GMPc (B) au niveau respectivement du domaine GAF-B des Ac CyaB2 et PDE-2A. L'Arg291 établit une liaison hydrogène avec le N1 de l'AMPc. Pour des raisons de clarté, sont représentés en mode bâton les résidus impliqués dans la stabilisation/reconnaissance du nucléotide cyclique. La figure a été réalisée avec le logiciel Pymol d'après les fichiers PDB 1MC0 et 1YKD.

Dans les PDE et les AC, les domaines GAF en tandem sont précédés d'une séquence d'environ 50 à 220 acides aminés selon la protéine considérée, qui pourrait jouer un rôle dans la fixation des nucléotides cycliques. En effet, il a été montré que la fixation du GMPc sur les domaines GAF de PDE-5 et PDE-11A4, et donc l'activation du domaine catalytique, ne se fait que si le domaine N-terminal des domaines GAF est présent. Par contre, la délétion du domaine N-terminal des domaines GAF de la PDE-10 n'a aucun effet sur la régulation de la fixation du GMPc (Zoraghi *et al.*, 2004, Bruder *et al.*, 2006, Gross-Langenhoff *et al.*, 2008).

L'analyse des alignements de séquences des domaines GAF qui fixent les nucléotides cycliques et qui sont en tandem de la famille des PDE et des AC, ont permis de mettre en évidence la présence de résidus conservés NKFDE (Figure 20) qui pourraient être impliqués de manière indirecte dans la fixation des nucléotides. L'analyse des structures 3D connues montre que les résidus NKFDE ne font partie ni de la poche de liaison des nucléotides ni de l'interface du dimère ; par contre ils sont proches les uns des autres, et également proches d'un réseau conservé de liaisons hydrogène et d'un pont salin (Figure 21) (Martinez *et al.*, 2002 ; Martinez *et al.*, 2005). Plusieurs études ont montré que la substitution du résidu Asp de ce motif dans le domaine GAF-A de la PDE5 par un résidu Ala conduit à une baisse de l'affinité pour les NMPc pouvant aller jusqu'à une abolition de la fixation (Ho *et al.*, 2000 ; McAllister-Lucas *et al.*, 1995 ; Turko *et al.*, 1996, Sopory *et al.*, 2003 ; Wu *et al.*, 2004 ; Kanacher *et al.*, 2002 ; Granovsky *et al.*, 1998).

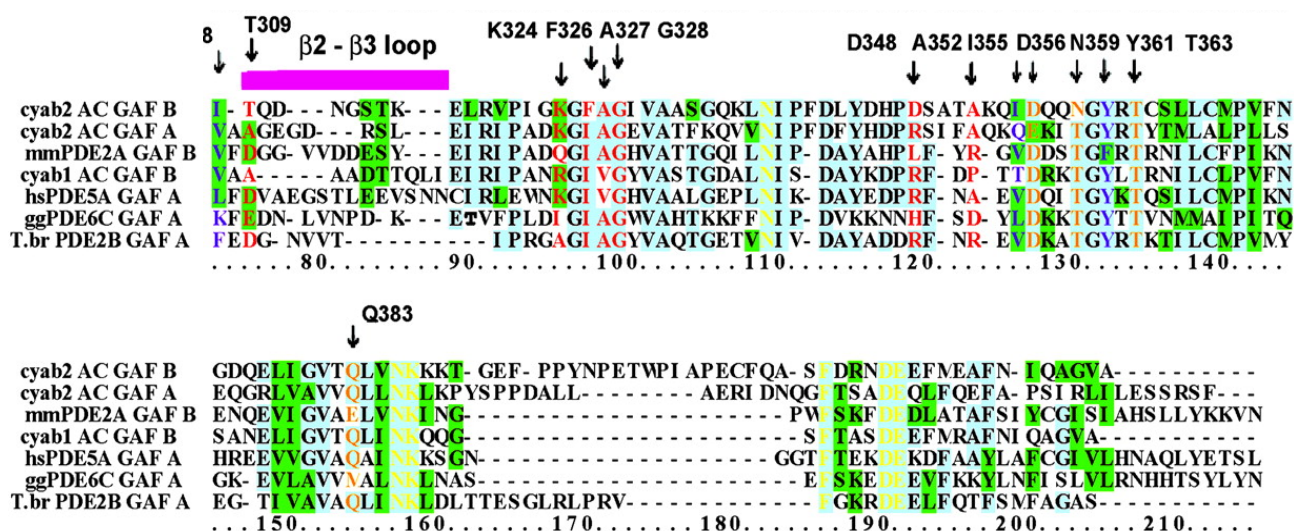
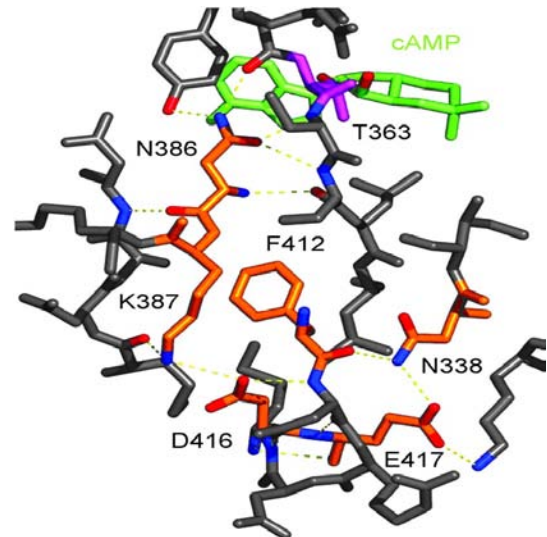


Figure 20 : Alignement des séquences des domaines GAF connus pour lier le GMPc ou l'AMPc (d'après Martinez *et al.*, 2005).

Les séquences correspondent aux domaines GAF-A et GAF-B de *cyab2* de l'AC, au domaine GAF-B de la PDE-2A de souris, au domaine GAF-B de *cyab1* de l'AC, au domaine GAF-A de la PDE-5A humaine, de la PDE-6C de poule, et de la PDE-2B de *Trypanosoma brucei*. La barre en magenta indique la position de la boucle β2-β3 qui est la partie la moins conservée des séquences des domaines GAF. Les 17 résidus du domaine GAF-B de *CyaB2* qui établissent un contact direct ou *via* une molécule d'eau avec l'AMPc, sont « listés » au dessus de l'alignement et indiqués en couleur dans l'alignement : orange pour les chaînes latérales polaires; pourpres pour des chaînes latérales hydrophobes, et rouge pour les groupements amides et carbonyles de la chaîne principale. Les résidus du motif NKFDE sont indiqués en jaune sur fond bleu clair. Le fond bleu désigne les résidus présentant une identité de séquences supérieure à 50%, et le vert une identité de séquence inférieure à 50%.

Figure 21 : Représentation schématique du motif NKFDE du domaine GAF-B de l'AC CyaB2 (d'après Martinez *et al.*, 2005).

Dans cette structure, le résidu Asn386 est enfoui et forme des liaisons hydrogène avec les résidus Tyr361, Arg362, Cys364, et Ser365 (non représentés). Le résidu Thr363 forme une liaison hydrogène avec le 2'-OH du ribose de l'AMPc. En pointillé sont représentées les interactions entre le motif NKFDE et les résidus proches de la protéine. Les résidus NKFDE sont représentés en orange, la Thr363 en rose, et l'AMPc en vert.



III.1.7 La signalisation vers le domaine effecteur

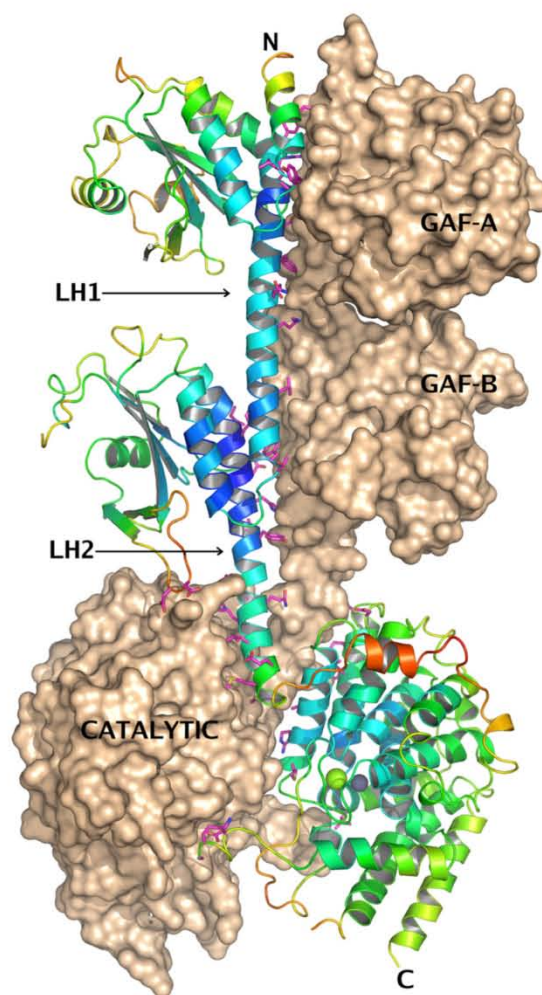
Plusieurs études ont montré que les domaines catalytiques isolés des PDE et des AC étaient parfaitement actifs, et que l'activation de ces domaines effecteurs nécessitait la présence des deux domaines GAF-A et -B. Ainsi, il semblerait que la fixation du nucléotide cyclique et la dimérisation soient nécessaires pour réguler l'activation du domaine effecteur. En 2009, la détermination de la structure 3D de la PDE-2A entière (domaine catalytique et domaines GAF-A et GAF-B mais sans la partie N-terminale) a permis de proposer un mécanisme de régulation de l'activité par les domaines GAF (Pandit *et al.*, 2009).

Dans le cristal, la PDE-2A est sous forme de dimère parallèle avec une interface de contact qui s'étend sur toute la longueur de la protéine. Il existe deux hélices « Linker » appelées LH 1 et 2, hydrophobes et parallèles à l'axe du dimère. L'hélice LH 1 relie les domaines GAF A et GAF B, et LH 2 le domaine GAF B et le domaine catalytique (Figure 22). De plus, deux conformations du site actif du domaine catalytique ont été observées : une conformation fermée dans laquelle l'accès au site de fixation du substrat est bloqué par une boucle flexible appelée boucle H, et une conformation ouverte dans laquelle l'accès au site actif est libre (Figure 23). Une deuxième boucle, appelée boucle M, subit également un changement conformationnel lors du passage de l'enzyme de la conformation fermée à la conformation ouverte. Ainsi, quand le domaine GAF-B est libre, l'accès au site actif du domaine catalytique est bloqué par l'autre partenaire du dimère, et la liaison du nucléotide à GAF-B provoquerait un changement conformationnel qui se propagerait au domaine catalytique et libérerait l'accès au site de liaison du substrat, permettant ainsi l'activation de la PDE (Figure 24). La PDE-2A étant active *in vitro* en absence de GMPC fixé sur le

domaine GAF B, il est possible que dans ce cas, la conformation du site actif soit en équilibre entre les deux conformations fermée et ouverte, et que la fixation du GMPC sur le domaine GAF-B soit un moyen de stabiliser le site actif dans la conformation ouverte. Par contre, le mécanisme moléculaire par lequel la fixation du GMPC sur le domaine GAF-B, induirait un changement conformationnel des boucles H et M reste encore inconnu.

Figure 22: Structure de PDE-2A humaine (d'après Pandit *et al.*, 2009).

L'orientation du dimère place l'axe de symétrie, reliant les deux molécules A et B, plus ou moins parallèle au plan du papier dans la direction verticale. Les trois domaines sont marqués dans la figure, ainsi que les hélices LH1 et LH2 qui les relient. La molécule B est représentée en mode surface et la molécule A en mode ruban, la coloration représentant la valeur du facteur B cristallographique (bleu faible et rouge élevé). Toutes les chaînes latérales de la molécule A qui sont à moins de 3,5 Å de la molécule B sont représentées en mode bâton de couleur magenta. L'interface du dimère s'étend sur la surface de la molécule entière. Les deux sites catalytiques contenant des ions Zn^{2+} et Mg^{2+} (représentés comme des sphères grises et vertes) se ferment mutuellement au niveau de l'interface du dimère.



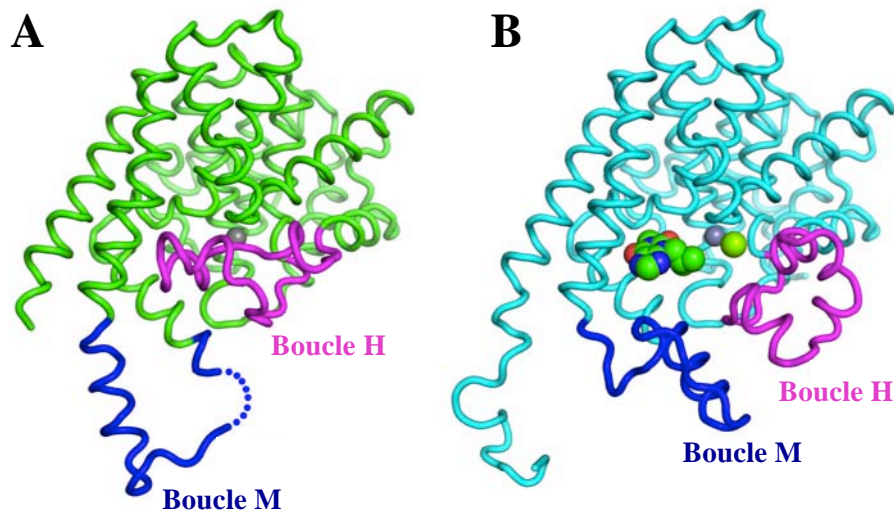


Figure 23 : Représentation des structures « cristallines » des conformations ouvertes (A) et fermées (B) de la PDE-2A (d’après Pandit *et al.*, 2009).

La chaîne principale du domaine catalytique isolé (579–919) est représentée en vert ou bleu clair. Les boucles H (700–723) et M (830–856) sont représentées respectivement en magenta et bleu foncé. (A) L’ion Zn^{2+} est représenté comme une sphère grise. (B) L’inhibiteur IBMX présent dans le site actif est représenté en mode sphère. Les ions Zn^{2+} et Mg^{2+} sont représentés respectivement par des sphères grise et verte.

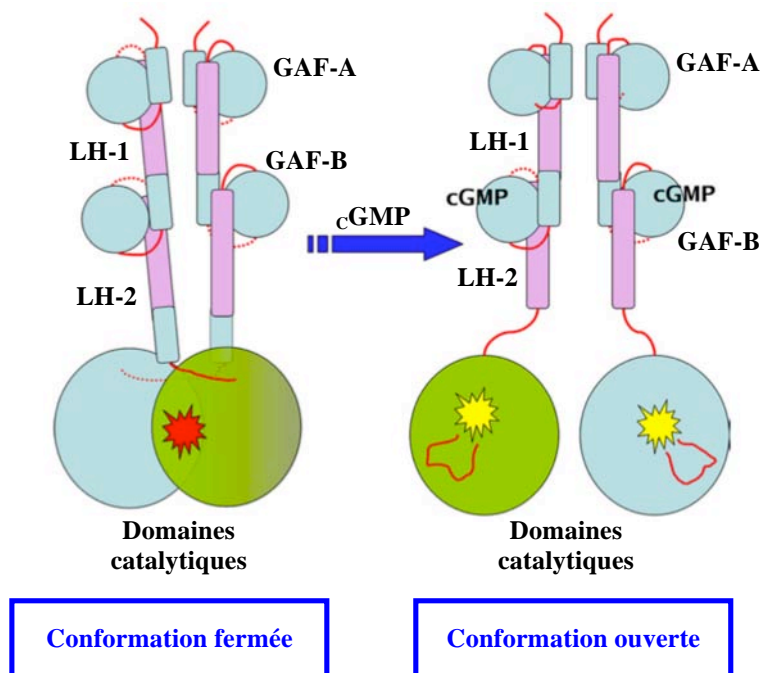


Figure 24: Mécanisme d’activation proposé (d’après Pandit *et al.*, 2009).

La liaison des nucléotides cycliques au domaine GAF-B s’accompagne d’une rotation des deux hélices LH2 provoquant un éloignement des deux domaines catalytiques l’un par rapport à l’autre. La boucle H, qui bloquait le site de fixation du substrat du domaine catalytique dans la conformation fermée, libère l’accès au site de fixation du substrat dans la conformation ouverte.

III.2 Les fRMsr

Lorsque ce travail a été initié, seule la fRMsr d'*E. coli* avait été identifiée. Bien que l'existence de cette activité Met-O réductase ait été suggérée dès 2002 suite à l'étude d'une souche d'*E. coli* MsrA⁻/MsrB⁻ (Etienne *et al.*, 2003), ce n'est qu'en 2007 que la protéine responsable de cette activité a été identifiée, après purification à partir d'extraits cellulaires de la souche d'*E. coli* MsrA⁻/MsrB⁻, et appelée fRMsr pour « free R Methionine sulfoxide reductase ». La caractérisation *in vitro* d'une forme recombinante de la fRMsr d'*E. coli* a montré qu'elle catalyse la réduction spécifiquement de la Met-R-O libre avec une activité spécifique comparable à celle de la MsrA, et qu'elle présente une activité Trx-dépendante (Lin *et al.*, 2007).

III.2.1 Distribution dans les trois domaines du vivant et rôle physiologique

L'analyse comparative des génomes d'organismes complètement séquencés (en utilisant les banques de données présentées sur le site de la NCBI) a permis de montrer que la présence du gène de la fRMsr est limitée aux organismes unicellulaires des trois domaines du vivant, à savoir les *Archaea*, les bactéries, et les eucaryotes unicellulaires, alors qu'il n'est pas retrouvé chez les organismes multicellulaires, comme les plantes et les mammifères. En outre, il a été observé que le nombre de séquences génomiques de fRMsr est très restreint par rapport au nombre de séquences de MsrA et MsrB ; et elles sont présentes en une seule copie dans les organismes contenant cette protéine (Le *et al.*, 2009).

D'autre part, cette analyse a également montré l'existence de protéines contenant un domaine fRMsr fusionné à un autre domaine, chez les bactéries et les eucaryotes unicellulaires. C'est le cas des *Kinetoplastidae* (*Leishmania* et *Trypanosoma*) chez lesquelles le domaine fRMsr est fusionné du côté C-terminal à un domaine homologue à la protéine TIP41. Cette dernière serait impliquée dans la régulation des voies d'assimilation de l'azote par la voie TOR (TOR pour « Target Of Rapamycin ») mais également dans la régulation de la phosphatase 2A dans le cas de la levure (Smetana *et al.*, 2007). Cependant, la relation fonctionnelle entre TIP41 et la fRMsr n'est pas claire. C'est le cas également des *Thermotogae* chez lesquelles le domaine fRMsr est fusionné du côté N-terminal à un domaine d'une centaine de résidus d'acides aminés de fonction inconnue.

Enfin, l'étude de mutants de *S. cerevisiae* n'exprimant plus la fRMsr a permis de suggérer que la fRMsr a un rôle antioxydant (Le *et al.*, 2009). Ainsi, le système d'oxydation

de la Met et de réduction des Met-O par la fRMsR pourrait constituer un système antioxydant cellulaire permettant de piéger les ERO. De plus, la fRMsR aurait un rôle important dans le métabolisme de la Met en conditions de stress oxydant. En effet, elle permettrait, à partir de Met-O, de fournir de la Met pour le métabolisme cellulaire.

D'un point de vue évolution, le fait que les fRMsR, qui font partie de la famille structurale des protéines de type GAF, ont évolué pour acquérir une activité Msr, spécifique des Met-R-O, suggère que les fRMsR pourraient être impliquées dans des mécanismes de régulation liée au stress oxydant et médiée par les Met-O.

III.2.2 Mécanisme

L'activité fRMsR nécessitant la présence de la Trx, Lin *et al.* ont proposé un mécanisme comparable à celui de la MsrA d'*E. coli* impliquant : 1) la formation d'un intermédiaire acide sulfénique sur la Cys94, 2) la formation d'un premier pont disulfure Cys84-Cys94, puis 3) la formation d'un second pont disulfure Cys84-Cys118 réduit par la Trx (Lin *et al.*, 2007).

III.2.3 Structure tridimensionnelle

La détermination en 2000 de la structure tridimensionnelle de la première protéine à domaine GAF a été celle de la protéine ykg9 de levure qui s'est révélée être une fRMsR (voir paragraphe III.1.5 page 27). Les fRMsR sont composées d'un seul domaine GAF d'environ 160 acides aminés, et sont sous forme dimérique en solution.

La détermination de la structure cristalline de la fRMsR d'*E. coli* en 2005, lors du projet "Structural GenomiX bacterial genomics", montrait que le site actif est occupé par une molécule de MES « 2-(N-Morpholino) Ethanesulfonic Acid », issue des conditions de cristallisation, et contient notamment les trois résidus Cys (84, 94 et 118), qui pourraient être impliqués dans le mécanisme (Badger *et al.*, 2005) (figure 25). La superposition des structures cristallines de la fRMsR d'*E. coli* et de la protéine YKG9 de *S. cerevisiae* indique que les résidus du site actif sont conservés, en particulier les 3 résidus Cys, suggérant que la protéine YKG9 serait elle aussi une Met-(R)-SO réductase. Enfin, l'inspection du site actif montre qu'il forme une cavité dont la taille ne permet pas l'entrée de peptides, ce qui pourrait expliquer la spécificité de la fRMsR pour la Met-O libre.

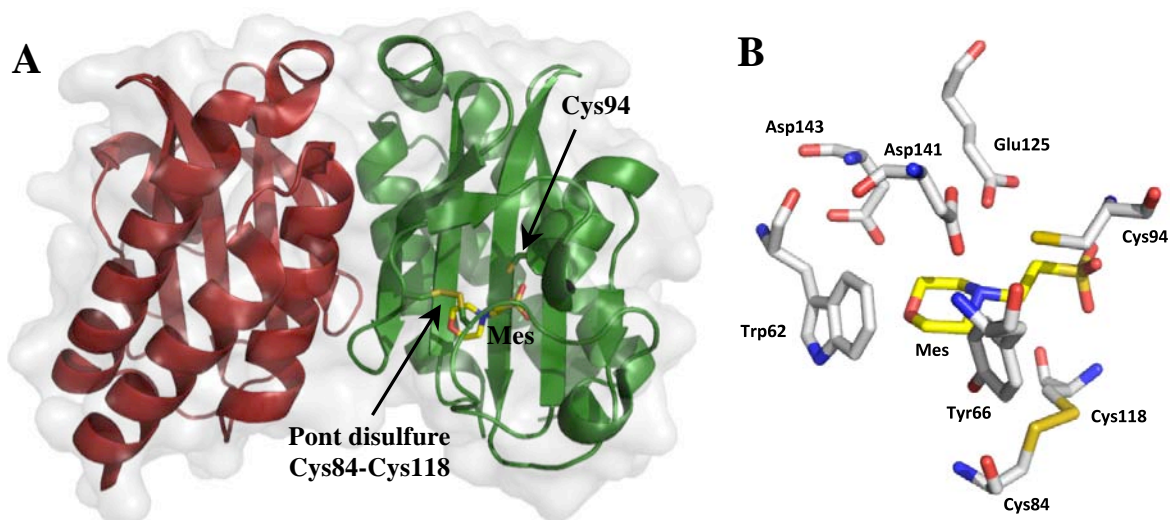


Figure 25 : Représentation schématique de la structure du dimère (A) et du site actif (B) de la fRMsR d'*E. coli* en complexe avec une molécule de Mes.

A) Architecture globale de la fRMsR d'*E. coli*. Les chaînes latérales des résidus Cys84, Cys94 et Cys118 sont représentées en mode bâton avec les atomes de carbone en vert, d'azote en bleu, d'oxygène en rouge et de soufre en jaune foncé. Le site actif est occupé par une molécule de Mes qui est représentée en mode bâton avec les atomes de carbone en vertjaune, d'azote en bleu, d'oxygène en rouge et de soufre en jaune.

B) Les résidus d'acides aminés et la molécule de Mes sont représentés en mode bâton avec les atomes de carbone respectivement en gris et jaune (N en bleu, O en rouge et S en jaune foncé). Sont représentés les résidus Cys (Cys94 et pont disulfure Cys84-Cys118) et les résidus acides (Glu125, Asp141 et Asp143) et hydrophobes (Trp62 et Tyr66).

La figure a été réalisée avec le logiciel Pymol d'après le fichier PDB 1VHM.

OBJECTIFS

Avant que je ne débute mes travaux de thèse, les travaux réalisés par l'équipe Enzymologie Moléculaire avaient porté sur l'étude des MsrA et MsrB au niveau catalyse, spécificités structurales et relations structure-fonction. L'équipe avait ainsi mis en évidence l'existence pour les deux classes de Msr d'un mécanisme nouveau de réduction d'une fonction sulfoxyde basée sur la réactivité des Cys, associée à la chimie de l'acide sulfénique, et impliquant une catalyse acide assistée par le substrat. Les deux classes de Msr, bien que structurellement distinctes, ont un mode de reconnaissance structurale reposant majoritairement sur des interactions de type van der Waals avec une poche hydrophobe dans laquelle se fixent les groupements méthyle ϵ et méthylène γ du substrat Met-O.

En 2007, une nouvelle activité Msr portée par une structure de type GAF a été mise en évidence chez *E. coli*. Cette nouvelle Msr, appelée fRMsr, reconnaît spécifiquement l'isomère *R* des Met-O libres, les Met-*R*-O incorporées dans un peptide et le DMSO n'étant pas réduits. Elle utilise la Trx comme réducteur pour régénérer l'activité. La structure cristallographique de la fRMsr d'*E. coli*, obtenue dès 2005 suite à un projet de génomique structurale alors que sa fonction était inconnue, montrait une structure dimérique de type GAF avec une molécule de MES située dans une poche pouvant être le site actif (Badger *et al.*, 2005). Rien n'était connu concernant le mécanisme, la catalyse et la spécificité structurale de cette nouvelle classe d'enzyme, lorsque j'ai débuté mes travaux de master 2.

Dans ce contexte, mon projet de thèse avait pour objectif la caractérisation du mécanisme et l'identification des facteurs moléculaires et/ou structuraux impliqués dans la catalyse et la spécificité de substrat de la fRMsr de *N. meningitidis*, utilisée comme modèle d'étude. Le projet s'articulait autour de trois axes :

- Le premier axe concernait la caractérisation du mécanisme de la fRMsr. Le mécanisme devant faire intervenir au moins un résidu Cys, et la fRMsr de *N. meningitidis* possédant quatre résidus Cys, l'implication éventuelle de chacun de ces résidus dans le mécanisme réactionnel a été étudiée par une approche de mutagenèse dirigée couplée à la caractérisation des propriétés catalytiques des différents mutants générés.

- Le second axe concernait la caractérisation cinétique des différentes étapes du mécanisme de la fRMsr, et l'identification de l'étape cinétiquement limitante de la réaction globale. Ceci a nécessité la mise au point de différentes approches de cinétique rapide.

Objectifs

- Le troisième axe était consacré à l'identification des acides aminés impliqués dans la catalyse de l'étape réductase et dans la spécificité de substrat. Ceci a été fait à partir de l'étude des propriétés cinétiques de mutants générés par mutagenèse dirigée.

Les résultats sont donc présentés en trois parties, chacune correspondant à un axe. Parallèlement à mes travaux, la structure tridimensionnelle de la fRMs_r de *N. meningitidis* en complexe avec son substrat *L*-Met-*R*-O a été résolue par le Dr A. Gruez de l'équipe. Les résultats concernant la caractérisation du mécanisme des fRMs_r et la structure tridimensionnelle ont été publiés dans *J. Biol. Chem.* en 2011 et l'article est joint en annexe. Et également jointe en annexe une Note publiée dans *J. Biol. Chem.* en 2011 suite à la publication par des concurrents de la structure de l'enzyme de *Staphylococcus aureus* en complexe avec le substrat.

L'ensemble de mes résultats, publiés et non publiés, est discuté dans la partie Discussion/Perspectives en relation avec les données structurales disponibles.

RESULTATS

Partie I : Caractérisation du mécanisme de la fRMsr de *N. meningitidis*

I.1 Stratégie utilisée

Au moment où le projet concernant les fRMsr a été initié au sein de l'équipe, peu d'informations étaient disponibles concernant le mécanisme. En effet, comme il a été indiqué dans l'Introduction, c'est l'équipe de T. Lowther qui a mis en évidence chez *E. coli* l'existence des fRMsr et montré qu'elles présentaient une activité dépendante de la Trx (Lin *et al.*, 2007). En termes de mécanisme, cela supposait l'existence d'un mécanisme comparable à celui des MsrA et MsrB avec formation d'un intermédiaire acide sulfénique sur la Cys catalytique, puis d'au moins un intermédiaire pont disulfure, entre une Cys catalytique et une Cys de recyclage, réduit par la Trx.

L'analyse de l'alignement des 358 séquences de fRMsr putatives associée à l'examen de la structure tridimensionnelle de la fRMsr d'*E. coli* en complexe avec une molécule de Mes montrait que les fRMsr possédaient 25 résidus conservés à plus de 90%, dont trois résidus Cys en position 84, 94 et 118 présents dans le site actif qui pouvaient être impliqués dans le mécanisme. La Cys118 est conservée dans toutes les fRMsr alors que les Cys84 et Cys94 le sont respectivement à 94,5% et 98,6% (voir Figure SI annexe 1). D'où les questions qui se posaient concernant le nombre (deux ou trois) et la position des résidus Cys de la fRMsr impliqués dans le mécanisme.

Afin de déterminer le mécanisme des fRMsr, une étude de relations structure-fonction par ingénierie a été entreprise sur la fRMsr de *N. meningitidis* utilisée comme modèle. Le choix de cette fRMsr a été dicté par la nécessité de pouvoir produire et purifier l'enzyme en grande quantité, ce qui est le cas pour l'enzyme de *N. meningitidis* alors que la fRMsr d'*E. coli* n'a pas pu être produite en grande quantité et ceci pour une raison inconnue. A noter que la fRMsr de *N. meningitidis* possède quatre Cys : Cys84, Cys94 et Cys118, et une quatrième non conservée en position 136 (Fig. 26 page 41). Afin de déterminer le rôle de chaque résidu Cys, ceux-ci ont été substitués individuellement ou non par des résidus Ala. Ainsi les mutants suivants ont été générés C84A, C94A, C118A, C136A, C84A/C136A, C94A/C136A, C84A/C94A/C136A et les propriétés des différents mutants générés ont été comparées à celles de l'enzyme sauvage.

Résultats : Mécanisme de la fRMs_r de *N. meningitidis*

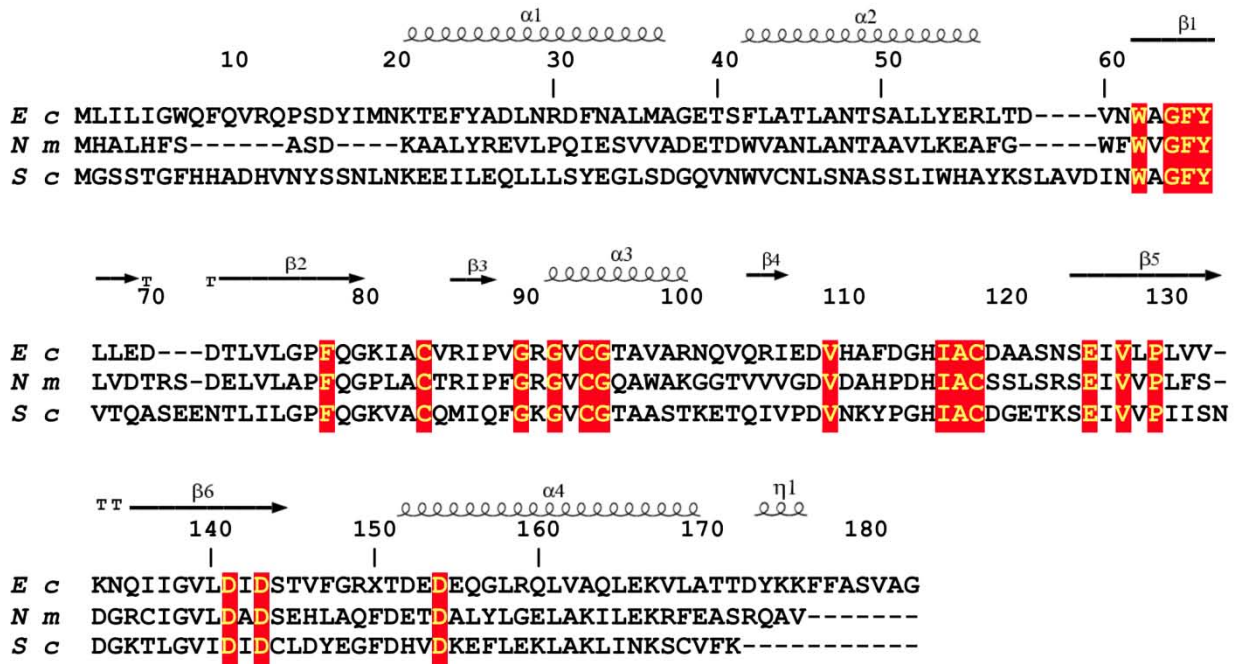


Figure 26 : Alignement des séquences des fRMs_r d'*E. coli* (E c), de *N. meningitidis* (N m), et de *S. cerevisiae* (S c).

La numérotation utilisée est basée sur celle de l'enzyme d'*E. coli*. Les structures secondaires indiquées sont celles correspondant à l'enzyme d'*E. coli* (Badger *et al.*, 2005) : les hélices α sont indiquées par des traits hélicoïdaux, les brins β par des flèches. Les résidus conservés à plus de 80 % dans les séquences de fRMs_r putatives (alignement réalisé sur 358 séquences de fRMs_r putatives extraites des banques protéiques (voir Figure SI annexe 1)) sont représentés en jaune sur fond rouge.

Pour étudier les propriétés des différents mutants, la stratégie utilisée a été d'étudier l'activité de réduction de la Met-O en absence de réducteur, en condition de cycle catalytique unique. Ainsi, la quantification de la Met formée par chromatographie hydrophobe a permis de montrer l'existence ou non d'une activité réductase des différents mutants, et en parallèle, le dosage des thiols libres a permis de montrer la formation de différents intermédiaires oxydés. De plus, les paramètres cinétiques (k_{cat} et K_M) de la fRMs_r sauvage et des différents mutants Cys ont été déterminés à l'état stationnaire en présence de Met-O et de la Trx d'*E. coli*, afin de déterminer quels intermédiaires oxydés étaient réduits efficacement par la Trx. Enfin, des études complémentaires ont été réalisées afin d'essayer de caractériser les différents intermédiaires réactionnels.

Il est à noter que le substrat appelé Met-O est en fait un mélange des 4 diastéréoisomères : *D*-Met-*R*-O, *D*-Met-*S*-O, *L*-Met-*R*-O et *L*-Met-*S*-O.

I.2 Production et purification des fRMsr sauvage et mutées de *N. meningitidis*

Le vecteur permettant la production de la fRMsr sauvage de *N. meningitidis* a été obtenu par sous-clonage aux sites *NdeI* et *SacI* de produits d'amplification, issus de la technique de polymérisation de chaîne (PCR) réalisée sur l'ADN génomique de *N. meningitidis*, dans un phagemide bluescript. Ceci a permis de placer la séquence codante sous la dépendance du promoteur P_{Lac}. La construction plasmidique obtenue, appelée pSK-fRMsr, a été vérifiée par séquençage du fragment inséré. A partir de cette construction, les différents plasmides mutants ont été générés par mutagenèse dirigée.

Les fRMsr sauvage et mutées ont été produites dans des cellules d'*E. coli* HB101 transformées respectivement par le plasmide recombinant pSK-fRMsr, ou un plasmide recombinant pSK-fRMsr muté portant la(les) mutation(s) désirée(s) au niveau de la séquence codante correspondante du gène *fRMsr* de *N. meningitidis*. Le taux de production des mutants, estimé à partir d'analyses par électrophorèse SDS-PAGE d'extraits bruts d'échantillons de culture (gels non présentés), était plus faible que celui de l'enzyme sauvage (d'un facteur 2 à 5), mais restait suffisant pour permettre leur étude.

Les fRMsr sauvage et mutées ont été purifiés selon un protocole identique, mis au point pour l'enzyme sauvage, et qui consiste en : 1) une sonication des cellules permettant de récupérer l'enzyme dans la fraction soluble, 2) une précipitation à 50% de saturation en sulfate d'ammonium, 3) une chromatographie par interactions hydrophobes Phényl-Sépharose, 4) une chromatographie échangeuse d'anions Q-Sépharose, et 5) une concentration sur cellule Amicon des fractions issues de la purification. Les profils d'élution des différents mutants sur les différentes colonnes de chromatographie sont identiques à ceux de l'enzyme sauvage, sauf pour la fRMsr C118A. En effet, la fRMsr C118A n'est pas retenue sur la colonne de chromatographie échangeuse d'anions Q-Sépharose contrairement à l'enzyme sauvage dans les mêmes conditions. Cependant, le spectre de dichroïsme circulaire de la fRMsr C118A étant parfaitement superposable à celui de la fRMsr sauvage et des autres mutants purifiés, ceci suggère que la fRMsr C118A est correctement structurée (résultat non présenté).

La pureté des protéines obtenues a été vérifiée par électrophorèse SDS-PAGE 12,5% (Fig. 27 page 43). Après concentration sur cellule Amicon des fractions issues de la purification, un rendement moyen de 50 mg d'enzyme /litre de culture est obtenu avec un rapport DO₂₈₀/DO₂₆₀ de l'ordre de 1,8.

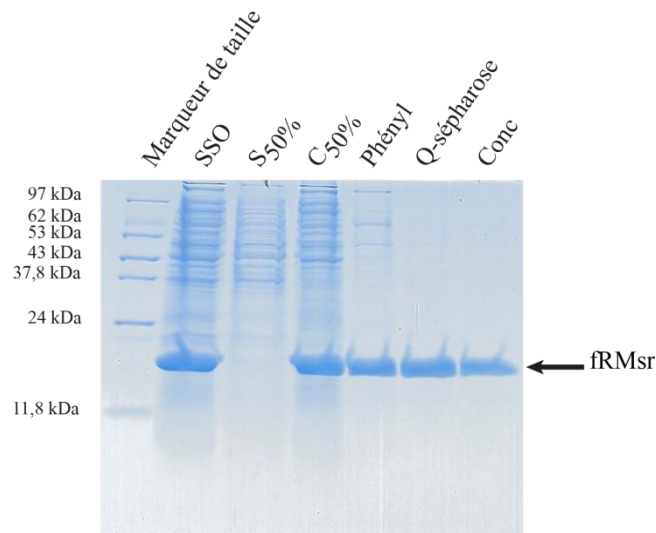


Figure 27 : Electrophorèse SDS-PAGE 12,5% des diverses étapes de la purification de la fRMs_r sauvage de *N. meningitidis*.

Marqueur de taille protéique contenant l'ARN polymérase T7 (97 kDa), l'albumine de sérum bovin (62 kDa), la rétinol déshydrogénase (53 kDa), la protéine TruD (43 kDa), la GAPDH (37,8 kDa), la MsrA (24 kDa) et la Trx (11,8 kDa). SSO, S50%, C50%, Phényl, Q-sépharose et Conc représentent respectivement, le surnageant de sonication, le surnageant de précipitation à 50% de saturation en sulfate d'ammonium, le culot de précipitation à 50% en sulfate d'ammonium, l'éluat de la chromatographie par interactions hydrophobes Phényl-Sépharose, l'éluat de la chromatographie échangeuse d'anions Q-Sépharose et la protéine purifiée après concentration.

D'autre part, les analyses par spectrométrie de masse en conditions dénaturantes des différentes fRMs_r ont permis de montrer la présence d'une population majoritaire de masse correspondant à la masse attendue (résultats non montrés). Enfin, les analyses de la fRMs_r sauvage d'une part par spectrométrie de masse en conditions natives (résultats non montrés) et d'autre part par SLS ont montré que l'enzyme purifiée est comme attendu dimérique

I.3 Dosage des groupements thiols libres avant et après ajout de Met-O, détermination de la stœchiométrie de formation de Met en absence de réducteur et détermination des pK_{app} des Cys

Les groupements thiols ont été titrés avant et après ajout de Met-O en conditions natives et dénaturantes en utilisant une sonde chimique spécifique des résidus Cys, le 5,5'-dithio-bis (2-nitro) benzoate (DTNB). Le DTNB réagit avec les fonctions thiolates et la réaction libre

du thionitrobenzoate (TNB⁻) qui est un chromophore absorbant à 412 nm ($\epsilon_{\text{TNB}^-}_{412\text{nm}} = 13600 \text{ M}^{-1} \cdot \text{cm}^{-1}$).

En conditions natives, deux Cys par monomère ont réagi avec le DTNB pour l'enzyme sauvage et le mutant C94A/C136A (tableau 1 page 45). La réaction des deux Cys avec le DTNB peut conduire à la formation soit de deux adduits thionitrobenzoate (TNB), soit d'un pont disulfure intramoléculaire. Pour différencier ces deux possibilités, un réducteur (DTT) a été ajouté aux fRMsr oxydées, après élimination du TNB⁻ et de l'excès de DTNB par chromatographie d'exclusion de taille. Aucune augmentation de l'absorbance à 412 nm correspondant à la libération de TNB⁻ a été observée (résultats non montrés). Ces résultats indiquent que les Cys118 et Cys84 sont accessibles au DTNB dans l'enzyme sauvage, et qu'elles forment un pont disulfure en présence de DTNB. Pour les mutants C84A/C136A et C118A/C136A, deux Cys étaient également titrés en conditions natives, correspondant respectivement aux Cys94 et Cys118 et Cys84 et Cys94 (tableau 1 page 45). Contrairement à l'enzyme sauvage, une augmentation de l'absorbance à 412 nm correspondant à la libération de deux moles de TNB⁻ par monomère a été observée après ajout de DTT. Ces résultats indiquent que la Cys94 devient accessible au DTNB lorsque la Cys84 ou la Cys118 est absente, et que les ponts disulfures Cys94-Cys118 et Cys84-Cys94 ne se forment pas en présence de DTNB, même en absence respectivement de la Cys84 et de la Cys118.

La Met formée après 10 min de réaction entre 150 à 300 μM d'enzyme et 100 mM de Met-O a été quantifiée par chromatographie en phase inverse sur une colonne C18 à partir d'une gamme étalon de Met. Les résultats montrent la libération d'environ 1 mole de Met par monomère d'enzyme pour l'enzyme sauvage et tous les mutants, à l'exception du mutant C118A pour lequel aucune Met n'est détectée (tableau 1 page 45).

| fRMsr | Nombre de thiols titrés par monomère d'enzyme | Nombre de thiols titrés par monomère d'enzyme | | | | | | Stœchiométrie de formation de la Met (mol Met/mol monomer) |
|-----------------|-----------------------------------------------|-----------------------------------------------|-------------------------|----------------------|-------------------------|---------------------------------------------|-------------------------|------------------------------------------------------------|
| | | Avant ajout de Met-O | | Après ajout de Met-O | | Différence du nombre de thiols ^a | | |
| | | Conditions natives | Conditions dénaturantes | Conditions natives | Conditions dénaturantes | Conditions natives | Conditions dénaturantes | |
| Sauvage | 4 | 2,2 | 3,9 | 0 | 2,1 | 2,2 | 1,8 | 1,0 |
| C84A | 3 | 1,9 | 2,9 | 0 | 1,0 | 1,9 | 1,9 | 1,0 |
| C94A | 3 | 2,1 | 3,0 | 0 | 1,0 | 2,1 | 2,0 | 1,1 |
| C118A | 3 | N.D. | 3,1 | N.D. | 3,0 | N.D. | 0,1 | 0 |
| C136A | 3 | 2,1 | 3,1 | 0 | 1,1 | 2,1 | 2 | 1,1 |
| C84A/C136A | 2 | 1,8 | 2,1 | 0 | 0,1 | 1,8 | 2 | 0,9 |
| C94A/C136A | 2 | 1,9 | 1,9 | 0 | 0 | 1,9 | 1,9 | 0,8 |
| C84A/C94A/C136A | 1 | 1,0 | 1,0 | 0 | 0 | 1,0 | 1,0 | 0,9 |

Tableau 1 : Titration des groupements thiols libres de la fRMsr sauvage et de différents mutants avant et après ajout de Met-O, et stœchiométrie de formation de la Met.

Les dosages ont été effectués en tampon Tris-HCl 50 mM, pH 8,0, EDTA 2 mM, dans des conditions natives ou dénaturantes (10% de SDS), pour deux concentrations en enzyme (7,35 μ M et 14,7 μ M). La concentration en DTNB était de 300 μ M. Dans les expériences de titration du nombre de thiols sur la forme oxydée, la Met-O est ajoutée à l'enzyme sous forme réduite à une concentration finale de 100 mM. Le mélange est incubé 10 min à 25°C. Les stœchiométries sont calculées à partir de la quantité de Met formée mesurée par chromatographie. L'erreur expérimentale sur les dosages est de $\pm 10\%$. ^a Représente la différence du nombre de thiols par monomère d'enzyme après ajout de Met-O. N.D. : non déterminé.

Parallèlement, le dosage des groupements thiols libres des enzymes a été réalisé en conditions dénaturantes avant et après ajout de Met-O. Comme indiqué dans le tableau 1, deux Cys par monomère n'étaient plus titrés après ajout de Met-O pour toutes les fRMsr, sauf pour le mutant C118A, pour lequel aucune différence n'est observée, et le mutant C84A/C94A/C136A pour lequel une seule Cys n'est plus titrée.

L'ensemble de ces résultats montre que 1) la fRMsr sauvage et le mutant C136A présentent une activité réductase en présence de Met-O avec la formation d'un intermédiaire pont disulfure, 2) le mutant C118A ne présente pas d'activité réductase en présence de Met-O, alors que le mutant C84A/C94A/C136A, qui ne contient que la Cys118, présente une activité Met-O réductase avec une stœchiométrie d'environ 1 mole de Met par monomère d'enzyme. La Cys118 est donc la Cys catalytique, et 3) les fRMsr C84A, C94A, C84A/C136A et C94A/C136A présentent une activité réductase en présence de Met-O avec la formation d'un pont disulfure Cys94-Cys118 ou Cys84-Cys118. Ainsi, les Cys84 et Cys94 peuvent former un pont disulfure avec la Cys118 au cours de l'acte catalytique.

Les pK_a des deux Cys titrables en conditions natives de la fRMsr sauvage ont été déterminés en utilisant une autre sonde chimique spécifique des résidus Cys, le 2,2'-dithiodipyridine (2PDS) pour des pH allant de 5,5 à 9,5 pour lesquels l'enzyme est stable. Cette sonde réagit avec les fonctions thiolates et la réaction libère de la pyridine-2-thione qui est un chromophore absorbant à 343 nm ($\epsilon_{343\text{nm}} = 8800 \text{ M}^{-1} \cdot \text{cm}^{-1}$). La réaction du 2PDS avec fRMsr suit une cinétique de pseudo-premier-ordre avec libération de 2 moles de pyridine-2-thione par mole de monomère d'enzyme. Ce résultat est similaire à celui obtenu en présence de DTNB. De plus, aucune augmentation de l'absorbance à 343 nm correspondant à la libération de pyridine-2-thione a été observée lorsque du DTT a été ajouté à la fRMsr oxydé (résultats non montrés). Ceci montre que, comme en présence de DTNB, un pont disulfure est formé. La courbe k_2 en fonction du pH, présentée dans la figure 28 page 47, s'ajuste à une équation de type monosigmoïde avec un pK_{app} de 8,5 et un $k_{2\text{max}}$ de $1,5 \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$. De plus, le fait que les valeurs de pK_{app} et de $k_{2\text{max}}$ obtenues soient proches de celles d'une Cys libre comme dans le glutathion (Marchal et Branlant, 1999), montre qu'au moins une des deux Cys de la fRMsr est accessible dans l'enzyme libre.

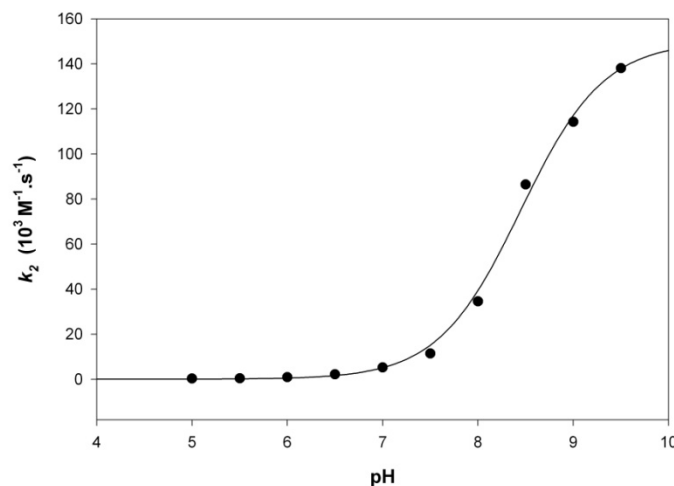


Figure 28 : Variation de la constante de seconde-ordre k_2 de la réaction des groupements thiols de la fRMSr sauvage de *N. meningitidis* avec le 2PDS, en fonction du pH.

Les expériences ont été réalisées à 25°C avec un polytampon composé de 120 mM Tris-HCl, 30 mM acide acétique, et 30 mM imidazole de force ionique constante de 0,15 M et de pH allant de 5 à 9,5. La concentration en fRMSr est de 12,5 μM , celle de 2PDS de 200 μM . Les valeurs de k_2 ont été ajustées à l'équation d'une monosigmoïde ($y = k_{2\text{max}} / (1 + 10^{(pK_a - \text{pH})})$) et les paramètres obtenus sont les suivants : $pK_{\text{app}} = 8,5 \pm 0,1$, et $k_{2\text{max}} = (1,50 \pm 0,04) \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$. Les points représentent les valeurs expérimentales et le trait continu l'ajustement. Chaque point représente la moyenne de cinq mesures, et l'expérience a été réalisée deux fois de façon indépendante.

I.4 Détermination des paramètres cinétiques des fRMSr sauvage et mutées à l'état stationnaire

Les paramètres cinétiques k_{cat} et K_M à l'état stationnaire ont été déterminés en présence de Met-O, de la Trx d'*E. coli*, et du système couplé thiorédoxine réductase / NADPH en suivant la diminution de l'absorbance à 340 nm, suite à l'oxydation du NADPH. Les résultats obtenus sont présentés dans la figure 29 page 48 et la figure 30 page 49, et les valeurs des constantes cinétiques correspondantes sont reportées dans le tableau 2 page 50.

Le mutant C118A ne présente pas d'activité Trx-dépendante, en accord avec l'absence d'activité réductase observée. Comme indiqué dans le tableau 2, tous les autres mutants y compris le triple mutant C94A/C84A/C136A présentent une activité en présence de Met-O comme substrat et de Trx comme réducteur (Fig. 29 page 48). Les valeurs de k_{cat} varient de 8,5 à 18 s^{-1} pour l'enzyme sauvage et les mutants dans lesquels la Cys84 est présente, et de 1 à 2,6 s^{-1} , lorsque la Cys84 est absente. Les valeurs de K_M pour la Met-O varient de 2 pour l'enzyme sauvage jusqu'à 32 mM pour le mutant C84A, et celles pour la Trx de 15 pour le

mutant C94A/C136A jusqu'à 50 μM pour le mutant C84A (Fig. 29 page 48 et Fig. 30 page 49).

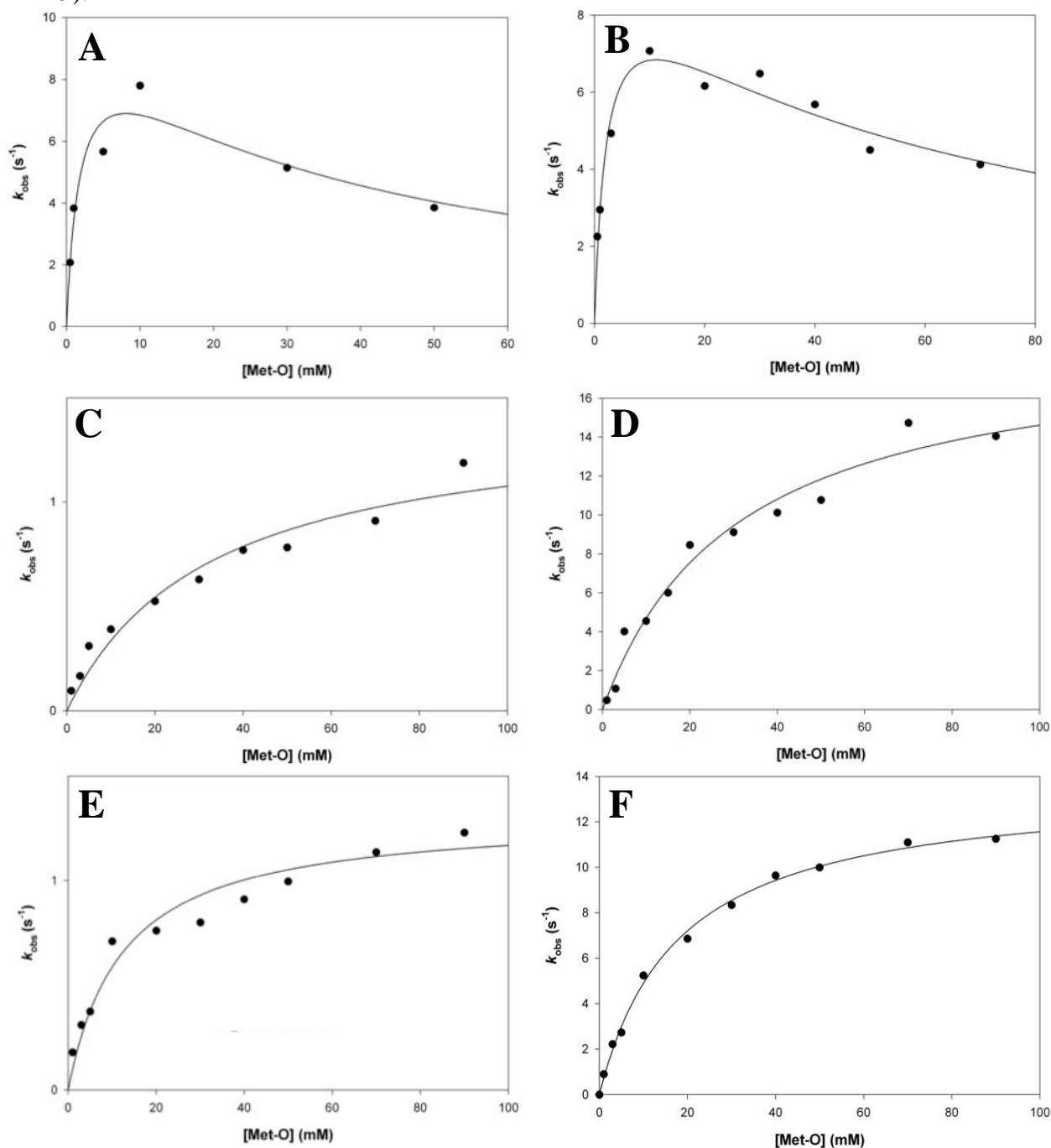


Figure 29 : Détermination des paramètres cinétiques, k_{cat} et K_{M} pour le substrat Met-O, des fRMsR sauvage (A), C136A (B), C84A (C), C94A (D), C84A/C136A (E) et C94A/C136A (F) de *N. meningitidis*.

Les activités enzymatiques ont été mesurées à l'état stationnaire à 25°C en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. Le mélange réactionnel comporte : Met-O (concentration variable), Trx d'*E. coli* (100 μM), Trx réductase d'*E. coli* (1,25 μM), NADPH (0,3 mM) et fRMsR entre 0,1 et 1 μM . Les valeurs expérimentales ont été ajustées par régression non linéaire à l'équation de Michaelis-Menten ou à l'équation correspondant au modèle d'inhibition compétitive par le substrat : $k_{\text{obs}} = (k_{\text{cat}} \cdot [\text{S}]) / K_{\text{M}} \cdot ([\text{S}] + [\text{S}]^2 / K_{\text{I}})$. Les valeurs des

constantes obtenues sont reportées dans le tableau 2 page 50. Les points représentent les données expérimentales et les traits les ajustements. Chaque expérience a été réalisée au moins deux fois de façon indépendante.

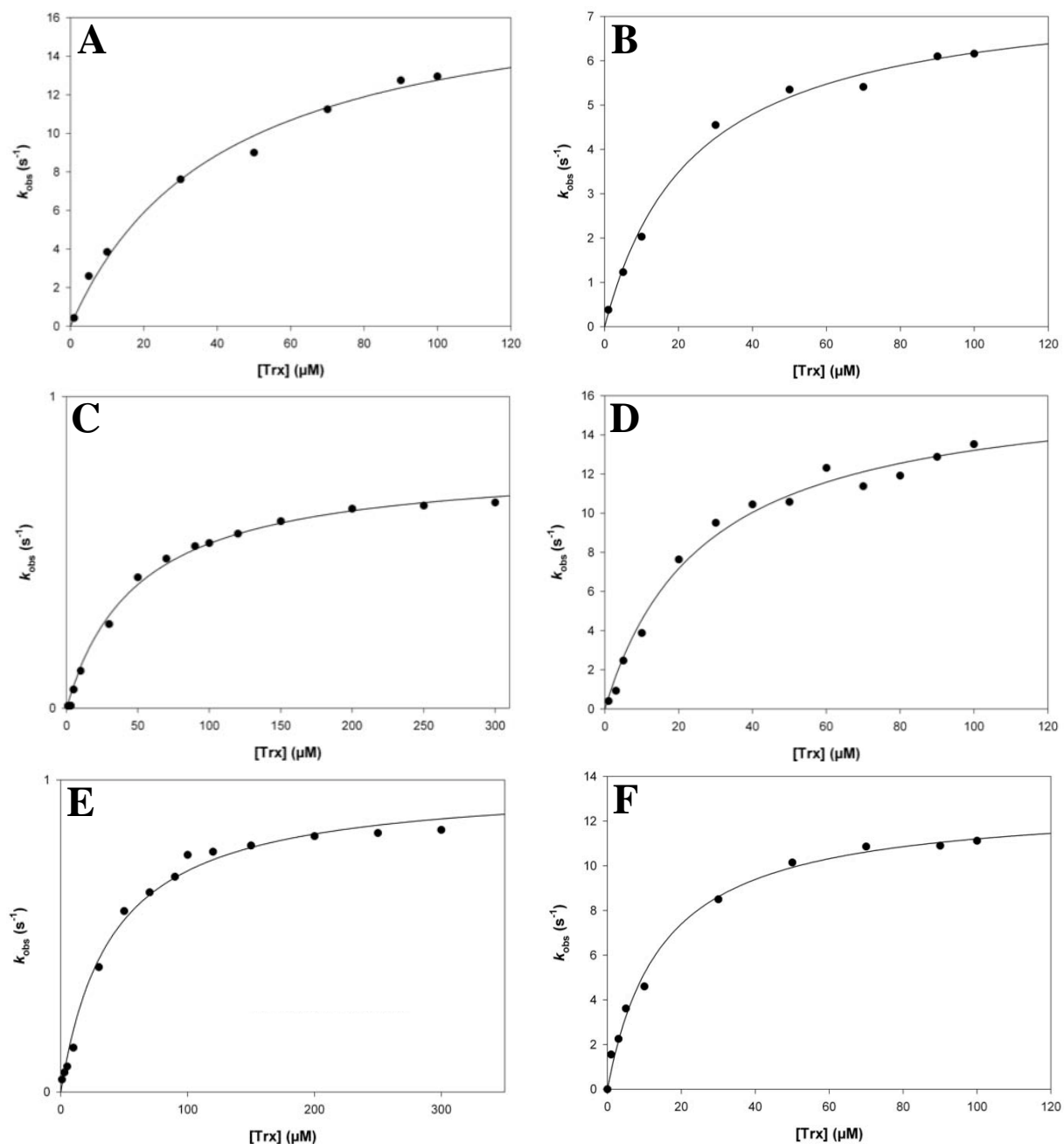


Figure 30 : Détermination des paramètres cinétiques, k_{cat} et K_M pour la Trx, des fRMs_r sauvage (A), C136A (B), C84A (C), C94A (D), C84A/C136A (E) et C94A/C136A (F) de *N. meningitidis*.

Les activités enzymatiques ont été mesurées à l'état stationnaire à 25°C en tampon Tris-HCl 50 mM, pH 8,0, EDTA 2 mM. Le mélange réactionnel comporte : Met-O (10 mM pour A et B, et 100 mM pour C, D, E et F), Trx d'*E. coli* (concentration variable), Trx réductase d'*E. coli* (1,25 μM), NADPH (0,3 mM) et fRMs_r (entre 0,1 et 1 μM). Les valeurs expérimentales ont été ajustées par régression non linéaire à l'équation : $k_{obs} = (k_{cat} \cdot [S]) / (K_M + [S])$. Les valeurs des constantes obtenues sont reportées dans le tableau 2 page 50. Les points représentent les

données expérimentales et les traits les ajustements. Chaque expérience a été réalisée au moins deux fois de façon indépendante.

Une inhibition par excès de substrat est observée, mais seulement avec l'enzyme sauvage et le mutant C136A. Cette observation reste à expliquer, mais semble être corrélée avec la présence des Cys84 et Cys94. Ces résultats montrent que le mutant C136A a un comportement cinétique parfaitement équivalent à celui de l'enzyme sauvage (inhibition par le substrat en excès et valeurs des constantes cinétiques). Les autres mutants ne présentant pas cette inhibition par le substrat en excès, il n'est pas possible de comparer les valeurs des constantes apparentes d'affinité pour la Met-O.

| | k_{cat} | K_M Met-O | K_I | K_M Trx |
|-----------------|---------------|---------------|-------------|------------|
| fRMsR | (s^{-1}) | (mM) | (mM) | (μ M) |
| Sauvage | 14 ± 2 | 2 ± 1 | 34 ± 19 | 40 ± 7 |
| C136A | $8,5 \pm 0,6$ | $2,2 \pm 0,5$ | 56 ± 16 | 24 ± 3 |
| C84A | $1,0 \pm 0,1$ | 32 ± 9 | - | 50 ± 4 |
| C94A | 18 ± 1 | 30 ± 6 | - | 27 ± 4 |
| C84A/C136A | $2,6 \pm 0,8$ | 12 ± 4 | - | 40 ± 6 |
| C94A/C136A | 13 ± 1 | 18 ± 1 | - | 15 ± 2 |
| C84A/C94A/C136A | $1,0 \pm 0,1$ | 30 ± 8 | - | 16 ± 4 |

Tableau 2: Paramètres cinétiques déterminés à l'état stationnaire de la fRMsR sauvage et des différents mutants.

Les valeurs des constantes de vitesse et des constantes apparentes d'affinité ont été obtenues par régression non linéaire des données expérimentales présentées dans les figures 29 et 30 pages 48 et 49.

D'autre part, afin d'étudier la spécificité de substrat de la fRMsR par rapport à la configuration du C α de la Met-O, les paramètres cinétiques pour la *L*-Met-*R,S*-O et la *D*-Met-*R,S*-O ont été déterminés pour le double mutant C94A/C136A qui ne présente pas d'inhibition par excès de substrat. Ainsi, la *D*-Met-*R,S*-O n'est pas un bon substrat. En effet, aucune

saturation a été observée jusqu'à 500 mM de substrat et la valeur de la constante k_2 , qui représente la valeur du k_{cat}/K_M , est d'environ $0,4 \text{ M}^{-1} \cdot \text{s}^{-1}$ (Fig. 31, panneau A page 51). Par conséquent, la valeur du K_M pour la *D*-Met-*R,S*-O est d'au moins 0,25 M et celle du k_{cat} de $0,16 \text{ s}^{-1}$. En revanche, les valeurs de k_{cat} et K_M pour la *L*-Met-*R,S*-O sont dans la gamme de celles observées avec la *D,L*-Met-*R,S*-O, c'est-à-dire respectivement 18 s^{-1} et 10 mM (Fig. 31, panneau B page 51). Il est à noter que la valeur de K_M pour la *L*-Met-*R,S*-O est plus faible d'un facteur environ 2 par rapport à celle du K_M pour la *D,L*-Met-*R,S*-O, comme attendu puisque la *D*-Met-*R,S*-O est un très mauvais substrat. L'efficacité catalytique en présence de *L*-Met-*R,S*-O est donc supérieure d'un facteur maximum de $4 \cdot 10^4$ comparée à celle observée en présence de *D*-Met-*R,S*-O, et ceci est lié à des effets k_{cat} et K_M . Enfin, la fRMsr ne réduisant que l'isomère *R* de la Met-O, la *D*-Met-*R,S*-O étant un très mauvais substrat et la Met-O utilisée étant un mélange des quatre diastéréoisomères, les valeurs de K_M obtenues en présence de Met-O doivent donc être divisées par 4 pour obtenir les valeurs de K_M pour la *L*-Met-*R*-O. L'efficacité catalytique (k_{cat}/K_M) de l'enzyme sauvage par rapport à la *L*-Met-*R*-O est donc meilleure d'au moins un facteur 1000 par rapport aux MsrB de *N. meningitidis* (Olry *et al.*, 2002), et ce en raison d'effets K_M et k_{cat} .

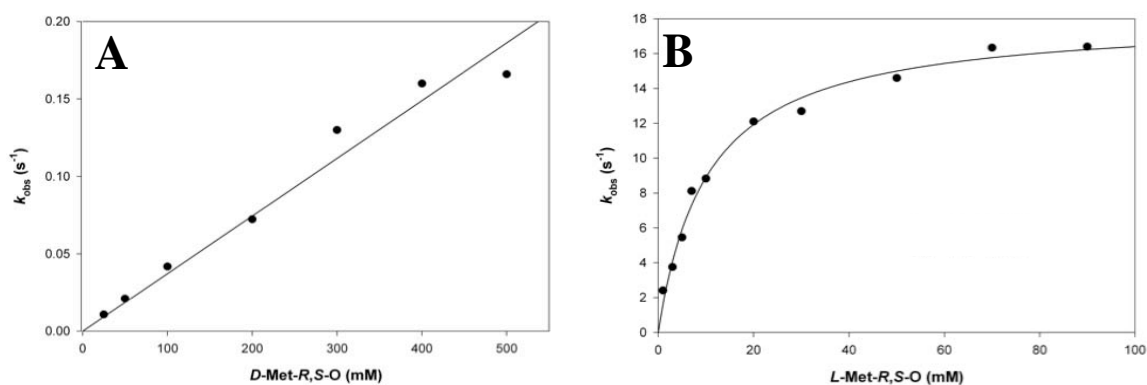


Figure 31 : Détermination des paramètres cinétiques, k_{cat} et K_M pour les substrats *D*-Met-*R,S*-O (A) et *L*-Met-*R,S*-O (B), de la fRMsr C94A/C136A de *N. meningitidis*.

Les activités enzymatiques ont été mesurées à l'état stationnaire à 25°C en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. Le mélange réactionnel comporte : *D*-Met-*R,S*-O (concentration variable), Trx d'*E. coli* (100 μM), Trx réductase d'*E. coli* (1,25 μM), NADPH (0,3 mM) et fRMsr (10 μM). Les valeurs expérimentales ont été ajustées par régression linéaire, pour le substrat *D*-Met-*R,S*-O, et non linéaire à l'équation : $k_{obs} = (k_{cat} \cdot [S]) / (K_M + [S])$ pour le substrat *L*-Met-*R,S*-O. Pour le panneau A, la saturation n'ayant pas été atteinte pour une concentration en substrat de 500 mM, la valeur de K_M pour la *D*-Met-*R,S*-O est supérieure à 250 mM et la valeur du k_{obsmax} est supérieure ou égale à $0,16 \text{ s}^{-1}$, la constante de vitesse k_2 est de $0,4 \text{ M}^{-1} \cdot \text{s}^{-1}$. Pour le panneau B, les valeurs des constantes k_{cat} et K_M pour le substrat *L*-

Met-*R,S-O*, sont de $18,2 \pm 0,5 \text{ s}^{-1}$ et $10 \pm 1 \text{ mM}$. Les points représentent les données expérimentales et les traits les ajustements.

I.5 Caractérisation des intermédiaires réactionnels

Le fait que tous les mutants des résidus Cys présentaient une activité dépendante de la Trx, à l'exception du mutant C118A, posait la question de la nature des formes oxydées réduites par la Trx. En particulier, l'intermédiaire acide sulfénique peut-il être réduit efficacement par la Trx ? Et qu'en est-il du pont disulfure Cys84-Cys94 ?

Pour essayer de répondre à ces questions, deux études ont été initiées. La première concernait la caractérisation de la forme oxydée du mutant monocystéinique C84A/C94A/C136A. La seconde concernait la formation et la réduction du pont disulfure Cys84-Cys94 *via* l'étude du mutant C118A/C136A.

I.5.1 Etudes du mutant monocystéinique C84A/C94A/C136A

La réduction de la Met-O par le mutant monocystéinique doit conduire à la formation d'un intermédiaire acide sulfénique sur la Cys118 catalytique. Cet intermédiaire n'a pas pu être caractérisé à l'aide d'un réactif chimique de second ordre spécifique, le 3-carboxy 4-nitrobenzenethiol (TNB). Afin de caractériser la forme oxydée de ce mutant, la stratégie a consisté à utiliser une approche de spectrométrie de masse, qui est l'autre méthode classique de caractérisation des intermédiaires acides sulféniques. Cette étude a été complétée par la caractérisation de l'état oligomérique du mutant oxydé en solution.

I.5.1.1 Détermination de la masse des formes oxydées du mutant monocystéinique en conditions dénaturantes

Cette étude a été réalisée en collaboration avec Stéphanie Petiot et le Dr Sarah Sanglier du Laboratoire de Spectrométrie de Masse Bio-Organique à Strasbourg dirigé par le Dr. A. Van Dorsselaer. La technique utilisée a été la LC/MS (« Liquid Chromatography coupled to Mass Spectrometry ») sur une colonne de phase inverse C8 en tampon acétonitrile, 0,1 % TFA couplée à une analyse sur un spectromètre de masse MicrOTOF-Q (Bruker Daltonics, Bremen, Germany). Les premières analyses ont révélé l'existence d'une

hétérogénéité au niveau de l'extrémité N-terminale de la protéine résultant d'un clivage partiel de la Met N-terminale et du résidu suivant. Pour éviter cette hétérogénéité nous avons alors substitué par mutagenèse dirigée les résidus His2 et Ala3 respectivement par une Gly et une Ser (Frottin *et al.*, 2006), et vérifié que ces substitutions ne modifiaient pas le comportement cinétique de l'enzyme. Ainsi l'étude a été réalisée sur les formes réduites et oxydées du mutant H2G/A3S/C84A/C94A/C136A mais également du mutant H2G/A3S/C136A utilisé comme représentant le comportement de l'enzyme sauvage. Les formes oxydées ont été obtenues en incubant 50 μ M d'enzyme et 50 mM de Met-O pendant 10 s avant ajout de 1% de TFA pour arrêter la réaction.

Dans ces conditions dénaturantes, on s'attendait à observer 1) des masses correspondant aux formes monomériques pour tous les échantillons, 2) une diminution de 2 Da de la masse de la fRMsr H2G/A3S/C136A après réaction avec la Met-O, correspondant à la formation d'un pont disulfure, et 3) une augmentation de 16 Da de la masse de la fRMsr H2G/A3S/C84A/C94A/C136A après réaction avec la Met-O, correspondant à la formation d'un intermédiaire acide sulfénique.

Les résultats obtenus pour le mutant fRMsr H2G/A3S/C136A montrent, comme attendu, la présence uniquement de formes monomériques avec une diminution de 2 Da de la masse du monomère de la fRMsr H2G/A3S/C136A après réaction avec la Met-O (Fig. 32, panneau B page 54). Par contre, pour la fRMsr H2G/A3S/C84A/C94A/C136A, non seulement aucune masse correspondant à une augmentation de 16 Da de la masse du monomère après réaction avec la Met-O a été observée, mais en plus environ la moitié de l'enzyme oxydée est présente sous forme dimérique. En effet, alors que sur le chromatogramme, un seul pic correspondant à des espèces monomériques est observé pour la forme réduite (Fig. 32 panneau C page 54, Pic 1), deux pics sont observés pour le mutant oxydé (Fig. 32 panneau D page 54, Pic 1 et 2). L'analyse des spectres de masse correspondant à ces deux pics montre l'existence 1) d'une espèce monomérique majoritaire dans le premier pic (Pic 1) avec une masse correspondant à la masse du monomère réduit - 2 Da, et 2) de deux espèces dimériques majoritaires dans le deuxième pic (Pic 2) avec des masses correspondant à la masse d'un dimère - 2 Da et + 14 Da (Fig. 34 page 57) (Tableau 3 page 55). Les analyses ayant été réalisées en conditions dénaturantes et la fRMsr étant dimérique en solution, la présence de dimères observés par LC-MS ne peut résulter que de la formation de liaisons covalentes, intra- ou interdimères.

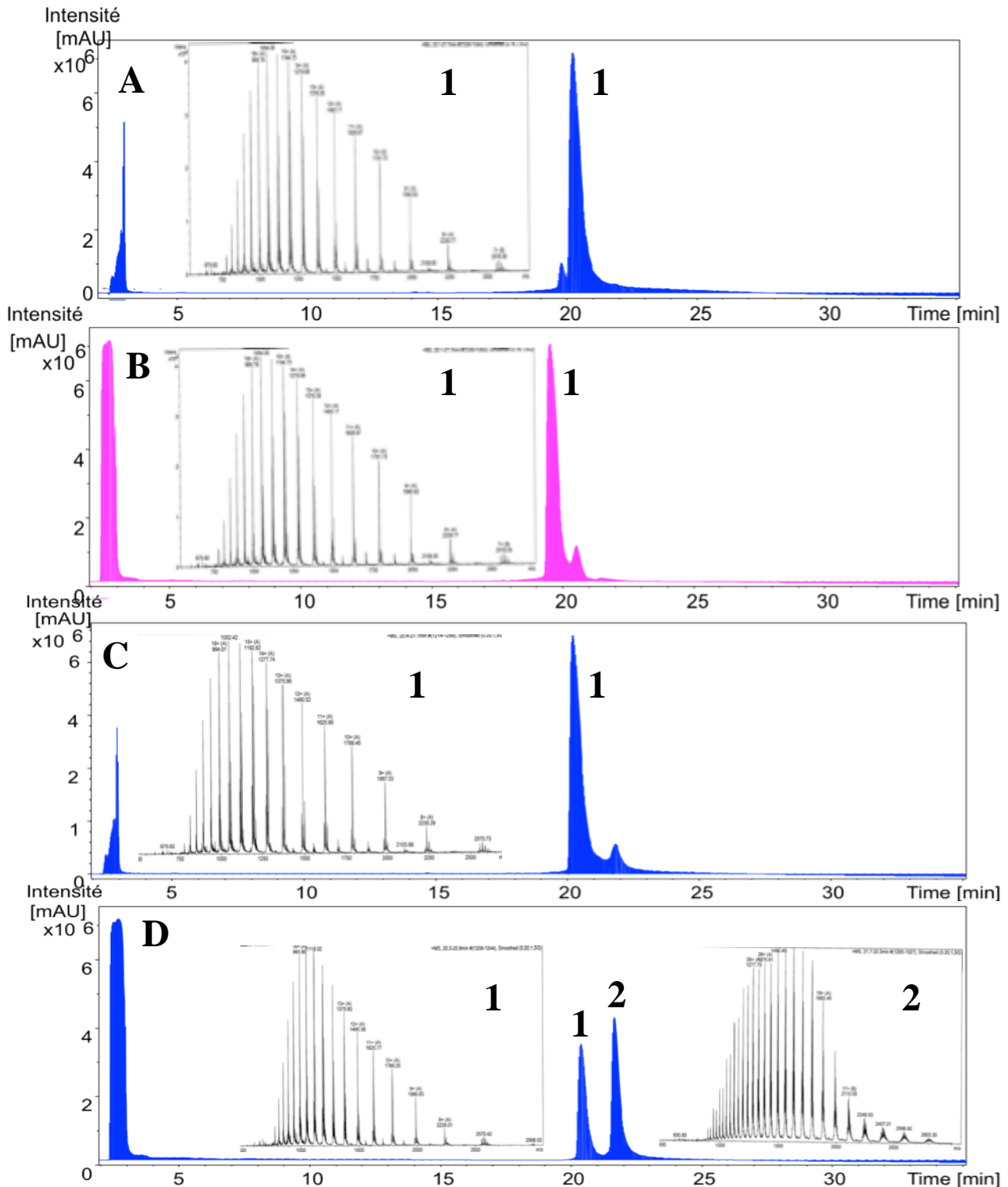


Figure 32 : Profils d'éluion des formes réduite et oxydée de la fRMsr H2G/A3S/C136A (respectivement A et B) et de la fRMsr H2G/A3S/C84A/C94A/C136A (respectivement C et D). La concentration des enzymes est de 50 μ M. Les formes oxydées ont été obtenues après incubation de la fRMsr avec 50 mM de Met-O pendant 10 s. Les protéines ont été séparées par chromatographie en phase inverse sur colonne C8 en tampon acétonitrile, 0,1 % TFA couplée à l'analyse sur un spectromètre de masse MicroTOF-Q (Bruker Daltonics, Bremen,

Germany). L'élution des protéines est suivie par la mesure de la DO à 215 nm. Les régions m/z 600-3000 des spectres de masse ESI obtenus sont associées aux profils de chromatographie. Les inscriptions 1 et 2 font référence au nombre de pics observés sur les chromatogrammes. Les masses mesurées obtenues sont reportées dans le tableau 3 page 55.

| fRMs _r | Pic | Espèce | Masse théorique (Da) | Masse mesurée (Da) |
|---------------------------------------|-----|----------------------------------|----------------------|--------------------|
| H2G/A3S/C136A réduit | 1 | Monomère réduit | 17938,03 | 17937,9 ± 0,3 |
| H2G/A3S/C136A oxydé | 1 | Monomère oxydé pont disulfure | 17936,03 | 17936,2 ± 0,4 |
| H2G/A3S/C84A/C94A/C136A réduit | 1 | Monomère réduit | 17873,90 | 17874,2 ± 0,2 |
| H2G/A3S/C84A/C94A/C136A oxydé | 1 | Monomère – 2 Da | 17871,90 | 17872,2 ± 0,3 |
| | 2 | Dimère – 2 Da | 35745,80 | 35747,3 ± 0,5 |
| | | Dimère + 14 Da | 35761,80 | 35762,0 ± 0,3 |

Tableau 3 : Résultats des analyses par LC-MS des fRMs_r H2G/A3S/C136A et H2G/A3S/C84A/C94A/C136A sous forme réduite et oxydée en conditions dénaturantes. Les masses théoriques sont calculées à partir de la séquence des différents mutants en tenant compte de l'état redox des Cys. Les masses mesurées correspondent aux espèces majoritaires présentes dans les différents échantillons.

Il est à noter que des espèces minoritaires supplémentaires ont été observées correspondant à la présence de formes non mutées à l'extrémité N-terminale et ceci lié à un problème de mélange de plasmides lors des productions. Les productions ont été refaites et les analyses devront être refaites.

I.5.1.2 Détermination de l'état oligomérique du mutant C84A/C94A/C136A sous forme oxydée en solution

L'état oligomérique a été déterminé en utilisant la technique de SLS (« Static Light Scattering ») qui permet par mesure de l'intensité de la lumière dispersée de déterminer la masse moléculaire d'une protéine en solution. L'étude a été réalisée sur les formes réduites et oxydées des mutants C84A/C94A/C136A et C136A, ce dernier servant de référence. Les

formes oxydées ont été obtenues par incubation de l'enzyme sous forme réduite en présence d'une concentration saturante de Met-O (100 mM) en tampon Tris-HCl 50 mM EDTA 2 mM, pH 8, suivie d'une étape de gel filtration pour éliminer l'excès de substrat et la Met produite. L'utilisation de DTNB a permis de vérifier la formation d'un pont disulfure pour le mutant C136A et l'absence de thiol libre pour le mutant C84A/C94A/C136A.

Dans ces conditions non dénaturantes, on s'attendait à observer 1) des masses correspondant à l'état dimérique en solution des fRMsr réduites, 2) une masse correspondant à un état dimérique pour la fRMsr C136A après réaction avec la Met-O, le pont disulfure formé étant intradimère, et 3) pour la fRMsr C84A/C94A/C136A après réaction avec la Met-O, une masse correspondant soit à un état dimérique, si les liaisons covalentes sont intradimères, soit à un état multimérique, si les liaisons covalentes sont interdimères.

Les résultats montrent que, comme attendu, le mutant C136A est présent en solution uniquement sous forme dimérique aussi bien à l'état réduit qu'oxydé, avec une masse moléculaire de l'ordre de 37 kDa (Fig. 33, Panneau A et B page 56). Ces résultats confirment que le pont disulfure présent dans le mutant oxydé est un pont disulfure intradimère.

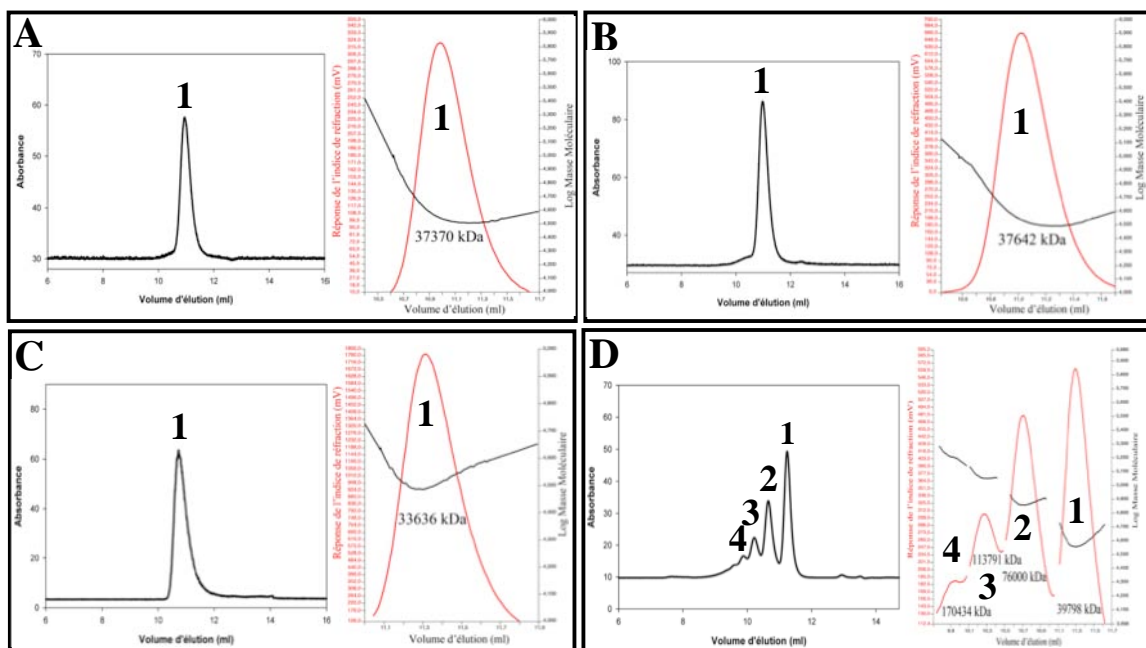


Figure 33 : Profils d'élution des fRMsr C136A et C84A/C94A/C136A sous formes réduites (respectivement A et C) et oxydées (respectivement B et D).

La concentration des enzymes est de 200 μ M. Les formes oxydées ont été obtenues par incubation de l'enzyme avec 100 mM de Met-O. Les protéines ont été séparées par chromatographie d'exclusion stérique sur une colonne Shodex KW-802,5 en tampon Tris-HCl 50 mM, KCl 150 mM, pH 8 couplée avec l'analyse par SLS. L'élution des protéines est suivie

par la mesure de la DO à 280 nm (Profils représentés à gauche des panels A, B, C et D), le RI « Refractive Index » (Profils représentés à droite des panels A, B, C et D) et le RALS « Right Angle Light Scattering » (Profils non représentés pour des raisons de clarté). La représentation logarithmique des masses moléculaires sur les profils RI est en traits noirs (échelle sur l'axe des ordonnées à droite). Les masses moléculaires calculées sont indiquées sur les profils RI en kDa.

Par contre, alors que le mutant C84A/C94A/C136A sous forme réduite est également dimérique (33 kDa) (Fig. 33, Panneau C page 56), il existe à l'état oxydé sous plusieurs états oligomériques en solution. On observe ainsi une forme dimérique majoritaire d'environ 39 kDa, une forme tétramérique d'environ 76 kDa, une forme hexamérique d'environ 114 kDa et une forme octamérique d'environ 170 kDa (Fig. 33, Panneau D page 56). Ces résultats suggèrent fortement l'existence de liaisons covalentes interdimères qui ne se forment qu'à partir de la forme oxydée, acide sulfénique, du dimère (Fig. 34 page 57).

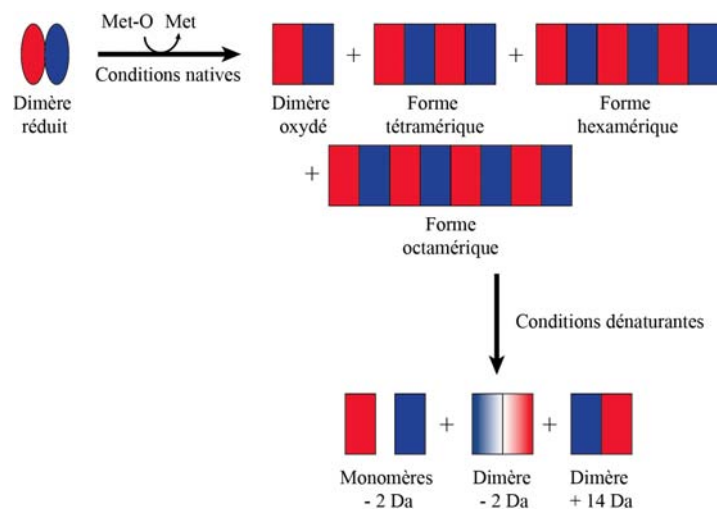


Figure 34 : Représentation schématique des formes oxydées du mutant C84A/C94A/C136A. En conditions natives (technique de SLS), des formes multimériques ont été observées. En conditions dénaturantes (technique LC-MS), un mélange de formes monomérique et dimérique a été observé.

I.5.1.3 Conclusion concernant les formes oxydées du mutant monocystéinique

L'ensemble des résultats montre que, dans les conditions utilisées, l'intermédiaire acide sulfénique n'est pas stable et qu'il évolue vers la formation de différentes espèces. Ces espèces correspondent pour une partie à des formes multimériques issues de la formation de

liaisons covalentes interdimeres (avec des différences de masse de $- 2$ Da et $+ 14$ Da par dimère / masse du dimère natif), et pour une autre partie de formes dimériques (avec une différence de masse de $- 2$ Da par monomère / masse du monomère natif) (Fig. 34 page 57) (voir interprétations dans la partie Discussion).

I.5.2 Etude de la fRMsr C118A/C136A

Alors que les études cinétiques faites sur les fRMsr C94A/C136A et C84A/C136A montraient que les ponts disulfures Cys84-Cys118 et Cys94-Cys118 pouvaient se former, aucune donnée expérimentale n'existait concernant la formation du pont disulfure Cys84-Cys94. L'étude de la fRMsr C118A/C136A, qui ne contient plus que les Cys84 et Cys94, a permis de montrer que le pont disulfure Cys84-Cys94 peut se former. En effet, l'incubation de $50 \mu\text{M}$ du mutant C118A/C136A avec 100 mM de 2PDS pendant 30 min conduit à la formation d'une forme oxydée qui ne contient plus de thiol titrable et aucune augmentation de l'absorbance à 343 nm correspondant à la libération de pyridine-2-thione a été observée lorsque du DTT a été ajouté.

Restait posée la question de la réduction de ce pont disulfure par la Trx. De plus, le fait que les mutants Cys94/Cys136 et Cys84/Cys136 étaient actifs en présence de Trx ne prouvait pas que c'était les ponts disulfure Cys84-Cys118 et Cys94-Cys118 qui étaient réduits par la Trx. Afin de le confirmer, les différentes formes oxydées ont été préparées et leur réduction par la Trx a été testée en identifiant par chromatographie en phase inverse l'apparition de Trx sous forme oxydée et de fRMsr mutées sous forme réduite après 10 min d'incubation.

Les formes oxydées de la fRMsr sous forme pont disulfure Cys84-Cys118 et Cys94-Cys118 ont été générés par incubation de $100 \mu\text{M}$ de double mutant respectivement C94A/C136A et C84A/C136A avec 100 mM de Met-O pendant 10 min . La forme pont disulfure Cys84-Cys94 a été obtenue par réaction de $50 \mu\text{M}$ du mutant C118A/C136A avec 100 mM de 2PDS pendant 30 min . Après dessalage sur une colonne de gel filtration (Econo-Pac 10-DG Biorad), les formes oxydées disulfures ont été incubées avec $100 \mu\text{M}$ de Trx réduite pendant 10 min et le mélange réactionnel a ensuite été analysé par chromatographie en phase inverse C8 en tampon acétonitrile, $0,1 \%$ TFA. Les résultats montrent la formation dans tous les cas de Trx oxydée et de fRMsr réduite (Fig. 35 page 59). Ainsi les trois ponts

disulfures Cys84-Cys118, Cys94-Cys118 et Cys84-Cys94 sont réduits par la Trx dans les conditions de temps de réaction utilisé.

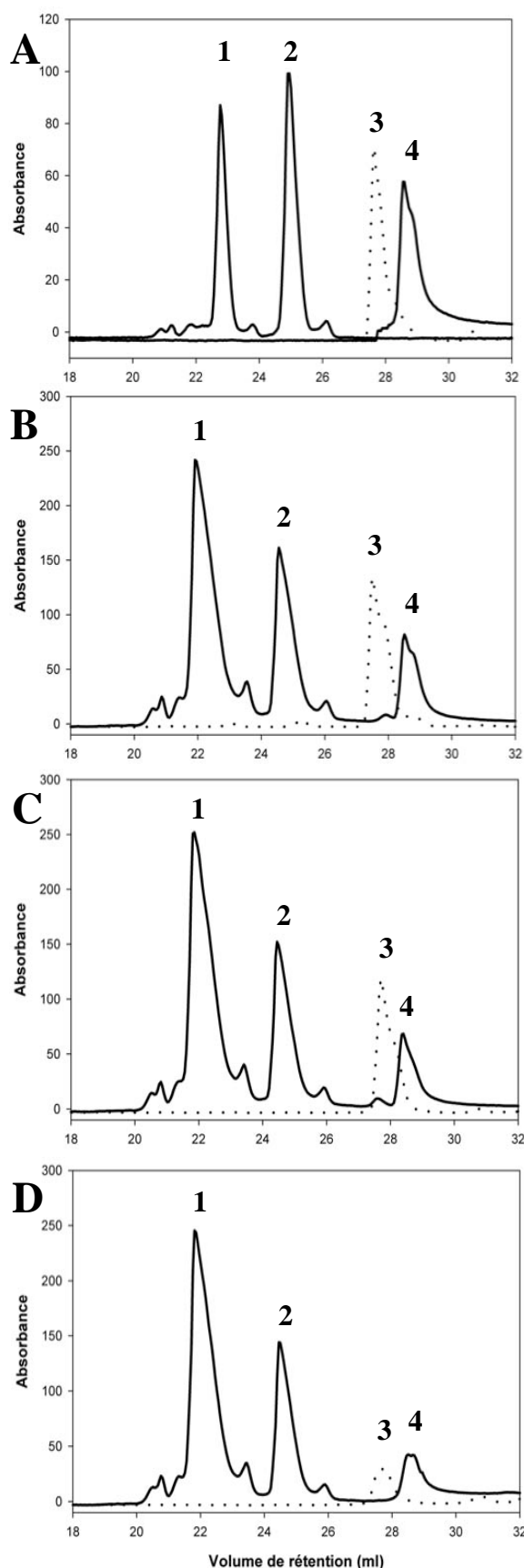


Figure 35 : Profils d'éluion des fRMsR C94A/C136A, C84A/C136A, et C118A/C136A et de la Trx sous formes réduites et oxydées.

Les formes oxydées Cys84-Cys118 et Cys94-Cys118 de la fRMsR ont été obtenues après incubation de la fRMsR avec 100 mM de Met-O. La forme Cys84-Cys94 a été obtenue par réaction du mutant C118A/C136A avec 100 mM de 2PDS pendant 30 min. Les mutants oxydés (à une concentration de 50 μ M pour le mutant C118A/C136A et 100 μ M pour les mutants C84A/C136A et C94A/C136A) ont été incubés en présence de 200 μ M de Trx sous forme réduite. Les mélanges réactionnels ont été séparés par chromatographie en phase inverse sur colonne C8 en tampon acétonitrile, 0,1 % TFA. L'éluion des différentes espèces a été suivie par la mesure de la DO à 280 nm.

A) Témoins : superposition des profils d'éluion de la Trx réduite (pic 1), de la Trx oxydée (pic 2), du mutant C136A oxydé (pic 3) et réduit (pic 4).

B), C) et D) correspondent respectivement aux profils d'éluion des fRMsR C94A/C136A, C84A/C136A et C118A/C136A sous forme oxydée (traits pointillés) ou après incubation avec de la Trx réduite (traits continus).

I.6 Conclusion

En combinant les résultats biochimiques et cinétiques sur l'enzyme de *N. meningitidis* sauvage et les mutants dans lesquels les résidus Cys ont été remplacés par des résidus Ala, nous avons démontré que 1) la Cys136 ne joue aucun rôle dans la réaction catalysée par la fRMsr, et 2) la Cys118 est le résidu catalytique sur lequel l'intermédiaire acide sulfénique doit se former. De plus, en se basant sur 1) l'existence d'un pont disulfure entre les Cys84 et Cys118 dans la structure de la fRMsr de *E. coli*, 2) le contenu en Cys libres dans l'enzyme sauvage et les différents mutants Cys en Ala avant et après réduction de Met-O, en conditions natives et dénaturantes, et 3) les valeurs des constantes catalytiques des mutants C94A et C94A/C136A comparables à celles de l'enzyme sauvage et du mutant C136A, nous avons proposé un scénario pour le processus de recyclage par la Trx dans lequel un pont disulfure Cys118-Cys84 se formerait et serait réduit par la Trx (Fig. 36, étape IIIA page 60).

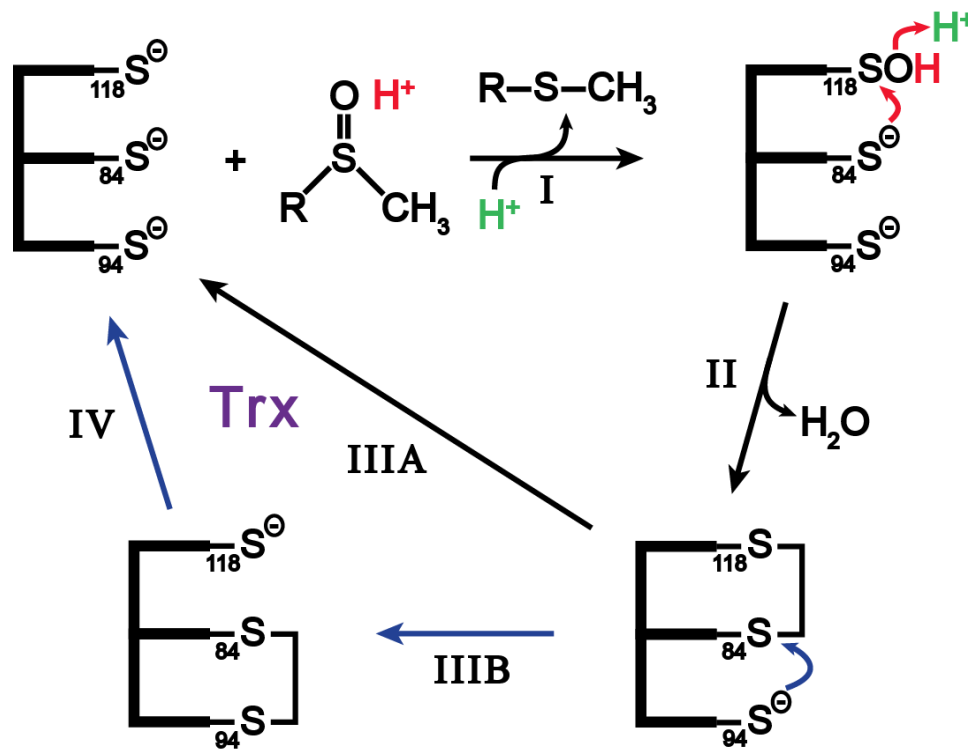


Figure 36 : Mécanisme catalytique proposé pour la fRMsr de *N. meningitidis*.

Le substrat sulfoxyde est attaqué par la Cys118 catalytique lors de l'étape I, conduisant à la formation de l'intermédiaire acide sulfénique. L'attaque de la Cys84 sur cet intermédiaire permet la formation d'un pont disulfure Cys84-Cys118 intramoléculaire et la libération concomitante d'une molécule d'eau, lors de l'étape II. Le pont disulfure Cys84-Cys118 peut être réduit par la Trx au cours de l'étape IIIA, ou un deuxième pont disulfure Cys84-Cys94 est formé et réduit par la Trx lors de l'étape IIIB.

Résultats : Mécanisme de la fRMsR de N. meningitidis

Cependant, nos résultats expérimentaux ne permettent pas de définitivement exclure un mécanisme faisant intervenir également la Cys94 avec la formation d'un deuxième pont disulfure Cys84-Cys94. Enfin, de façon surprenante, le mutant monocystéinique de la fRMsR de *N. meningitidis*, dans lequel seule la Cys118 catalytique est présente, s'est révélé présenter une activité dépendante de la Trx.

Partie II : Détermination de la vitesse des différentes étapes du mécanisme

Afin de déterminer l'étape cinétiquement limitante de la réaction globale, une étude cinétique des différentes étapes du mécanisme a été initiée. La stratégie utilisée a consisté d'une part à étudier l'activité réductase de l'enzyme en absence de réducteur, afin de ne permettre qu'un seul cycle catalytique de la fRMsR (« single turn-over ») et d'arrêter la réaction au niveau de l'intermédiaire pont disulfure, et d'autre part d'étudier la réaction de réduction de l'enzyme oxydée par la Trx réduite.

L'étude a été réalisée sur les fRMsR C94A/C136A et C136A (Paragraphe II.1). La réaction de la fRMsR C94A/C136A avec le substrat Met-O s'arrête à la formation d'un seul intermédiaire pont disulfure Cys84-Cys118. Quant à la fRMsR C136A, son comportement cinétique est équivalent à celui de la fRMsR sauvage, son oxydation s'accompagne probablement de la formation du même intermédiaire pont disulfure Cys84-Cys118, et l'absence de la Cys136 simplifie les analyses des thiols libres.

L'étude a également été entreprise sur les fRMsR C84A/C136A et C84A/C94A/C136A, qui ne forment pas le pont disulfure Cys84-Cys118, afin de déterminer les cinétiques d'apparition des différentes formes oxydées, si possible (Paragraphe II.2).

II. 1 Mise en évidence de l'étape cinétiquement limitante

II. 1.1 Etudes en absence de réducteur

II.1.1.1 Stratégie utilisée

Afin d'atteindre de façon spécifique la vitesse de réduction de la Met-O et celle de la formation de l'intermédiaire pont disulfure, et en se basant sur les approches développées pour l'étude des MsrA/B, deux techniques de cinétique rapide ont été utilisées.

La première est une technique de mélange rapide résolu en temps « quenched-flow » (KinTek Corporation). Cette étude a pu être réalisée grâce à l'aide du Pr Kai Tittman (Bioanalytik Albrecht-von-Haller-Institut Georg-August-Universität, Göttingen, Allemagne) qui nous a permis de réaliser les premières expériences de quench-flow dans son laboratoire, avant que nous ayons accès à un appareil de quench-flow à Nancy, acheté par le Service Commun de Biophysicochimie des Interactions macromoléculaires (SCBIM) de l'UL. Cette technique permet d'atteindre des valeurs de temps de réactions comprises entre 5 et 100 ms,

Résultats : Vitesse des différentes étapes du mécanisme

ce qui correspond à des valeurs de constantes de vitesse mesurables comprises entre ≈ 7 et $\approx 140 \text{ s}^{-1}$. Lors de la première étape de la réaction appelée « étape réductase », la Cys118 réalise une attaque nucléophile sur l'atome de soufre de la fonction sulfoxyde du substrat Met-O, conduisant à la formation de l'intermédiaire acide sulfénique sur la Cys catalytique et à la libération concomitante d'une mole de Met par mole d'enzyme. Ainsi, la vitesse de l'étape réductase a pu être déterminée en suivant la libération de Met au cours du temps. Pour cela, la réaction a été bloquée à différents temps par acidification avec du TFA. La quantification de la Met formée a ensuite été réalisée par chromatographie en phase inverse sur une colonne C18. Lors de la deuxième étape, la Cys84 réalise une attaque sur le soufre de la fonction acide sulfénique, conduisant à la formation d'un intermédiaire pont disulfure. Le dosage des groupements thiols libres de l'enzyme à l'aide d'un réactif spécifique, le DTNB (en conditions dénaturantes, en présence d'urée), à partir des échantillons collectés précédemment, a permis d'atteindre la vitesse des étapes réductase et de formation du pont disulfure. En effet, lors de la première étape conduisant à la formation de l'intermédiaire acide sulfénique, la Cys118 n'est plus titrable alors que dans la sous-étape de formation du pont disulfure, les Cys118 et Cys84 ne sont plus titrables (les formes acide sulfénique et pont disulfure ne réagissant pas avec le DTNB). Ainsi, on peut observer soit 1) la disparition concomitante des deux Cys titrables (allure monoexponentielle décroissante), si l'intermédiaire acide sulfénique ne s'accumule pas, ce qui montre que l'étape de formation de l'intermédiaire acide sulfénique est l'étape cinétiquement déterminante du mécanisme à deux étapes qui conduit à la formation de l'intermédiaire pont disulfure, soit 2) la disparition successive des deux Cys titrables (allure bi-exponentielle décroissante), si l'intermédiaire acide sulfénique s'accumule, ce qui montre que l'étape réductase est plus rapide que l'étape de formation du pont disulfure qui elle est cinétiquement déterminante.

La deuxième approche consiste, toujours en absence de réducteur, à utiliser les propriétés de variation de l'intensité d'émission de fluorescence de résidus Trp de l'enzyme lors de son passage de l'état réduit à l'état oxydé. L'utilisation de cette approche nécessitait d'avoir une variation de l'intensité d'émission de fluorescence de l'enzyme lors de son passage de l'état réduit à l'état oxydé, et cela a permis de déterminer les paramètres cinétiques en faisant varier la concentration en substrat.

II.1.1.2 Détermination des vitesses de réduction de la Met-O, étape réductase, et de la formation du pont disulfure

Les vitesses de réduction de la Met-O, étape réductase, et de la formation du pont disulfure étant forcément supérieure ou égale au k_{cat} qui est de l'ordre de 10 s^{-1} , des temps de réaction allant de 5 à 100 ms ont été utilisés. Les concentrations dans le mélange réactionnel étaient de $300 \mu\text{M}$ d'enzyme et 150 mM de Met-O, et la réaction a été arrêtée par ajout de 1% de TFA.

La variation de la quantité de Met formée au cours du temps présente une allure monoexponentielle croissante pour les fRMsr C94A/C136A et C136A, avec l'apparition d'environ une mole de Met par mole de monomère d'enzyme (Fig. 37 page 64, courbe bleue). L'ajustement des données expérimentales a permis de déterminer une valeur de l'ordre de 40 s^{-1} pour la constante de vitesse de formation de la Met (k_{obsMet}). De plus, les dosages des thiols titrables ont permis de mettre en évidence la perte concomitante des deux groupements thiols de l'enzyme (Fig. 37 page 64, courbe rouge). L'ajustement des données expérimentales à l'équation d'une monoexponentielle décroissante a permis de déterminer une valeur de l'ordre de 30 s^{-1} pour la constante de vitesse de formation du pont disulfure (k_{obsS}) pour les deux fRMsr. Les valeurs des différentes constantes sont reportées dans le tableau 4 page 65.

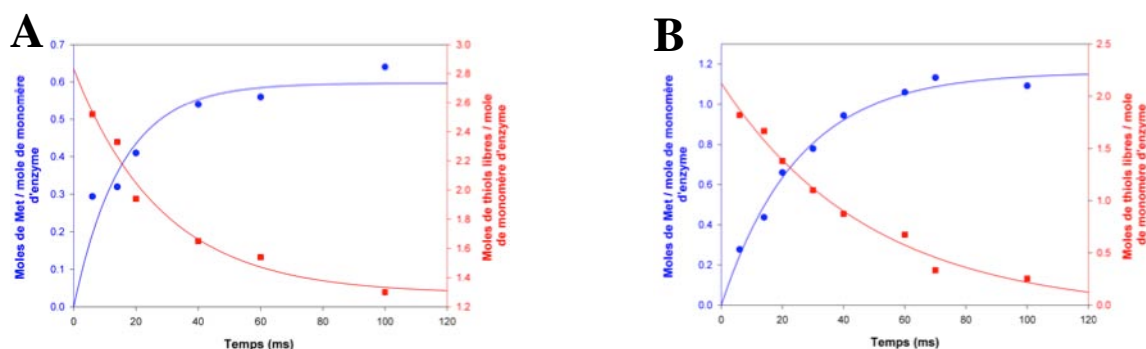


Figure 37 : Détermination des vitesses de formation de la Met et du pont disulfure Cys84-Cys118 pour les fRMsr C136A (A) et C94A/C136A (B).

La réaction a été réalisée dans du tampon Tris-HCl 50 mM , EDTA 2 mM , pH 8.0 à 25 °C en présence de $300 \mu\text{M}$ d'enzyme et de 150 mM de Met-O et elle a été arrêtée à différents temps par l'ajout de TFA à 1% final. La Met formée au cours du temps a été quantifiée par chromatographie en phase inverse sur une colonne C18 (●), et les groupements thiols libres de l'enzyme ont été dosés à l'aide de DTNB en présence d'urée (■). L'ajustement des données expérimentales par régression non linéaire aux équations $y = a(1 - e^{-k_{obs,t}t})$ pour la quantification de la Met et $y = y_0 + a.e^{-k_{obs,t}t}$ pour le dosage des thiols libres a permis d'obtenir les valeurs des constantes de vitesse et des amplitudes. Ces valeurs sont reportées dans le tableau 4 page 65. Les points représentent les données expérimentales et les traits les ajustements.

| Enzyme | moles de Met / mole de monomère d'enzyme | $k_{\text{obsMet}} \text{ (s}^{-1}\text{)}$ | moles de thiols libres / mole de monomère d'enzyme | $k_{\text{obsS}} \text{ (s}^{-1}\text{)}$ | Moyenne des $k_{\text{obs}} \text{ (s}^{-1}\text{)}$ |
|------------|------------------------------------------|---------------------------------------------|----------------------------------------------------|-------------------------------------------|------------------------------------------------------|
| C136A | $0,6 \pm 0,1$ | 65 ± 15 | $1,5 \pm 0,2$ | 36 ± 10 | 50 |
| C94A/C136A | $1,2 \pm 0,1$ | 40 ± 3 | $2,2 \pm 0,1$ | 22 ± 2 | 31 |

Tableau 4 : Constantes de vitesses des deux premières étapes du mécanisme de la fRMsr de *N. meningitidis*.

Les valeurs des constantes de vitesse et des amplitudes ont été obtenues par régression non linéaire des données expérimentales présentées dans la figure 37 page 64.

L'analyse des résultats montre que les valeurs des constantes k_{obsMet} et k_{obsS} sont du même ordre de grandeur pour les deux fRMsr (environ 40 s^{-1}). Ainsi, dans les conditions utilisées (150 mM en Met-O), la formation du pont disulfure est concomitante à la réduction de la Met-O, et l'étape réductase est l'étape cinétiquement déterminante du mécanisme à deux étapes qui permet en absence de réducteur de passer de l'enzyme sous forme réduite à l'enzyme sous forme pont disulfure.

II.1.1.3 Paramètres cinétiques des deux premières étapes du mécanisme

Cette approche nécessitait d'avoir une variation de l'émission de fluorescence pour la fRMsr de *N. meningitidis*, qui possède quatre résidus Trp. L'examen de la structure de la fRMsr d'*E. coli* et des alignements de séquence montrait la présence dans le site actif des fRMsr d'un résidu Trp62 conservé, qui pourrait servir de sonde de fluorescence. Pour le vérifier, les propriétés de fluorescence des formes réduites et oxydées des fRMsr C136A et C94A/C136A ont été étudiées. Les spectres de fluorescence obtenus pour les deux formes présentent un profil similaire, avec une longueur d'onde maximale d'excitation centrée à 284 nm et une longueur d'onde maximale d'émission centrée à 330 nm (Fig. 38, Panneau A page 66). De plus, une diminution de 30% de l'intensité d'émission de fluorescence est observée entre le spectre de l'enzyme sous forme réduite et celui de la forme oxydée. Enfin, le remplacement du résidu Trp62 par une Phe éteint cette variation du signal d'émission de fluorescence lié au passage de l'enzyme de l'état réduit à l'état oxydé (inférieure à 5%) (Fig. 38, Panneau B page 66). Ainsi, la variation de l'émission de fluorescence du Trp62 a été

utilisée comme sonde de fluorescence spécifique pour suivre, en absence de réducteur, la vitesse de formation du pont disulfure Cys84-Cys118 de la fRMsr de *N. meningitidis*.

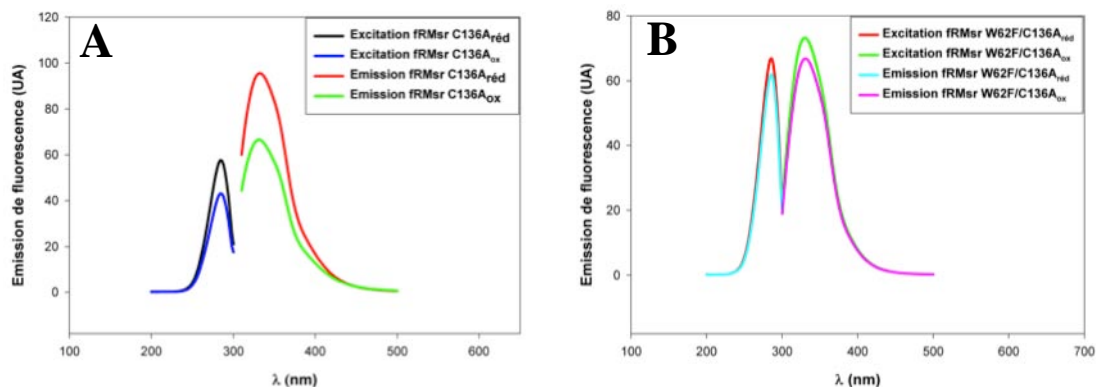


Figure 38 : Spectres d'excitation et d'émission de la fluorescence des fRMsr C136A (A) et W62F/C136A (B) sous formes réduites et oxydées.

Les mesures ont été effectuées dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8.0, avec une concentration en enzyme de 10 μ M. Pour les spectres d'excitation, la longueur d'onde d'excitation varie de 200 à 300 nm et la longueur d'émission est fixée à 340 nm. Pour les spectres d'émission, la longueur d'onde d'émission varie de 310 à 500 nm et la longueur d'excitation est fixée à 284 nm. Les spectres d'excitation et d'émission de la fRMsr C94A/C136A sont comparables à ceux présentés pour la fRMsr C136A.

Les cinétiques ont été suivies avec un appareil de mélange rapide (« stopped-flow »), dans des conditions de cycle catalytique unique. Les concentrations utilisées étaient de 10 μ M en enzyme et de 0,1 à 150 mM en Met-O. Pour chaque concentration en Met-O, une diminution de l'intensité d'émission de fluorescence en fonction du temps a été observée, avec une allure bi-exponentielle pour les deux fRMsr C94A/C136A et C136A (voir exemple présenté dans la Figure 39, page 67). L'ajustement des données expérimentales au modèle cinétique $A \Rightarrow B \Rightarrow C$ à l'aide du logiciel Pro Kineticist (Applied Photophysics Ltd) a permis de déterminer les valeurs des deux constantes de vitesse observées appelées k_{obs1} et k_{obs2} .

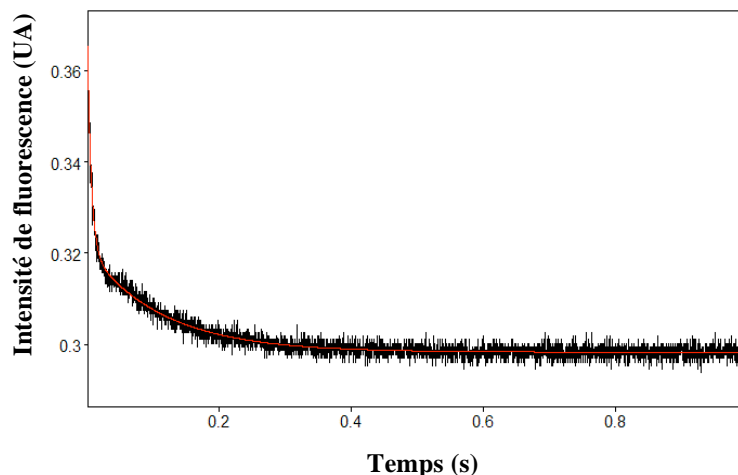


Figure 39 : Cinétique de l'extinction du message d'émission de fluorescence de la fRMsr C136A en présence de 10 mM de Met-O et en absence de réducteur.

Les concentrations en enzyme et en Met-O étaient respectivement de 10 μM et 10 mM. Les cinétiques ont été suivies avec un appareil de mélange rapide à 25°C en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. La longueur d'onde d'excitation fixée à 284 nm. Le message d'émission de fluorescence est mesuré pendant 1 s. Le message d'émission de fluorescence correspond à la somme des intensités de fluorescence détectées pour des longueurs d'onde supérieures à 320 nm grâce à l'utilisation d'un filtre. Les données expérimentales ont été ajustées au modèle $A \Rightarrow B \Rightarrow C$. Les résultats de l'ajustement sont $k_{\text{obs}1} = 141,8 \text{ s}^{-1}$ et $k_{\text{obs}2} = 13,3 \text{ s}^{-1}$ ($k_{\text{obs}1}$ et $k_{\text{obs}2}$ représentent respectivement les constantes de vitesse observées des étapes $A \Rightarrow B$ et $B \Rightarrow C$). Les valeurs expérimentales sont représentées par les points noirs et l'ajustement est représenté par la ligne continue rouge.

Les variations des valeurs de k_{obs} en fonction de la concentration en Met-O ont été ajustées par régression non linéaire à l'équation de Michaelis-Menten ou à l'équation d'inhibition par excès de substrat, afin de déterminer les valeurs des paramètres cinétiques que sont la constante de vitesse observée maximale (k_{obsmax}), la constante apparente d'affinité (K_S) et la constante apparente d'inhibition (K_I). L'ajustement des valeurs de $k_{\text{obs}1}$ et $k_{\text{obs}2}$ a permis la détermination respectivement des valeurs des constantes $k_{\text{obsmax}1}$ et $k_{\text{obsmax}2}$ et K_{S1} et K_{S2} (Fig. 40 page 68). Les valeurs des différentes constantes sont reportées dans le tableau 5 page 68. Ainsi, pour la fRMsr C94A/C136A, la valeur de $k_{\text{obsmax}2}$ est 4 fois plus faible que celle du $k_{\text{obsmax}1}$ (21 s^{-1} par rapport à 88 s^{-1}), alors que les valeurs de K_{S1} et K_{S2} sont du même ordre de grandeur (respectivement 54 et 56 mM). De même, pour la fRMsr C136A, la valeur de $k_{\text{obsmax}2}$ est 24 fois plus faible que celle du $k_{\text{obsmax}1}$ (30 s^{-1} par rapport à 711 s^{-1}), alors que les valeurs de K_{S1} et K_{S2} sont du même ordre de grandeur (respectivement 40 et 10 mM).

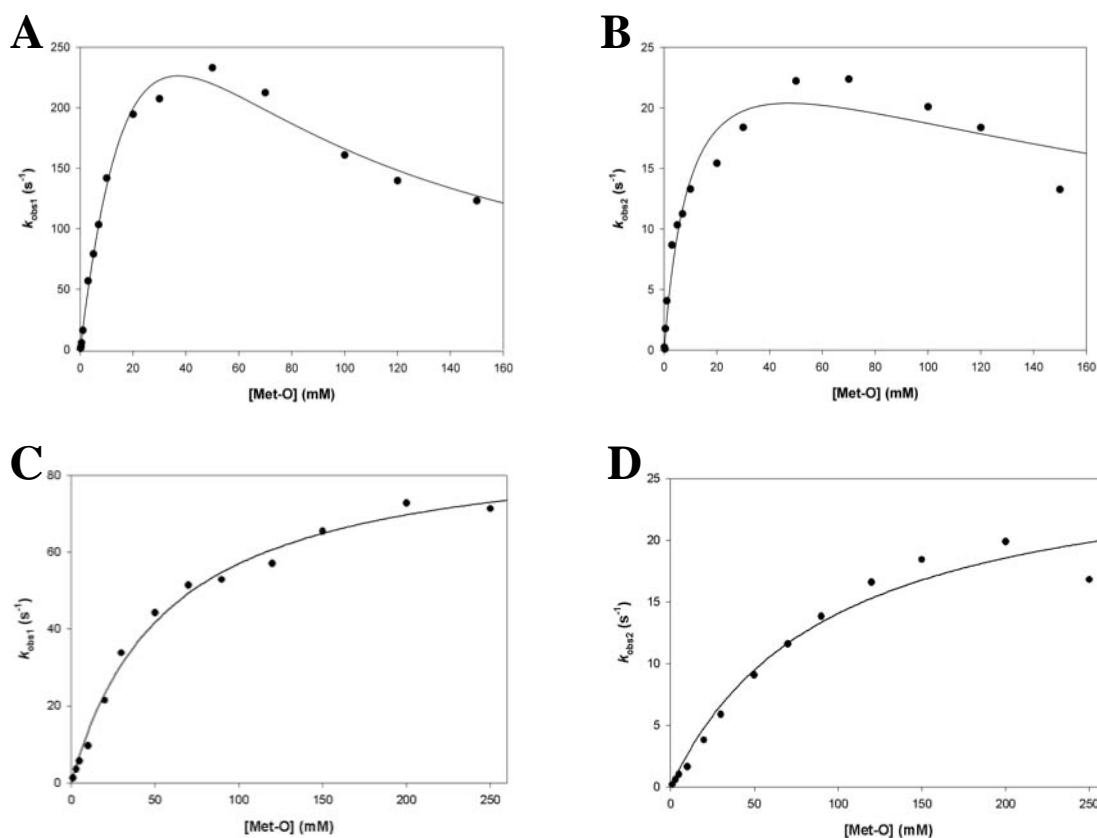


Figure 40 : Détermination des paramètres cinétiques de l'étape réductase des fRMsr C136A (A et B) et C94A/C136A (C et D) en présence de Met-O et en absence de réducteur.

Les constantes de vitesse k_{obs} ont été mesurées en présence de concentrations variables de Met-O. Pour chaque concentration en Met-O, les données expérimentales ont été ajustées au modèle cinétique $A \rightleftharpoons B \rightleftharpoons C$. Les panneaux A et C correspondent aux valeurs des constantes de vitesse k_{obs1} de l'étape $A \rightleftharpoons B$ respectivement pour les fRMsr C136A et C94A/C136A. Les panneaux B et D correspondent aux valeurs des constantes de vitesse k_{obs2} de l'étape $B \rightleftharpoons C$ respectivement pour les fRMsr C136A et C94A/C136A. Les valeurs de k_{obs} ont été ajustées par régression non linéaire à l'équation $k_{obs} = (k_{obsmax} \cdot [S]) / (K_S + [S])$ ou $k_{obs} = (k_{obsmax} \cdot [S]) / (K_S \cdot ([S] + [S]^2 / K_I))$ en fonction respectivement de l'absence ou de la présence d'un effet inhibiteur. Les paramètres obtenus sont reportés dans le tableau 5 page 68. Les points représentent les valeurs expérimentales et les lignes continues les ajustements. Chaque point représente la moyenne de trois mesures, et l'expérience a été réalisée trois fois de façon indépendante.

| Enzyme | $k_{obsmax1}$ (s^{-1}) | K_{S1} (mM) | K_{I1} (mM) | $k_{obsmax2}$ (s^{-1}) | K_{S2} (mM) | K_{I2} (mM) |
|------------|-------------------------------|---------------|------------------|-------------------------------|------------------|------------------|
| C136A | 711 ± 143 | 40 ± 11 | 35 ± 10 | 30 ± 5 | 10 ± 3 | 215 ± 108 |
| C94A/C136A | 88 ± 4 | 54 ± 7 | - | 21 ± 4 | 56 ± 29 | - |

Tableau 5 : Paramètres cinétiques de l'étape réductase des fRMsr C136A et C94A/C136A de *N. meningitidis*.

Les valeurs des paramètres ont été obtenues par régression non linéaire des données expérimentales présentées dans la figure 40 page 68. - : non observé.

Ainsi, cette approche par suivi des variations d'émission de fluorescence de la fRMs_r a permis de caractériser cinétiquement deux étapes de la réaction en absence de réducteur. Afin d'identifier ces deux étapes, les valeurs des constantes k_{obsmax} ont été comparées à la valeur de la constante k_{obsMet} déterminée précédemment. En effet, les valeurs des constantes apparentes d'affinité pour le substrat étant de l'ordre de 50 mM, la concentration de 150 mM utilisée dans l'approche « quenched-flow » peut être considérée comme saturante et la valeur de la constante de vitesse observée k_{obsMet} comme représentative de la valeur de la constante de vitesse maximale. La comparaison des valeurs montre que, quel que soit le mutant, la valeur de la constante $k_{\text{obsmax}2}$ est du même ordre de grandeur que la valeur de la constante k_{obsMet} , alors que la valeur de la constante $k_{\text{obsmax}1}$ est plus élevée. Ceci montre que la seconde étape cinétique correspond à l'étape réductase, et que la première étape cinétique correspond à une étape qui précède l'acte catalytique et probablement associée à un changement conformationnel de l'enzyme.

II. 1.2 Etude en présence de réducteur : Détermination de la vitesse de réduction du pont disulfure Cys84-Cys118 de la fRMs_r par la Trx réduite

Afin d'étudier la vitesse de réduction de la fRMs_r par la Trx, deux techniques ont été utilisées. La première approche a consisté à utiliser une technique comparable à celle utilisée pour l'étude des MsrA/B, à savoir de suivre en conditions de cycle catalytique unique la cinétique de variation de l'intensité d'émission de fluorescence liée au passage de la Trx de sa forme réduite à sa forme oxydée pont disulfure. En effet, la formation du pont disulfure Cys32-Cys35 dans le site actif de la Trx entraîne l'extinction de l'émission de fluorescence du Trp31 voisin. La variation de l'émission de fluorescence du Trp de la Trx constituerait ainsi une sonde spectroscopique indirecte de l'état redox du site actif de la fRMs_r en permettant de suivre spécifiquement la vitesse de formation de la Trx oxydée lors de la réduction de la fRMs_r. Afin de suivre sélectivement l'oxydation de la Trx, la longueur d'onde d'excitation est fixée à 310 nm. Dans ces conditions, la variation de l'intensité d'émission de fluorescence de la fRMs_r est négligeable par rapport à la contribution de la Trx (Spectres non montrés). L'étude a été réalisée sur les fRMs_r C94A/C136A et C136A. Les cinétiques ont été suivies avec un appareil de mélange rapide (« stopped-flow ») en présence de 10 μM de fRMs_r sous forme oxydée et de 100 μM de Trx sous forme réduite. Une diminution de l'intensité d'émission de fluorescence a été observée et l'ajustement des données à l'équation d'une mono-exponentielle décroissante a permis de déterminer une valeur de constante de vitesse

k_{obs} de l'ordre de $2 \cdot 10^{-3} \text{ s}^{-1}$. Cette valeur de k_{obs} est 7000 fois plus faible que celle du k_{cat} , ce qui n'est pas compatible avec l'acte catalytique. De plus, l'amplitude du signal observé n'est pas proportionnelle à la concentration de fRMsr oxydée utilisée. Ceci montre que la vitesse de réduction de la fRMsr oxydée par la Trx ne peut pas être mesurée par cette approche, soit parce que la réaction n'est pas efficace dans le complexe binaire, soit parce que le message de fluorescence de la Trx est « masqué » dans le complexe avec la fRMsr.

Une deuxième approche expérimentale a donc été utilisée. Elle a consisté à utiliser le système couplé Trx réductase / NADPH afin de suivre l'apparition de Trx oxydée en suivant la disparition du NADPH en absorbance à 340 nm. L'étude a été réalisée en présence de 10 μM de fRMsr C94A/C136A oxydée (pont disulfure Cys84-Cys118), 2,5 μM de Trx réductase, 0,3 mM de NADPH et des concentrations variables de Trx réduite (de 10 à 100 μM). Dans ces conditions, une diminution de l'absorbance à 340 nm est observée avec une amplitude proportionnelle à la quantité de fRMsr et une valeur de constante de vitesse de l'ordre de $1,2 \text{ s}^{-1}$ quelle que soit la concentration en Trx réduite utilisée (Fig. 41 page 71). De plus, cette valeur de constante de vitesse est équivalente à la valeur de la constante de vitesse de réduction de la Trx oxydée par la Trx réductase mesurée dans les mêmes conditions (Cinétique non montrée). Ceci montre que la vitesse observée en présence de fRMsr oxydée est la vitesse du système couplé Trx réductase / NADPH. Cette approche a donc permis de montrer que la vitesse de réduction du pont disulfure Cys84-Cys118 de la fRMsr par la Trx est supérieure ou égale à $1,2 \text{ s}^{-1}$, et donc que la réduction est bien efficace dans le complexe binaire.

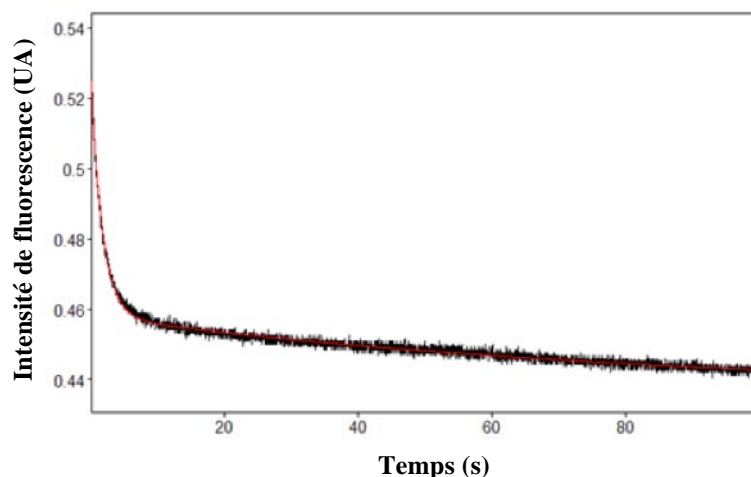


Figure 41 : Cinétique de réduction du pont disulfure Cys84-Cys118 de la fRMsr C94A/C136A par la Trx en présence du système couplé Trx réductase / NADPH.

Les concentrations en enzyme sous forme oxydée pont disulfure, en Trx sous forme réduite, en Trx réductase et en NADPH étaient respectivement de 10 μ M, 10 μ M, 2,5 μ M et 0,3 mM. Les cinétiques ont été suivies à 340 nm avec un appareil de mélange rapide « Stopped-flow » en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0 et à 25°C. L'ajustement par régression non linéaire a été réalisé sur un modèle cinétique de type $A \rightleftharpoons B$. Les résultats de l'ajustement sont $k_{\text{obs}} = 1,2 \text{ s}^{-1}$. Les valeurs expérimentales sont représentées par les points noirs et l'ajustement est représenté par la ligne continue rouge.

II. 1.3 Nature de l'étape cinétiquement limitante du mécanisme global

L'ensemble des paramètres obtenus est reporté dans le tableau 6 page 72.

Les paramètres obtenus pour l'étape réductase et pour l'état stationnaire pour les fRMsr C136A et C94A/C136A sont comparables, à l'exception de l'existence d'un effet inhibiteur par excès de substrat dans le cas de la fRMsr C136A qui n'est pas observé dans le cas de la fRMsr C94A/C136A. Il est à noter que cet effet d'inhibition par excès de substrat est également observé dans le cas de la fRMsr sauvage (résultats non montrés) sans qu'aucune explication n'ait pu être avancée à ce jour. D'autre part, le fait que les paramètres cinétiques des fRMsr C136A et C94A/C136A soient comparables constitue un argument supplémentaire en faveur de l'existence d'un seul pont disulfure Cys84-Cys118 dans la réaction catalysée par la fRMsr C136A, et par conséquent dans celle de la fRMsr sauvage.

| | Enzyme | C136A | C94A/C136A |
|-------------------|---------------------------------|------------------------|----------------------|
| Etat stationnaire | k_{cat} (s ⁻¹) | 8,5 ± 0,6 ^a | 13 ± 1 ^a |
| | K_M (mM) | 2,2 ± 0,5 ^a | 18 ± 1 ^a |
| | K_I (mM) | 56 ± 16 ^a | - |
| Etape I | k_{obsmax} (s ⁻¹) | 30 ± 5 ^b | 21 ± 4 ^b |
| | K_S (mM) | 10 ± 3 ^b | 56 ± 29 ^b |
| | K_I (mM) | 215 ± 108 ^b | - |
| Etape II | k_{obsmax} (s ⁻¹) | ≥ 30 ^b | ≥ 21 ^b |
| Etape III | k_{obsmax} (s ⁻¹) | ND | ≥ 1,2 ^c |

Tableau 6 : Paramètres cinétiques des trois étapes et de la réaction des fRMsr C136A et C94A/C136A de *N. meningitidis*.

^a : résultats décrits dans la partie I.4 ; ^b : résultats décrits dans le paragraphe II.1.1.3 ; ^c : résultat décrit dans le paragraphe II.1.2 ; -, non observé ; ND : non déterminé.

Les études présentées précédemment n'ont pas permis de déterminer les paramètres cinétiques de l'étape de réduction de la fRMsr oxydée par la Trx. Cependant, les paramètres cinétiques de l'étape réductase et de l'état stationnaire des fRMsr C136A et C94A/C136A peuvent être considérés comme identiques. En effet, les valeurs des constantes de vitesse et des constantes apparentes d'affinité ne sont que légèrement supérieures (d'un facteur 1,6 à 4,5) pour l'étape réductase par rapport à l'état stationnaire. Ceci suggère fortement que l'étape réductase pourrait être l'étape cinétiquement limitante de la réaction globale de la fRMsr.

II. 2 Etude en absence de réducteur des fRMsr C84A/C136A et C84A/C94A/C136A

Les vitesses de formation de Met et de disparition des thiols libres ont été mesurées par l'approche de quench-flow couplée à la quantification de Met et des thiols libres comme

décrit dans le paragraphe II.1.1.2. Les concentrations dans le mélange réactionnel étaient de 300 μM d'enzyme et de 150 mM de Met-O, et la réaction a été arrêtée par ajout de 1% de TFA.

La variation de la quantité de Met formée au cours du temps présente une allure monoexponentielle croissante pour les fRMsr C84A/C136A et C84A/C94A/C136A, avec l'apparition d'environ 0,5 mole de Met par mole de monomère d'enzyme (Fig. 42 page 73, courbe bleue). L'ajustement des données expérimentales a permis de déterminer une valeur de l'ordre de 20 s^{-1} pour la constante de vitesse de formation de la Met (k_{obsMet}). De plus, les dosages des thiols titrables ont permis de mettre en évidence la perte de la totalité des thiols libres de la fRMsr C84A/C94A/C136A et la perte concomitante des deux groupements thiols de la fRMsr C84A/C136A (Fig. 42 page 73, courbe rouge). L'ajustement des données expérimentales à l'équation d'une monoexponentielle décroissante a permis de déterminer une valeur de l'ordre de 15 s^{-1} pour la constante de vitesse de disparition des thiols libres pour les deux fRMsr. Les valeurs des différentes constantes sont reportées dans le tableau 7 page 74.

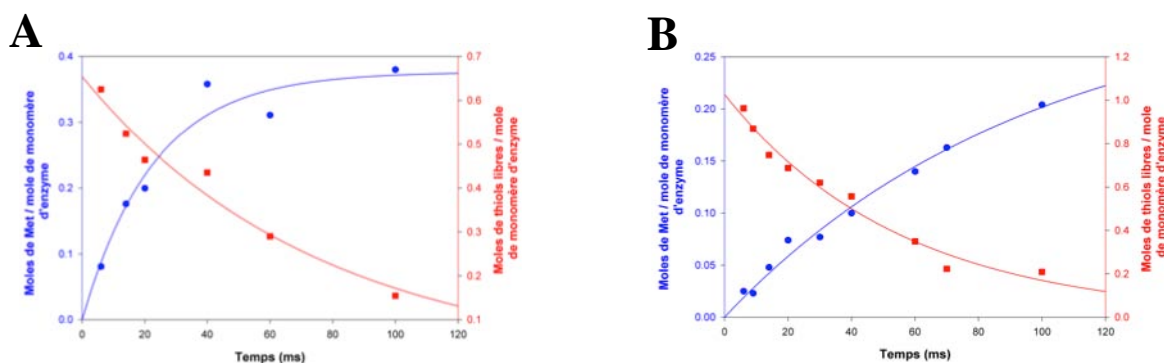


Figure 42 : Détermination des vitesses de formation de la Met et de la disparition des thiols libres pour les fRMsr C84A/C136A (A) et C84A/C94A/C136A (B).

La réaction a été réalisée dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8.0 à 25 °C en présence de 300 μM d'enzyme et de 150 mM de Met-O et elle a été arrêtée à différents temps par l'ajout de TFA à 1% final. La Met formée au cours du temps a été quantifiée par chromatographie en phase inverse sur une colonne C18 (●), et les groupements thiols libres de l'enzyme ont été dosés à l'aide de DTNB en présence d'urée (■). L'ajustement des données expérimentales par régression non linéaire aux équations $y = a(1 - e^{-k_{\text{obs.t}}})$ pour la quantification de la Met et $y = y_0 + a.e^{-k_{\text{obs.t}}}$ pour le dosage des thiols libres a permis d'obtenir les valeurs des constantes de vitesse et des amplitudes. Ces valeurs sont reportées dans le tableau 7 page 74. Les points représentent les données expérimentales et les traits les ajustements.

| | Enzyme | C84A/C136A | C84A/C94A/C136A |
|--------------------------|----------------------------------------------------|------------------------|------------------------|
| Etat stationnaire | k_{cat} (s ⁻¹) | 2,6 ± 0,8 ^a | 1,0 ± 0,1 ^a |
| | K_M (mM) | 12 ± 4 ^a | 30 ± 8 ^a |
| Etape I | k_{obsMet} (s ⁻¹) | 44 ± 8 | 10 ± 2 |
| | Moles de Met / mole de monomère d'enzyme | 0,40 ± 0,02 | 0,30 ± 0,04 |
| | k_{obsS} (s ⁻¹) | 13,0 ± 1,6 | 18,0 ± 1,5 |
| | Moles de thiols libres / mole de monomère d'enzyme | 0,60 ± 0,03 | 1,00 ± 0,04 |
| | Moyenne des k_{obs} (s ⁻¹) | 28,5 | 14 |

Tableau 7 : Paramètres cinétiques des fRMsr C84A/C136A et C84A/C94A/C136A de *N. meningitidis*.

Les valeurs des constantes de vitesse et des amplitudes ont été obtenues par régression non linéaire des données expérimentales présentées dans la figure 42 page 73. ^a: résultats décrits dans la partie I.

L'analyse des résultats montre que les valeurs des constantes k_{obsMet} et k_{obsS} sont du même ordre de grandeur pour la fRMsr C84A/C136A (environ 28 s⁻¹). Ainsi, dans les conditions utilisées (150 mM en Met-O), la formation du pont disulfure est concomitante à la réduction de la Met-O, et l'étape réductase est l'étape cinétiquement déterminante du mécanisme à deux étapes qui permet en absence de réducteur de passer de l'enzyme sous forme réduite à l'enzyme sous forme pont disulfure. De plus, les valeurs des constantes k_{obsMet} sont du même ordre de grandeur pour les deux fRMsr, et comparables à la valeur de la constante k_{obsMet} de la fRMsr C136A (voir paragraphe II.1.1.2 page 64). Ainsi, l'étape réductase ne semble pas modifiée pour les fRMsr C84A/C136A et C84A/C94A/C136A par rapport à la fRMsr C136A (voir paragraphe précédent). Par contre, les valeurs des constantes k_{obsMet} des deux fRMsr sont supérieures d'un facteur au moins de 10 par rapport à la valeur des k_{cat} . Ceci montre que l'étape cinétiquement limitante du mécanisme des fRMsr

Résultats : Vitesse des différentes étapes du mécanisme

C84A/C136A et C84A/C94A/C136A se situe après l'étape réductase, et même après la formation du pont disulfure Cys94-Cys118 pour la fRMsr C84A/C136A.

Reste posée la question de comment expliquer la stoechiométrie de 0,5 moles de Met formée par mole de monomère de fRMsr observée pour ces deux fRMsr, alors que l'on observe une disparition complète des thiols titrables (voir partie discussion).

Partie III : Caractérisation des acides aminés impliqués dans la catalyse de l'étape réductase et dans la spécificité de substrat

Pour les MsrA et MsrB, l'équipe enzymologie Moléculaire a montré que l'étape réductase est très rapide grâce à une catalyse assistée par le substrat et à la présence d'un catalyseur acide. Les fRMsR doivent également utiliser un catalyseur acide afin de protoner la fonction sulfoxyde et permettre la formation de l'état de transition de type sulfurane. D'autre part, la reconnaissance du substrat sulfoxyde par les MsrA et MsrB repose majoritairement sur des interactions de type liaisons de van der Waals entre une poche hydrophobe et les groupements méthyle ϵ et méthylène γ du substrat Met-O. Qu'en est-il pour les fRMsR ?

L'examen des alignements de séquence des fRMsR disponibles dans les banques génomiques ainsi que de la structure tridimensionnelle de la fRMsR d'*E. coli* en complexe avec une molécule de MES montrait la présence dans le site actif de résidus conservés donc potentiellement impliqués dans la catalyse et/ou la reconnaissance du substrat sulfoxyde (Fig. 43 page 77 + Figure SI annexe 1). Ainsi trois résidus acides invariants sont présents dans le site actif et pourraient participer à la stabilisation de la forme polarisée de la fonction sulfoxyde ou à la catalyse *via* un rôle d'acide/base général (Glu125, Asp141, Asp143). D'autre part, deux autres résidus, Trp62 et Tyr66, sont présents dans le site actif et conservés à plus de 80 %. En particulier, le résidu Trp62 pourrait participer à la formation d'une poche hydrophobe pour accommoder le groupement ϵ -méthyle du substrat. Ces deux résidus étaient donc de bons candidats pour jouer un rôle dans la fixation et/ou le positionnement du substrat.

Afin de valider ou non ces hypothèses, nous avons entrepris une étude de relations structure-fonction sur la fRMsR de *N. meningitidis* en utilisant une approche d'ingénierie protéique par mutagenèse dirigée couplée à la caractérisation des paramètres cinétiques des différents mutants générés.

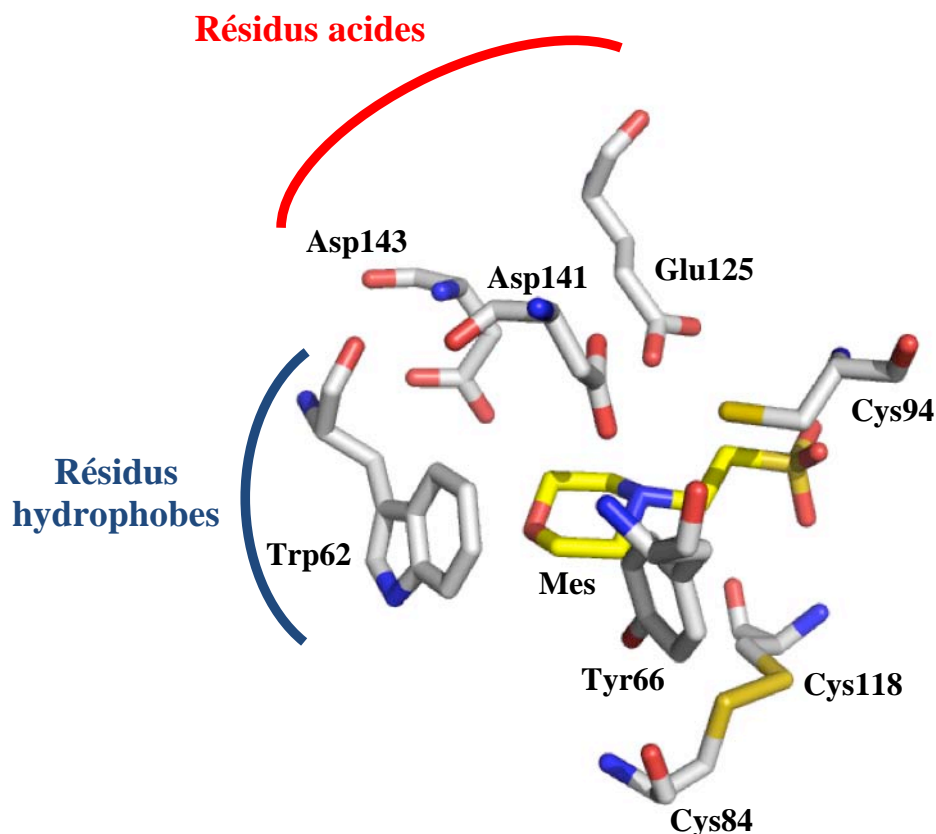


Figure 43 : Représentation schématique du site actif de la fRMsR d'*E. coli* en complexe avec une molécule de MES.

Les résidus d'acides aminés et la molécule de MES sont représentés en mode bâton avec les atomes de carbone respectivement en gris et jaune clair (N en bleu, O en rouge et S en jaune foncé). Sont représentés les résidus Cys (Cys94 et pont disulfure Cys84-Cys118) et les résidus acides (Glu125, Asp141 et Asp143) et hydrophobes (Trp62 et Tyr66). La figure a été réalisée avec le logiciel Pymol d'après le fichier PDB 1VHM.

III.1 Stratégie utilisée

La substitution de la Cys136 par une Ala ne modifiant en rien les paramètres cinétiques de la fRMsR, les différentes substitutions ont été réalisées sur la fRMsR C136A, ce qui a permis de simplifier les analyses du contenu en thiols des différents mutants. Ainsi, pour évaluer l'effet des substitutions, il conviendra de comparer les résultats cinétiques des mutants étudiés à l'état stationnaire et pré-stationnaire avec ceux de la fRMsR C136A. Afin de mettre en évidence le rôle des résidus cités auparavant, les résidus acides ont été substitués un à un par des résidus Ala et Asn ou Gln, et les résidus hydrophobes par des résidus Phe. Ainsi, les mutants suivants ont été générés W62F/C136A, Y66F/C136A, E125A/C136A, E125Q/C136A, D141A/C136A, D141N/C136A, D143A/C136A, et D143N/C136A. Pour

cela, les cellules d'*E. coli* HB101 ont été transformées par le plasmide recombinant pSK-fRMs_r portant la séquence codante mutée correspondante du gène *fRMs_r* de *N. meningitidis*. Les mutants fRMs_r ainsi produits ont été purifiés selon le protocole précédemment utilisé pour l'enzyme sauvage. Un rendement de l'ordre de 20 mg d'enzyme pure par litre de culture a été obtenu pour les différents mutants. Il est à noter que les mutants W62A/C136A et Y66A/C136A n'ont pas pu être étudiés car les protéines étaient produites sous forme insoluble.

Les paramètres cinétiques de l'état stationnaire et de l'étape réductase ont été déterminés, en utilisant les approches précédemment utilisées pour caractériser la fRMs_r C136A. A savoir, les paramètres k_{cat} et K_M ont été obtenus en présence de Met-O, de Trx, et du système couplé Trx réductase/ NADPH. Les paramètres k_{obsmax} et K_S de l'étape réductase ont été déterminés en présence de Met-O et en absence de réducteur, lorsque cela a été possible.

III.2 Dosage des groupements thiols libres et détermination de la stœchiométrie de formation de Met en absence de réducteur

Le dosage des groupements thiols libres des différents mutants a été réalisé en conditions dénaturantes avant et après ajout de Met-O et en absence de réducteur. En absence de Met-O, trois Cys sont dosées pour les différents mutants, ce qui est en parfait accord avec leur séquence protéique. Après ajout de Met-O, une seule Cys reste titrable pour l'ensemble des mutants (tableau 8 page 79). Parallèlement, la quantification de la Met formée en présence de Met-O montre la libération d'environ 1 mole de Met par monomère d'enzyme pour les différents mutants. L'ensemble de ces résultats montre que tous les mutants présentent une activité réductase en présence de Met-O avec la formation d'un pont disulfure, comme pour l'enzyme sauvage.

| Enzyme | Nombre de thiols théorique par monomère d'enzyme | Nombre de thiols titrés par monomère d'enzyme | Nombre de thiols titrés par monomère d'enzyme après ajout de Met-O | Différence de thiols par monomère après ajout de Met-O | Mole de Met formée / Monomère d'enzyme après ajout de Met-O |
|-------------|--------------------------------------------------|-----------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------|
| C136A | 3 | 3,1 | 1,1 | 2,0 | 1,1 |
| W62F/C136A | 3 | 3,1 | 1,1 | 2,0 | 0,9 |
| Y66F/C136A | 3 | 2,9 | 1,1 | 1,8 | 0,8 |
| E125A/C136A | 3 | 3,3 | 1,1 | 2,2 | 0,9 |
| E125Q/C136A | 3 | 2,9 | 0,9 | 2,0 | 1,2 |
| D141A/C136A | 3 | 3,2 | 1,1 | 2,1 | 0,8 |
| D141N/C136A | 3 | 2,9 | 0,9 | 2,0 | 0,8 |
| D143A/C136A | 3 | 3,2 | 1,1 | 2,1 | 0,7 |
| D143N/C136A | 3 | 2,9 | 0,9 | 2,0 | 1,1 |

Tableau 8 : Titration des groupements thiols libres des différents mutants de la fRMsR de *N. meningitidis* et stœchiométrie de formation de Met avant et après incubation avec de la Met-O.

Les dosages ont été effectués dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8 et ce pour deux concentrations d'enzymes (7,35 et 14,7 μ M). La concentration en DTNB est de 200 μ M. Dans les expériences de titration du nombre de thiols sous la forme oxydée, la Met-O a été ajoutée à une concentration finale de 100 mM. Le mélange a été incubé 10 minutes à température ambiante. La quantité de Met formée a été mesurée par chromatographie en phase inverse. La stœchiométrie a été calculée en rapportant la quantité de Met formée à la quantité d'enzyme utilisée pour la réaction. L'erreur expérimentale sur les dosages est de 10%.

III.3 Détermination des paramètres cinétiques à l'état stationnaire

Les paramètres cinétiques k_{cat} et K_M des différents mutants ont été déterminés en présence de Met-O, de Trx, et du système couplé Trx réductase/ NADPH, en suivant la disparition de l'absorbance du NADPH à 340 nm. Tous les mutants présentent, comme attendu, une activité Trx-dépendante. Les résultats obtenus sont présentés dans la figure 44 page 80 et les valeurs des constantes cinétiques correspondantes sont reportées dans le tableau 9 page 87.

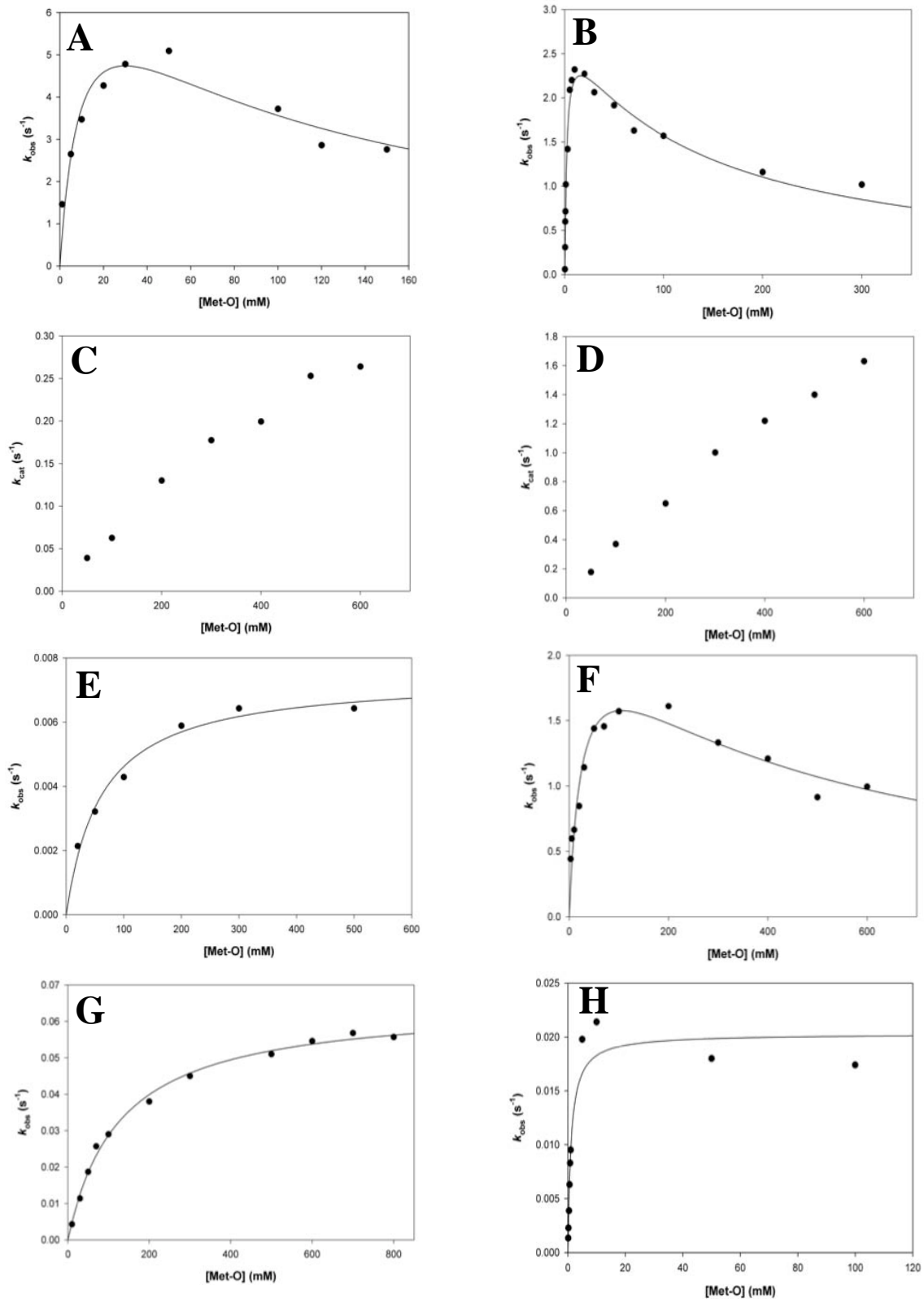


Figure 44 : Détermination des paramètres cinétiques k_{cat} et K_M , pour le substrat Met-O, des fRMsr W62F/C136A (A), Y66F/C136A (B), E125A/C136A (C), E125Q/C136A (D), D141A/C136A (E), D141N/C136A (F), D143A/C136A (G) et D143N/C136A (H) de *N. meningitidis*.

Les activités enzymatiques ont été mesurées à l'état stationnaire à 25°C en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. Le mélange réactionnel comporte : La Met-O (concentration

variable), Trx d'*E. coli* (100 μM), Trx réductase d'*E. coli* (1,25 μM), NADPH (0,3 mM) et fRMsr entre 0,1 et 15 μM . Les valeurs expérimentales ont été ajustées par régression non linéaire à l'équation de Michaelis-Menten ou à l'équation correspondant au modèle d'inhibition compétitive par le substrat : $k_{\text{obs}} = (k_{\text{cat}} \cdot [\text{S}]) / K_{\text{M}} \cdot ([\text{S}] + [\text{S}]^2 / K_{\text{I}})$. Les valeurs des constantes obtenues sont reportées dans le tableau 9 page 87. Pour (C) et (D), la saturation n'ayant pas été atteinte pour une concentration en substrat de 600 mM, les valeurs des K_{M} pour la Met-O sont supérieures à 300 mM et les valeurs des k_{obsmax} sont supérieures ou égale à 0,25 s^{-1} pour (C) et à 1,6 s^{-1} pour (D). Les points représentent les données expérimentales et les traits les ajustements.

La comparaison des valeurs des constantes catalytiques des différents mutants avec celles de la fRMsr C136A montre que la constante catalytique n'est pas significativement modifiée pour les fRMsr W62F/C136A, Y66F/C136A, E125Q/C136A et D141N/C136A alors qu'elle est diminuée d'un facteur 1100, 130 et 420 respectivement pour les fRMsr D141A/C136A, D143A/C136A et D143N/C136A.

De plus, les mutants D143A/C136A, D143N/C136A, D141A/C136A, E125A/C136A et E125Q/C136A ne présentent pas d'inhibition apparente par le substrat en excès, ce qui ne permet pas de comparer les valeurs des constantes apparentes d'affinité pour la Met-O de ces mutants avec celles de la fRMsr C136A. Cependant, le fait que la saturation en substrat n'a pas été atteinte pour une concentration en Met-O de 500 mM pour les fRMsr E125A/C136A et E125Q/C136A suggère l'existence d'une affinité altérée pour ces mutants.

Enfin, la comparaison des valeurs des constantes apparentes d'affinité pour la Met-O des fRMsr W62F/C136A et D141N/C136A avec celle de la fRMsr C136A montre une augmentation respectivement d'un facteur 5 et 10, alors qu'aucune variation n'est observée pour la fRMsr Y66F/C136A.

III.4 Détermination des propriétés de fluorescence

Les propriétés de fluorescence des formes réduites et oxydées des différents mutants ont été étudiées. Tous les spectres de fluorescence obtenus pour les deux formes présentaient un profil similaire, avec une longueur d'onde maximale d'excitation centrée à 286 nm et une longueur d'onde maximale d'émission centrée à 330 nm (Pour exemple Fig. 45 A page 83). Une diminution d'au moins 20% de l'intensité d'émission de fluorescence a été observée pour les fRMsr Y66F/C136A, D141A/C136A, D141N/C136A, E125Q/C136A et D143A/C136A entre le spectre de leur forme réduite et celui de leur forme oxydée, comparable à celle précédemment observée pour la fRMsr C136A. Par contre, aucune variation significative

(inférieure à 10 %) de l'intensité d'émission de fluorescence entre les formes réduites et oxydées n'a été observée pour les fRMsr W62F/C136A, et E125A/C136A (Pour exemple Fig. 45 B page 83).

Le suivi du message d'émission de fluorescence avec un appareil de mélange rapide dans des conditions de cycle catalytique unique en présence de 100 mM Met-O a permis de montrer l'existence d'une diminution de l'intensité d'émission de fluorescence, certes faible mais facilement mesurable, pour la fRMsr E125A/C136A mais pas pour les fRMsr W62F/C136A et Y66F/C136A (résultats non montrés). L'approche de cinétique rapide étant beaucoup plus sensible que l'approche sur le fluorimètre (mesures sur le même échantillon de concentration constante et mesure de l'ensemble de l'émission de fluorescence pour les longueurs d'onde supérieures à 320 nm), ces résultats montrent que c'est bien le résidu Trp62 (parmi les 4 Trp de la fRMsr) qui sert de sonde de fluorescence pour les études cinétiques. Le fait que les fRMsr Y66F/C136A et E125A/C136A présentent des propriétés de variation de fluorescence différentes de celles de l'enzyme sauvage montre que le Trp62 n'est pas dans le même environnement.

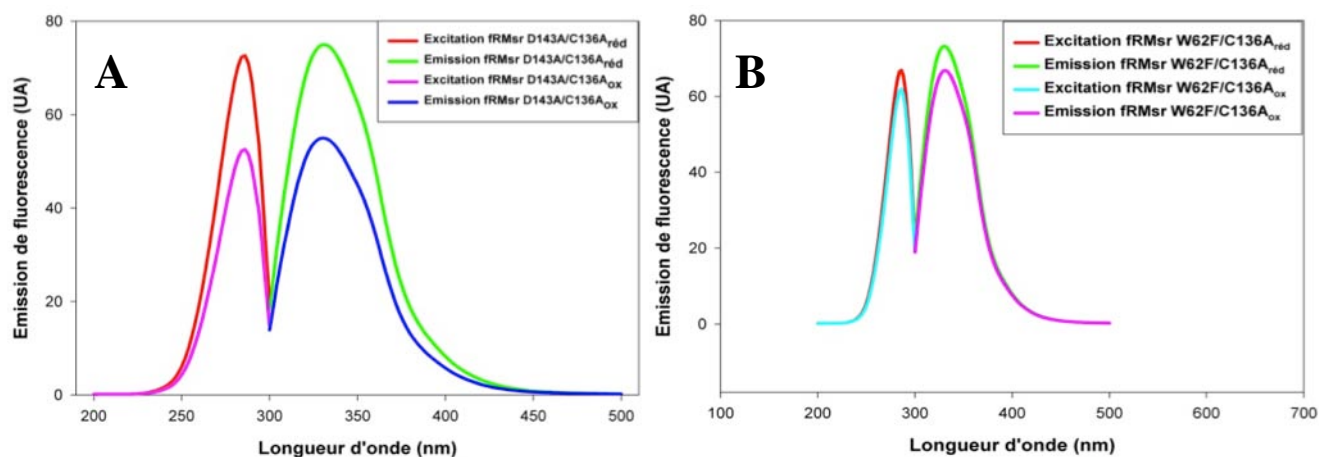


Figure 45 : Spectres d'excitation et d'émission de la fluorescence des fRMsr D143A/C136A (A) et W62F/C136A (B) sous forme réduite et oxydée. Les mesures ont été effectuées dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0, avec une concentration en enzyme de 10 μ M. Pour les spectres d'excitation, la longueur d'onde d'excitation varie de 200 à 300 nm et la longueur d'onde d'émission est fixée à 340 nm. Pour les spectres d'émission, la longueur d'onde varie de 300 nm à 500 nm et la longueur d'onde d'excitation est fixée à 286 nm. Les analyses ont été réalisées à une valeur du photomultiplicateur (PM) de 520V.

III.5 Détermination des paramètres cinétiques de l'étape réductase

Comme décrit dans la partie II, deux techniques complémentaires ont été mises au point pour étudier l'étape réductase de la fRMsr. A ce jour, pour les différents mutants, nous n'avons utilisé que la technique de « stopped-flow » en fluorescence en faisant l'hypothèse que l'étape cinétiquement déterminante du mécanisme à deux étapes (passage d'une forme réduite à une forme pont disulfure en absence de réducteur) reste associée à l'étape réductase, comme pour la fRMsr C136A. Cette hypothèse, bien que fortement probable, devra être validée en utilisant la technique de « quench-flow ». Ceci nous a permis de déterminer les paramètres cinétiques de l'étape réductase des fRMsr D141N/C136A, D143A/C136A, D143N/C136A, E125A/C136A et E125Q/C136A, mais pas des fRMsr W62F/C136A, Y66F/C136A et D141A/C136A.

En effet, les vitesses de variation du message d'émission de fluorescence ont été suivies avec un appareil de mélange rapide (« stopped-flow »), dans des conditions de cycle catalytique unique (« single turn-over »). Les concentrations utilisées étaient de 10 μ M en fRMsr et de 5 à 500 mM en Met-O. Pour chaque concentration en Met-O, une diminution de l'intensité d'émission de fluorescence en fonction du temps a été observée pour tous les mutants sauf pour les fRMsr W62F/C136A (le Trp62 étant la sonde de fluorescence), Y66F/C136A et D141A/C136A (résultats non montrés).

L'ajustement des données expérimentales obtenues a été effectué à l'aide du logiciel Pro Kineticist (Applied Photophysics Ltd) au modèle cinétique $A \rightleftharpoons B \rightleftharpoons C$, comme cela avait été réalisé pour les cinétiques de la fRMsr C136A. Contrairement à ce qui avait été observé pour la fRMsr C136A, les constantes de vitesse de la seconde étape cinétique sont très lentes, du même ordre de grandeur que celles observées en absence de substrat et ne varient pas en fonction de la concentration en Met-O (résultats non montrés). Ceci montre que 1) la première étape cinétique représente l'étape réductase de la réaction, 2) la deuxième étape lente doit refléter une modification de l'environnement du Trp62 lié à la pression que subit l'enzyme dans la cellule du « stopped-flow », et 3) que, contrairement à ce qui a été observé pour la fRMsr C136A, aucune étape cinétique rapide précédant l'étape réductase n'a été observée (voir paragraphe II.1.1.3 page 65).

Les variations des valeurs de k_{obs} de la première étape cinétique en fonction de la concentration en Met-O ont ensuite été ajustées par régression non linéaire à l'équation de Michaelis-Menten afin de déterminer les valeurs de k_{obsmax} et de K_S pour les fRMsr

D141N/C136A, D143A/C136A et D143N/C136A (Figure 46 E, F et G page 85). Pour les fRMsr E125A/C136A et E125Q/C136A, les données expérimentales n'ont pas pu être ajustées à l'équation de Michaelis-Menten car la saturation n'était pas atteinte à une concentration de 500 mM de Met-O. Les valeurs des constantes de second ordre k_2 ont alors été déterminées en conditions de sous-saturation en substrat (Figure 46 B et D pages 84 et 85).

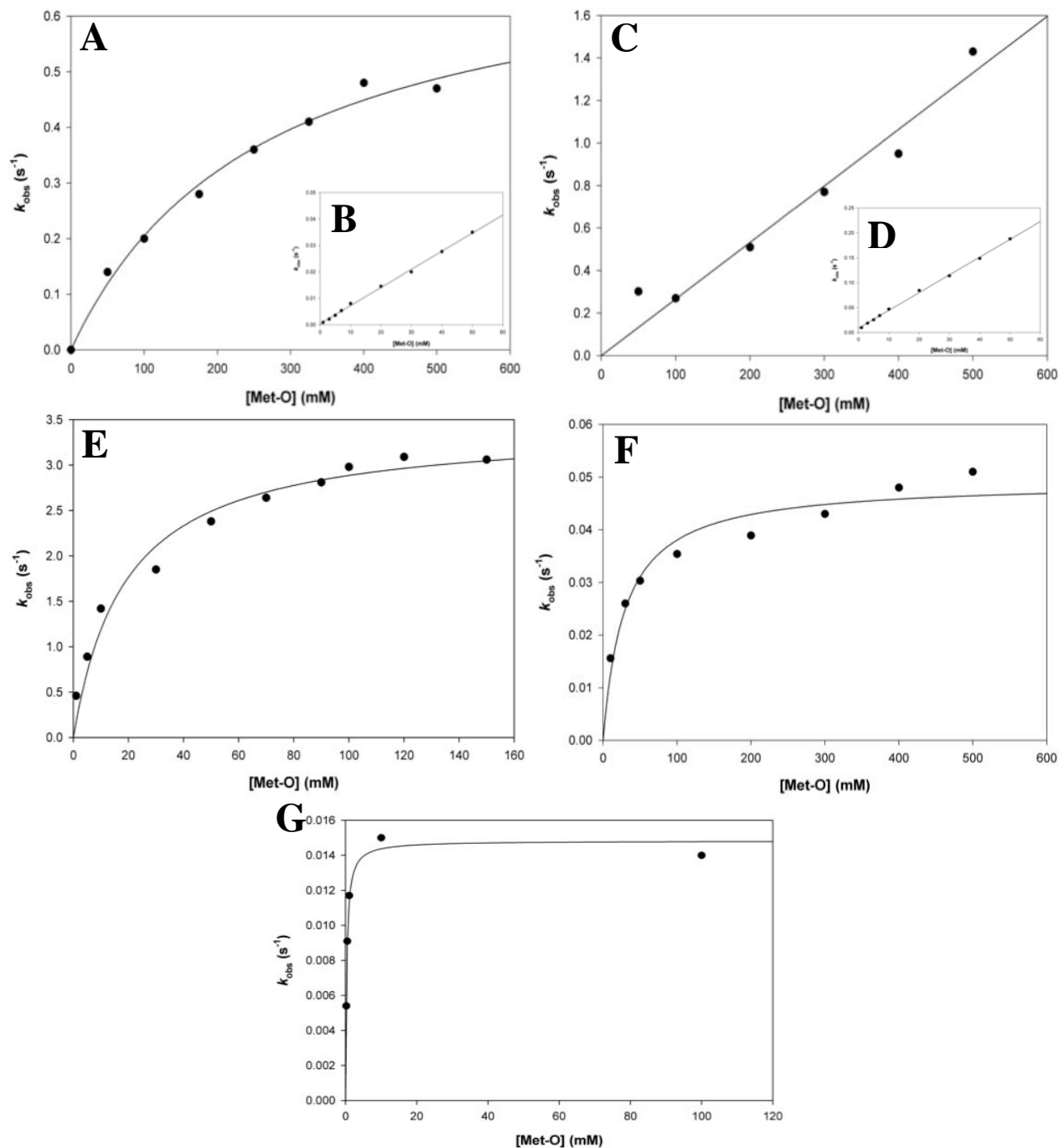


Figure 46 : Détermination des paramètres cinétiques de l'étape réductase pour les fRMsr E125A/C136A (A), E125Q/C136A (C), D141N/C136A (E), D143A/C136A (F), D143N/C136A (G) et de la constante k_2 pour les fRMsr E125A/C136A (B), et E125Q/C136A (D).

Le mélange réactionnel comporte : Met-O (en concentration variable) et fRMsR (10 μ M). Les vitesses ont été mesurées en suivant la variation du message d'émission de fluorescence de l'enzyme qui passe de la forme réduite à la forme oxydée pont disulfure grâce à un appareil de mélange rapide (Applied Photophysics SX18MV-R, Leatherhead, UK). Les mesures ont été réalisées à 25°C dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. La longueur d'onde d'excitation est fixée à 286 nm et la somme des intensités de fluorescence détectées pour des longueurs d'onde supérieures à 320 nm est mesurée, grâce à l'utilisation d'un filtre. Pour (A) et (C), la saturation n'ayant pas été atteinte pour une concentration en substrat de 500 mM, les valeurs des K_M pour la Met-O sont supérieures à 250 mM et les valeurs des k_{obsmax} sont supérieures ou égale à 0,5 s⁻¹ pour (A) et à 1,4 s⁻¹ pour (C). Les valeurs de la constante k_2 pour les mutants E125A/C136A et E125Q/C136A ont été déterminées à sous-saturation en substrat (B et D). Les valeurs des constantes obtenues sont reportées dans le tableau 9 page 87. Les points représentent les valeurs expérimentales et la ligne continue l'ajustement. Chaque point représente la moyenne de cinq mesures.

Ainsi tous les mutants présentent une constante de vitesse maximale diminuée d'un facteur variant de 10 à 2000 par rapport à celle de la fRMsR C136A, et du même ordre de grandeur que les constantes catalytiques mesurées à l'état stationnaire. De plus, les mutants ne présentent pas d'inhibition apparente par le substrat en excès, ce qui ne permet pas de comparer directement les valeurs des constantes apparentes d'affinité pour la Met-O de ces mutants avec celles de la fRMsR C136A. Cependant, le fait que la saturation en substrat n'a pas été atteinte pour une concentration en Met-O de 500 mM pour les fRMsR E125A/C136A et E125Q/C136A suggère l'existence d'une affinité fortement altérée pour ces mutants, contrairement aux fRMsR D141N/C136A, D143A/C136A et D143N/C136A.

III.6 Conclusion

Les résultats obtenus sur les différentes fRMsR montrent que tous les mutants présentent une activité en présence de Trx, et suggèrent que l'étape cinétiquement limitante de la réaction globale est associée à l'étape réductase ($k_{obsmax} \sim k_{cat}$).

Les données obtenues à l'état stationnaire et pour l'étape réductase pour les différents mutants, et leur comparaison avec celles de la fRMsR C136A, ont permis de montrer que les résidus acides jouent un rôle important dans la catalyse et la spécificité de substrat.

Ainsi, le résidu D143 joue un rôle majeur dans la catalyse de l'étape réductase. En effet, l'absence de ce résidu entraîne une diminution drastique de la constante de vitesse maximale de la première étape (facteur d'au moins 600), sans altération majeure de l'affinité. Le fait que la présence d'une Asn conduit aux mêmes effets drastiques qu'une Ala suggère un rôle de catalyseur acide/base pour la fonction carboxylate.

Résultats : Catalyse de l'étape réductase et spécificité de substrat.

De plus, un effet drastique de la substitution du Glu125 par une Ala ou une Gln sur l'affinité apparente pour le substrat Met-O (facteur 10^3 sur le k_2) a été observé, ce qui suggère que le résidu Glu125 pourrait être impliqué dans la reconnaissance du substrat Met-O. Le fait que la présence d'une Gln conduit aux mêmes effets drastiques qu'une Ala suggère un rôle non pas par établissement de liaisons hydrogène, mais plutôt *via* des interactions électrostatiques avec la fonction carboxylate.

Enfin, la substitution de l'Asp141 par une Ala conduit à une diminution drastique de la constante de vitesse k_{cat} (facteur 10^3) alors qu'aucune modification significative des constantes catalytiques k_{cat} et k_{obsmax} n'est observée lorsque c'est une Asn qui est présente. Ceci implique que le résidu Asp141 joue un rôle important dans l'étape réductase, probablement *via* l'établissement de liaisons hydrogènes avec sa fonction acide carboxylique.

En ce qui concerne les résidus Trp62 et Tyr66, le fait qu'aucune modification significative des constantes catalytiques à l'état stationnaire ait été observé pour les fRMsr W62F/C136A et Y66F/C136A ne permet pas de conclure quant à leur rôle.

| Enzyme | Etat stationnaire | | | | Etape réductase | | | |
|--------------------|-------------------------------|------------------------|------------------------|-------------------------------------------|---------------------------------|--------------------------|------------------------|-------------------------------------------|
| | k_{cat} (s ⁻¹) | K_M (mM) | K_I (mM) | k_2 (M ⁻¹ .s ⁻¹) | k_{obsmax} (s ⁻¹) | K_S (mM) | K_I (mM) | k_2 (M ⁻¹ .s ⁻¹) |
| C136A | 8,5 ± 0,6 ^a | 2,2 ± 0,5 ^a | 56 ± 16 ^a | 3864 ^c | 30 ± 5 ^b | 10 ± 3 ^b | 215 ± 108 ^b | 3000 ^c |
| D141A/C136A | (7,5 ± 0,3).10 ^{-3e} | 61 ± 10 ^e | - | 0,12 ^c | ND | ND | ND | ND |
| D141N/C136A | 2,3 ± 0,3 ^e | 25 ± 7 ^e | 440 ± 130 ^e | 92 ^c | 3,4 ± 0,2 ^e | 23 ± 5 ^e | - | 148 ^c |
| D143A/C136A | (6,5 ± 0,1).10 ^{-2e} | 126 ± 8 ^e | - | 0,52 ^c | (4,9 ± 0,2).10 ^{-2e} | 29 ± 6 ^e | - | 1,7 ^c |
| D143N/C136A | (2,0 ± 0,1).10 ^{-2e} | 1,1 ± 0,3 ^e | - | 18,2 ^c | (1,50 ± 0,05).10 ^{-2e} | 0,30 ± 0,05 ^e | - | 50 ^c |
| E125A/C136A | > 0,25 ^e | > 300 ^e | - | - | > 0,5 ^e | > 300 ^e | - | 2,0 ± 0,2 ^d |
| E125Q/C136A | > 1,6 ^e | > 300 ^e | - | - | > 1,5 ^e | > 300 ^e | - | 2,7 ± 0,1 ^d |
| W62F/C136A | 7,9 ± 1,8 ^f | 10 ± 5 ^f | 89 ± 42 ^f | 790 ^c | ND | ND | ND | ND |
| Y66F/C136A | 2,8 ± 0,1 ^f | 2,0 ± 0,3 ^f | 131 ± 20 ^f | 1400 ^c | ND | ND | ND | ND |

Tableau 9 : Paramètres cinétiques des fRMsr C136A, D141A/C136A, D141N/C136A, D143A/C136A, D143N/C136A, E125A/C136A, E125Q/C136A, W62F/C136A et Y66F/C136A de *N. meningitidis* déterminés à l'état stationnaire et pour l'étape réductase.

^a: résultats décrits dans la partie I ; ^b : résultats décrits dans le paragraphe II.1.3 (chapitre II des résultats) ; ^c : les valeurs de k_2 obtenues par le calcul de $k_2 = k_{cat}/K_M$; ^d : les valeurs de k_2 obtenues expérimentalement à sous-saturation en substrat ; ^e : résultats décrits dans le paragraphe III.3 ; ^f : résultats décrits dans le paragraphe III.5 ; -, non observé ; ND : non déterminé.

*DISCUSSION
PERSPECTIVES*

Les études de biochimie, d'enzymologie et de biologie structurale menées depuis une dizaine d'années par l'équipe Enzymologie Moléculaire et Structurale sur les Msr de classe A et de classe B avaient permis, avant que je ne débute mon projet de thèse, de mettre en évidence un mécanisme nouveau de réduction des fonctions sulfoxydes par les thiols passant par la chimie de l'acide sulfénique et d'identifier les facteurs moléculaires et structuraux impliqués dans la catalyse et la reconnaissance structurale (Boschi-Muller *et al.*, 2005 ; Boschi-Muller *et al.*, 2008).

Les résultats présentés dans le cadre de mon mémoire ont permis de montrer que la fRMsr, troisième classe de Msr, présente un mécanisme similaire de réduction d'une fonction sulfoxyde par un thiol, qui passe par la chimie de l'acide sulfénique. Ce mécanisme comprend au moins trois étapes avec : 1) la formation d'un intermédiaire acide sulfénique sur la Cys118 catalytique avec relargage d'une mole de Met par mole d'enzyme, 2) la formation d'au moins un intermédiaire pont disulfure intramoléculaire entre la Cys118 et la Cys84 et; 3) la réduction du pont disulfure par la Trx. Cependant restent posées de nombreuses questions concernant 1) le mécanisme, et en particulier les étapes de régénération de la forme réduite de la fRMsr, 2) la nature des facteurs structuraux et moléculaires impliqués dans la catalyse de l'étape réductase et la reconnaissance du substrat Met-O par la fRMsr et de la fRMsr_{ox} par la Trx, et 3) le mécanisme de l'activité de la fRMsr C84A/C94A/C136A en présence de Trx.

Ces différents points vont être discutés en intégrant les données de la littérature et les données structurales. En effet, parallèlement à mon travail, la structure tridimensionnelle de la fRMsr de *N. meningitidis* en complexe avec le substrat *L*-Met-*R*-O a été obtenue par le Dr. A. Gruez de l'équipe EMS, à une résolution de 1,25 Å (Fig. 47 page 89). Comme attendu, la fRMsr est présente sous forme dimérique et possède un repliement de type domaine GAF. Un pont disulfure est présent entre la Cys84 et la Cys118 catalytique. Comme dans toutes les autres structures de fRMsr, un clapet formé des résidus 107 à 124 couvre le site actif. Le substrat *L*-Met-*R*-O est fortement lié dans une petite cavité coiffée par ce clapet, qui contient la Cys118 catalytique et dont la liberté conformationnelle est limitée par le pont disulfure Cys84-Cys118. Il est important de noter que le pont disulfure s'est formé au cours du processus de cristallisation, l'addition de Met-O ayant été faite après l'obtention des cristaux et juste avant la congélation.

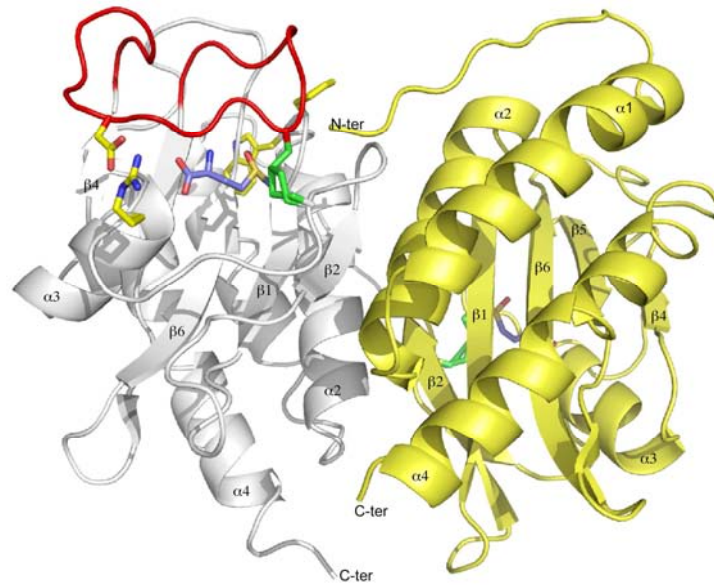


Figure 47 : Représentation de la structure tridimensionnelle du dimère de la fRMSr de *N. meningitidis* en complexe avec la *L-Met-R-O* (d'après Gruez *et al.*, 2010a).

Les deux sous-unités A et B sont représentées en mode ruban respectivement en blanc et jaune. Le clapet (résidus 107 à 124) de la sous-unité A est représenté en rouge. Les chaînes latérales des résidus Cys84 et Cys118 sont représentées en mode bâton et colorées en vert. La *L-Met-R-O* est représentée en mode bâton avec les atomes de carbone en violet, les atomes d'oxygène en rouge, les atomes d'azote en bleu et ceux de soufre en jaune. Les chaînes latérales de Arg91, Asp114, Phe61 et Trp62 sont représentées en mode bâton de couleur jaune pour les atomes de carbone, rouge pour les atomes d'oxygène et bleu pour les atomes d'azote.

I. Mécanisme des fRMSr

Les premiers travaux effectués sur les fRMSr d'*E. coli* et de *S. cerevisiae* avaient conduit les auteurs à proposer la Cys94 comme Cys catalytique, en se basant respectivement sur la structure de la fRMSr d'*E. coli* dans laquelle une molécule de MES est présente dans le site actif et sur un modèle de complexe enzyme-substrat de l'enzyme de *S. cerevisiae* (Lin *et al.*, 2007 ; Le *et al.*, 2009). Notre approche de caractérisation des propriétés de différents mutants dans lesquels chaque résidu Cys a été substitué individuellement ou non par un résidu Ala, et en particulier la caractérisation de l'activité réductase dans des conditions de cycle catalytique unique, nous a permis de démontrer clairement que la Cys118 est la Cys catalytique (Fig.48 page 90). Ceci est d'ailleurs en accord avec le mécanisme proposé pour les enzymes de *S. aureus* et *S. cerevisiae* (Bong *et al.*, 2010 ; Kwak *et al.*, 2010).

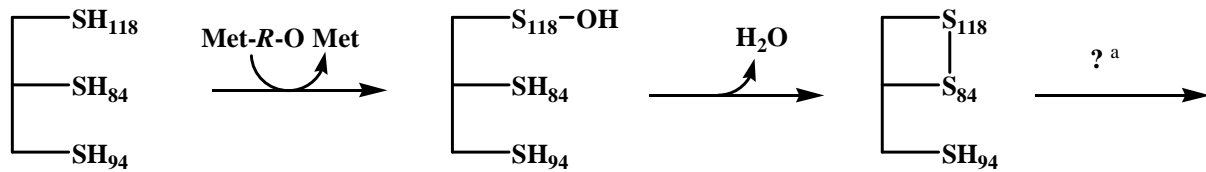


Figure 48 : Mécanisme proposé pour la réaction catalysée par la fRMsR de *N. meningitidis* jusqu'à la formation du pont disulfure Cys84-Cys118 (d'après Gruez *et al.*, 2010a).

L'étape réductase (la première) conduit à la formation d'un intermédiaire acide sulfénique sur la Cys118 catalytique et de Met. Dans une deuxième étape, l'attaque de la Cys84 sur l'atome de soufre de l'intermédiaire acide sulfénique conduit à la formation d'un pont disulfure entre les Cys118 et Cys84 et à la libération d'une molécule d'eau.^a Le retour du site actif à un état complètement réduit se fait par la réduction par la Trx du pont disulfure Cys84-Cys118, ou du pont disulfure Cys94-Cys84 s'il se forme.

I.1 Etape réductase

Ainsi, au cours de l'étape réductase, la Met-O est réduite en Met et un intermédiaire acide sulfénique doit se former sur la Cys118. Pourtant, cet intermédiaire acide sulfénique n'a pas pu être caractérisé directement par spectrométrie de masse, comme cela avait été fait pour la MsrA d'*E. coli* par exemple (Boschi-Muller *et al.*, 2000). Ceci est probablement dû à une durée de vie relativement faible de cet intermédiaire dans les conditions utilisées. Les acides sulféniques sont des composés généralement instables, cette instabilité provenant de leur très forte réactivité. En effet, la fonction acide sulfénique, au niveau de son soufre, peut exprimer à la fois un caractère nucléophile et un caractère électrophile, selon les conditions expérimentales. Différentes études sur des modèles chimiques ou enzymatiques ont montré qu'un nucléophile peut facilement attaquer le soufre de la fonction acide sulfénique, faisant ainsi appel à son caractère électrophile (Jacob *et al.*, 2006 ; Sarma et Mugesh, 2007). Ces réactions conduisent à la formation d'une liaison covalente *via* le départ du groupement OH sous forme de molécule d'eau, et sont d'autant plus rapide en présence d'une catalyse assistée pour le départ de la molécule d'eau. Ainsi, un acide sulfénique peut évoluer vers la formation 1) d'un disulfure *via* l'attaque d'un atome de soufre d'un thiol sous forme thiolate, 2) d'un sulfénylamide *via* l'attaque du NH peptidique du résidu adjacent, ou 3) d'un thiosulfinate *via* l'attaque d'une deuxième molécule d'acide sulfénique. Or la caractérisation par des approches de SLS et de LC-MS de la forme oxydée de la fRMsR ne contenant plus que la Cys118 catalytique a montrée, en absence de réducteur, la formation de multimères covalents avec des masses compatibles avec des formes disulfure (Dimère -2 Da), thiosulfinate (Dimère +14 Da) et sulfénylamides (Monomère -2 Da) qui ne peuvent être

obtenues qu'à partir de la formation préalable d'un intermédiaire acide sulfénique sur la Cys118 (Fig. 49 page 91). Cependant l'existence de ces réactions secondaires impliquerait d'avoir des stœchiométries de formation de moles de Met par mole de monomère d'enzyme inférieures à 1, ce qui a été observé dans les conditions de « quench flow », mais pas dans les conditions classiques sur des temps de réaction de plusieurs secondes ou minutes. Il semblerait donc que les conditions de pression modifient les vitesses des différentes réactions secondaires. Pour le confirmer, il faudra quantifier les différentes espèces monomériques et dimériques formées après des temps courts (technique quench-flow) par l'approche de spectrométrie de masse LC-MS.

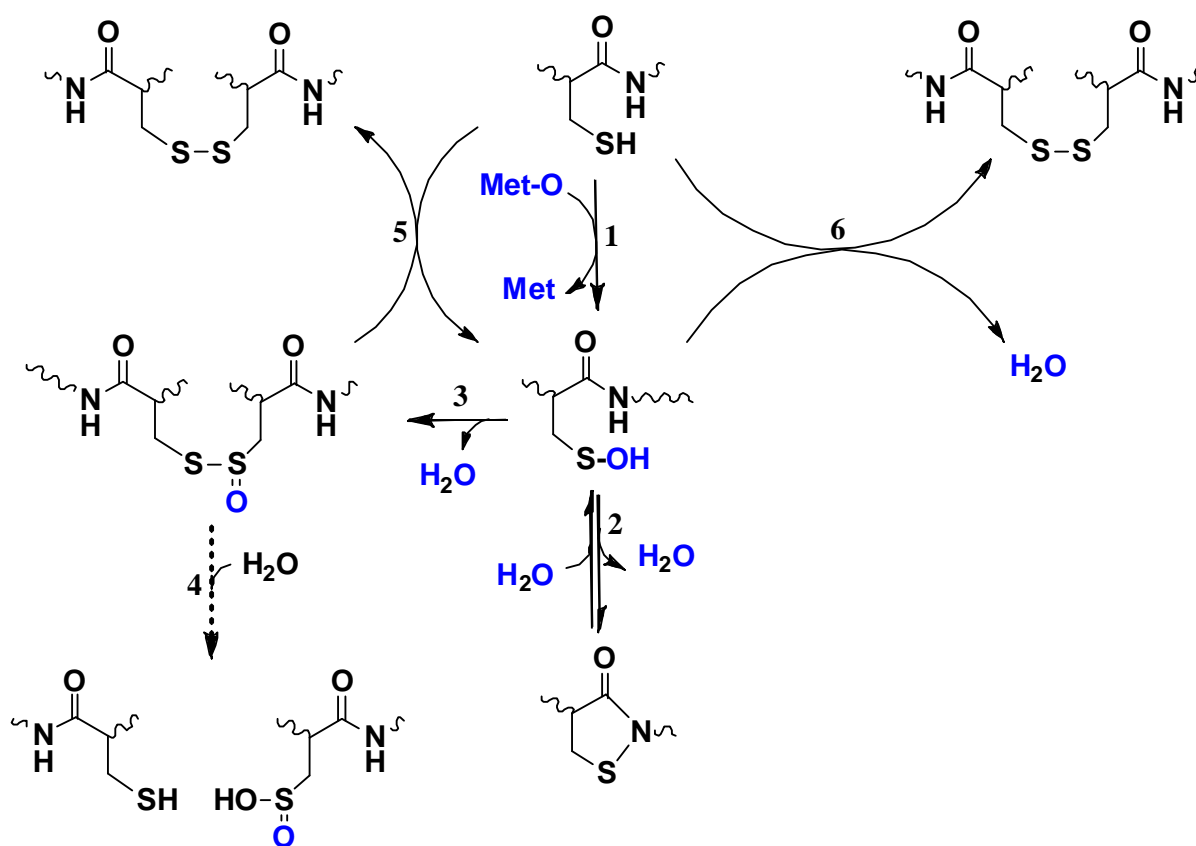


Figure 49 : Schéma des réactions secondaires possibles à partir de l'intermédiaire acide sulfénique du triple mutant de la fRMsR. L'oxydation de la Cys118 de l'enzyme par la Met-O conduit à la formation d'un intermédiaire acide sulfénique sur la Cys118 (1). L'intermédiaire acide sulfénique peut évoluer vers la formation réversible d'un intermédiaire sulfénylamide (Masse du monomère -2 Da) obtenu *via* l'attaque du NH peptidique du résidu 118 sur l'atome de S de la fonction acide sulfénique (2). Deux molécules d'acide sulfénique peuvent se condenser pour former un dimère covalent, par liaison thiosulfinate (3)(Masse dimère + 14 Da), qui peut s'hydrolyser sur des temps longs pour libérer l'enzyme réduite et l'enzyme suroxydée sous forme acide sulfinique (4) (la forme SO₂ n'apparaît qu'après plusieurs jours en absence de TFA). Enfin, un dimère covalent par pont disulfure (Masse dimère - 2 Da) peut être formé *via* l'attaque de la Cys118 sous forme thiolate sur l'atome de S de l'intermédiaire acide sulfénique (5) ou thiosulfinate (6).

I.2 Formation du(des) pont(s) disulfure(s)

Au niveau de la seconde étape, reste posée la question de la nature et du nombre du(des) pont(s) disulfure formé(s). En effet, les résultats expérimentaux montrent bien la formation pour l'enzyme sauvage d'un seul pont disulfure concomitant à la libération d'une seule mole de Met par monomère, mais ils ne permettent pas pour autant d'exclure la formation de deux ponts disulfures successifs. L'intermédiaire acide sulfénique se formant sur la Cys118, un premier pont disulfure doit se former entre la Cys118 et la Cys84 ou la Cys94. Nos résultats expérimentaux montrent que les Cys84 et Cys94 peuvent toutes les deux former un pont disulfure avec la Cys118 catalytique, tout du moins en absence de réducteur respectivement dans les fRMsr C94A/C136A et C84A/C136A. Une façon d'identifier le premier pont disulfure qui se forme dans l'enzyme sauvage serait d'atteindre la vitesse de formation de ces deux ponts disulfure, le premier qui se forme devant présenter une constante de vitesse supérieure à l'autre. Ceci n'est pas possible sur l'enzyme sauvage mais l'utilisation des fRMsr C84A/C136A et C94A/C136A pourrait apporter des éléments de réponse. Cependant, les résultats montrent que la vitesse de formation des ponts disulfure Cys118-Cys84 et Cys118-Cys94 est limitée par celle de l'étape réductase (de l'ordre de 15 s^{-1}), ce qui ne permettra pas de conclure.

L'examen de la structure du complexe enzyme-substrat de l'enzyme de *N. meningitidis* et du complexe enzyme-MES de l'enzyme d'*E. coli* montre d'une part la présence d'un unique pont disulfure Cys118-Cys84, et d'autre part que la Cys94 est située à l'extrémité N-terminale de l'hélice α_3 à une distance d' $\sim 8 \text{ \AA}$ de la Cys118. Ainsi, la Cys94 n'est pas bien positionnée pour former un pont disulfure avec la Cys118, sauf à considérer une flexibilité de l'enzyme. Ceci suggère que le pont disulfure Cys118-Cys84 est probablement le premier formé. Enfin, l'étude d'une fRMsr qui naturellement ne possède que deux Cys pourrait permettre de répondre. En effet, l'analyse des séquences de fRMsr montre que 5,6 % et 1,4 % des séquences contiennent uniquement la Cys118 et respectivement la Cys94 et la Cys84 (Fig. SI Annexe 1). Si ces enzymes sont actives et efficaces, cela montrerait que le site actif de la fRMsr peut utiliser les deux ponts disulfure pour la catalyse.

Reste posée la question de la suite du mécanisme ? Y a-t-il un échange de pont disulfure impliquant la Cys94 ? En effet, bien que dans la structure cristallographique les distances entre les Cys118 et Cys84 et la Cys94 soient trop élevées pour permettre la formation d'un pont disulfure, les résultats montrent que le site actif de la fRMsr est flexible puisqu'en solution les ponts disulfure Cys118-Cys94 et Cys84-Cys94 peuvent se former.

Cependant, la stœchiométrie de formation de Met en absence de réducteur étant de un pour la fRMsR de type sauvage, les formes oxydées pont disulfure Cys118-Cys94 et Cys84-Cys94, si elles se forment, ne sont donc pas compétentes pour la catalyse. Ceci s'explique facilement pour le pont disulfure Cys118-Cys94 qui ne permettrait pas de régénérer la Cys118 sous forme réduite, contrairement au pont Cys84-Cys94. Il faudrait dans ce cas que la formation du pont disulfure perturbe la conformation du site actif, qui en retour ne permettrait plus la formation d'un complexe compétent avec la *L*-Met-*R*-O. Ce comportement de la fRMsR de type sauvage est clairement différent de celui de la MsrA d'*E. coli*, qui possède également trois Cys et forme deux ponts disulfure successifs. Dans ce cas, la formation du second pont disulfure libère la Cys51 catalytique sous forme thiolate et permet la réduction d'une deuxième molécule de Met-O, en absence de réducteur (Boschi-Muller *et al.*, 2000). Une autre façon de répondre serait d'étudier les vitesses de réduction des différents ponts disulfure par la Trx en utilisant les différents mutants Cys générés. Or pour le moment, nous n'avons pas pu atteindre ces vitesses par une approche basée sur la différence de fluorescence de la Trx sous forme réduite et oxydée. Nous avons cependant pu montrer que les trois ponts disulfures (118-84, 118-94 et 84-94) peuvent être réduits par la Trx en identifiant les formes oxydées de la Trx par chromatographie. Ainsi, l'utilisation de l'approche par quench-flow couplée à la quantification des formes réduites et oxydées de la Trx par chromatographie en phase inverse devrait permettre d'atteindre ces vitesses, en se plaçant dans des conditions de vitesse inférieures à 1 s^{-1} (vitesse maximale mesurable en quench-flow avec un temps minimum de 5 ms).

I.3 Le devenir de l'intermédiaire acide sulfénique en l'absence de la Cys84

Enfin, le fait que la fRMsR ne possédant que les Cys118 et Cys94 soit active en présence de Trx pourrait suggérer que le pont disulfure Cys118-Cys94 se forme et est réduit par la Trx. Ces résultats sont d'ailleurs en accord avec les données publiées sur les fRMsR de *S. aureus* et *S. cerevisiae*, qui montrent que l'absence de la Cys84 n'abolit pas l'activité en présence de Trx (Bong *et al.*, 2010 ; Kwak *et al.*, 2010). En fait, cela dépend de la nature de l'intermédiaire de la fRMsR C84A/C136A qui est réduit par la Trx, à savoir l'intermédiaire acide sulfénique sur la Cys118 (ou une forme dérivée obtenue par réaction secondaire) ou l'intermédiaire pont disulfure Cys118-Cys94.

D'un point de vue réactivité chimique, la formation d'un acide sulfénique peut être facilement réversée par sa réaction avec des réducteurs à thiol *via* le passage par un intermédiaire pont disulfure qui sera ensuite réduit. Dans le cas d'un site actif enzymatique, la

réactivité d'une fonction acide sulfénique peut être fortement modulée par l'environnement et l'accessibilité du site actif aux réducteurs. En ce qui concerne les MsrA et MsrB, notre équipe a clairement montré que les intermédiaires acides sulféniques formés au cours de la catalyse sur la Cys catalytique des MsrA et MsrB d'*E. coli* et de *N. meningitidis* ne sont pas réduits efficacement par la Trx. Par contre, l'intermédiaire acide sulfénique formé sur la MsrB à une Cys de *X. campestris* est réductible par le système GSH/glutarédoxine, mais la réaction est très lente (thèse Fabrice Neiers). Il est important de noter que dans la littérature plusieurs études montrent que des MsrA et B à une seule Cys sont actives en présence de Trx et /ou du système GSH/Grx mais les réactions sont lentes et la forme oxydée réduite par ces réducteurs à thiol n'a jamais été identifiée (Kim and Kim, 2008 ; Lee and Kim, 2008 ; Neiers, 2007 ; Tarrago *et al.*, 2009 ; Kim *et al.*, 2011).

En ce qui concerne la nature de l'intermédiaire de la fRMsr C84A/C136A qui est réduit par la Trx, plusieurs arguments s'opposent. Le fait que le mutant ne possédant plus que la Cys118 est actif en présence de Trx avec une valeur de k_{cat} identique à celle du mutant C84A/C136A pourrait laisser penser que le même intermédiaire est réduit par la Trx dans les deux cas et le seul intermédiaire commun est l'intermédiaire acide sulfénique. Cependant, l'intermédiaire acide sulfénique ne s'accumule pas pour la fRMsr C84A/C136A puisque l'étape cinétiquement limitante se situe après l'étape de formation du pont disulfure Cys94-Cys118. De plus, l'intermédiaire acide sulfénique du mutant qui ne contient plus que la Cys118 n'est pas stable, tout du moins sur une échelle de temps de 10 s. La détermination des constantes de vitesse d'apparition des différentes espèces oxydées par une approche de quench-flow couplée à la LC-MS pourrait permettre d'apporter des éléments de réponse.

II. Catalyse et spécificités structurales des fRMsr

II.1 Catalyse de l'étape réductase

En termes de catalyse, l'étape réductase étant rapide, cela implique une déprotonation de la Cys catalytique et une protonation de l'atome d'oxygène de la fonction sulfoxyde du substrat afin de faciliter l'attaque nucléophile du thiolate et la formation de l'état de transition de type sulfurane. Ceci doit se faire *via* un acide aminé accepteur et donneur de proton, éventuellement par un relais *via* une molécule d'eau. Les résultats cinétiques suggèrent que la Cys118 catalytique est activée lors de la formation du complexe Michaélien et que l'Asp143, qui se trouve sur brin $\beta 6$, joue le rôle de catalyseur acide. En effet, la Cys118 doit être sous forme thiolate pour attaquer le soufre de la fonction sulfoxyde du substrat et comme le pK_{app}

de la Cys118 est de 8,5 dans l'enzyme libre, la liaison du substrat doit contribuer à la diminution de son pK_{app} , comme c'est d'ailleurs le cas pour les MsrA et MsrB. Ceci est en accord avec la structure cristalline du complexe enzyme-substrat dans laquelle, si la Cys118 est modélisée sous forme réduite, le thiolate peut adopter une orientation favorable pour attaquer l'atome de soufre du substrat et conduire à la formation d'un sulfurane avec les groupements SH et OH en positions apicales. De plus, le fait que le remplacement de l'Asp143 par une Ala ou une Asn conduise à une diminution drastique des constantes catalytiques, d'un facteur au moins 10^2 pour l'étape réductase et l'état stationnaire, suggère fortement un rôle de la fonction carboxylate du résidu Asp143 dans la catalyse de l'étape réductase (Fig. 50 page 95).

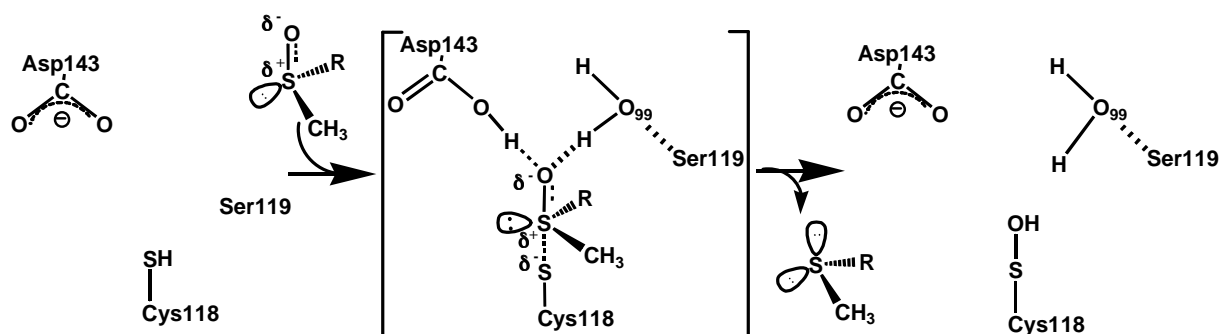


Figure 50: Mécanisme catalytique proposé pour l'étape réductase de la fRMsR (d'après Gruez *et al.*, 2010a)

Le substrat se lie au site actif avec sa fonction sulfoxyde largement polarisée, conduisant à un état de transition compatible avec un sulfurane de géométrie bipyramidale. Le réarrangement de l'état de transition conduit à la formation d'un intermédiaire acide sulfénique sur le résidu Cys118 catalytique et de Met. Dans ce scénario, le résidu Asp143 et une molécule d'eau (H_2O_{99}), elle-même liée par liaison hydrogène avec l'azote de la chaîne principale de la Ser119, stabilisent l'état de transition sulfurane *via* des interactions avec le groupement OH, une situation semblable à celle décrite pour les MsrB.

Nos études ne permettent cependant pas de savoir si l'Asp143 participe, directement ou indirectement, à l'activation de la Cys118. En absence de substrat, le pK_a de la Cys118 étant de 8,5, on peut exclure une interaction directe entre la Cys118 et l'Asp143. Ceci est en accord avec la structure tridimensionnelle de l'enzyme de *N. meningitidis* en complexe avec le substrat qui montre une distance entre la Cys118 et l'oxygène de l'Asp143 qui semble incompatible avec une interaction directe entre les deux acides aminés. Par contre, l'oxygène du sulfoxyde interagit fortement non seulement avec l'Asp143 mais aussi une molécule d'eau (H_2O_{99}), qui forme elle-même une liaison hydrogène avec l'azote de la chaîne principale de la

Ser119 (Fig. 51 page 96). Un scénario peut donc être proposé dans lequel le résidu Asp143 et la molécule d'eau doivent interagir non seulement avec l'oxygène de la fonction sulfoxyde sous forme polarisée, mais probablement aussi avec le groupement OH de l'état de transition sulfurane, et directement ou indirectement avec le groupement OH de l'intermédiaire acide sulfénique (Fig. 50). Il est intéressant de noter que la catalyse de l'étape réductase des fRMs est semblable à celle des MsrA et MsrB, à l'exception de la nature des résidus qui interviennent, excepté bien sûr la Cys catalytique. En effet, les résultats obtenus par l'équipe montrent que dans les MsrA, le rôle de relais en tant que catalyseur acide serait assuré par le résidu Glu94 et que l'abaissement du pK_{app} de la Cys catalytique fait intervenir, de façon directe ou indirecte, non seulement ce résidu Glu mais également les résidus Tyr82 et Tyr134. Pour les MsrB, c'est un résidu His invariant qui joue le rôle de catalyseur acide et qui permet également l'abaissement du pK_{app} de la Cys catalytique dans le complexe enzyme-substrat. Ceci montre que différentes solutions ont été retenues au cours de l'évolution pour catalyser de façon efficace la réduction d'une fonction sulfoxyde par une Cys d'un site actif.

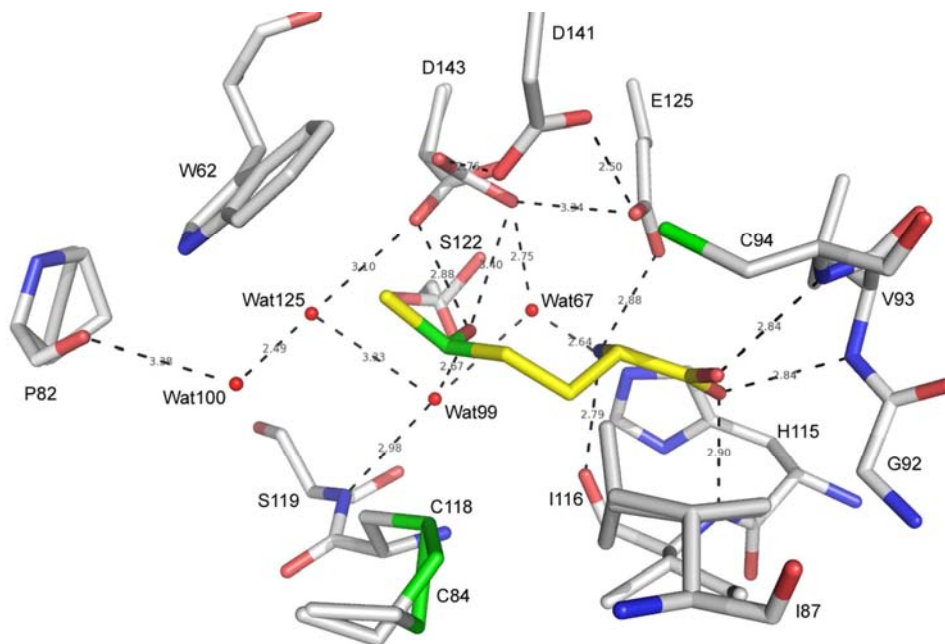


Figure 51 : Représentation en mode bâton du site actif de la fRMs de *N. meningitidis* avec la *L*-Met-*R*-O fixée (d'après Gruez *et al.*, 2010a).

Les atomes de carbone de la *L*-Met-*R*-O sont représentés en jaune, tandis que ceux des acides aminés environnants sont en gris. Les atomes d'oxygène sont représentés en rouge, ceux d'azote en bleu et ceux de soufre en jaune. Les lignes pointillées indiquent les interactions polaires. Les molécules d'eau sont représentées en mode boule. Les résidus Pro82 et Cys84 sont présents dans deux conformations. La chaîne latérale de la Ser122 a été affinée en trois conformations. La chaîne latérale de l'Asp143 a été affinée en conformation double. La conformation mineure de la chaîne latérale de l'Asp143 (occupation 0,3) interagit avec la

molécule d'eau Wat125 qui pourrait interagir avec la chaîne latérale de la Cys118 après une simple rotation.

II.2 Spécificités structurales

En termes de spécificité, les études cinétiques ont montré que la fRMsr est sélective de l'isomère *L* du carbone C α du substrat Met-O. De plus, seule la forme libre des Met-O est réduite, les Met-O dans un contexte de peptides ou de protéines et le DMSO n'étant pas substrat. Ces résultats suggèrent un rôle important dans la reconnaissance des groupements COO⁻ et NH₃⁺, contrairement à ce qui est observé pour les MsrA et MsrB. L'inspection de la structure cristalline de la fRMsr de *N. meningitidis* en complexe avec le substrat montre que 1) la fonction carboxylate chargée négativement est localisée dans un site de reconnaissance oxyanion formé par les NH peptidiques des résidus Val93, Cys94 et Ile116, et 2) le groupement NH₃⁺ est fortement stabilisé par un réseau de liaisons hydrogène impliquant l'oxygène de la chaîne principale du résidu Ile116 et la chaîne latérale du résidu Glu125 invariant, elle-même stabilisée par des interactions avec les résidus invariants Asp141 et Asp143 (Fig.52 page 98). L'ensemble de ces données suggère un rôle important des chaînes latérales des résidus invariants Asp141 et Glu125 dans la spécificité des fRMsr, ce qui est conforté par l'étude des propriétés des mutants de ces deux positions. Ainsi, le résidu Glu125 est impliqué dans la reconnaissance et/ou le positionnement du substrat Met-O probablement *via* la stabilisation du groupement NH₃⁺ qui est pris en sandwich entre d'un côté les deux oxygènes de la fonction acide carboxylique du résidu Glu125 et de l'autre la fonction carbonyle de la chaîne principale du résidu Ile116. Le résidu Asp141 serait lui impliqué dans le positionnement des résidus Asp143 et Glu125, d'où un rôle indirect dans la spécificité et la catalyse. Toutes ces interactions stabilisantes favorisent la liaison de l'épimère *L*-Met-O, empêchent la liaison d'une Met-O dans un contexte de peptides ou de protéines en raison de l'encombrement stérique, et sont responsables de l'absence d'activité avec le DMSO en raison, en particulier, de l'absence de groupes COO⁻ et NH₃⁺ dans le DMSO.

De plus, les fRMsr sont spécifiques de l'isomère *R* de la fonction de sulfoxyde de la Met-O. Cette préférence stéréochimique peut s'expliquer par la stabilisation de la chaîne latérale de la Met-O dans le site actif. Ainsi les groupes méthylène- β et γ sont stabilisées par des interactions de van der Waals avec les chaînes latérales des résidus invariants Ile87, Ile116 et Tyr66, ce qui introduit des contraintes stériques qui empêchent tout renversement du groupe SOCH₃ autour de l'angle χ . Les chaînes latérales des résidus Ile87 et Leu75 sont situées de part et d'autre de la chaîne latérale aromatique de la Tyr66. Face à la chaîne latérale de la Leu75, le noyau indole du Trp62 invariant assure la stabilisation du méthyle ϵ du

substrat, qui pointe perpendiculairement à l'indole. La chaîne latérale aromatique de la Phe78 forme une plateforme sur laquelle les résidus Leu75 et Trp62 sont empilés, ce qui détermine la conformation de leur chaîne latérale. Enfin, les chaînes latérales d'Ile87, Tyr66 et Trp62 créent une surface hydrophobe à l'intérieur du site actif qui est nécessaire à la liaison du substrat. Le fait que le remplacement du Trp62 par un résidu Phe ne modifie pas les paramètres cinétiques de l'enzyme montre que soit la seule stabilisation du groupement méthyle- ϵ n'est pas un élément majeur pour la reconnaissance et le positionnement efficace du substrat, contrairement à ce qui se passe pour les MsrA et MsrB, soit la Phe introduite est capable de positionner correctement le groupement méthyle- ϵ . Dans ce contexte, il serait intéressant d'étudier la contribution des interactions de van der Waals entre le substrat et les chaînes latérales des résidus Ile87, Ile116 et Tyr66. Ainsi, l'étude de l'effet du remplacement des résidus Ile87, Ile116 et Tyr66 par des résidus Ala, et l'utilisation d'analogues de substrat décarboxylé ou désaminé en position α pourrait permettre de mieux comprendre la spécificité de substrat des fRMsr.

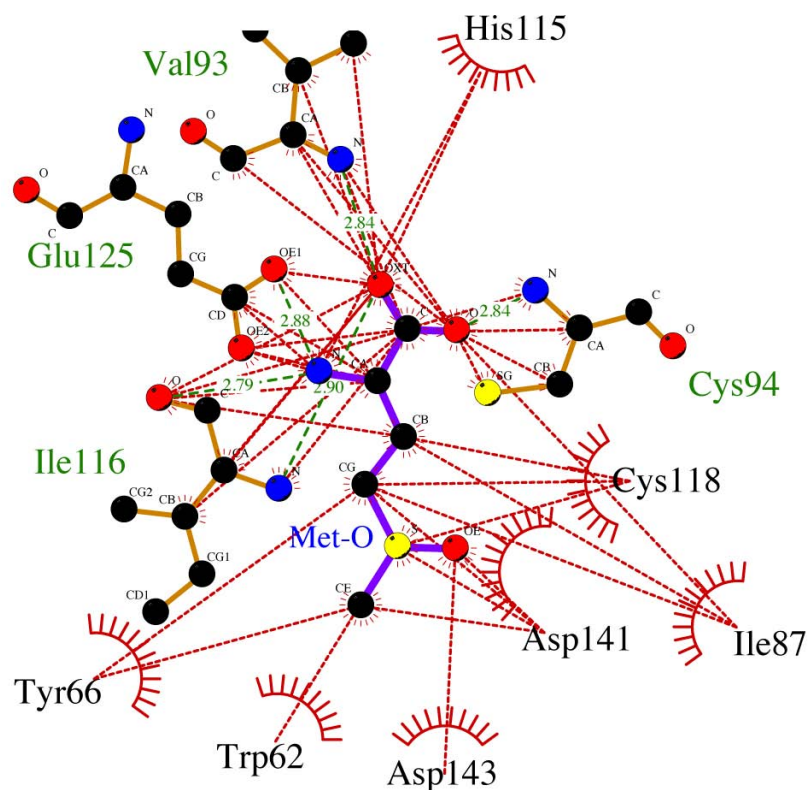


Figure 52 : Diagramme des interactions existant entre le site actif de la fRMsr de *N. meningitidis* et le substrat *L-Met-R-O*.

La molécule de Met-O et les résidus Cys94, Glu125 et Ile116 sont représentés en mode bâton. Les atomes de carbone, d'oxygène, d'azote et de soufre sont représentés respectivement par des sphères noires, rouges, bleues, et jaunes. Les liaisons hydrogène sont représentées en

traits pointillés rouges et les arcs de cercle avec des rayons rouges représentent les contacts de van der Waals. La figure a été réalisée avec le logiciel LigPlot+ à partir du fichier pdb 3MMH.

Parallèlement à nos travaux, le groupe de Chi et collaborateurs a publié les structures cristallines des formes oxydée, réduite et complexée au substrat de la fRMsr de *Staphylococcus aureus* (Bong *et al.*, 2010). Or la superposition des structures cristallines des complexes enzyme-substrat de *N. meningitidis* et de *S. aureus* montre que le substrat n'est pas positionné de la même façon dans les deux structures (Fig. 53 page 99). Ainsi, on observe une inversion de la position des groupements sulfoxyde et carboxylate du substrat entre les deux structures, ce qui montre que la structure du complexe de l'enzyme de *S. aureus* n'est pas un complexe efficace pour la catalyse. En effet, la réduction de la fonction sulfoxyde nécessite la protonation de l'oxygène qui n'est pas possible si le groupement sulfoxyde est stabilisé dans le site de reconnaissance oxyanion. De plus, dans la structure du complexe de *S. aureus*, la Cys118 n'est pas bien positionnée pour réaliser une attaque nucléophile sur l'atome de soufre de la fonction sulfoxyde. Ceci nous a conduits à publier une lettre à l'éditeur dans *J. Biol. Chem.* (voir annexe 2).

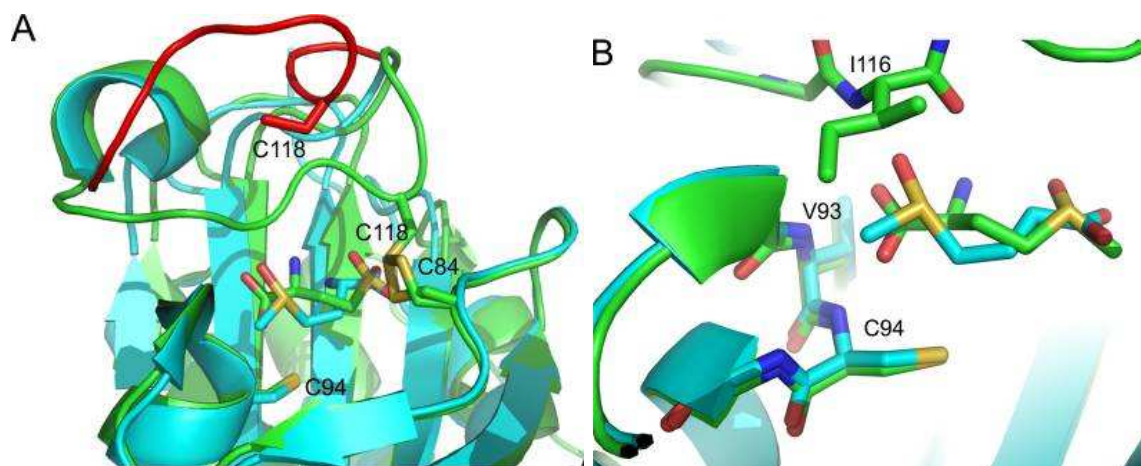


Figure 53 : Superposition des structures des fRMsr de *N. meningitidis* (en vert) et de *S. aureus* (en cyan) en complexe avec la *L-Met-R-O* (d'après Gruez *et al.*, 2010b).

A La conformation d'une partie du clapet (résidus 113 à 120, qui contient la Cys118 catalytique), n'est pas superposable (en rouge sur la structure de la fRMsr de *S. aureus*, en vert sur celle de *N. meningitidis*). Les chaînes latérales des résidus Cys84, Cys94 et Cys118 et la *L-Met-R-O* sont représentées en mode bâton avec les atomes d'oxygène en rouge, l'azote en bleu, les atomes de soufre en jaune et les atomes de carbone en vert (*N. meningitidis*) ou en cyan (*S. aureus*). B Sont représentés en plus les azotes peptidiques des résidus invariants

Val93, Cys94 et Ile116 en mode bâton. Le rôle du site de reconnaissance oxyanion serait de stabiliser la liaison du carboxylate de la *L*-Met-*R*-O.

Des travaux présentés dans ce mémoire, seule une partie a pu être publiée à ce jour, celle concernant la caractérisation du mécanisme de la fRMsr. La caractérisation de l'étape réductase ainsi que des résidus impliqués dans la catalyse et la spécificité de substrat s'étant révélée techniquement particulièrement complexe, par rapport à celle des MsrA et MsrB réalisée précédemment par l'équipe Enzymologie Moléculaire et Structurale, ces résultats n'ont pas pu être publiés pour le moment. Pour pouvoir être publiés, les interprétations des cinétiques par suivi de fluorescence des différents mutants devront être confortées par la détermination de la vitesse de formation de Met et de disparition des thiols libres par la technique de « quench-flow ». En ce qui concerne les étapes qui suivent l'étape réductase, la situation est plus complexe et nécessite de poursuivre le travail pour essayer notamment d'obtenir les cinétiques de formation des différentes espèces par « quench-flow » couplé à une analyse par LC-MS. Enfin, ces travaux ouvrent des perspectives intéressantes en ce qui concerne en particulier la chimie des réactions impliquant la formation d'intermédiaires acides sulféniques.

MATERIELS

ET

METHODES

I. Matériel biologique et milieux de culture

I.1. Souches bactériennes

- *E. coli* DH5 α : supE44, Δ lacU169 (ϕ 80 lacZ Δ M15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1. Cette souche a été utilisée pour l'obtention d'ADN plasmidique en quantité importante et pour la production de la fRMsr sauvage de *N. meningitidis*.

- *E. coli* HB101 : supE44, hsdS20 (r^-_B , m^-_B), recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1. Cette souche a été utilisée pour la production des mutants de la fRMsr de *N. meningitidis*.

I.2. Vecteurs plasmidiques

Le vecteur recombinant pSK-fRMsr est construit à partir d'un phagemide recombinant pBluescript II KS (+) portant la séquence codante de la fRMsr de *N. meningitidis* sous le contrôle du promoteur constitutif P_{Lac} (pSK-Msr, Boschi-Muller *et al.*, 2000). Ce vecteur porte également le gène de résistance à l'ampiciline (Fig. 54 page 101). Tous les mutants étudiés sont issus de cette construction.

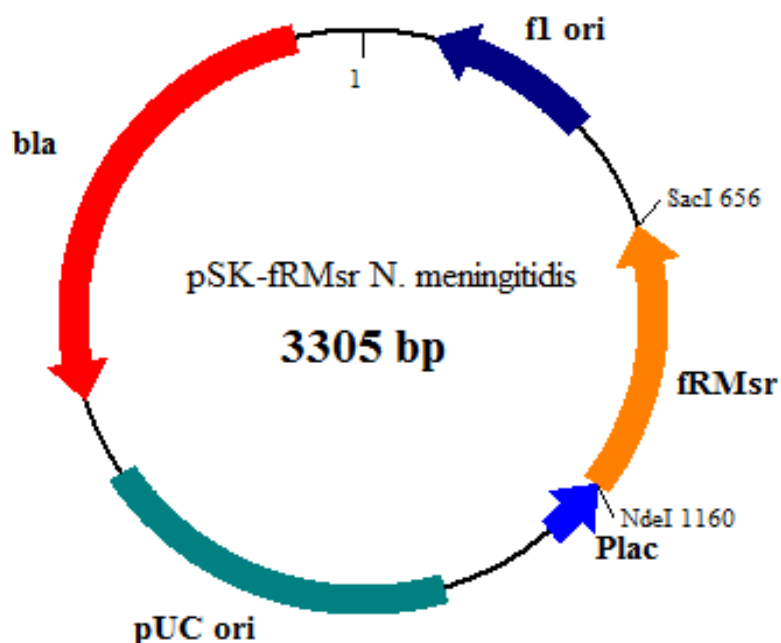


Figure 54 : Représentation schématique du vecteur qui permet la production de la fRMsr de *N. meningitidis* dans les cellules d'*E. coli*.

Les sites de restriction *NdeI* et *SacI* ayant servis à l'insertion du gène d'intérêt sont représentés.

P_{Lac} = promoteur de l'opéron lactose ; f1 ori = origine de réplication du phage filamenteux f1 ; bla = gène de résistance à l'ampicilline ; pUC ori = origine de réplication du plasmide ; fRMsr = ORF de la fRMsr de *N. meningitidis*.

I.3. Milieux de culture bactérienne

I.3.1. Milieu riche Luria Bertani (LB)

Composition : Tryptone (10 g/L), extrait de levure (5 g/L), NaCl (5 g/L). Le pH est ajusté à 7,5 par ajout de NaOH et le milieu est autoclavé pendant 20 min à 120 °C. Ce milieu est utilisé pour la production en grande quantité de toutes les protéines recombinantes utilisées dans ce travail. Pour les milieux gélosés, de l'agar est ajouté à raison de 17 g/L de milieu de culture avant stérilisation. Afin de sélectionner les clones recombinants, l'ampicilline est ajoutée à une concentration de 200 mg/L après stérilisation.

I.3.2. Milieu riche SOB

Composition : Tryptone (20 g/L), extrait de levure (5 g/L), NaCl (10 mM), KCl (2,5 mM), MgCl₂ (10 mM), MgSO₄ (10 mM). Ce milieu est utilisé lors de la préparation de cellules compétentes. Le milieu est autoclavé 20 min à 120 °C.

II. Techniques de génie génétique

II.1. Minipréparation et analyse d'ADN plasmidique

La minipréparation d'ADN est une technique fiable pour la préparation rapide de l'ADN plasmidique à partir de 5ml de culture bactérienne. Les minipréparations ont été réalisées à l'aide du kit Nucleospin plasmid (Macherey Nagel).

Un séquençage des constructions plasmidiques a été réalisé (GATC Biotech) afin de vérifier à la fois l'intégrité de la séquence d'ADN, l'introduction des mutations désirées et l'absence de mutations parasites.

II.2. Clonage de la fRMsr de *N.meningitidis*

Le clonage de l'ORF de la fRMsr a été réalisé à partir de l'ADN génomique de la souche 8013 de *N. meningitidis*. Les oligodésoxyribonucléotides utilisés permettent d'introduire un site *Nde*I (5'-CATATG-3') au niveau du codon ATG de la première Met et un site *Sac*I (5'-GAGCTC-3') après le codon stop. Le fragment de 504 bp obtenu par PCR

est ensuite purifié par électrophorèse sur gel d'agarose 0,8 % contenant du bromure d'éthidium suivie d'une extraction (kit Nucleospin[®] Extract II, Macherey Nagel).

Le fragment d'ADN qui correspond à l'ORF de la fRMs_r et le vecteur pSK-Msr sont ensuite hydrolysés par les enzymes de restriction *Nde*I et *Sac*I pendant 2h à 37°C dans les conditions préconisées par les fournisseurs. Les fragments hydrolysés sont ensuite purifiés par électrophorèse sur gel d'agarose 0,8 % suivie d'une extraction à l'aide du kit Nucleospin Extract II (Macherey Nagel).

Le fragment d'ADN qui correspond à l'ORF de la fRMs_r de *N. meningitidis* est inséré par ligation aux sites de restriction *Nde*I et *Sac*I du vecteur, en utilisant un rapport molaire insert/vecteur d'environ 3/1. Le mélange de ligation est incubé en présence d'ADN ligase, qui catalyse la formation de liaisons phosphodiester, pendant une nuit à 16°C. Le produit de la réaction est alors utilisé lors de la réaction de transformation de cellules compétentes DH5 α d'*E.coli*. Les cellules transformées sont reprises dans 200 μ l de milieu SOC et incubées à 37°C pendant 1h avant d'être étalées sur milieu LB solide additionné d'ampicilline à 200 μ g/ml.

L'ADN plasmidique de quelques clones bactériens est extrait et analysé afin de contrôler la qualité des vecteurs obtenus.

II.3. Mutagenèse dirigée du matériel plasmidique

La mutagenèse dirigée a été réalisée en utilisant l'ADN polymérase thermostable *Pfu*, qui est plus fidèle que la *Taq* polymérase et permet grâce à son activité exonucléasique 3' \rightarrow 5' de corriger les erreurs commises lors de la réplication de l'ADN plasmidique. Cette méthode nécessite la présence de deux oligodésoxyribonucléotides, chacun complémentaire d'un des deux brins du plasmide, qui serviront d'amorces 3'OH dans la réaction de PCR pour la synthèse de nouveaux brins.

Le produit de la réaction de PCR obtenu est ensuite traité par l'endonucléase *Dpn*I (qui a comme séquence cible 5'-G^{m6}ATC-3'), qui est spécifique des brins d'ADN méthylés et hémiméthylés, et qui va permettre d'hydrolyser l'ADN parental afin de ne sélectionner que l'ADN nouvellement synthétisé portant la mutation introduite. Le mélange obtenu est alors utilisé pour transformer des cellules compétentes DH5 α d'*E.coli*. Les cellules transformées sont reprises dans 200 μ l de milieu SOC et incubées à 37°C pendant 1h avant d'être étalées sur milieu LB solide additionné d'ampicilline à 200 μ g/ml.

L'ADN plasmidique de quelques clones bactériens est extrait et analysé afin de contrôler la présence des mutations introduites.

III. Production et purification de la fRMsr sauvage et des différents mutants

III.1. Production

Les cellules DH5 α d'*E. coli* sont transformées à l'aide du plasmide pSK-fRMsr afin de produire la forme sauvage de la fRMsr, tandis que les cellules HB101 d'*E. coli* sont transformées à l'aide des plasmides pSK-fRMsr mutés afin de produire les différents mutants de la fRMsr.

Des cultures en milieu LB additionné de 200 mg/l d'ampicilline sontensemencées au 1/100^{ème} à partir d'une pré-culture. Les cultures sont incubées pendant 24h à 37°C sous agitation, puis collectées par centrifugation à 4000 g pendant 20 min.

III.2. Purification

Les culots cellulaires sont repris dans 15 ml de tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, additionné de 50 mM de DTT.

III.2.1. Sonication

Une sonication a été réalisée en 2 fois 2 min, par période de 1 seconde d'ultrasons et 1 seconde d'arrêt pour lyser un maximum de cellules. Les débris cellulaires sont éliminés par une centrifugation à 17000 g pendant 50 min. Le surnageant de sonication est amené à 50% de saturation en sulfate d'ammonium et précipité pendant 2h sous légère agitation à 4°C. A cette concentration en sulfate d'ammonium, l'enzyme est précipitée et elle est récupérée dans le culot après centrifugation pendant 45 minutes à 12000 g. Le culot est repris dans 20 ml de tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, additionné de 50 mM de DTT.

III.2.2. Filtration sur gel ACA 54

La solution protéique récupérée est alors déposée sur une colonne de filtration sur gel ACA 54 (acrylamide 5 %, agarose 4 %, BIOSEPRA), dont le domaine de fractionnement est situé entre 5 et 70 kDa. L'élution se fait avec le tampon Tris-HCl 50 mM, EDTA 2 mM, KCl 150 mM, pH 8. L'analyse des fractions obtenues est réalisée en déposant une quantité

aliquote de chacune de ces fractions sur gel SDS-PAGE 12,5 %. Les fractions contenant la protéine sont ensuite rassemblées. Une grande partie des contaminants protéiques est éliminée.

III.2.3. Chromatographie échangeuse d'anions

La solution protéique issue de l'étape précédente est filtrée puis injectée sur une colonne de chromatographie échangeuse d'anions Q-Sépharose (agarose 6 % sur laquelle sont greffés des groupements triméthylammonium) couplée à un système FPLC (Amersham Biosciences). Cette colonne est préalablement équilibrée avec le tampon Tris-HCl 50 mM, EDTA 2mM, pH 8, et l'élution est réalisée par un gradient croissant de KCl (entre 150 et 400 mM). L'analyse des fractions obtenues est réalisée par la lecture du chromatogramme (DO à 260 et 280 nm), et les fractions contenant la protéine sont rassemblées. Cette étape permet d'éliminer en plus des protéines contaminantes, des acides nucléiques présents dans les fractions issues de la filtration sur gel.

III.2.4. Chromatographie par interactions hydrophobes

Le mélange protéique issu de la chromatographie échangeuse d'anions est doucement amené à une concentration de 1M en sulfate d'ammonium, puis injecté sur une colonne de chromatographie hydrophobe Phényl-Sépharose couplée à un système FPLC (Amersham Biosciences). Cette colonne est préalablement équilibrée avec le tampon Tris-HCl 50 mM, EDTA 2 mM, sulfate d'ammonium 1 M, pH 8, et l'élution est réalisée par un gradient décroissant de sulfate d'ammonium (entre 1 et 0 M). L'analyse des fractions obtenues est réalisée par la lecture du chromatogramme (DO à 260 et 280 nm). Les fractions contenant la protéine sont rassemblées. Cette étape permet d'éliminer les derniers contaminants protéiques.

III.2.5. Concentration et précipitation au sulfate d'ammonium

Le mélange protéique est concentré par ultra-filtration sur une membrane Amicon YM 10 jusqu'à une concentration de l'ordre de 15 à 20 DO₂₈₀/ml. La protéine est ensuite précipitée en présence de 70% de saturation en sulfate d'ammonium et de 50 mM de DTT, et est conservée à -20°C.

III.2.6. Gels Bilans de purification des protéines

L'électrophorèse en conditions dénaturantes (SDS-PAGE 12,5 %) permet de suivre l'évolution du degré de pureté de l'enzyme au cours de chaque étape de purification. Les protéines sont colorées spécifiquement par le bleu de Coomassie R-250 à 0,2 % (p/v).

III.2.7. Détermination de la masse moléculaire par spectrométrie de masse

La spectrométrie de masse est une technique qui nécessite l'utilisation d'un échantillon sans aucune trace de sels. Pour cela 500 µg de protéines sont dessalés sur une colonne de filtration sur gel "HiPrep Desalting" couplée à un système FPLC. L'élution est réalisée par un tampon Bicarbonate d'ammonium (50 mM) fraîchement préparé. Les fractions contenant la protéine seront récupérées puis lyophilisées. Les masses moléculaires des protéines ont été déterminées au laboratoire de Spectrométrie de Masse Bio-organique de Strasbourg (UMR CNRS-ULP 7509 dirigée par le Dr A. Van Dorsselaer).

IV. Etudes enzymatiques de la fRMs_r sauvage et des différents mutants

Pour la préparation des enzymes sous forme réduites, les protéines précipitées en sulfate d'ammonium sont incubées en présence de 50 mM de DTT pendant 30 min à 4 °C. Deux dialyses successives sont réalisées en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8 sous un flux d'azote continu afin d'obtenir des enzymes sous forme réduites et sans trace de DTT. Pour la préparation des enzymes sous forme oxydée, les protéines sont incubées en présence de 100 mM de Met-O pendant 10 min à température ambiante en absence de réducteur. L'enzyme est ensuite dessalée sur une colonne de filtration sur gel Econo-Pac 10 DG (Bio-Rad). La concentration en monomères d'enzyme est calculée à partir de la DO mesurée à 280 nm en utilisant le coefficient d'extinction molaire théorique déterminé par la méthode de Gill & Von Hippel qui est de 26720 M⁻¹.cm⁻¹/monomère (Gill et Von Hippel, 1989).

Le substrat appelé Met-O est de la *D,L*-Met-*R,S*-O, c'est-à-dire un mélange des 4 diastéréoisomères.

IV.1. Dosage des groupements thiols

IV.1.1. Dosage des groupements thiols libres par le 5,5'-dithiobis (2-nitro) benzoate (DTNB)

Le dosage des groupements thiols libres des enzymes sous forme réduites ou oxydées a été réalisé en utilisant le DTNB (0,3 mM final). Le suivi de l'apparition du thionitrobenzoate (TNB^-) est mesuré à 412 nm en conditions natives (Tris-HCl 50 mM, EDTA 2 mM, pH 8) ou dénaturantes (SDS à 10%). Les mesures sont effectuées à deux concentrations en enzyme (7,35 et 14,7 μM) qui correspondent respectivement à un $\Delta\text{DO}_{412\text{nm}}/\text{SH}$ libre de 0,1 et 0,2. La quantité de TNB^- formé est calculée en utilisant un coefficient d'extinction molaire de $13600\text{ cm}^{-1}\cdot\text{M}^{-1}$ à 412 nm.

IV.1.2. Titration des résidus Cys par le 2,2'-dipyridyl disulfure (2PDS)

Le 2PDS est une sonde chimique spécifique des Cys, qui permet d'appréhender la réactivité et l'accessibilité des Cys d'une protéine. En effet, l'attaque du thiolate d'un résidu Cys sur une molécule de 2PDS, conduit à la formation d'un adduit pyridine disulfure et à la libération concomitante d'une molécule de pyridine 2 thiolate qui absorbe à 343 nm ($\epsilon_{343\text{nm}} = 8080\text{ M}^{-1}\cdot\text{cm}^{-1}$). Les mesures sont réalisées à l'aide d'un appareil de cinétique rapide Applied photophysics SX18MV-R (« Stopped-flow »), où une première seringue contient la fRMs_r réduite à 12,5 μM (qui correspond à un saut 0,1 $\text{DO}_{343\text{nm}}/\text{SH}$ libre), la deuxième seringue contient le 2PDS à 200 μM , dans un polytampon composé d'acide acétique 30 mM, d'imidazole 30 mM, et de Tris-HCl 120 mM de force ionique constante (0,15 M) et de pH variable (de 5 à 9,5). Pour chaque pH, les données expérimentales ont été ajustées à l'équation d'une monoexponentielle croissante ($y = a \cdot e^{-k_{\text{obs}} \cdot t}$, où a représente l'amplitude et k_{obs} la constante de vitesse) afin de déterminer la valeur de la constante k_{obs} . Les valeurs de k_{obs} ont ensuite été ajustées par régression non linéaire à l'aide du logiciel SigmaPlot 2000 (Jandel Scientific) à l'équation $k_2 = k_{2\text{max}} / (1 + 10^{(\text{pK}_{\text{app}} - \text{pH})})$, afin de déterminer les valeurs de pK_{app} et de $k_{2\text{max}}$.

IV.2. Détermination des paramètres cinétiques K_M et k_{cat} en présence du système réducteur Trx /Trx réductase/NADPH à l'état stationnaire

La disparition du NADPH au cours du temps est suivie par la diminution de l'absorbance à 340 nm au cours du temps. L'activité de l'enzyme est exprimée en μmoles

de NADPH oxydé par seconde et par μ moles d'enzyme. Le mélange réactionnel contient la fRMs_r, le substrat et la Trx en concentrations variables, 1,25 μ M de Trx réductase d'*E. coli*, 0,3 mM de NADPH, dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8.

La thiorédoxine et la thiorédoxine réductase d'*E. coli* ont été produites et purifiées au laboratoire selon les protocoles publiés (Mulrooney, 1997 ; Mossner *et al.*, 1998)

Les activités sont mesurées en faisant varier soit la concentration en substrat en présence de concentration saturante en Trx, soit la concentration en Trx en présence de concentration saturante en substrat. Le substrat utilisé est de la Met-O, de la *L*-Met-*R,S*-O ou de la *D*-Met-*R,S*-O. Les points expérimentaux sont ajustés par régression non linéaire à l'aide du logiciel SigmaPlot 2000 (Jandel Scientific) à l'équation d'une droite, de Michaelis-Menten ($v = (k_{cat} \cdot [E_0] \cdot [S]) / (K_M + [S])$) ou à l'équation du modèle d'inhibition compétitive par excès de substrat ($v = (k_{cat} \cdot [E_0] \cdot [S]) / (K_M \cdot ([S] + [S]^2/K_I))$).

IV.3. Etudes en absence de réducteur

IV.3.1. Détermination de la stœchiométrie de formation de Met

Un mélange réactionnel contenant 50 mM de Met-O et 100 ou 200 μ M d'enzyme dans du tampon Tris-HCl 50 mM, EDTA 2mM, pH 8 est incubé à 25°C durant 10 min. La réaction est arrêtée par ajout de 1% final de TFA. La quantification de la Met est effectuée sur une colonne en phase inverse Atlantis dC18 (4.6x250 mm, Waters) couplée au système ÄKTA explorer (Amersham Biosciences). La colonne est équilibrée en présence de TFA 0,1% et d'acétonitrile 7,5%. Une boucle d'injection de 100 μ l est utilisée et la Met est éluée en condition isocratique, avec un temps d'éluion supérieur à celui de la Met-O. L'intégration du pic est réalisée à 215 nm. Avant chaque série de mesure, un étalonnage de la colonne est effectué avec trois solutions de Met de concentration connue. La stœchiométrie est alors calculée en rapportant la quantité de Met formée à la quantité de monomères d'enzyme présents.

IV.3.2. Détermination de la vitesse d'apparition de la Met et de disparition des thiols libres

Un appareil de cinétique rapide à 3 seringues (RQF3, KinTek Corporation, Austin), appelé « Quenched-Flow », permet de mélanger rapidement deux solutions (enzyme et substrat) initialement contenues dans deux seringues, et après un temps d'incubation compris entre 5 et 100 ms, une 3^{ème} solution (acide) est ajoutée au premier mélange pour

arrêter la réaction. Le mélange final a un volume de 500 μl , et contient en volume 1/3 de chaque solution. La première seringue contient de l'enzyme sous forme réduite (600 μM) dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, la deuxième le substrat Met-O (300 mM) et la troisième le TFA (3%).

La quantification de la Met formée et la titration des thiols libres de l'enzyme sont réalisés sur les différents mélanges obtenus. La quantification de la Met est réalisée par chromatographie en phase inverse (voir paragraphe IV.3.1 page 109) après précipitation de l'enzyme en présence de TFA 2%. La titration des thiols libres est réalisée à l'aide de DTNB (voir paragraphe IV.1.1 page 107) après pré-incubation de l'échantillon en présence d'urée 8M pendant 15 min.

La vitesse de formation de Met a été déterminée par l'ajustement des données expérimentales à l'équation d'une mono-exponentielle croissante, $y = a (1 - e^{-k_{\text{obs}}t})$ dans laquelle « a » représente le nombre de mole de Met formée par mole de monomère de fRMsR et k_{obs} la constante de vitesse. La vitesse de formation de pont disulfure a été déterminée par l'ajustement des données à l'équation d'une monoexponentielle décroissante, $y = y_0 + a.e^{-k_{\text{obs}}t}$ dans laquelle « y_0 » représente le nombre de moles de thiols libres par mole de monomère d'enzyme, « a » le nombre de moles de thiols oxydés par mole de monomère d'enzyme, et k_{obs} la constante de vitesse.

IV.3.3. Techniques de fluorescence

IV.3.3.1. Spectres de fluorescence des enzymes

Les spectres d'excitation et d'émission de fluorescence des fRMsR sous forme réduite et oxydée, ont été réalisés à l'aide d'un fluorimètre (SAFAS) à une concentration de 10 μM en enzyme, dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, dans des cuves en quartz et sous légère agitation. Pour les spectres d'excitation, la longueur d'onde d'excitation varie de 200 à 300 nm par incrément de 1 nm, et la longueur d'onde d'émission est de 340 nm. Pour les spectres d'émission, la longueur d'onde d'excitation est de 284 nm, et la longueur d'onde d'émission varie de 310 à 500nm par incrément de 1 nm.

IV.3.3.2. Etudes cinétiques

L'étude est basée sur la différence d'amplitude d'émission de fluorescence des fRMsR entre la forme réduite et oxydée. Les mesures sont réalisées à l'aide d'un appareil de cinétique rapide Applied photophysics SX18MV-R (« Stopped-flow »), qui permet de

mélanger rapidement deux solutions contenues dans deux seringues. Un filtre permet de ne mesurer que l'émission de fluorescence pour les longueurs d'onde supérieures à 320 nm. La première seringue contient la fRMs_r sauvage ou un mutant sous forme réduite (10 µM) dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, et la deuxième seringue contient le substrat Met-O à différentes concentrations dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8. La longueur d'onde d'excitation est de 284 nm.

Les traces expérimentales obtenues sont ajustées au modèle cinétique $A \rightarrow B \rightarrow C$ à l'aide du logiciel Pro Kineticist (Applied Photophysics Ltd) afin de déterminer les constantes de vitesse correspondantes. Les variations des constantes de vitesse en fonction de la concentration en Met-O sont ensuite ajustés par régression non linéaire à l'aide du logiciel SigmaPlot 2000 (Jandel Scientific) à l'équation d'une droite ($k_{\text{obs}} = k_2 \cdot [S]$), de Michaelis-Menten ($k_{\text{obs}} = (k_{\text{obsmax}} \cdot [S]) / (K_S + [S])$) ou à l'équation du modèle d'inhibition compétitive par excès de substrat ($k_{\text{obs}} = (k_{\text{obsmax}} \cdot [S]) / (K_S \cdot ([S] + [S]^2/K_I))$).

IV.3.4. Caractérisation des formes oxydées par spectrométrie de masse

Cette étude a été réalisée au laboratoire de Spectrométrie de Masse Bio-organique à Strasbourg dirigé par le Dr A. Van Dorsselaer. Le mélange réactionnel contient 50 µM de fRMs_r et 50 mM de Met-O dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, et 0,1 % de TFA est ajouté après 10 s d'incubation pour arrêter la réaction. Le mélange réactionnel est séparé par chromatographie sur une colonne de phase inverse C8 en tampon acétonitrile 0,1% TFA et en tampon aqueux 0,1 % TFA avec un débit de 1ml/min, couplée à une analyse sur un spectrophotomètre de masse MicrOTOF-Q (Bruker Daltonics, Bremen). Le gradient appliqué était d'abord de 0 à 30 % d'acétonitrile 0,1% TFA en 1 volume colonne, suivi par un gradient de 30 à 90 % d'acétonitrile 0,1% TFA en 10 volumes colonne. L'élution des protéines est suivie par la mesure de la DO à 215 nm.

IV.3.5. Détermination de l'état oligomérique

La détermination de l'état oligomérique a été réalisée en utilisant l'approche de chromatographie d'exclusion stérique couplée à un appareil de SLS (TDA Viscotek, Houston). Les protéines sous formes oxydées sont obtenues après incubation de 200 µM de fRMs_r avec 100 mM de Met-O. Les protéines sous formes réduite ou oxydée sont séparées sur une colonne de type Shodex KW-802,5 en tampon Tris-HCl 50 mM, KCl 150 mM, pH 8. L'élution des protéines a été suivie par la mesure de la DO à 280 nm, le RI « Refractive Index » pour indice de réfraction et le RALS « Right Angle Light Scattering » pour

diffusion de la lumière à angle droit. Le programme OmniSEC est utilisé pour l'acquisition et l'analyse des données est réalisée selon l'équation de Rayleigh.

IV.4. Réduction des ponts disulfure par la Trx

IV.4.1. Approche chromatographique

La méthodologie utilisée consiste à séparer les formes réduites et oxydées à l'état disulfure des fRMs_rA et Trx par chromatographie analytique en phase inverse. Les mélanges réactionnels contenaient 50 ou 100 µM de fRMs_r oxydée et 200 µM de Trx réduite dans un volume final de 300 µl, en présence de tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8,0. Après 10 minutes d'incubation, la réaction a été arrêtée par acidification en ajoutant du TFA à une concentration finale de 0,1 %. Les différentes formes rédox de la fRMs_r et de la Trx ont ensuite été séparées par chromatographie analytique en phase inverse dans un tampon aqueux contenant 0,1 % de TFA avec un débit de 1 ml/min sur une colonne BrownLee Aquapore RP300 C₈ (Applied Biosystems) couplée à un système AKTA explorer (Amersham Bioscience) en utilisant un gradient de 0 à 80 % d'acétonitrile contenant 0,1 % de TFA. L'élution des protéines est suivie par la mesure de la DO à 280.

IV.4.2. Cinétique rapide

L'étude a été réalisée par deux approches :

La première approche est basée sur la différence d'amplitude d'émission de fluorescence de la Trx entre sa forme réduite et oxydée. Les mesures sont réalisées à l'aide d'un appareil de cinétique rapide Applied photophysics SX18MV-R (« Stopped-flow »). La première seringue contient la fRMs_r oxydée (10 µM) dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8. La deuxième seringue contient la Trx réduite à différentes concentrations en tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8. La longueur d'onde d'excitation est fixée à 310 nm pour n'observer que la différence d'intensité de fluorescence de la Trx. Un filtre permet de ne mesurer que l'émission de fluorescence pour les longueurs d'onde supérieures à 320 nm. Les tracés expérimentaux sont ajustés à l'équation d'une monoexponentielle décroissante ($y = a \cdot e^{-k_{obs} \cdot t}$ dans laquelle a représente l'amplitude et k_{obs} la constante de vitesse) afin de déterminer la valeur de la constante de vitesse k_{obs} .

La deuxième approche est basée sur l'utilisation du système couplé Trx réductase/NADPH, et en suivant la disparition du NADPH à 340 nm. La première seringue contient la fRMs_r oxydée (10 µM) dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8, et la deuxième seringue contient la Trx réduite (de 10 à 100 µM), la Trx réductase (2,5 µM) et

Matériels et Méthodes

le NADPH (0,3 mM) dans du tampon Tris-HCl 50 mM, EDTA 2 mM, pH 8. Les tracés expérimentaux sont ajustés à l'équation d'une monoexponentielle croissante ($y = a(1 - e^{-k_{\text{obs}} \cdot t})$) dans laquelle a représente l'amplitude et k_{obs} la constante de vitesse) afin de déterminer la valeur de la constante de vitesse k_{obs} .

ANNEXES

Annexe 1:

“Structural and Biochemical Characterization of free Methionine-*R*-sulfoxide Reductase from *Neisseria meningitidis*” Gruez, A.*, **Libiad, M.***, Boschi-Muller, S. et Branlant, G. (2010) *J. Biol. Chem.* **285**, 25033-43. * Co-premiers auteurs

Et le « Supplemental Data » : Alignement de 358 séquences de fRMsrS putatives.

Annexe 2:

“Substrate Binding in free methionine-*R*-sulfoxide reductase” Gruez, A., **Libiad, M.**, Boschi-Muller, S. et Branlant, G. (2010) *J. Biol. Chem.* **285**, Letter comment on, *J. Biol. Chem.* **285**, 25044-52.

Structural and Biochemical Characterization of Free Methionine-*R*-sulfoxide Reductase from *Neisseria meningitidis**[§]

Received for publication, April 16, 2010, and in revised form, May 6, 2010. Published, JBC Papers in Press, May 19, 2010, DOI 10.1074/jbc.M110.134528

Arnaud Gruez¹, Marouane Libiad^{1,2}, Sandrine Boschi-Muller, and Guy Branlant³

From AREMS, UMR CNRS-UHP 7214, Nancy Université, Faculté des Sciences et Technologies, Bld. des Aiguillettes, BP 70239, 54506 Vandoeuvre-les-Nancy, France

A new family of methionine-sulfoxide reductase (Msr) was recently described. The enzyme, named fRMsr, selectively reduces the *R* isomer at the sulfoxide function of free methionine sulfoxide (Met-*R*-O). The fRMsrs belong to the GAF fold family. They represent the first GAF domain to show enzymatic activity. Two other Msr families, MsrA and MsrB, were already known. MsrA and MsrB reduce free Met-*S*-O and Met-*R*-O, respectively, but exhibit higher catalytic efficiency toward Met-O within a peptide or a protein context. The fold of the three families differs. In the present work, the crystal structure of the fRMsr from *Neisseria meningitidis* has been determined in complex with *S*-Met-*R*-O. Based on biochemical and kinetic data as well as genomic analyses, Cys¹¹⁸ is demonstrated to be the catalytic Cys on which a sulfenic acid is formed. All of the structural factors involved in the stereoselectivity of the *L*-Met-*R*-O binding were identified and account for why Met-*S*-O, DMSO, and a Met-O within a peptide are not substrates. Taking into account the structural, enzymatic, and biochemical information, a scenario of the catalysis for the reductase step is proposed. Based on the thiol content before and after Met-O reduction and the stoichiometry of Met formed per subunit of wild type and Cys-to-Ala mutants, a scenario of the recycling process of the *N. meningitidis* fRMsr is proposed. All of the biochemical, enzymatic, and structural properties of the *N. meningitidis* fRMsr are compared with those of MsrA and MsrB and are discussed in terms of the evolution of function of the GAF domain.

Methionine (Met)⁴ is one of the two amino acids in proteins that are the most susceptible to oxidation by reactive oxygen

* This work was supported by the CNRS, the University of Nancy I, and the IFR 111 Bioingénierie.

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Fig. S1.

The atomic coordinates and structure factors (code 3MMH) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ Both authors contributed equally to this work.

² Supported by the French Ministry of Research.

³ To whom correspondence should be addressed: AREMS, UMR CNRS-UHP 7214, Nancy Universités, Faculté des Sciences et Technologies, Bld. des Aiguillettes, BP 70239, 54506 Vandoeuvre-les-Nancy, France. Tel.: 33-3-83-68-43-04; Fax: 33-3-83-68-43-07; E-mail: Guy.Branlant@maem.uhp-nancy.fr.

⁴ The abbreviations used are: Met, methionine; Met-O, methionine sulfoxide (DL-Met-*R*,*S*-O); MES, 2-(*N*-morpholino)ethanesulfonic acid; 2-PDS, 2,2'-dithiodipyridine; DTNB, 5,5'-dithiobis(2-nitro)benzoate; fRMsr, free methionine-*R*-sulfoxide reductase; DTT, dithiothreitol; MPD, 2-methylpentane-2,4-diol; r.m.s., root mean square.

species, forming Met-O (1). Prior to 2007, two families of methionine-sulfoxide reductase (Msr) enzymes, called MsrA and -B were known to reduce Met-O back into Met (2, 3). The MsrA family reduces the *S* isomer at the sulfoxide function, whereas MsrB is specific for the *R* isomer. Both Msrs, which reveal distinct unrelated folds, were shown to reduce more efficiently a Met-O within a polypeptide chain. Therefore, Msrs are repair enzymes that play important roles in the protection of cells against oxidative stress (3–5).

The catalytic mechanism of MsrA and MsrB is now well documented. For most MsrAs and MsrBs (*i.e.* those possessing a recycling Cys), the mechanism comprises three steps: a reductase step that leads to formation of a sulfenic acid intermediate on a catalytic Cys (6), a second step in which a disulfide bond is formed between the catalytic Cys and a recycling Cys, and finally, a step in which the intradisulfide bond is reduced by thioredoxin (Trx) or a Trx-like protein (7). A catalytic residue (*i.e.* Glu⁹⁴ in MsrA and His¹⁰³ in MsrB) was characterized, one of whose major roles is to protonate the oxygen of the sulfoxide function (8, 9). In all of the MsrAs and MsrBs studied to date, the reduction of the disulfide bond is rate-limiting, whereas the formation of the sulfenic acid intermediate is rate-determining in the process leading to intradisulfide bond formation. In other words, in the absence of Trx, only the intradisulfide intermediate accumulates.

Recently, a third family of Msr, named fRMsr, was discovered (10). The fRMsrs exhibit a GAF-type fold. GAF domains are one of the largest and most widespread domains found in all kingdoms of life. They are dimeric and are generally arranged in tandem in modular proteins to provide a large variety of regulation functions. However, most of the functions of GAF domains remain to be studied in detail. The fRMsr is the first case of a GAF domain that bears a catalytic activity. The fRMsrs are present in eubacteria and unicellular eukaryotes (11). The family displays a methionine-sulfoxide reductase activity, reducing selectively free Met-O with an *R* configuration at the sulfoxide (10). Such a function led Lowther and co-workers (10) to propose that Met-*R*-O can represent a signaling molecule in response to oxidative stress.

Only the crystal structures of fRMsr from *Escherichia coli* and *Saccharomyces cerevisiae* have been solved to date, without substrate (12, 13). Both structures are described as being composed of six β -strands, four α -helices, and two prominent loops, loop 1 and loop 2, located on the surface of the protein

Structure and Catalytic Mechanism of *N. meningitidis* fRMsr

between $\beta 2$ and $\beta 3$ and between $\beta 4$ and $\beta 5$, respectively. Formation of a disulfide bond between Cys⁸⁴ of loop 1 and Cys¹¹⁸ of loop 2 is assumed to close off the cavity in which a molecule of MES, which derives from the crystallization buffer, is bound in the *E. coli* enzyme (Protein Data Bank entry 1vhm). The sulfonic acid moiety is closed to another Cys⁹⁴, which is located at the N terminus of an α -helix. The rest of the cavity is lined by several invariant residues. From the inspection of the crystal structure of *E. coli* fRMsr in complex with a MES buffer molecule and by carrying out computational docking of Met-*R*-O into the cavity of the *S. cerevisiae* enzyme, a reductase mechanism was proposed in which the Cys⁹⁴ located at one of the sides of the cavity is assumed to be the catalytic Cys (10). The sulfenic acid postulated to be formed on Cys⁹⁴ is then attacked by Cys¹¹⁸, forming a disulfide bond. A disulfide bond exchange then occurs with Cys⁸⁴, which finally leads to the Cys¹¹⁸-Cys⁸⁴ disulfide intermediate.

In the present study, we report the crystal structure of the fRMsr from *Neisseria meningitidis* in complex with L-Met-*R*-O, solved at 1.25 Å resolution. By combining genomic analyses, structural information, and biochemical and kinetic data from the wild type and mutants in which Cys was substituted by Ala, Cys¹¹⁸ is shown to be the catalytic residue on which the sulfenic acid is formed.⁵ All of the molecular and structural factors involved in the stereoselective binding of the L-Met-*R*-O stereoisomer are identified from the crystal structure and explain why neither a Met-*S*-O, a Met-*R*-O in a peptide or a protein context, nor a DMSO molecule are substrates. Scenarios for both substrate binding and catalysis are proposed for the reductase step.

A disulfide bond between Cys⁸⁴ and Cys¹¹⁸, present in all of the fRMsr crystal structures described so far, including the *N. meningitidis* fRMsr, stabilized the loop 2 as a flap that is an integral part of the active site. Formation of this disulfide bond seems to be a prerequisite to obtaining crystals. However, it is the free reduced form that is catalytically relevant to the reductase step. Based on the content of 1) the Cys in wild type and Cys-to-Ala mutants before and after Met-O reduction and 2) the Cys titrated under native conditions, a scenario for the Trx-recycling process is proposed in which a disulfide Cys¹¹⁸-Cys⁸⁴ bond is formed first. All of the results are compared with those obtained for MsrA and MsrB and discussed in terms of the evolution of function of the GAF domain.

EXPERIMENTAL PROCEDURES

Plasmid Constructions, Site-directed Mutageneses, Production, and Purification of Wild Type and Cys-to-Ala Mutants of fRMsr from *N. meningitidis*—Expression plasmid pSKfRMsr for fRMsr was obtained by cloning the corresponding open reading frame by PCR from *N. meningitidis* 8013 genomic DNA, into the plasmid pSKMsr (6) between the NdeI and SacI sites. Site-directed mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Stratagene).

⁵ The numbering of fRMsr amino acid residues is based on the numbering of the *E. coli* fRMsr sequence according to sequence alignment presented in Fig. 1.

The *E. coli* strain HB101 used for production of fRMsr was transformed with a pSKfRMsr plasmid. All wild type and mutants were produced in soluble form. For enzyme purification, cells were harvested by centrifugation, resuspended in buffer A (50 mM Tris-HCl, 2 mM EDTA, pH 8) containing 50 mM dithiothreitol (DTT) and sonicated. The enzyme was then precipitated at 50% (NH₄)₂SO₄ saturation. The contaminating proteins were removed by applying the enzymatic solutions, after extensive dialysis against buffer A, onto a phenyl-Sepharose column (Amersham Biosciences) equilibrated with buffer A containing 1 M (NH₄)₂SO₄, using a fast protein liquid chromatography system (Amersham Biosciences). Wild type and all mutants eluted at 150 mM (NH₄)₂SO₄ from a linear gradient from 1 to 0 M (NH₄)₂SO₄ in buffer A. Purified fractions were then pooled and applied to a Q-Sepharose column equilibrated with buffer A, followed by a linear gradient of KCl (0–0.4 M). All fRMsr eluted at 250 mM KCl and were pure at this stage. The yield was ~80 mg of fRMsr/liter of culture. An additional step was added to prepare samples for crystallization. Before (NH₄)₂SO₄ precipitation, a treatment using 0.05% polyethyleneimine was carried out to remove most of the nucleic acids. The purity of the wild type and all mutants was checked by electrophoresis on a 12.5% SDS-polyacrylamide gel followed by Coomassie Brilliant Blue R-250 staining and by electrospray ionization mass spectrometry analyses. Purified enzymes were stored at –20 °C in the presence of 50 mM DTT and 70% (NH₄)₂SO₄. Under these conditions, the wild type and all mutants were stable for several weeks. Their molar concentrations were determined spectrophotometrically, using extinction coefficients at 280 nm of 26,720 M^{–1} cm^{–1} for wild type and the mutants.

Mass Spectrometry Analyses—Samples were desalted by applying the enzymatic solutions onto an exclusion size column (HIPREP 26/10 Desalting, Amersham Biosciences) equilibrated in 50 mM ammonium bicarbonate buffer, pH 7.6. Electrospray ionization mass spectrometry measurements were performed on an electrospray ionization time-of-flight mass spectrometer (LCT, Micromass, Altrincham, UK) fitted with a standard Z-spray source. Measurements were performed in nondenaturing conditions by diluting enzymes to 10 μM in 100 mM ammonium acetate buffer, pH 6.8.

Quantification of the Free Cys Content of Wild Type and Mutants with 5,5'-Dithiobis(2-nitro)benzoate (DTNB) and Determination of the pK_{app} of Cys¹¹⁸ and Cys⁸⁴ in the Wild Type with 2,2'-Dipyridyl Disulfide (2-PDS)—The Cys content of the wild type and mutants was determined using DTNB under native and denaturing conditions (final concentration of 9% SDS in buffer A) in the absence and presence of incubation with DL-Met-*R,S*-O at a concentration of 100 μM, as described previously (6).

Due to the high chemical reactivity of Cys⁸⁴ and Cys¹¹⁸ and of the thionitrobenzoate (TNB[–]; pK_{app} of 7), fast kinetic measurements for determining the Cys pK_{app} were carried out on an Applied Photophysics SX18MV-R stopped flow apparatus with 2-PDS instead of DTNB, because the pK_{app} of pyridine-2-thione is 4.3. The experimental conditions were as previously described (14), with wild type and 2-PDS concentrations after mixing of 12.5 μM/subunit and 200 μM, respectively. The pseu-

do-first-order rate constant was determined at each pH by fitting the absorbance (A) at 343 nm *versus* time (t) to monoexponential Equation 1, where a is the burst amplitude and c is the end point. The second-order rate constants k_2 were calculated by dividing k_{obs} by 2-PDS concentration and then fitted to Equation 2, in which $k_{2\text{max}}$ represents the second rate constant for the thiolate form.

$$A = a(1 - e^{-k_{\text{obs}}t}) + c \quad (\text{Eq. 1})$$

$$k_2 = \frac{k_{2\text{max}}}{1 + 10^{(\text{p}K_a - \text{pH})}} \quad (\text{Eq. 2})$$

Quantification of Met Formed in the Absence of Trx—The reaction mixture containing DL-Met-*R,S-O* (100 mM) and the wild type or mutants (50–100 μM) was incubated at 25 °C for 10 min in buffer A, and then 250 μl of each assay was injected onto a Sephasil C18 column for Met quantification, as described earlier (15).

Steady-state Kinetics of Wild Type and Mutants with the Trx-regenerating System and with DL-Met-*R,S-O* as a Substrate—Steady-state kinetic parameters were determined with DL-Met-*R,S-O* as a substrate and *E. coli* Trx1 as a reductant in the presence of a Trx-regenerating system (1.25 μM *E. coli* Trx reductase and 0.3 mM NADPH), as previously described (15). Initial rate measurements were carried out at 25 °C in buffer A on a SAFAS UVmc² spectrophotometer by following the decrease in the absorbance at 340 nm due to the oxidation of NADPH. Initial rate data were fitted to the Michaelis-Menten Equation 3 or to Equation 4, which fits the model of competitive inhibition by excess substrate. Using least squares analysis, k_{cat} , K_m , and K_i values were determined using the program Sigmaplot (Jandel Scientific Software).

$$k_{\text{obs}} = \frac{k_{\text{cat}}S}{K_m + S} \quad (\text{Eq. 3})$$

$$k_{\text{obs}} = \frac{k_{\text{cat}}S}{K_m(S + S^2/K_i)} \quad (\text{Eq. 4})$$

E. coli Trx1 and Trx reductase were prepared following published experimental procedures (15).

Steady-state Kinetics of the C84A/C136A fRMs with the Trx-regenerating System and with L-Met-*R,S-O* or D-Met-*R,S-O* as Substrate—The kinetic experimental conditions were the same as those used with DL-Met-*R,S-O*. The experimental protocol used to prepare D-Met-*R,S-O* was adapted from Sharov *et al.* (16). Briefly, 0.3 g of D-Met was dissolved in 2 ml of H₂O in a boiling water bath. After slight cooling, oxidation was carried out at 60 °C overnight with a 1.1-fold molar excess of H₂O₂. The D-Met-O was then precipitated by the addition of 200 ml of acetone and collected by centrifugation. The pellet was washed twice with acetone and dried under vacuum. Therefore, the D-Met-O was a mixture of *R* and *S* isomers at the sulfoxide function.

Crystallization and X-ray Diffraction Data Collection—The fRMs used for crystallization was concentrated to 20 mg/ml in buffer (20 mM Tris-HCl, pH 8). Initial crystallization conditions were determined in 96-well Greiner plates using a nanoliter

TABLE 1
Summary of data collection and refinement statistics

| Parameters | Values |
|----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Crystal | |
| Space group | P1 |
| Unit cell | $a = 39.32 \text{ \AA}$, $b = 39.48 \text{ \AA}$, $c = 52.25 \text{ \AA}$, $\alpha = 104.17^\circ$, $\beta = 105.16^\circ$, $\gamma = 100.50^\circ$ |
| Data collection | |
| Beamline | Proxima I |
| Wavelength (\AA) | 0.886 |
| Resolution (\AA) | 20-1.25 (1.30-1.25) ^a |
| Reflections observed | 624,099 (30,697) |
| Unique reflections | 73,065 (7048) |
| Redundancy | 8.54 (4.35) |
| Completeness (%) | 92.8 (80.3) |
| $I/\sigma I$ | 22.2 (3.97) |
| R_{merge} (%) ^b | 5.3 (47.0) |
| Refinement | |
| Resolution (\AA) | 20-1.25 |
| No. of atoms | 5809 |
| $R_{\text{work}}/R_{\text{free}}$ (%) ^c | 11.78/15.68 |
| r.m.s. deviation bonds (\AA) | 0.022 |
| r.m.s. deviation angles (degrees) | 2.286 |
| Mean B values (\AA^2) | |
| Wilson B overall | 19.3 |
| Overall | 18.68 |
| Protein | 17.08 |
| Water molecules | 33.68 |

^a Highest resolution shell data are shown in parenthesis.

^b $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_{hkl_i} - \langle I_{hkl_i} \rangle|}{\sum_{hkl} \sum_i \langle I_{hkl_i} \rangle}$.

^c $R_{\text{cryst}} = \frac{\sum_{hkl} |F_o| - |F_c|}{\sum_{hkl} |F_o|}$.

drop Mosquito robot (TTP LabTech) and the commercial Structure Screens I and II (Molecular Dimension Ltd.) and Wizard Screens 1 and 2 (Emerald Biostructure). Protein solution droplets of 200 nl were mixed with precipitant droplets of 100, 200, and 300 nl and then equilibrated against 80 μl of well buffer. Small crystal plates appeared after 2 days of sitting drop vapor diffusion using 30% MPD, 200 mM Mg(CH₃COO)₂, and 100 mM cacodylate, pH 6.5, at 20 °C. Diffraction quality crystals (0.3 × 0.4 × 0.5 mm) were obtained at 20 °C by the hanging drop vapor diffusion by mixing 4 μl of protein solution and 2 μl of 23% MPD, 120 mM Mg(CH₃COO)₂, and 100 mM cacodylate, pH 6.1–6.4, in 24-well plates filled with 1 ml of well buffer. Crystals were soaked for 30 s in a cryoprotectant solution containing 40% MPD, 120 mM Mg(CH₃COO)₂, 100 mM cacodylate, pH 6.5, and 10 mM DL-Met-*R,S-O*. The crystals were then flash-frozen under a gaseous nitrogen stream at 100 K. A data set was collected on the Proxima-1 beamline at SOLEIL (SOLEIL Synchrotron, Gif-sur-Yvette, France) using an ADSC Q315r detector. Diffraction data were indexed, integrated, and merged using XDS (17). The crystal diffracts at 1.25 \AA resolution and belongs to the P1 space group with two molecules per unit cell. Data collection statistics are summarized in Table 1.

Structure Determination and Refinement—Initial phases were obtained by the molecular replacement method using the crystal structure of the dimeric protein YebR from *E. coli* (Protein Data Bank entry 1vhm) (12) as a search model and the program MOLREP (18). The structure refinement was initiated by a rigid body refinement of both subunits between 20 and 3 \AA resolution. Refinements using REFMAC5 (19) were then carried out between 20 and 1.25 \AA resolution. After each cycle of refinement, the structure was manually built using the graphics program COOT (20). The quality of the $2F_o - F_c$ maps and the

TABLE 2

Free sulfhydryl content and stoichiometry of Met formed in the absence of Trx-regenerating system of wild type and Cys-to-Ala fRMsr

The values indicated represent the average of two independent measurements of at least two enzyme concentrations (S.D. range 10%). Cys content was determined spectrophotometrically using DTNB in the presence or absence of 9% SDS, in the absence or in the presence of 100 mM DL-Met-R,S-O without any regenerating system (see "Experimental Procedures").

| fRMsr | No. of Cys | No. of Cys measured under native conditions | No. of Cys measured under denaturated conditions | | | Stoichiometry of Met formed |
|-----------------|------------|---------------------------------------------|--------------------------------------------------|----------------------------|-------------------------------------|-----------------------------|
| | | | Without DL-Met-R,S-O | With DL-Met-R,S-O | Decrease in free thiol ^a | |
| | | <i>mol Cys/mol subunit</i> | | <i>mol Cys/mol subunit</i> | | <i>mol Met/mol subunit</i> |
| Wild type | 4 | 2.2 | 3.9 | 2.1 | 1.8 | 1.0 |
| C84A | 3 | 1.9 | 2.9 | 1.0 | 1.9 | 1.0 |
| C94A | 3 | 2.1 | 3.0 | 1.0 | 2.0 | 1.1 |
| C118A | 3 | 2.8 | 3.1 | 3.0 | 0.1 | 0.0 |
| C136A | 3 | 2.1 | 3.1 | 1.1 | 2.0 | 1.1 |
| C84A/C136A | 2 | 1.8 | 2.1 | 0.1 | 2.0 | 0.9 |
| C94A/C136A | 2 | 1.9 | 1.9 | 0.0 | 1.9 | 0.8 |
| C84A/C94A/C136A | 1 | 1 | 1.0 | 0.0 | 1.0 | 0.9 |

^a The difference in the number of free Cys thiols upon treatment with DL-Met-R,S-O versus no treatment.

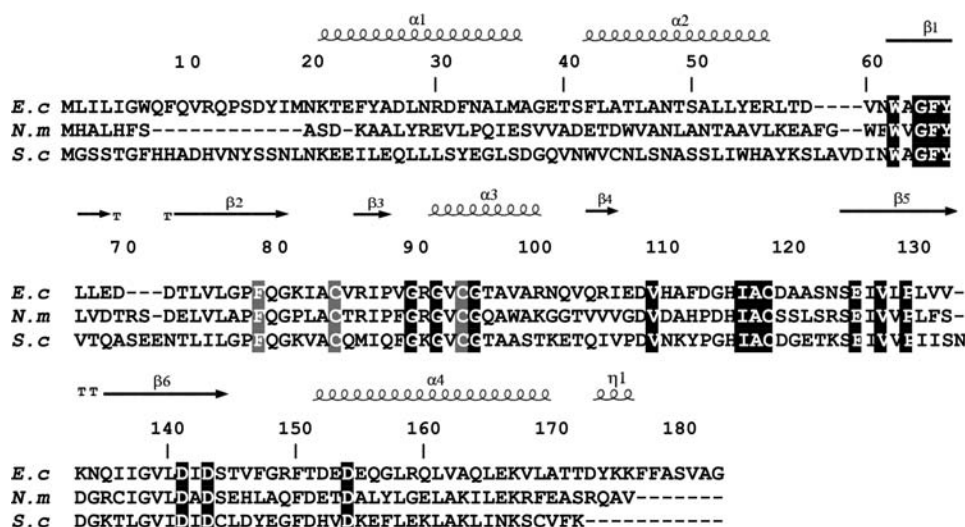


FIGURE 1. Alignment of sequences of fRMsr from *E. coli* (Ec), *N. meningitidis* (Nm), and *S. cerevisiae* (Sc). The sequences were aligned with Bioedit software. The numbering of the amino acid residues indicated is based on the numbering of the *E. coli* fRMsr sequence. The conserved residues in the 358 putative fRMsr sequences are indicated in white on black boxes, whereas residues Cys⁸⁴, Cys⁹⁴, and Phe⁷⁸ are in white on gray boxes. Cys⁸⁴ and Cys⁹⁴ are present in 94.5 and 98.6%, respectively, of the sequences of all putative fRMsr. In terms of signatures, Cys¹¹⁸/Cys⁸⁴/Cys⁹⁴, Cys¹¹⁸/Cys⁹⁴, and Cys¹¹⁸/Cys⁸⁴ are present in 93, 5.6, and 1.4% of the putative fRMsr sequences. Phe⁷⁸ is present in 92.7% of the sequences but is replaced by Tyr in the other sequences. The secondary structures depicted above the alignment are those from the fRMsr of *E. coli* (12).

$F_o - F_c$ residual maps during the refinement were excellent and allowed the construction of a protein model between the ninth residue and the C-terminal residue, as well as localization of two substrate molecules (*i.e.* L-Met-R-O). Water molecules as well as MPD and acetate molecules were added to the model using the available options in COOT. Hydrogen atoms were then automatically generated, and temperature factors were refined anisotropically using REFMAC. During the last steps of refinement, weak $2F_o - F_c$ and $F_o - F_c$ residual maps appeared at the N-terminal parts and allowed the first six amino acids of both subunits to be constructed with low occupancy. On the basis of the $F_o - F_c$ residual maps and their temperature factors, 11 atoms were refined as magnesium ions. Atomic ligands located within the active site were modeled and refined as water molecules but not as magnesium ions. The final model includes residues 1–167 of chains A and B, two L-Met-R-O molecules, four MPD molecules, two acetates, 11 magnesium ions, and 318 water molecules. The quality of the refined structure, which was checked using MolProbity, indicates that 98.1% of the residues lie in the favored regions of the Ramachandran plot, and

99.5% of all of the residues are in the allowed region. 0.61% are present in the disallowed region and concern the Ala⁷⁶ residue of the A and B subunits that are well defined in the electron $2F_o - F_c$ map. Within the dimer, both Ala⁷⁶ residues that precede in sequence Pro⁷⁷ have a constraint geometry that leads to ϕ, ψ angles of $(-123.4^\circ, -160.2^\circ)$ and $(-125.3^\circ, -162.6^\circ)$, respectively.

RESULTS

Biochemical and Kinetic Properties of Wild Type and fRMsr Mutants—The fRMsr were over-expressed in *E. coli*. About 10% of the soluble proteins were fRMsr. The molecular mass determined by mass spectrometry under non-denaturing conditions ($36,333 \pm 2$ Da) is in accord with the dimeric state (theoretical value: 36,330 Da) of a GAF domain. Size exclusion chromatography coupled to static light scattering also showed that the *N. meningitidis* fRMsr is in a dimeric form (data not shown). The use of DTNB reagent revealed four Cys residues under denaturing conditions (Table 2). This result was expected from the fRMsr sequence of *N. meningitidis*, which contains four Cys at positions 84, 94, 118, and 136 (Fig. 1).

To evaluate the role of these Cys in the catalytic mechanism, they were substituted, individually, in tandem, and three at once (Cys⁸⁴, Cys⁹⁴, and Cys¹³⁶), by Ala. The C118A mutant was inactive, whereas all of the other mutants were active under steady-state conditions, including the triple mutant C84A/C94A/C136A. In all of the mutants except for the C118A mutant, 1 mol of Met/subunit was formed in the absence of Trx.

To investigate how the recycling process occurs, the free thiol content of the wild type and the different mutants was determined under denaturing conditions by DTNB, before and after incubation with Met-O and in the absence of reductant. As shown in Table 2, 2 Cys residues/subunit were no longer titrated with DTNB in all of the fRMsr except for the

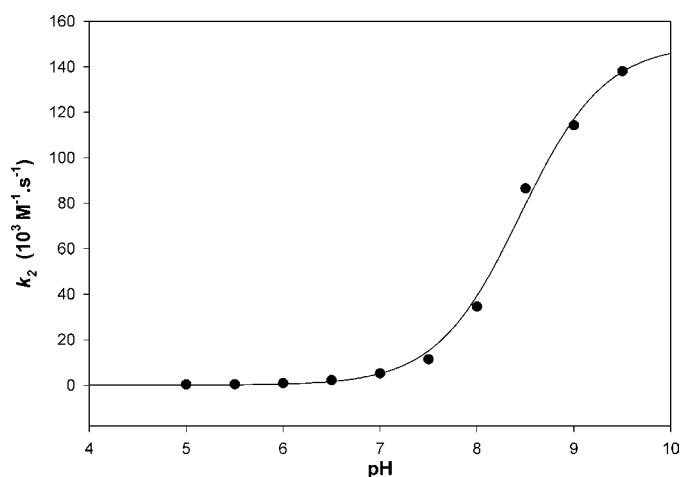


FIGURE 2. pH dependence of the second-order rate constant k_2 for the reaction of the thiol group of wild type fRMs with 2-PDS. Reaction kinetics were performed at 25 °C in 120 mM Tris-HCl, 30 mM acetic acid, 30 mM imidazole buffer at a constant ionic strength of 0.15 M over a pH range of 5–9.5. The concentrations of wild type fRMs and 2-PDS were 12.5 and 200 μM , respectively. The values represent the average of three independent measurements (S.D. range <2%). Values of second-order rate constants k_2 (●) were fitted to Equation 2 (solid line), leading to $k_{2\text{max}}$ and $\text{p}K_{\text{app}}$ values of $(1.50 \pm 0.04) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ and 8.5 ± 0.1 , respectively.

C118A mutant, in which no difference in Cys titration was observed, and the C84A/C94/C136 mutant, in which only one Cys was lost.

Under native conditions, two Cys reacted with DTNB in the wild type and C94A/C136A mutant (Table 2). This observation indicates that either these two Cys reacted individually and formed an interdisulfide bond with the TNB[−] moiety or one of the two Cys reacted more rapidly and the second formed a disulfide bond with the first via an intramolecular attack. To differentiate between these two possibilities, the two oxidized fRMs were isolated, and DTT was added. No absorbance at 412 nm corresponding to release of TNB[−] was observed (data not shown). In mutants in which either Cys⁸⁴ and Cys¹³⁶ or Cys¹¹⁸ and Cys¹³⁶ were substituted by Ala, two Cys residues were also titrated under native conditions, therefore corresponding to either Cys¹¹⁸ and Cys⁹⁴ or Cys⁹⁴ and Cys⁸⁴, respectively. In contrast to the wild type, an absorbance at 412 nm corresponding to release of 2 mol of TNB[−] was observed when DTT was added. Altogether, these results indicate that only Cys¹¹⁸ and Cys⁸⁴ are accessible in the wild type and form a disulfide bond under DTNB titration.

Reaction of 2-PDS with the wild type obeyed pseudo-first-order kinetics, with formation of 2 mol of pyridine-2-thione per subunit of fRMs, as shown from the absorbance increase at 343 nm. This result is similar to that found with DTNB. Moreover, no absorbance at 343 nm corresponding to release of pyridine-2-thione was observed when DTT was added to the oxidized fRMs (data not shown). Again, similar to DTNB, a disulfide bond is thus formed, probably between Cys¹¹⁸ and Cys⁸⁴. For all pH values used, stopped-flow traces fitted to monoexponential Equation 1, with amplitude corresponding to the release of 2 mol of pyridine-2-thione/subunit. The pH- k_2 profile fitted to monosigmoidal Equation 2 with a $\text{p}K_{\text{app}}$ value of 8.5 and $k_{2\text{max}}$ value of $1.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ (see Fig. 2). This finding supports a $\text{p}K_{\text{app}}$ of 8.5 for both Cys¹¹⁸ and Cys⁸⁴. The fact that

TABLE 3
Kinetic parameters of wild type and Cys-to-Ala fRMs under steady-state conditions

Reactions were carried out at 25 °C in 50 mM Tris-HCl, 2 mM EDTA buffer at pH 8.0, as described under “Experimental Procedures.” Enzyme concentrations were 0.1, 0.5, or 1 μM . The substrate used was DL-Met-*R,S*-O. For determining the $K_m(\text{Met-O})$, Trx concentration was 100 μM , and Met-O concentration varied from 0.1 to 500 mM. For $K_m(\text{Trx})$ determination, Trx concentration varied from 1 to 200 μM , and Met-O concentration was 100 mM, except for wild type and C136A fRMs (20 mM). Data presented were obtained by fitting the experimental data to the Michaelis-Menten equation, except data of wild-type and C136A fRMs which were obtained by using the equation of competitive inhibition by excess substrate. At least two independent determinations were performed to determine each constant. The K_m values must be divided by 4, considering the fact that the *S* isomer at the sulfoxide function is not a substrate, the *D*-Met-O is not efficient as a substrate, and the Met-O used is a mixture of the four stereoisomers. ND, activity not detectable.

| fRMs | k_{cat} s^{-1} | $K_m(\text{Met-O})$ mM | K_I mM | $K_m(\text{Trx})$ μM |
|-----------------|-------------------------------------|---------------------------|-------------|------------------------------------|
| Wild type | 14 ± 2 | 2 ± 1 | 34 ± 19 | 40 ± 7 |
| C118A | ND | | | |
| C136A | 8.5 ± 0.6 | 2.2 ± 0.5 | 56 ± 16 | 24 ± 3 |
| C84A | 1.0 ± 0.1 | 32 ± 9 | | 50 ± 4 |
| C94A | 18 ± 1 | 30 ± 6 | | 27 ± 4 |
| C84A/C136A | 2.6 ± 0.8 | 12 ± 4 | | 40 ± 6 |
| C94A/C136A | 13 ± 1 | 18 ± 1 | | 15 ± 2 |
| C84A/C94A/C136A | 1.0 ± 0.1 | 30 ± 8 | | 16 ± 4 |

the k_2 value is high also indicates that both Cys residues are highly accessible.

As shown in Table 3, all of the mutants (except the C118A mutant), including the triple mutant C84A/C94A/C136A, are active under steady-state conditions using Trx as the reductant and using DL-Met-*R,S*-O as substrate. The k_{cat} values vary from 8.5 to 18 s^{-1} for wild type and mutants in which Cys⁸⁴ is present and from 1 to 2.6 s^{-1} when Cys⁸⁴ is absent. The K_m values for Met-O vary from 2 for the wild type up to 32 mM for the C84A mutant. Because the *L*-Met-O is preferentially reduced (see below), the *S* isomer at the sulfoxide function is not active, and the Met-O used is a mixture of four stereoisomers, the K_m value must be divided by 4. The catalytic efficiency (k_{cat}/K_m) of the wild type relative to *L*-Met-*R*-O is better by a factor of at least 1000 when compared with MsrB from *N. meningitidis* (15), due to both k_{cat} and K_m effects. Inhibition by excess substrate is observed but only with the wild type and the C136A mutant. This observation remains to be explained but seems to be correlated with the presence of both Cys⁸⁴ and Cys⁹⁴. To determine whether the *D*-Met-O was a substrate, the kinetic parameters for both *L*-Met-O and *D*-Met-O were determined in the presence of the Trx-regenerating system not with the wild type but with the double mutant C94A/C136A because this mutant does not show substrate inhibition. The *D*-Met-O is not efficient as substrate. No saturating effect was observed up to 500 mM, and the k_2 constant, which represents the k_{cat}/K_m value, is $\sim 0.4 \text{ M}^{-1} \text{ s}^{-1}$. Therefore, the K_m value is at least 0.5 M, and thus the k_{cat} value is lower than 0.8 s^{-1} . In contrast, k_{cat} and K_m values for *L*-Met-O are in the range of those found with DL-Met-O (*i.e.* 18 s^{-1} and 10 mM, respectively), with a K_m value decreased by a factor of ~ 2 , as expected. The catalytic efficiency, which is better by a factor of $4 \cdot 10^4$ compared with *D*-Met-O, is the consequence of both k_{cat} and K_m effects.

Overall Structure—The crystal structure of fRMs of *N. meningitidis* in complex with *L*-Met-*R*-O was solved at 1.25 Å resolution. As expected, the fRMs structure shares the overall topology of the GAF domains (Fig. 3). Structural superimpo-

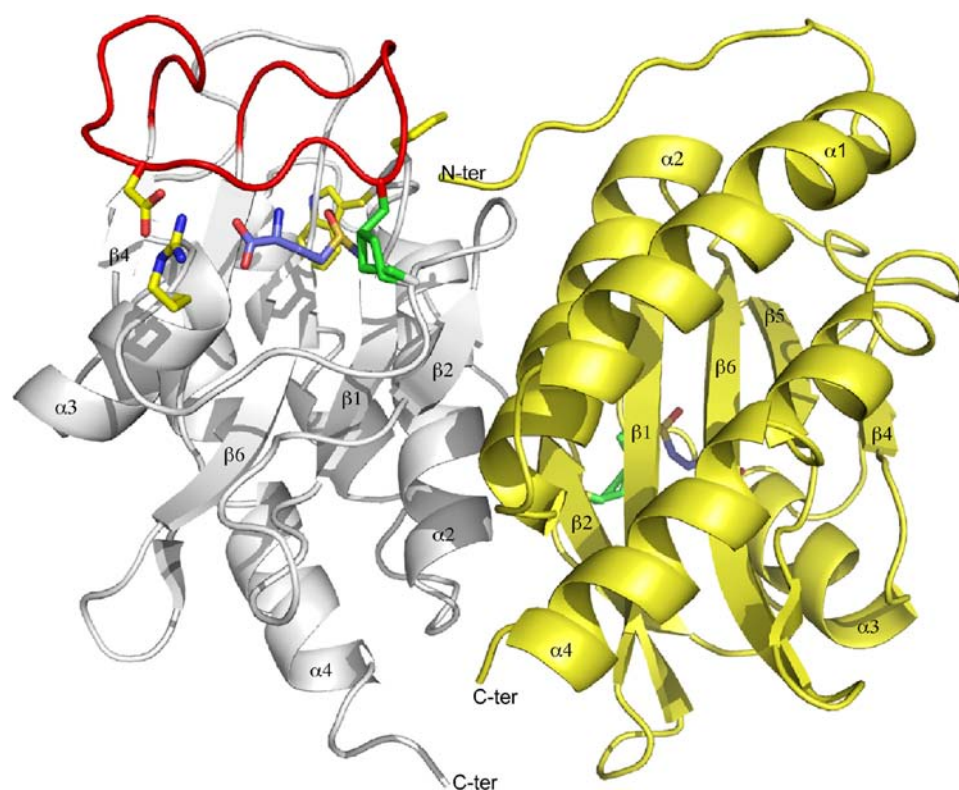


FIGURE 3. Ribbon representation of fRMSr from *N. meningitidis* in complex with L-Met-R-O. The two subunits are in white and yellow. The flap of the subunit A is in red. The side chains of Cys¹¹⁸ and Cys⁸⁴ are shown in stick representations and colored in green. The L-Met-R-O is represented in a stick representation with the carbon atoms in purple, the oxygen atoms in red, the nitrogen in blue, and sulfur atoms in yellow. The side chains of Arg⁹¹, Asp¹¹⁴, Phe⁶¹, and Trp⁶² are represented in sticks colored yellow for the carbon atoms, red for oxygen atoms, and blue for nitrogen atoms.

sition of the dimeric fRMSr of *N. meningitidis* with two known structures of fRMSrs from *S. cerevisiae* (13) and *E. coli* (12) (Protein Data Bank entries 1f5m and 1vhm, respectively), reveals that most of the differences are located within the protein loops (r.m.s. deviations of 1.218 and 1.151 Å, respectively, calculated with 266 C α and 245 atoms). In the following, the residue numbering is that from the sequence of the *E. coli* enzyme according to the sequence alignment (Fig. 1).

The interface between the two subunits involves the β -strand β 2, the α -helix α 2, the loop (Gly⁸⁰–Arg⁸⁶), and the first 6 amino acids that are not well defined in the $2F_o - F_c$ electron density maps. The 2125-Å² area between the two interacting subunits, which is similar to that found for the *E. coli* enzyme, represents 25% of the overall area of each subunit and provides a strong dimer association. Molecular contacts between both subunits mainly involve direct van der Waals interactions, hydrogen bonds, and two salt bridges between Arg⁸⁶ and Glu³⁹ of each subunits. In addition, indirect hydrogen bonds and salt bridges are mediated via water molecules and magnesium ions.

The structure of a subunit is organized around a central twisted antiparallel β -sheet consisting of six strands (β 3- β 2- β 1- β 6- β 5- β 4) flanked on one side by a three- α -helix bundle that encompasses the helices α 1, α 2, and α 4. On the other side, the “loop- β 3- α 3-loop” motif encompassing residues Gly⁸⁰–Thr¹⁰³ links the strands β 2 to β 4 and girdles the overall β -sheet and sits on the short loop located between β 1 and β 2 strands. In

addition, the loop between β 4 and β 5 strands (residues Gly¹⁰⁷–Ser¹²⁴) acts as a capping flap that shields the active site.

A disulfide bond is observed between Cys⁸⁴ and Cys¹¹⁸. The disulfide bond is formed during the crystallization process similarly to the two other crystal structures of fRMSr described so far. Indeed, the addition of Met-O, which is substrate, was done after obtaining the crystals and just before flash-freezing.

The Flap—Like the other fRMSr structures, the flap is locked within the crystal structure by a disulfide bridge between Cys⁸⁴ and Cys¹¹⁸ (Fig. 3). The side chain of Cys⁸⁴ adopts a double conformation. The conformation of the flap is stabilized by van der Waals interactions with main and side chain atoms of residues located in the loop between the strand β 6 and the helix α 4. A loop located between the helix α 2 and the strand β 1 contains 3 residues that also participate in the stabilization of the loop β 6- α 4 via van der Waals interactions with the side chain atom of Trp⁶². Indeed, at one side of the active site, Phe⁶¹ directly

interacts with 3 residues of the flap (Ser¹¹⁹, Leu¹²¹, and Ser¹²²). At the other side of the active site, the Arg⁹¹ that precedes the helix α 3 is involved in a salt bridge with Asp¹¹⁴ of the flap. As a consequence, the disulfide bridge (Cys⁸⁴–Cys¹¹⁸), the van der Waals interactions of Phe⁶¹ with residues 119–122, and the salt bridge (Arg⁸⁶–Asp¹¹⁴) maintain the conformation of the flap. As shown below, residues of the flap also interact with the substrate.

The mean isotropic thermal values were calculated with the atoms of the flap of both subunits A and B (20.27 and 24.37 Å², respectively), which are significantly higher than the overall B factors of the subunits A and B (15.54 and 18.62 Å², respectively). The temperature factors of the atoms of the flaps of each subunit reveal and assess the dynamic properties as well as the flexibility of the flap. Finally, the superimposition of the overall dimer with the structure of the fRMSr from *E. coli* shows r.m.s. deviations of 1.84 and 2.07 Å between residues of the flap (residues 99–113), as calculated with 14 C α of the subunits A and B, respectively. This structural difference could be related to the flexibility of the flap.

By contrast, the r.m.s. deviations with the *S. cerevisiae* enzyme reveal that the flap adopts the same conformation (r.m.s. deviations of 1.16 and 1.00 Å calculated on 14 C α atoms for subunits A and B, respectively).

Active Site—The two active sites within the dimer are separated from each other by 23.4 Å. Both bind an L-Met-R-O substrate and water molecules. The L-Met-R-O is bound in a small

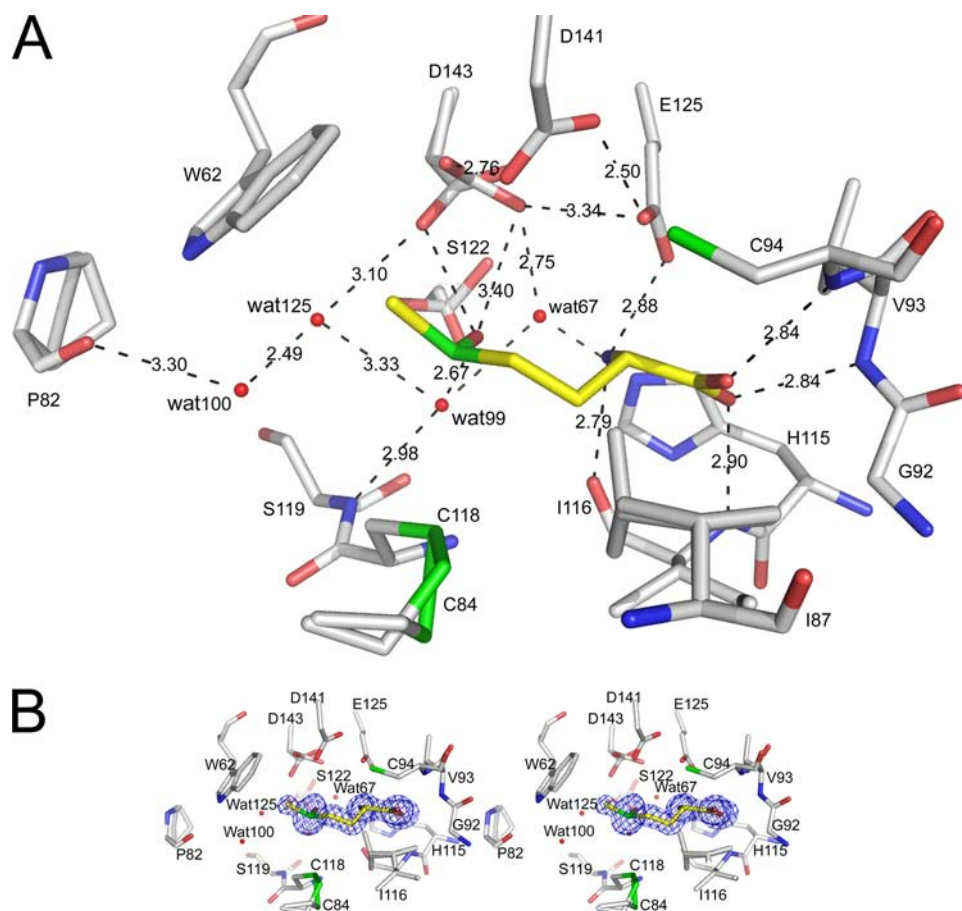


FIGURE 4. **Stick representation of the active site of fRMs_r from *N. meningitidis* with L-Met-R-O bound.** A, the L-Met-R-O is colored according to atom type with carbon atoms in yellow, whereas the surrounding amino acids are colored according to atom type with carbon atoms in white. Dashed lines, polar interactions. Water molecules are represented in a ball representation. The Pro⁸² and Cys⁸⁴ are in double conformation. The side chain of Ser¹²² was refined in three conformations. The side chain of Asp¹⁴³ was refined in double conformation. The minor conformation of Asp¹⁴³ side chain (occupancy 0.3) interacts with Wat¹²⁵ that could interact with the side chain of Cys¹¹⁸ by rotation of its side chain. B, wide eye stereo view of the active site, showing the selected residues (labeled) and omit map. The omit map was calculated around L-Met-R-O. The electron density map was contoured at 4 σ and calculated with all of the reflections between 20 and 1.25 Å. The colors of the atoms are the same as were used in A.

cavity that excludes solvent accessibility. The substrate lies down full-length on the central antiparallel β -sheet and interacts with the β 1-, β 6-, and β 5-strands. The “loop- β 3- α 3-loop” motif surrounds the substrate, forming a cavity that ensures tight binding, whereas the N-terminal part of α 3 helix plays a major role in substrate recognition and positioning via its carboxylic group. Finally, the flap that caps the active site shields the substrate molecule within the cavity and assumes the active site architecture.

Substrate Binding—The L-Met-R-O is tightly bound within the active site (Fig. 4A). The partial positive charge of the helix α ₃ dipole points toward the carboxylate of the L-Met-R-O, which makes strong interactions with invariant Val⁹³ and Cys⁹⁴ as well as an additional residue, invariant Ile¹¹⁶, all via their main chain peptidic nitrogens. Indeed, the oxygen O1 of the carboxylate is located 2.84 Å from the NH of Cys⁹⁴, whereas the oxygen O2 atom points toward the NH of Val⁹³ and Ile¹¹⁶ (2.84 and 2.90 Å, respectively). Val⁹³ and Cys⁹⁴, which are both situated at the N terminus of helix α ₃, as well as Ile¹¹⁶ are positioned on one side of the active site and form an oxyanion hole. In

addition, the NH₃⁺ is sandwiched between the oxygen of the carbonyl of Ile¹¹⁶ (2.79 Å) and the O ϵ 1 and O ϵ 2 atoms of Glu¹²⁵ (2.88 and 3.01 Å, respectively), to which it forms salt bridges. In addition, the imidazole ring of His¹¹⁵ which is located at the protein surface, acts like a wall for both the NH₃⁺ group and the O2 atom of the carboxylate of the substrate, which both point perpendicularly toward the side chain.

The C β and the C γ of the substrate are shape-stabilized within the active site. In particular, the side chains of Ile⁸⁷ that were refined in a double conformation, of Tyr⁶⁶, and of Ile¹¹⁶ participate in van der Waals interactions with these two CH₂ groups, introducing steric constraints that prevent any tipping of the SOCH₃ around the χ angle. The side chains of Ile⁸⁷ and Leu⁷⁵ are located on either side of the aromatic side chain of Tyr⁶⁶. Facing the side chain of Leu⁷⁵, the indole ring of the invariant Trp⁶² ensures the stabilization of the ϵ -methyl of the sulfoxide that points perpendicularly toward the indole moiety (Fig. 4B). The aromatic side chain of Phe⁷⁸ (or a Tyr in ~7% of the fRMs_r sequences; see the alignment shown in supplemental Fig. S1) acts like a platform upon which both Leu⁷⁵ and Trp⁶² are stacked, strictly determining their side chain conformations. Finally, the side chains of

Ile⁸⁷, Tyr⁶⁶, and Trp⁶² create a hydrophobic step shape arrangement within the active site that is necessary for substrate binding. The oxygen atom of the sulfoxide is hydrogen-bonded to a water molecule (Wat⁹⁹) (*i.e.* 2.67 Å), which is itself hydrogen-bonded to the main chain NH of Ser¹²² (*i.e.* 2.98 Å). Therefore, the oxygen atom of the sulfoxide is located at a distance of 2.88 Å from the O δ 1 atom of Asp¹⁴³. In the later stages of the refinement, the quality of the 2F_o - F_c maps as well as the F_o - F_c residual maps around the side chain of Asp¹⁴³ allowed us to refine the side chain to two conformations. Whereas the minor conformation (occupancy 0.3) of the side chain of Asp¹⁴³ is able to interact directly with the oxygen atom of the sulfoxide, this interaction does not occur in the major conformation (occupancy 0.7), and instead a contact is formed with Asp¹⁴¹. The major conformation of Asp¹⁴³ allows a network of hydrogen bonds to form between two other important and conserved acidic residues, Asp¹⁴¹ and Glu¹²⁵. The distance between the O δ 2 of both aspartates 143 and 141 is 2.76 Å, whereas the oxygen O δ 1 of Asp¹⁴¹ strongly interacts with the O ϵ 2 atom of Glu¹²⁵ (2.50 Å), which, as mentioned previously, is involved in a

Structure and Catalytic Mechanism of *N. meningitidis* fRMsr

salt bridge with NH_3^+ . The carboxylates of these residues that belong to the central β sheet give rise to a relatively flat acidic surface. In addition, another water molecule (Wat⁶⁷) is located in a distal position with respect to the oxygen atom of the sulfoxide (3.22 Å). The side chains of Asp¹⁴³, His¹¹⁵, Ser¹²², and another water molecule, Wat⁹⁹, are closely hydrogen-bonded to the water molecule Wat⁶⁷ (2.75, 2.64, 2.70, and 3.03 Å, respectively). Finally, all of the residues of the active site impose geometrical constraints on the substrate, illustrated by the proximity of the amino group to the sulfoxide (distance N–O, 2.79 Å). All of the enzyme-substrate interactions raise the question of how these geometrical constraints affect the electronic state of the sulfoxide that could be catalytically relevant. The crystal structure reveals that the sulfur atom of the sulfoxide is placed almost linearly between the O δ 2 atom of Asp¹⁴¹ and the thiol of Cys¹¹⁸ (3.77 and 3.64 Å, respectively). In addition, the final residual $F_o - F_c$ maps reveal a small negative peak at 5.5σ on the thiol of Cys¹¹⁸. In that context, according to the 1.25 Å resolution, it is difficult to deduce whether the crystal suffered from radiation damages during the data collection or whether this peak revealed a lack of electron density on the thiol of Cys¹¹⁸ with respect to electronic effects. In addition, a small positive peak at 4σ very close to the sulfur atom of the substrate (1.5 Å) but distinct from that in the $2F_o - F_c$ map is located between the sulfur atom of the substrate and Asp¹⁴¹.

The Cys⁸⁴ adopts two conformations and makes a disulfide bridge with Cys¹¹⁸ in the crystal, as observed in the two other crystal structures solved so far. Although the thiols of Cys⁹⁴ and Cys⁸⁴ are distant from the sulfur atom of the L-Met-R-O (5.73 and 4.84 Å, respectively), the thiol of Cys¹¹⁸ is located only 3.64 Å away. In contrast to the other Cys within the active site, the positioning of the side chain of Cys¹¹⁸ is appropriate to allow a nucleophilic attack on the sulfur atom of the sulfoxide by side chain rotation, when the disulfide bond between Cys¹¹⁸ and Cys⁸⁴ is reduced. In addition, the angle of attack is consistent with a catalytic trigonal bipyramidal geometry of the transition state, with the oxygen atom of the sulfoxide positioned in a distal location. A search of other rotamers of the Cys¹¹⁸ side chain reveals that one rotamer allows for the thiol to be brought close to the sulfur atom of the sulfoxide (2.51 Å).

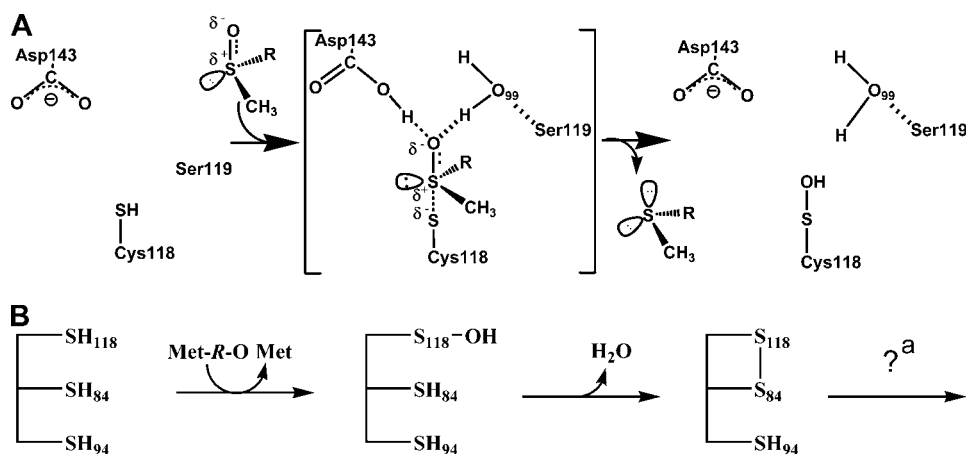
DISCUSSION

A scenario of the catalytic mechanism, similar to that of MsrA and MsrB, can be proposed for fRMsr based on the genomic, biochemical, kinetic, and structural analyses carried out in the present study. First, a sulfenic acid intermediate is formed on the catalytic Cys¹¹⁸. This conclusion is based on four types of data. First, only the substitution of Cys¹¹⁸ by Ala leads to a complete loss of activity, whereas the mutants in which the other three Cys are individually substituted with Ala, in particular the C84A and C94A mutants, remain active under steady-state conditions. These latter results contradict those described for fRMsr from *S. cerevisiae*, for which substitution of Cys⁹⁴ and Cys⁸⁴ by Ser led to a complete loss of activity (11). Second, the C84A/C94A/C136A mutant is active when tested under steady-state conditions with the Trx-regenerating system. Third, in the absence of Trx, several species with the C84A/

C94A/C136A mutant are efficiently formed, including a covalent intermediate, which migrates as a dimer of fRMsr on denaturing SDS-polyacrylamide gel. Its mass corresponds to a dimer linked by a thiosulfinate bond (data not shown). Because formation of a thiosulfinate bond can only proceed via condensation of two sulfenic acids, this result demonstrates that reduction of Met-O passes through the sulfenic acid chemistry. Fourth, inspection of the alignment of the 358 protein sequences of putative fRMsrs deduced from the DNA sequences available to date shows that Cys¹¹⁸ is invariant, whereas Cys⁸⁴ and Cys⁹⁴ are 94.5 and 98.6% conserved (see the alignment shown in supplemental Fig. S1). Taken together, these data definitively prove that 1) Cys¹¹⁸ is the catalytic Cys and 2) the catalytic mechanism involves the formation of a sulfenic acid intermediate on Cys¹¹⁸. Therefore, these results exclude Cys⁹⁴ as the catalytic Cys as postulated from molecular modeling on the *S. cerevisiae* fRMsr (11) and from inspection of the *E. coli* fRMsr in complex with a MES molecule trapped from the crystallization buffer (10).

The crystal structure of the *N. meningitidis* fRMsr in complex with L-Met-R-O illuminates not only the catalysis of the reductase step but also how fRMsr binds Met-O with a strong preference for the L-Met-R-O isomer. As already indicated, the catalysis implies a proton transfer from the catalytic Cys to the oxygen of the sulfoxide. Such a function in fRMsr is probably played by the invariant Asp¹⁴³, which is located on strand β_6 , possibly via Wat¹²⁵. In accord with this assumption, a D143A fRMsr shows a 350-fold decrease in activity (data not shown), similar to that observed for E94A MsrA and H103A MsrB, when determined under similar steady-state conditions (8, 9). The crystal structure also shows that the oxygen of the sulfoxide is located in a hydrophilic cavity, where it strongly interacts not only with Asp¹⁴³ but also with Wat⁹⁹, which is itself hydrogen-bonded with the main chain nitrogen of Ser¹¹⁹. Wat⁹⁹ recalls the water molecule observed in MsrB that is stabilized through a hydrogen-bonding network with the side chains of Thr²⁶, His¹⁰⁰, and Asn¹¹⁹ (21). In the scenario, a sulfurane intermediate/transition state is formed that rearranges into a sulfenic acid intermediate (see Scheme 1A). Therefore, Asp¹⁴³ and Wat⁹⁹ should interact not only with the oxygen of the sulfoxide and its polarized form but probably also with the OH of the sulfurane transition state and directly or indirectly with the OH of the sulfenic acid, as observed in the crystal structure of the MsrA sulfenic acid intermediate (22). To form a sulfurane, the Cys¹¹⁸ has to be under the thiolate form to attack the sulfur of the sulfoxide, the geometry of which is tetrahedral. Clearly, because the pK_{app} of Cys¹¹⁸ in the free enzyme is 8.5, binding of the substrate should contribute to the decrease of the pK_{app} of Cys¹¹⁸. This situation is similar to that observed in MsrA and MsrB. As shown by the crystal structure, the sulfur is closer to Cys¹¹⁸ than to Cys⁸⁴, and consequently it is in a better position to form a covalent bond with Cys¹¹⁸. Moreover, if Cys¹¹⁸ is reduced, the thiolate can adopt a favorable orientation that permits formation of the sulfurane with the SH and the OH groups in apical positions.

As shown previously (10), fRMsrs only reduce the R isomer at the sulfoxide function. This stereochemical preference is con-



SCHEME 1. Proposal for the catalysis of the fRMsr reductase step (A) and the mechanism up to the formation of the disulfide Cys¹¹⁸-Cys⁸⁴ intermediate (B). A, the substrate binds to the active site with its sulfoxide function largely polarized, leading to a transition state compatible with a sulfuran of bipyramidal geometry. The rearrangement of the transition state leads to the formation of a sulfenic acid intermediate on the catalytic Cys¹¹⁸ residue and of Met. In this scenario, Asp¹⁴³ and Wat⁹⁹, which is itself hydrogen-bonding with the main chain nitrogen of Ser¹¹⁹, stabilize the sulfuran transition state via interaction with the OH group, a situation similar to that described for MsrB (7). B, the reductase step (the first one) leads to the formation of a sulfenic acid intermediate on the catalytic Cys¹¹⁸ and of Met. In the second step, attack of Cys⁸⁴ on the sulfur atom of the sulfenic acid intermediate leads to the formation of a disulfide bond between Cys¹¹⁸ and Cys⁸⁴ and the release of a water molecule. a, return of the active site to a fully reduced state can proceed via reduction of the Cys¹¹⁸-Cys⁸⁴ disulfide bond formed (see "Discussion") by Trx or if the Cys⁹⁴-Cys⁸⁴ disulfide bond forms (which is rather unlikely), via the reduction of the Cys⁹⁴-Cys⁸⁴ disulfide bond by Trx.

firmed by the structure of *N. meningitidis* fRMsr in complex with L-Met-R-O. The methyl group of the sulfoxide points perpendicularly toward the indole ring of the invariant Trp⁶², whose conformation is strongly stabilized within the active site. Therefore, the indole ring is the major factor involved in the binding and stabilization of the ϵ -methyl group, as already observed in MsrA and MsrB, where at minimum a Trp residue is involved. The β - and δ -methylene groups are stabilized by van der Waals interactions with the side chains of the invariant Ile⁸⁷, Ile¹¹⁶, and Tyr⁶⁶. Together, these structural data explain why only the R isomer at the sulfoxide function is recognized by the fRMsrs.

As shown in the present study from kinetic studies, fRMsr is highly selective for the L-isomer at the C α carbon. Moreover, as already described, it is the free form of Met-O that is active. Inspection of the crystal structure of fRMsr shows the following. 1) The negative charge of the carboxylate function, which is completely delocalized, is buried and points toward the positive helix dipole at the N-terminal end of the helix α 3. In addition, the backbone amides of Val⁹³, Cys⁹⁴, and Ile¹¹⁶ are within hydrogen distance and form an oxyanion hole. A similar situation is observed with the negative charge of the sulfonic acid function of the MES molecule in the *E. coli* fRMsr (10), the negatively charged 3',5'-cyclic phosphate of cGMP and cAMP in the phosphodiesterase 5A (23), and adenylate cyclase from *Anabaena* (24), respectively. 2) The NH₃⁺ group is strongly stabilized by a hydrogen-bonding network involving (i) the side chain of the invariant Glu¹²⁵, which itself is stabilized by interactions with the invariant Asp¹⁴¹ and Asp¹⁴³, and (ii) the main chain oxygen of Ile¹¹⁶. All of these stabilizing interactions 1) favor binding of the L-Met-O epimer, 2) prevent the binding of a Met-R-O within a peptide or protein context due to steric hindrance, and 3) are responsible for the absence of activity

with DMSO because of the absence of COO⁻ and NH₃⁺ groups in DMSO.

As clearly revealed by the crystal structure, the L-Met-R-O is strongly bound in a small cavity that is capped by the flap (residues 107–124, including Cys¹¹⁸), which thus protects the active site from the solvent. In the structure, the Cys⁸⁴-Cys¹¹⁸ disulfide bond should restrict the conformational freedom of the flap, which is situated between strands β 4 and β 5, respectively. Because Cys¹¹⁸ has to be free to play its catalytic role, the disulfide-bonded Cys¹¹⁸-Cys⁸⁴ form is not catalytically relevant for the reductase step. Nevertheless, the positioning of Cys¹¹⁸ is compatible with a nucleophilic attack of the sulfoxide function of the substrate. The disulfide bond Cys¹¹⁸-Cys⁸⁴ is, however, essential to obtaining crystals because the addition of

DTT dissolves them (data not shown) and thus prevents obtention of crystals of the reduced form.

Formation of a sulfenic acid intermediate on Cys¹¹⁸ implies a recycling process in the wild type that can involve formation of an intradisulfide bond, which is then reduced by Trx. Two Cys, Cys⁹⁴ or Cys⁸⁴, could be implicated. All of the fRMsr mutants in which either Cys⁸⁴ or Cys⁹⁴ was substituted by Ala were active. In agreement with these results is the fact that 5.6 and 1.4% of the putative sequences of fRMsrs only contain the signatures Cys¹¹⁸/Cys⁹⁴ and Cys¹¹⁸/Cys⁸⁴, respectively (see the alignment shown in [supplemental Fig. S1](#)). When Cys¹¹⁸ and Cys⁹⁴ were individually mutated to Ala in the wild type from *N. meningitidis*, a stoichiometry of 1 mol of Met/subunit was observed in the absence of Trx with formation of a disulfide bond between Cys¹¹⁸ and Cys⁹⁴ or between Cys¹¹⁸ and Cys⁸⁴, depending on the mutant considered. Such results are based on the difference in thiol content observed before and after the addition of Met-O. For the wild type and the C136A fRMsr, a stoichiometry of 1 mol of Met/subunit was also observed with formation of a disulfide bond, the nature of which remains to be determined. Because 1) the disulfide bond in the crystal is between Cys¹¹⁸ and Cys⁸⁴ when Met-O is bound, 2) Cys⁹⁴ is located at the N terminus of helix α 3 on one of the two sides of the active site far away from Cys¹¹⁸ and Cys⁸⁴ (*i.e.* \sim 8 Å), and 3) Cys¹¹⁸ is located within the flap, it is probable that the first disulfide formed for recycling the activity of the fRMsr from *N. meningitidis* is Cys¹¹⁸-Cys⁸⁴ (see Scheme 1B). This is in accord with the DTNB titration results obtained under native conditions, which showed that Cys¹¹⁸ and Cys⁸⁴ are titrated by DTNB and formed a disulfide bond in the wild type and C136A/C94A mutant. In fact, one of the two Cys residues, probably Cys¹¹⁸ that is located on the flexible flap

Structure and Catalytic Mechanism of *N. meningitidis* fRMsr

forms an adduct with the TNB⁻ moiety, which, by an intramolecular attack of Cys⁸⁴, leads to formation of the Cys⁸⁴-Cys¹¹⁸ disulfide bond. When one of these two Cys residues, either Cys⁸⁴ or Cys¹¹⁸, is substituted by Ala, Cys⁹⁴ is now titrated together with either Cys¹¹⁸ or Cys⁸⁴, respectively, and thus is accessible. But no disulfide bond between Cys⁹⁴ and either Cys¹¹⁸ or Cys⁸⁴ is formed, in contrast to that observed between Cys¹¹⁸ and Cys⁸⁴. Clearly, formation of the Cys⁸⁴-Cys¹¹⁸ disulfide bond in the wild type prevents accessibility of the DTNB to Cys⁹⁴, whereas in the absence of the Cys⁸⁴-Cys¹¹⁸ disulfide bond, as is the case for the C118A or C84A mutant, Cys⁹⁴ becomes accessible to DTNB.

The fact that only one Met molecule was formed in the wild type and in the C136A mutant requires explanation. Either Cys¹¹⁸ remains engaged in a disulfide bond with Cys⁸⁴, it participates in a new disulfide bond with Cys⁹⁴, or it is free. The last hypothesis, although less probable, implies formation of a Cys⁸⁴-Cys⁹⁴ disulfide bond and an active site that is no longer catalytically competent to reduce another Met-O molecule. In other words, formation of the Cys⁸⁴-Cys⁹⁴ disulfide bond would perturb the conformation of the active site, which in return would prevent efficient binding of L-Met-R-O. In any case, the behavior of the wild-type fRMsr is clearly different from that of the MsrA from *E. coli* (6). Indeed, the *E. coli* MsrA, which also possesses three Cys residues, forms two successive disulfide bonds, first between the catalytic Cys⁵¹ and the recycling Cys¹⁹⁸ and then between Cys¹⁹⁸ and Cys²⁰⁶. In that case, formation of the second disulfide bond renders the catalytic Cys⁵¹ free and renders the active site competent to reduce a second molecule of Met-O.

The fact that oxidized disulfide forms of fRMsr are formed (*i.e.* containing a Cys¹¹⁸-Cys⁸⁴ or Cys¹¹⁸-Cys⁹⁴ bond) or would be formed in the absence of reductant (*i.e.* Cys⁸⁴-Cys⁹⁴) does not mean that all of these oxidized forms are catalytically competent under steady-state conditions (*i.e.* in the presence of Trx). For instance, in the 5.6% of putative fRMsrs where only Cys⁹⁴ is present, a disulfide bond probably forms between Cys¹¹⁸ and Cys⁹⁴ in the absence of reductant, as shown for the C84A/C136A fRMsr from *N. meningitidis*. However, inspection of the crystal structure shows, as already pointed out, that Cys⁹⁴ is situated at one of the sides of the active site far away from Cys¹¹⁸. Thus, Cys⁹⁴ is not in a good position to form a Cys¹¹⁸-Cys⁹⁴ bond efficiently unless the enzyme is highly flexible. At present, we do not know the rate at which this disulfide bond is formed in the C84A mutant and whether it is this form that is recognized and reduced by Trx.

Like all GAF domains, fRMsr is dimeric. An obvious question is whether not only both subunits of fRMsr are active as shown in the absence of Trx but also express their activity with the same catalytic efficiency. Another question concerns the nature of the rate-limiting step. This has to be determined in order to validate the amino acids that play a role in the catalysis and in the substrate specificity in the reductase step.

The fRMsr represents a unique case in which a GAF fold behaves as an enzyme and moreover as an independent folded unit not included in a modular larger protein. Like MsrA and MsrB, it catalyzes the reduction of Met-O and therefore constitutes the third protein fold to exhibit a Met-O reductase activity

that involves the sulfenic acid chemistry. From a structural point of view, the active site of fRMsr is clearly different from those of MsrA and MsrB. All of the amino acids within the active site that underlie the L-Met-R-O reductase activity are not present in the other GAF domains. From an evolutionary point of view, this observation raises the question of how a GAF fold has evolved to acquire a broad range of different functions by binding various molecules, including cyclic nucleotides, heme, amino acids, or Met-O.

Acknowledgments—We thank Dr. A. Thompson, Dr. B. Guimaraes, and Dr. P. Legrand for help during data collection at beamline PROXIMA 1 at synchrotron Soleil (GIF-sur-YVETTE); Dr. G. Chevreux, Dr. S. Sanglier-Cianferani, and Dr. A. Van Dorsselear for mass spectrometry analyses; Sabrina Hecht for efficient technical help in fRMsr purification; and Dr. K. Weissman for careful reading of the manuscript. Access to the Bruker X8-Proteum and to the TDA 302 VISCOTEK (SEC-Light Scattering) (Service Commun de Biophysicochimie des Interactions Moléculaires, Nancy I) is deeply appreciated.

REFERENCES

- Schöneich, C. (2005) *Biochim. Biophys. Acta* **1703**, 111–119
- Boschi-Muller, S., Olry, A., Antoine, M., and Branlant, G. (2005) *Biochim. Biophys. Acta* **1703**, 231–238
- Moskovitz, J. (2005) *Biochim. Biophys. Acta* **1703**, 213–219
- Stadtman, E. R., Van Remmen, H., Richardson, A., Wehr, N. B., and Levine, R. L. (2005) *Biochim. Biophys. Acta* **1703**, 135–140
- Weissbach, H., Resnick, L., and Brot, N. (2005) *Biochim. Biophys. Acta* **1703**, 203–212
- Boschi-Muller, S., Azza, S., Sanglier-Cianferani, S., Talfournier, F., Van Dorsselear, A., and Branlant, G. (2000) *J. Biol. Chem.* **275**, 35908–35913
- Boschi-Muller, S., Gand, A., and Branlant, G. (2008) *Arch. Biochem. Biophys.* **474**, 266–273
- Antoine, M., Gand, A., Boschi-Muller, S., and Branlant, G. (2006) *J. Biol. Chem.* **281**, 39062–39070
- Neiers, F., Sonkaria, S., Olry, A., Boschi-Muller, S., and Branlant, G. (2007) *J. Biol. Chem.* **282**, 32397–32405
- Lin, Z., Johnson, L. C., Weissbach, H., Brot, N., Lively, M. O., and Lowther, W. T. (2007) *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9597–9602
- Le, D. T., Lee, B. C., Marino, S. M., Zhang, Y., Fomenko, D. E., Kaya, A., Hacıoglu, E., Kwak, G. H., Koc, A., Kim, H. Y., and Gladyshev, V. N. (2009) *J. Biol. Chem.* **284**, 4354–4364
- Badger, J., Sauder, J. M., Adams, J. M., Antonysamy, S., Bain, K., Bergseid, M. G., Buchanan, S. G., Buchanan, M. D., Batiyenko, Y., Christopher, J. A., Emtage, S., Eroshkina, A., Feil, I., Furlong, E. B., Gajiwala, K. S., Gao, X., He, D., Hendle, J., Huber, A., Hoda, K., Kearins, P., Kissinger, C., Laubert, B., Lewis, H. A., Lin, J., Loomis, K., Lorimer, D., Louie, G., Maletic, M., Marsh, C. D., Miller, I., Molinari, J., Muller-Dieckmann, H. J., Newman, J. M., Noland, B. W., Pagarigan, B., Park, F., Peat, T. S., Post, K. W., Radojicic, S., Ramos, A., Romero, R., Rutter, M. E., Sanderson, W. E., Schwinn, K. D., Tresser, J., Winhoven, J., Wright, T. A., Wu, L., Xu, J., and Harris, T. J. (2005) *Proteins* **60**, 787–796
- Ho, Y. S., Burden, L. M., and Hurley, J. H. (2000) *EMBO J.* **19**, 5288–5299
- Antoine, M., Boschi-Muller, S., and Branlant, G. (2003) *J. Biol. Chem.* **278**, 45352–45357
- Olry, A., Boschi-Muller, S., Marraud, M., Sanglier-Cianferani, S., Van Dorsselear, A., and Branlant, G. (2002) *J. Biol. Chem.* **277**, 12016–12022
- Sharov, V. S., Ferrington, D. A., Squier, T. C., and Schöneich, C. (1999) *FEBS Lett.* **455**, 247–250
- Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
- Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
- Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, P. (1998) *Acta Cryst. D* **54**, 905–921

Annexe 1 *Structure and Catalytic Mechanism of N. meningitidis fRMs*

- E. J. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55**, 247–255
20. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
21. Ranaivoson, F. M., Neiers, F., Kauffmann, B., Boschi-Muller, S., Branlant, G., and Favier, F. (2009) *J. Mol. Biol.* **394**, 83–93
22. Ranaivoson, F. M., Antoine, M., Kauffmann, B., Boschi-Muller, S., Aubry, A., Branlant, G., and Favier, F. (2008) *J. Mol. Biol.* **377**, 268–280
23. Heikaus, C. C., Stout, J. R., Sekharan, M. R., Eakin, C. M., Rajagopal, P., Brzovic, P. S., Beavo, J. A., and Klevit, R. E. (2008) *J. Biol. Chem.* **283**, 22749–22759
24. Martinez, S. E., Bruder, S., Schultz, A., Zheng, N., Schultz, J. E., Beavo, J. A., and Linder, J. U. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3082–3087

Supplemental Data

| | 20 | 30 | 40 | 50 | 60 |
|----------------|-------------------|--------------|---------------|------------|-----------|
| NP_288269.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002403065.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002397981.1 | MNK-----AEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002412850.1 | MNK-----IEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_001880631.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002391608.1 | MNK-----AEFYADLN | RDFNALMVGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_001730808.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----INWA |
| ZP_03003290.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_001743413.1 | MNK-----AEFYADFNR | DFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002407226.1 | MNK-----AEFYSDFN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002329476.1 | MNK-----AGFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_002382399.1 | MNK-----TEFYADLN | RDFRALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| ZP_03445135.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_688895.1 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| YP_669680.1 | MNK-----AEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----VNWA |
| NP_416346.4 | MNK-----TEFYADLN | RDFNALMAGETS | -----FLATLANT | SALLYERLTD | -----INWA |
| ZP_02903917.1 | MNK-----TEFYADLN | RDFNALMTGETS | -----FLATLANT | SALLYERLAD | -----VNWA |
| ZP_03282954.1 | MNK-----SEFYADLN | RDFQALMAGETS | -----FLATLANT | SALLFERLSG | -----VNWA |
| YP_001452718.1 | MSK-----TELYADLN | RDFQAVMAGETS | -----FLATLANT | SALLFERLSG | -----VNWA |
| YP_001570142.1 | MSE-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| NP_460803.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_150310.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_002141797.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_002243298.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| NP_456352.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| NP_804849.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_001587603.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02666379.2 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02832604.2 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_001177122.1 | MNK-----TEFYADLN | RDFQALMTGETS | -----FLATLANT | SALLFERLSD | -----VNWA |
| YP_002226291.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| NP_754135.1 | ----- | -----MAGETS | -----FLXTLANT | SALLYERLTD | -----VNWA |
| ZP_02347214.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02574893.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02659310.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02663512.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02681881.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| ZP_02696584.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_002041101.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |
| YP_002045892.1 | MSK-----TELYAALNR | DFQSLMAGETS | -----FLATLANT | SALLFERLTE | -----VNWA |

ZP_03075263.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 YP_002114879.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 YP_002146178.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03165709.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 YP_002215294.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03221812.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03342093.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03347154.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03352578.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03357060.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03371453.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03379701.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03383790.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03412299.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 ZP_03217260.1 MSK-----TEL YAALNRDFQS LMAGETS-----FL ATLANTSALLFERLTE-----VNWA
 YP_002237781.1 MSK-----TEFYADLNRFQALMAGETS-----FL AMIANTSALLFERLSE-----VNWA
 YP_001336004.1 MSK-----TEFYADLNRFQALMAGETS-----FL AMIANTSALLFERLSE-----VNWA
 YP_001437512.1 MNK-----TQFYAELNRDFHS LMAGETS-----FL ATLANTSALLFERLEG-----VNWA
 ZP_03403067.1 -----MATLANTSALLFERLTE-----VNWA
 YP_001907453.1 MTK-----AEFYTDLNRMNALLSGETS-----FL AVMSNCSALLYKRLDG-----VNWA
 YP_001006127.1 MKK-----AEFYAELKRDL SALIAGETN-----FI ATLANASALLYERLDG-----LNWA
 NP_669181.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 NP_993079.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_070893.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_651729.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_647856.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001163103.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_01888376.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001400647.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001607069.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02221459.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02227087.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02229339.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02237158.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02307206.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02314388.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 ZP_02318544.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001720525.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001872871.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_002346706.1 MKK-----TEFYAELKRDL CALVAGETN-----FI ATLANASALLYERLDG-----LNWA
 YP_001478346.1 MTK-----EQFYAELKRDL CALLDGETN-----FI AALSNASALINERLDD-----VNWV
 YP_454994.1 MKK-----QQFYAELLQDMSALLDEETH-----FI AALANASALLYERLEG-----VNWA

YP_001217052.1 EVFIFQMEVIMNLEQYQRLTKQAVALLLEGETN-----LIANLANLSALLNMELTE-----LNWV
 YP_002150764.1 MPQ-----LNFYQELTDNLSALLSGEND-----FIASLANTSALIYEKLDK-----INWV
 YP_856816.1 IEK-----QAFYRTLNRQQAQALLLEGEPD-----LIANLANLSALLNQELAD-----INWV
 YP_001093925.1 MST-----PCYTTLTRQVEALFSGEDN-----LVTAMANFSALLNEHLDD-----INWV
 YP_269525.1 MSN-----QAEFYHELLGQTKALINDESD-----IIANMANISALLFERLAD-----VNWV
 NP_718189.1 MKP-----QFYESLNRQVLAALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_001366642.1 MKS-----KFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_001554990.1 MKS-----KFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_737796.1 MKP-----QFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_869415.1 MKP-----QFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 NP_797987.1 -----MNIEQYQRLTKQAVALLIESEPD-----FIANLANLSALLFMELED-----LNWA
 YP_001501790.1 MVN-----AFYDTLNRQVLAALLDGEDD-----TIAALANFSALLNDNLTD-----LNWV
 YP_927460.1 MKH-----AFYDTLNRQAEALLSGEDD-----LVAAAMNFSALLNEHMDD-----LNWV
 YP_001474222.1 MIT-----EFYSSLDQRQAKALLDGEDD-----LIAAMANFSALINDALDD-----LNWV
 YP_963195.1 MKS-----QFYESLNRQALALLLEGEDD-----VVAAAMNFSALLNDNLTD-----LNWV
 YP_001183724.1 MKS-----QFYESLNRQALALLLEGEDD-----VVAAAMNFSALLNDNLTD-----LNWV
 YP_733804.1 MKP-----QFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_001050811.1 MKS-----KFYESLNRQALALLLEGEDD-----LVAAAMNFSALLNDNLTE-----LNWV
 YP_750914.1 MKT-----EFYQTLNRQAQALLLEGEDD-----LIAAMANFSALLNDSLED-----LNWA
 YP_002126297.1 MK-----EPQYSQLINQAKSLVSGEHD-----LIANMANISALLFNHLEE-----VNWA
 YP_130129.1 -----DEPD-----LIANISNISALLNMELED-----INWV
 YP_001445537.1 -----MNIEQYQRLTKQAVALLIESEKD-----YIANLSNISALLFMELEN-----LNWA
 YP_661562.1 MTL-----KTDLYQEVAQQAQALMAGEKD-----LIANLANLSALMYMRLPE-----LNWA
 YP_001674584.1 MVD-----AFYDTLNRQVQALFADEDD-----TIAALANFSALINDSLDD-----LNWV
 YP_002417027.1 -----MKIEHYQRLTKQAVALLIESEPD-----LTANLANISSLLFMELDE-----LNWA
 NP_934438.1 -----MTIEHYQTLTKQAVALLIESERD-----LIANLSNISALLNMGMSD-----LNWV
 NP_761466.1 -----MTIEHYQTLTKQAVALLIESERD-----LIANLSNISALLNMGMSD-----LNWV
 YP_001760510.1 MDL-----ECYDSLNRQAQALLSGEDD-----LVAAAMNFSALINDAVEQ-----LNWV
 YP_002312133.1 MTA-----SFYDTLNRQVIALLSGEDD-----QIAALANFSALLNDNLTD-----LNWV
 YP_562727.1 MKA-----DFYDTLDRQASALLAGEDD-----VIAAMANFSALLNESLED-----LNWV
 NP_693123.1 MFEQKNYSG-EKAKDYELLIKQLDALSEGENN-----KIALLSNASALLNQFLTE-----VNWV
 YP_001699747.1 MFTQINYEG-SIADQYLTLSKQLDALLTGEKD-----QIANLSNASALLHQFLTN-----INWV
 YP_001343242.1 TIETHS--TGSSEEFYQELNIQLAALLSGERD-----FICNAAQFSAFVMQTVTD-----LNWS
 YP_001141806.1 IEK-----QAFYRGLNRQQAQALLLEGEPD-----LIANLANLSALLALELTD-----INWV
 YP_204661.1 -----MELKHYARLTQQALGLIEGETD-----YIANLANISALLNMELDD-----INWV
 YP_301135.1 MTEVK-----ETNYNLLKQVSSLIEDESN-----LIAILSNVSALLNDSIDQ-----INWV
 YP_002156075.1 -----MELKHYARLTQQALGLIEGETD-----YIANLANISALLNMELDD-----INWV
 NP_244074.1 MFQPESYQG-TLEEKYSLVTKQLAALLLEGESD-----AIAANLANASALLYHFLEE-----VNWV
 YP_001559126.1 MDQIDSMYPENKVMYELLVQLKAMLEGEQD-----VIANLSNVSALLNQALKD-----INWV
 YP_080250.1 VEKTEG----SKEKKYELLKQIDALTEGEPD-----MIANTANASALLYQSLED-----VNWA
 YP_253122.1 MPLIK-----ETNYNLLTKQLSLLIEEQN-----LIAILSNVSALLNDNLTDQ-----INWV
 YP_261333.1 MIDLQS--AGQGLDGYGLLCAQLESLLADERD-----FIANAAQFSAFLYSQLED-----LNWA

YP_001187407.1 MIDLSQ--TGAGMDGYPLLAQLEALLAG^ERD-----FI^ANAAQ^FSAFLFHELAD-----LN^WA
 YP_001422246.1 VEKQSG---DKEKSYQLLLKQLEAMTE^ETD-----AI^ANYAN^NASALLYHSLHE-----VN^WA
 YP_001748360.1 MIDLNA--SGAGLDGYDLLAAQVQALFAD^ERD-----FI^ANAAQ^FSAFLYNQVDD-----LN^WA
 YP_155660.1 -----MSSQLYSDLLAQSRA^LTEG^EPD-----LI^ANLAN^NISALLYEHIED-----IN^WL
 NP_744032.1 MIDLNA--SGAGLDGYNLLAAQVQALFAD^ERD-----FI^ANAAQ^FSAFLYNQVDD-----LN^WA
 YP_001269146.1 MIDLNA--SGAGLDGYNLLAAQVQALFAD^ERD-----FI^ANAAQ^FSAFLYNQVDD-----LN^WA
 YP_236680.1 MIDLQN--TGAGLDGYGLLAAQLEALLAD^ERD-----FI^ANAAQ^FSAFLYTQLDD-----LN^WA
 YP_001667695.1 MIDINA--SGAGLDGYNLLAAQVQALFAD^ERD-----FI^ANAAQ^FSAFLFNQVDD-----LN^WA
 YP_188854.1 MTKNI-----PTNYKLISKQLAS^LIED^EKN-----LI^AILSN^NTSALLNDTIDQ-----IN^WV
 YP_002275225.1 ----MTTLPADHRVTLEDLLSTVESVLTG-----ERD-----LI^ANMAN^NISALLFEALPD-----IN^WA
 YP_804779.1 MSE-----PINSIIVEQIDALLFN^ETN-----LI^ANLSN^NASALLYSETKD-----VN^WC
 NP_470973.1 MIEIQKMTG-TKEENYALALKQVQAMIAG^EPN-----LI^ANLSN^NVSSILNQALSD-----IN^WV
 YP_607257.1 MIDLNA--SGQGLDGYDLLAAQVQALFAD^ERD-----FI^ANAAQ^FSAFLYNQVDD-----LN^WA
 YP_001860169.1 MFEATINTALPKAEFYRELTSQARS^LLEG^ESN-----QI^ANAAN^NLSALIFHSLPQ-----VN^WA
 NP_465120.1 MIEIKKMTG-TKEENYALALKQVQAMIHG^EPN-----LI^ANLSN^NVSSIINQALSD-----IN^WV
 NP_372242.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLNQ-----IN^WV
 NP_374828.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLNQ-----IN^WV
 YP_001247140.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLNQ-----IN^WV
 YP_001316940.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLNQ-----IN^WV
 YP_001442294.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLNQ-----IN^WV
 YP_002314836.1 MFHTQTYTG-TREENYTLVIEQLKALISD^EKS-----VV^ANLAN^NASALLNYFLHD-----IN^WV
 YP_001814701.1 MFETTTYDYG-DRTKQYEMLNKQLDALLT^GETN-----QI^ANLSN^NASALLGQFLDR-----TN^WV
 NP_646478.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_043764.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_186602.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_494359.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_500333.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_001332646.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_001575592.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 ZP_02759966.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_001487827.1 VEKQAG---DPSKDYQLLVKQVEAITDG^EPD-----LI^ANLAN^NAAALLYHSLPE-----VN^WA
 YP_417047.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSAILNDNLDQ-----VN^WV
 YP_041183.1 MTTIN-----PTNYTLLKKQAAS^LIED^EHH-----MI^AILSN^NMSALLNDNLDQ-----IN^WV
 YP_176251.1 MFSKLDYPT-EQEAA^YRLLEQQAVALL^EEPD-----AL^ANYAN^NLCALLGQFLDD-----IN^WV
 YP_002349929.1 MIEIQKMTG-SKEENYGLALKQVQAMIHC^EPN-----LI^ANLSN^NVSSILNQALSD-----IN^WV
 YP_001376537.1 MFTKESYVG-SREKQYETVIKQLDALLT^GESN-----TV^ANLAN^NASALLNQFLER-----IN^WV
 YP_784647.1 MFQASTLSTDFKQFYQELAAQARALLD^GEHD-----RI^ANAAN^NFAALI^WQSVPD-----IN^WS
 YP_557760.1 MFEVSTASHLPKAAHYEELVAQARS^LLAG^ETD-----WI^ANAAN^NFSAFVFHSLSD-----LN^WA
 YP_275769.1 MIDLQN--TGAGLNGYGLLAAQLEALLAD^ERD-----SI^ANAAQ^FSAFLYTQIDD-----LN^WA
 YP_001118914.1 MFTLSTDPHASKAELYATLAEQARS^LVES^EEPD-----LI^ANAAN^NFSALVYHSLDR-----LN^WA
 YP_001782545.1 MFALEDFKGLTEEQKYENMLLFLKGQLL^GEKD-----VI^ANLSN^NASAIMALIDN-----LN^WA
 NP_882192.1 MFDATPISTVSKPVFYAELAAQARALL^GEHD-----RI^ANAAN^NFSALVYQALPD-----IN^WA

YP_001894915.1 MFEVSSASHLPKAAHYEELVAQARSLLAGETD-----WIANAANFSAFVFHSLSD-----LNWA
 YP_001647268.1 MFTKESYAG-SREEQYKTVIKQLDALLTGEPN-----VVANLSNASALLNQFLDR-----VNWV
 NP_882465.1 MFDATPISTVSKPVFYAELAAQARALLAGEHD-----RIANAANFSALVYQALPD-----INWA
 YP_002448219.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 NP_834359.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_002369448.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_191278.1 ----MS--APIHRLSVDELVPMVRSVLQT-----ETD-----MIANMANIAALLFETLPQ-----INWA
 NP_886654.1 MFDATPISTVSKPVFYAELAAQARALLAGEHD-----RIANAANFSALVYQALPD-----INWA
 NP_847103.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----IVANLSNASALLNQFLDR-----VNWV
 YP_030799.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----IVANLSNASALLNQFLDR-----VNWV
 NP_981083.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_038702.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_002340711.1 MFTKESYTG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_085975.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_002453717.1 MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----VVANLSNASALLNQFLDR-----VNWV
 YP_772921.1 MFTLSNDPHASKAELYATLVEQARALVESERD-----LIANAANFSALVYHSLDR-----LNWA
 YP_001843726.1 KMMSK-----QADYQLLYQQALALFEGETD-----WIANASNLAALLYNGLAD-----VNFA
 YP_942902.1 -----MTFKDYKLLSQQAVALLIADETD-----LTANLANLSALLNMNLDN-----INWV
 YP_121651.1 SFTVAE-PTGDRVEDYRRLVAQAAALIEGEPD-----RVANAANLSALIYHSLPE-----VNWA
 YP_001633478.1 MFTATDLPTDDKPRFYAELAAQARALLDGETD-----RIANAANFAALAYQALPR-----INWA
 YP_001392189.1 MFALEDKFGLTEEQKYENMLLFLKGQLLGEKD-----VIANLSNASAIMALIDN-----LNWA
 YP_795394.1 MSE-----SLNPIITKQLDALLYGETN-----LVANLSNAAALLNQTLSD-----INWA
 YP_002284153.1 MYATKNIEQGDKRSFYRELADQLQGLLQDERD-----AVANAANLSALVFELVAD-----LNWA
 YP_001255404.1 MFALEDKFELTEEQKYENMLLFLKGQLLGEKD-----VIANLSNASAIMALIDN-----LNWA
 YP_001385170.1 MFALEDKFELTEEQKYENMLLFLKGQLLGEKD-----VIANLSNASAIMALIDN-----LNWA
 YP_001388639.1 MFALEDKFELTEEQKYENMLLFLKGQLLGEKD-----VIANLSNASAIMALIDN-----LNWA
 YP_001347703.1 MIDLQQ--TGAGLDGYALLAAQAEALFADERD-----FIANAAQFSAFLFHELDD-----LNWA
 NP_103622.1 MFAAKAIDTSDKAIFYSDLAAQLKALLDGESESD-----SIANAANTAALIFQMVPD-----LNWA
 YP_001807739.1 MFTLSNDPHASKAELYATLVEQARALVESERD-----LTANAANFSALVYHSLDR-----LNWA
 YP_001580337.1 MFTLSHDPDASKADQYATLVEQARALVESERD-----LIANAANFSALVYHSLDR-----LNWA
 YP_002230176.1 MFTLSNDPTASKADQYATLVEQARALVESERD-----LTANAANFSALVYHSLDR-----LNWA
 NP_251512.1 MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----FIANAAQFSAFLFHELGD-----LNWA
 YP_459505.1 MFDFAPDETLPAEERYRQLVDAADALTAGEGD-----GVANMANLAALMWEFIPD-----LNWA
 YP_790351.1 MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----FIANAAQFSAFLFHELGD-----LNWA
 YP_002439846.1 MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----FIANAAQFSAFLFHELGD-----LNWA
 YP_001970489.1 MFANAS-LTGSKPEQYAQLLEQARGLVYAESD-----RIANAANLSALVYHALPD-----LNWV
 YP_617023.1 MSYTLSFTATEAAERWAEAGEAAAA LVAGEHD-----GIANMANVAALIWQAIPD-----LNWA
 YP_620554.1 MFTLSNDPTASKADQYATLVEQARALVESERD-----LTANAANFSALVYHSLDR-----LNWA
 YP_834795.1 MFTLSNDPTASKADQYATLVEQARALVESERD-----LTANAANFSALVYHSLDR-----LNWA
 YP_148653.1 MFRPMAYNG-TREENYALVIEQLKSLISGEPN-----FIANLANAAALLNQFFTD-----INWV
 YP_442277.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_334263.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA

YP_001067073.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_368503.1 MFTLSNDPNASKADQYATLVEQARALVESERD-----LTANAANFSALVYHSLDR-----LNWA
 YP_109010.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_993758.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_001028775.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_002026852.1 MFANAS-LTGSKPEQYAQLLEQARGLVYAESD-----RIANAANLSALVYHALPD-----LNWV
 YP_001764423.1 MFTLSNDPTASKADQYATLVEQARSLVESERD-----LTANAANFSALVYHSLDR-----LNWA
 NP_970249.1 SINYS-----DKTKFYKELHSEAEGLAEK--EWFVN-----LANFSALLKQHLPEIN-----WV
 YP_001059788.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_001081316.1 MFALSEALPSSKPALYETLAAQARALVETETD-----IVANAANFASLVYHSLDG-----LNWA
 YP_001875000.1 -----MEIILKQIEMLLS-----PEID-----AVANMANTAAVIFNSLPR-----LNWA
 YP_278248.1 -----MLVQQYKSLIADVEK-----IYSILANTSAFIYQNFKN-----LNWA
 YP_001309672.1 MFDIKIFEGMSTENKLDNMLIMLEGLIQGDSD-----SITKLCNASALINSLDR-----INWC
 YP_001576818.1 MTAAK-----KEEYELLVEQAKALVAGESD-----WIANTANLSALLFNSLNN-----VNFA
 YP_193206.1 MAATT-----KENYELLVQQAQALIANESD-----WIANTANLSALLFNLSLSD-----VNFA
 YP_199405.1 MFATSS-LTGSKPEQYAQLTAQAQALVHGEPE-----RIANAANLAALIFHSLPS-----LNWA
 YP_001915608.1 MFATSS-LTGSKPEQYAQLTAQAQALVHGEPE-----RIANAANLAALIFHSLPS-----LNWA
 YP_002509654.1 MEQ-----KEEKLREINEKLEYLIGNERD-----WLANLANAAALLYNELDE-----VNWA
 YP_757341.1 -----MSSKADAYQDVRKQIAAVIAGETS-----RTLARLATASCLLREMFGER-----FFWT
 YP_810624.1 YMETK-----QTRYSLLNDAQSLLDGEHD-----LIANMSNLVSLIFNSDQN-----INGT
 YP_244708.1 MFATSS-LTGSKPEQYAQLTAQAQALVHGEPE-----RIANAANLAALIFNSLPS-----LNWA
 NP_635978.1 MFATSS-LTGSKPEQYAQLTAQAQALVHGEPE-----RIANAANLAALIFNSLPS-----LNWT
 XP_505859.1 YSRFSS---ESKSELYDTMASNVEALVDGQTNWVAN-----LSNASSIIWHGLKSLPEPSSD-----INWA
 NP_964546.1 MAATT-----EDNYQLLLKQAQALVEGEND-----WIANTANISALLFNSLNN-----VNFA
 YP_001261873.1 MYMFDIDRHGDKASLYADMAAALALTADEPE-----AIANMANAAALIWEYLPD-----LNWA
 YP_814311.1 MSATT-----ENNYQLLVKQAEALVDGESD-----WIANTANVSALLFNALND-----VNFA
 NP_298520.1 MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----RIANAANLAALIFHSLPG-----LNWV
 YP_001775202.1 MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----RIANAANLAALIFHSLPG-----LNWV
 YP_001829224.1 MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----RIANAANLAALIFHSLPG-----LNWV
 NP_778732.1 MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----RIANAANLAALIFHSLPG-----LNWV
 NP_246070.1 -----MMNYALLLKQLEQILEPETY-----IISRMANTSAFLYQHMPD-----LNWV
 XP_645637.1 AEDLEISKG-SKEEQYENLLPQIEGLLTGENN-----QIANLANVTAALKEQFN-----FFWV
 XP_456263.1 YSSFQT---TDRKKALEQLLISYEA LAEGQDNWVCN-----LANAASLIWHCYISLN---VD-----VNWA
 YP_001300883.1 ETSFKQTTS-NKVERYRLFLPQFELLISDEKE-----EISVLANTAAALREAFG-----FFWV
 YP_812434.1 MAAAS-----QAEYELLLEQAKALTAGESD-----WIANAGNVSALLKAGMRN-----VNFA
 NP_814796.1 QTREN-----KKAAYELCLQQLSGLLEIETD-----AIANLANSSALLAQTLPD-----TVFA
 NP_358362.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA
 YP_816246.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA
 YP_001694316.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA
 YP_001835526.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA
 YP_002037500.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA
 NP_345351.1 MLKSE-----KQSRYQMLNEELSFLLGETN-----VLANLSNASALIKSRFPN-----TVFA

XP_001385863.2 YSNVTG---LSKEETLQHIIDSYEALASDTTNWVAN-----LSNCSLLVWHAYHSLN---VP-----VNWAWA
NP_985662.1 YSNFSA---SDRSEALIYMIESYQALAEHQDNMVCN-----LANASSLLWHVYRNLN---VK-----VNWAWA
YP_759928.1 -----MDKAEAYRQLKEEIDSVVAGETS-----DTARYASAACLLAQAFKPR-----FFWT
YP_002123053.1 MNKTT-----KIEHYQLLLAQAKALFSNESN-----ALANLSNASALLKMTLPN-----SVFA
YP_002131343.1 EAFNEIVLPASKAERYAAVAEEIASVLDGEPN-----LVARMATVASMLANSFEH-----YFWT
XP_001683506.1 ---TPAPTFSSKKELVDWLVKQVEALVHDMSKRFTPQAN-----LVIGLSNTAALCFYELNRFNPEAALEKLRVNWFW
NP_012854.1 HVNYSSN--LNKEEILEQLLLSYEGLSDGQVNWVCN-----LSNASSLIWHAYKSLA---VD-----INWAWA
XP_460154.1 YTSYSN--DGGKEEILQSVVDSYEALSYDTRNWVAN-----LANCSSLLVWHAYHSMK---IN-----VNWWS
XP_001801028.1 ASAFA--EGLSKKEVYAQVLEQARILFDGQRNWVCN-----FSNTSSLLWHALHSLPSPSHS-----VNWAWA
YP_001201197.1 MTKQE-----KISNYQLLLAQLEALLDGETN-----ALANLSNASALLNQLPN-----SVFT
XP_001484046.1 YSNYKG--DETKEETLQSVVDSYEALSSDTQNWVAN-----LANCSSLIWHAYHSLN---IP-----VNWAWA
NP_274882.1 MHALHFSAS-DKAALYREVLQPQIESVVAD EAD-----WVANLANTA AVLKEAFG-----WFWVWV
YP_001302571.1 AEELRIKNG-DKQEMYETLLPQIASLVGN ETD-----LIANMANVAAALKQTFG-----FFWVWV
YP_002000650.1 MHALHFSAS-DKAALYREVLQPQIESVVAD EAD-----WVANLANTA AVLKEAFG-----WLVWVWV
YP_001682978.1 EAFNDKTLTGDKAARYAELAQEIASVLDG ETD-----LTARMATVASMLASSFDH-----YFWT
YP_002342048.1 MHALHFSAS-DKAALYREVLQPQIESVVAD ETD-----WVANLANTA AVLKEAFG-----WFWVWV
YP_207199.1 MHALHFSAS-DKAALYREVLQPQIESVVAD EAD-----WVANLANTA AVLKEAFG-----WLVWVWV
XP_844799.1 GASTPTGGRSDKVTFYSRLVKEVQGIVEFPSKRLAPAAN-----AIINLSNVAASLFYGLNRHANPEAVLQSPPVNWL
YP_001598496.1 MHALHFSAS-DKAALYREVLQPQIESVVAD ETD-----WVANLANTA AVLKEAFG-----WFWVWV
YP_974447.1 MHALHFSAS-DKAALYREVLQPQIESVVAD ETD-----WVANLANTA AVLKEAFG-----WFWVWV
XP_001465858.1 ---TPAPTFSSKKELVDWLVKQVEALVHDMSKRFTPQAN-----LVIGLSNVAALCFYELNRFNPEAALEKLVNWF
YP_001280815.1 QAINPIDPNANKAEKYQLLLKQAEATISS ESD-----LIANMANLTALLKDTFN-----WFWWI
NP_664843.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
NP_802083.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
YP_280574.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
YP_001128311.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
YP_002286076.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
XP_390238.1 ASNFA--EGVTKEDAYEQVLWQAEGLVTDQRNWVCN-----LSNAAASLLWHAYKSLGSPSKD-----VNWAWA
XP_001565241.1 ---LPAPIFTSKKELVDWLVKQVEALVDGMSKRFTPQAN-----LIIGLSNVAALCFYELNRFNCPGAALEKLVNWF
YP_596891.1 MNKSK-----KIEQYQLMIEQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
YP_600772.1 MNKSK-----KIEQYQLMIEQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFA
YP_001351938.1 AEVLSFNKLGKADIYAELLPQIEALLA EEPD-----LIANQANIAAVLHQAFG-----FHWVWV
XP_001910908.1 ---K--IELTK-----GSN-----LSNTSSLLWHLYRSLPSPSSS-----VNWAWA
XP_446236.1 YTNFNA---TDRSEALQLLIDSYEALAMDQRNWVCN-----LANAASLLWHAYKSLN---VN-----VNWWT
NP_607473.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
YP_060405.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
YP_602707.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
NP_269471.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
YP_282476.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
YP_598792.1 MNKSK-----KIEQYQLMIAQAKELFANESN-----ALANLSNASALLNMTLPN-----SVFT
YP_099090.1 AEELTFISG-SKEEQYLSLLPQVRS LIEGEVD-----LVANLANVAAALKEAFD-----FFWVWV
NP_421082.1 EAFNDITLSADKATRYAEVADEIASVLDG EPN-----LTARMATVASMLANSFDS-----YFWAWA

YP_001034962.1 MNKTE-----KVSNYQILLAQLEALLLEGETN-----ALANLSNASALLNQALPH-----SVFT
 YP_211509.1 AEELTFISG-SKEEQYLSLLPQVRS LIEGEVD-----LVANLANVAAALKEAFD-----FFWV
 XP_001940256.1 -----MSKKEVYAQVLEQARS L-----N-----FSNTASLLWHAYHSLPSPSSS-----VNWA
 YP_001450196.1 MNTKD-----KISNYQILLAQLEALLLEGETN-----ALANLSNASALLNQALPH-----SVFT
 XP_680937.1 SSYFG---SGSKAEIYAQVIQQAQGLVTVGQRNWVRI-----LISIVSNFNSVASLLWHAYAALPAPSSS-----VNWA
 XP_001265696.1 SSYFSTSQEPTKAEVYAQVLEQAKGLVTVGQRNWATQ-----PSNFNSIASLLWHAYAALPPPSSS-----VNWA
 XP_367405.1 ASNFG--ANVTKEEAYKQVLEQAEG LFDGQRNWVCN-----LANTASLLWHAYHSLPSPSSA-----VNWA
 YP_001033719.1 MNKQE-----KIEGYELLNLQLKALLSEQNY-----TLNLANASSLLWSFLPE-----QVYT
 YP_812051.1 MNKQE-----KIEGYELLNLQLKALLSEQNY-----TLNLANASSLLWSFLPE-----QVYT
 NP_810723.1 AENLLIHTG-NKEEKYRELLPQLQALVSSSETN-----RIANLANIAAALKQTFH-----FFWV
 YP_141537.1 MTNEE-----KKIAYDLMLAQANVLFANEDN-----ALANFANASALLNTTLPN-----SVFT
 NP_721920.1 MKNQE-----KISAYQILLAQQSMFANETN-----ALANLANATAMLNKTLPN-----AIFT
 YP_820528.1 MTNEE-----KKIAYELMLAQAKVLFANEDN-----ALANFANASALLNTTLPN-----SVFT
 XP_001246405.1 -----MTCG--IWKAY-----FPSNLASVSSLLWHAYASLPAPSNA-----VNWA
 YP_612063.1 -----MRVNYEDLSKVVAALTEGETD-----EVALMATVVCCELHHADDR-----FDWT
 YP_444835.1 -PASSTATASDKAERYADTRRRIDALLDGPTD-----WVSAMATVACELHHSFDH-----YDWT
 NP_943356.1 -----MCWRRIS-----
 YP_631535.1 AEVTLDLRGMPKAEAYAEKQHVRAVLEGIDD-----DITGMATMSCLLHHAFG-----HLWT
 NP_687842.1 MNKSK-----KIENYQLLLLQAQALFSDETN-----ALANLSNASAMLNAMLPN-----SVFT
 YP_329573.1 MNKSK-----KIENYQLLLLQAQALFSDETN-----ALANLSNASAMLNAMLPN-----SVFT
 XP_001692446.1 PPLPVTASNEEREARYELVLESVKSVLEGEDD-----WVAVMATVACLLHEAFEY-----YHWT
 NP_390842.2 VEKQSG---DKEKDYQLLLKQLEAMTEDET-----QIANYANASALLYHSLPE-----VNWA
 YP_511991.1 APHSCPQRMDLPMADYATLTKTIASLTEGETD-----TITLMATVACELHHADPR-----FDWT
 XP_719233.1 YSNLKSQ-DFSKEETLQHVLDSDYKALETD--NWVAN-----LSNCSLLWHAYKSLN---IN-----VNWT
 NP_735295.1 MNKSK-----KIENYQLLLLQAQALFSDETN-----ALANLSNASAMLNAMLPN-----SVFT
 YP_168601.1 -----MTVDYPSLAKTIAALTEGETD-----PVALMATVTCVHHSDDR-----FNWT
 YP_001169098.1 -----MDYDRLDQTIDALCHGERD-----TVALMATVACELHHAHPL-----SDWT
 XP_001216799.1 SSYFG--GSGSKAEIYTQVLEQAKGLVTVGQRNWVTHNA-----PRHPSELTINSNFSNASLLWHAYAALPSPSSS-----VNWA
 YP_437091.1 DSLPLLSSEEEKSAHYQHTLRAIRHAIEGETD-----WTAILSTVVCCELHHRFEY-----FHWT
 XP_001390389.1 SSYFGA-GNGSKSEVYTQVLEQARGLVYDQRNWLTN-----HFSHRQQVSNFNSVASLLWHAYAALPAPSSS-----VNWA
 XP_001550500.1 SSTFK--ADISKAEAN--LSNAASLLWHAYKSLPS-----PSSEVNWAGFYTLDPSTP-----SKPS
 XP_001273080.1 SSYFASSGQASKADIYEQLLEQAKGLVTVGQRNWVHLP SHPTLQPPPHKPSHLTHIQPSNLSNIASLLWHAYAALPPPSSS-----VNWA
 YP_001042044.1 -----MDYDQLDQTIEALCHGEQD-----TVALMATVACEIHHAHPL-----SDWT
 YP_916587.1 -----MPDYDDLAARIQALTHDETD-----EVALMATLACEIHHS DER-----FDWT
 XP_961529.1 ASNFA--ANVTKEEAYQQVLDQAEG LFYEQRNWVCN-----LSNCAALLWHAFKSLPAPSNO-----VNWA
 YP_351544.1 -----MDYDQLDQTIEALCHGEQD-----TVALMATVACEIHHAHPL-----SDWT
 XP_001588803.1 ASTFK--ADISKEEAYAQVLEEARG LFDGQRNWIA-----IEIEIS---FYTLDPSSST-----KPS
 XP_762195.1 AAALPS-YVVTKADFYDHLESSLISLIDPGTDWISA-----LSNASSLVFNSMNRFPAWTSKR-----INWA
 XP_001537789.1 -----MVSNLANVSLLWHAYAALPAPSSA-----VNWA
 YP_001198994.1 MTKQE-----KISNYQLLLAQLEALLDGETN-----ALANLSNASALLNPALPN-----SVFT
 NP_390842.1 VEKQSG---DKEKDYQLLLKQLEAMTEDET-----QIANYANASALLYHSLPE-----VNWA
 YP_001469780.1 KSFRRG---ETMGSYTSKKQLFEE LTNQIEEIIIFQNNI-----HESMEEICNLLYENISY-----YNWV

```

NP_623333.1 -----VFERLIDEIKAKARDITDMDT-----FYKEVVRIILSENVPI-----YNWT
YP_001567589.1 MKFINSTPPNKKKADFWILYDKIKTLLISDKNYNDKQS-----ILKLISQLLAERIPY-----YNWV
NP_228329.1 IEFRKG---KEIGNYGGKKEYFDSLKEIEKILKRG-----KEALGDLCEFLRNHVSY-----YDWW
YP_001244004.1 IEFRKG---KEIGNYGGKKEYFDSLKEIEKILERG-----KEALGDLCKFLRNHVSY-----YDWW
YP_001738457.1 IEFRKG---KEIGNYGGKKEYFDSLKEIEKILKRG-----KEALGDLCEFLRNHVSY-----YDWW
YP_591728.1 -----TREQLTAQLKTMETAPTVED-----LMTDIVGRLQQYLPH-----YNWV
YP_001193798.1 -----MTFQELQPKISAIISDNNKVRDEK-----LLAVCQLLNENVEY-----YNWV
YP_001296902.1 -----MTFEALKPKVTEITTNVAFSRDEK-----LLKVCQLEANITY-----YTWW
XP_001439866.1 MNQPN----KVQYVEKVIDLLEHFLKKYQN-----LGRNAIETSIVSLLKNNVSG-----LFFV
XP_001438779.1 MNQHN----KVQYVEKVLDDLDFLKKYQN-----LGRNAIETSIVSLLKNNVTG-----LFFV
YP_001319456.1 -----KKELFLRLEKKAFALISKEEQLDN-----VYQGIVELLDDNNIPY-----YNWT
XP_001022412.1 IDQEA----NKIIEYKAFKQLQDFLEYNTKNK-----LTFEAKLTSIASVVKNLIPD-----FFFV
YP_001306613.1 VHFKN-----EKEGNFYDKDEIFEEILMELN--SVDD-----ESYMRKACEILYDKIPY-----YNWV
YP_002335308.1 IHFRNG---EMNGIYDYKDELFDKLLDKLEKESKDD--P-----VKYMRKICQYLYDEIPY-----YDWW
NP_394308.1 -----MDLKEFARSQ-----MQAACQYLKEKNPK-----YDWW
YP_862969.1 -----MTFDKLPQIEVLLKKEEDLEVKDK-----LTQVCELLEQNIPY-----YDWW
YP_023758.1 -----MAILEENLYLKIRDY-----LNDLCSYLRSENEK-----YNWV
NP_111339.1 -----MDLNDFSIVK-----LQEVCRILKESNKK-----YDWW
YP_001380982.1 -----MFGFTMDSQRAGALEALRAAAHRPYR-----EALAEAAAAILQRLPR-----YTGV
YP_806472.1 -MT-----TLDPIIVDQLDGLLTGETN-----PITVLANASALLNDLMTD-----LNWA
YP_001987415.1 -MT-----TLDPIIVDQLDGLLTGETN-----PITVLANASALLNDLMTD-----LNWA
NP_785802.1 MAT-----TETSLMNQQLDALLFQETN-----LVANLANASALLNSTYDN-----LNWA
YP_818384.1 -MS-----KVISLVGQQLDALLENETN-----IVSNLANAAALLFQNYEN-----LNWA
YP_001621174.1 -----MDYKVLLESAEGLLSTEPT-----NLSILSNASAFLNDAIPD-----LNWV

```

70 80 90 100 110 120

```

NP_288269.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002403065.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002397981.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002412850.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_001880631.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002391608.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_001730808.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
ZP_03003290.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_001743413.1 GFYLL--D-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002407226.1 GFYLL--D-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002329476.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_002382399.1 GFYLL--E-G-----ETLVLCPEFGK-IACVRIP--VGRVCGTAVARN--EVQRIEDVHAFDGHIACDAAASNAEIV
ZP_03445135.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_688895.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
YP_669680.1 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV
NP_416346.4 GFYLL--E-D-----DTLVLCPEFGK-IACVRIP--VGRVCGTAVARN--QVQRIEDVHAFDGHIACDAAASNSEIV

```

ZP_02903917.1 GFYLL--E-D-----DTLVLPFFQK-IACVRIP--VGRGCGTAVARN--EVQRIEDVHAFDGHIAFDAASNSEIV
ZP_03282954.1 GFYLL--E-D-----NTLVLPFFQK-IACVRIP--VGRGCGTAVATN--QVQRVEDVHAFDGHIAFDAASNSEIV
YP_001452718.1 GFYLL--E-N-----DTLVLPFFQK-LACVRIP--VGRGCGTAVAQN--QVQRIEDVHAFDGHIAFDSASNAEIV
YP_001570142.1 GFYLL--E-G-----DTLVLPFFQK-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
NP_460803.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_150310.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002141797.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002243298.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
NP_456352.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
NP_804849.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_001587603.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02666379.2 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02832604.2 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_001177122.1 GFYLL--E-G-----NTLVLPFFQK-LACVRIP--VGRGCGTAVAQN--QVQRVEDVHAFDGHIAFDSASNAEIV
YP_002226291.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASSAEIV
NP_754135.1 GFYLL--E-D-----DTLVLPFFQK-IACVRIP--VGRGCGTAVARN--QVQRIEDVHAFDGHIAFDAASNSEIV
ZP_02347214.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02574893.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02659310.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02663512.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02681881.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_02696584.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002041101.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002045892.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03075263.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002114879.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002146178.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03165709.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002215294.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03221812.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03342093.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03347154.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03352578.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03357060.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03371453.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03379701.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03383790.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03412299.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
ZP_03217260.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIAFDAASNAEIV
YP_002237781.1 GFYLL--E-G-----DTLVLPFFQK-LACVRIP--VGRGCGAAVAQA--QVQRVEDVHAFDGHIAFDAASNSEIV
YP_001336004.1 GFYLL--E-G-----DTLVLPFFQK-LACVRIP--VGRGCGAAVAQA--QVQRVEDVHAFDGHIAFDAASNSEIV
YP_001437512.1 GFYLL--E-D-----KTLVLPFFQK-IACVRIP--YGRGCGTAVATG--EVQRVDVHAFDGHIAFDASSNAEIV

ZP_03403067.1 GFYLL--E-G-----DTLVLPFFQGR-IACVRIP--VGRGCGAAVAQN--KVQRIDVHAFDGHIACD AASNAEIV
YP_001907453.1 GFYLLT-E-E-----KTLVLGPFQGR-IACVRIP--VGRGCGTAVSEN--RVQRVDVHAFPGHIACD AASNAEIV
YP_001006127.1 GFYLL--N-G-----NQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGDVHAFPGHIACD AASNAEIV
NP_669181.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
NP_993079.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_070893.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_651729.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_647856.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001163103.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_01888376.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001400647.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001607069.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02221459.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02227087.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02229339.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02237158.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02307206.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02314388.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
ZP_02318544.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001720525.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001872871.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_002346706.1 GFYLV--D-G-----KQLVLGPFQGR-IACVRIP--VKGKCGTAVAEN--RVQRVGNVHAFPGHIACD AASNAEIV
YP_001478346.1 GFYLM--D-G-----TQLVLGPFQGR-IACVRIP--VKGKCGSAVAEN--RVQRVGDVHEFPGHIACD AASNAEIV
YP_454994.1 GFYLL--D-G-----DILVLGPFQGR-VACVRIP--LKGKCGSAAAEN--RIQRIGDVHAFQGHIACD AASNAEIV
YP_001217052.1 GFYLM--QEN-----ELVLGPFQGR-PACVRIP--VGRGCGTAVAEN--KVQRVYDVHQFEGHIACD AASNSEIV
YP_002150764.1 GFYLN--D-G-----NSLVLGPFQGR-VACVRIG--FKGKCGTAFSEN--KILRVNDVHQFSGHIACD SASQSEIV
YP_856816.1 GFYLL--Q-G-----ETLVLGPFQGR-PACVRIP--VGRGCGTAVAEG--KTQLIDVHQFEGHIACD GASNAEIV
YP_001093925.1 GFYML--Q-G-----EQLVLGPFQGR-VACTRIP--MKGKCGTAAQSN--TTQRIADVHAFDGHIACD AASNSEIV
YP_269525.1 GFYRL--I-D-----GELVLGPFQGR-VACIRIP--VKGKCGTCSFTG--EVQRISDVHQFDGHIACD ASSNAELV
NP_718189.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--MKGKCGTAAALAN--QTQRVADVHQFDGHIACD AASNSEIV
YP_001366642.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--LKGKCGTAAFTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_001554990.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--LKGKCGTAAFTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_737796.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--MKGKCGTAAQTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_869415.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--MKGKCGTAAQTN--QTQRVADVHQFDGHIACD SASNSEIV
NP_797987.1 GFYLT--KGD-----ELVLGPFQGR-PACVRIP--MGRGCGTAAKTN--TTQRVYDVHEFEGHIACD AASNSEIV
YP_001501790.1 GFYLL--K-N-----DELVLGPFQGR-VACSRIP--LKGKCGTAVATN--STQRVDVHQFAGHIACD SASNSEIV
YP_927460.1 GFYIR--R-E-----DTLVLGPFQGR-VACTRIP--WKGKCGTAAATD--TTQRVADVHQFDGHIACD SASNSEIV
YP_001474222.1 GFYLL--K-N-----KQLVLGPFQGR-VACTRIP--LKGKCGTAASEG--TTQRIADVHQFDGHIACD AASNSEIV
YP_963195.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--LKGKCGTAAAYTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_001183724.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--LKGKCGTAAAYTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_733804.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--MKGKCGTAAQTN--QTQRVADVHQFDGHIACD SASNSEIV
YP_001050811.1 GFYVM--R-G-----EQLVLGPFQGR-VACTRIP--LKGKCGTAAFTN--QTQRVADVHQFDGHIACD SASNSEIV

YP_750914.1 GFYIK--R-G-----EQLVLGPFQ GK-VACTRIN--MCKG CGTAAAEN--KTQRVADVHQFDGHIACDSASNSEIV
YP_002126297.1 GFYLY--K-E-----EQLVLGPFQ GQ-PACIRIP--LCKG CGTSASTR--TVQRIADVHQFEGHIACDAASNSEIV
YP_130129.1 GFYLL--K-G-----DQLVLGPFQ GK-VACVRIP--VGRG CGTAISEN--KVQRVSDVHAFPGHIACDAVSNSEIV
YP_001445537.1 GFYLT--KGD-----ELVLGPFQ GK-PACVRIP--MGRG CGTAAQTN--SVQRVYDVHFEFEGHIACDAASNSEIV
YP_661562.1 GFYLL--K-D-----NELVLGPFQ GK-PACIRIP--MGRG CGKAAETG--LSQCIHDVHAFDGHIAADADSNAEAV
YP_001674584.1 GFYLL--K-N-----DQLVLGPFQ GK-LACSRIP--VCKG CGTAVALN--STQRIGDVHQFAGHIACDSASNSEIV
YP_002417027.1 GFYLM--KQDQAKEK-----DELVLGPFQ GQ-PACVRIP--VGRG CGTAVATN--TVQRIHDVHFEFEGHIACDAASNSEIV
NP_934438.1 GFYLY--KED-----ELVLGPFQ GK-VACVRIP--MCKG CGTAAQTY--TVQRVHDVHFEFEGHIACDAASNSEIV
NP_761466.1 GFYLY--KED-----ELVLGPFQ GK-VACVRIP--MCKG CGTAAQTY--TVQRVHDVHFEFEGHIACDAASNSEIV
YP_001760510.1 GFYLL--K-D-----EQLVLGPFQ GK-VACSRIP--LCKG CGTSAAEK--KTMRVDVHQFDGHIACDSASNSEIV
YP_002312133.1 GFYLL--K-N-----EQLVLGPFQ GK-VACTRIP--ICKG CGTAVEKN--ETQRIADVHQFVGHIAADSAVSEIV
YP_562727.1 GFYIK--R-E-----DQLVLGPFQ GK-VACTRIQ--LCKG CGNAAAQN--ITLRVADVHQFEGHIACDSASNSEIV
NP_693123.1 GFYLV--EEN-----ELILGPFQ GL-PACIRIP--YCKG CGTAVSEQ--QTQRVDVHAFPGHIACDGATNSEIV
YP_001699747.1 GFYLL--QEN-----ELVLGPFQ GL-PACVRIP--IGRG CGTAVATK--ETIVVKDVHFEFPGHIACDAASQSEIV
YP_001343242.1 GFYFARG-----EELVLGPFQ VQK-VACTRIP--LGRG CGKAAITL--QTQRIEDVHAFDGHIAADAAASEVIV
YP_001141806.1 GFYLL--Q-G-----DTLVLGPFQ GK-PACVRIP--VGRG CGTAVAEG--KTQLIDVHQFEGHIACDGASNAEIV
YP_204661.1 GFYLF--KEN-----ELVLGPFQ GK-PACIRIS--VCKG CGTAFATK--TVQRIDVHQFAGHIACDAESNSEIV
YP_301135.1 GFYLI--ENE-----ALILGPFQ GH-PACVHIA--ICKG CGTAVSEE--QTQLVDVNAFPGHIACDANSKSEIV
YP_002156075.1 GFYLF--KEN-----ELVLGPFQ GK-PACIRIS--VCKG CGTAFATE--TVQRIDVHQFAGHIACDAESNSEIV
NP_244074.1 GFYLI--KEG-----ELVLGPFQ GL-PACVRIP--IGRG CGTAAKEE--QTVRVEDVHQFPGHIACDAASRSEIV
YP_001559126.1 GFYLMKKGS-----LILGPFQ GK-VACIHIP--VGRG CGTAVNEG--KTQIVKDVHFEFPGHIACDSASNSEIV
YP_080250.1 GFYFAK--D-----GELVLGPFQ GL-PACVRIP--FKG CGTAYTNG--RVERVEDVNAFPGHIACDAASQSEIV
YP_253122.1 GFYLI--ENE-----ELILGPFQ GH-PACVHIA--ICKG CGTAVAED--KTQLVKDVHAFPGHIACDANSKSEIV
YP_261333.1 GFYLNRN-----QELVLGPFQ GQ-IACVRIP--FGRG CGTAAATL--QTQRVEDVHAFAGHIACDSASNSEIV
YP_001187407.1 GFYLVKD-----GELVLGPFQ GK-VACVRIA--FGRG CGAAAASL--QTQRVEDVHAFAGHIACDSASNSEIV
YP_001422246.1 GFYFAK--E-----GELVLGPFQ GL-PACVRIP--FGRG CGTAYSNG--DVQRIADVNAFPGHIACDAASQSEIV
YP_001748360.1 GFYLNRN-----EELVLGPFQ GQ-VACVRIP--FSRG CGAAAATR--QTQRVEDVHAFAGHIACDSASNSEIV
YP_155660.1 GFYLTRSD-----NELVLGPFQ GR-VACVRIP--FKG CGVAAETQ--TCQRVEDVHQFAGHIACDSASNSEIV
NP_744032.1 GFYLNRN-----EELVLGPFQ GQ-VACVRIP--FSKG CGAAAATR--QTQRVEDVHAFPGHIACDSASNSEIV
YP_001269146.1 GFYLNRN-----EELVLGPFQ GQ-VACVRIP--FSKG CGAAAATR--QTQRVEDVHAFPGHIACDSASNSEIV
YP_236680.1 GFYLNRN-----EELVLGPFQ GQ-IACVRIP--FGRG CGAAAQSR--QTQRVQDVHFEFPGHIACDSASNSEIV
YP_001667695.1 GFYLNRN-----EELVLGPFQ GQ-VACVRIP--FKG CGAAAATR--QTQRVEDVHAFPGHIACDSASNSEIV
YP_188854.1 GFYLI--ENN-----ELILGPFQ GH-PACVHIA--ICKG CGTAVSSG--ETQRVKDVHQFPGHIACDANSQSEIV
YP_002275225.1 GFYILRDGE-----LVLPFQ GR-VACTRIA--VGRG CGTAAARR--STIVVPDVHAFPGHIACDTASESEIV
YP_804779.1 GFYLFDSKKD-----ELVLGPFQ GN-VACMHIK--NGSG CGTAFKEN--KVLRVENVHFEFPGHIACDSASNSEIV
NP_470973.1 GFYLLEKETN-----QLVLGPFQ GL-PACIRIP--LCKG CGSAASDQ--KTYIVENVHDFPGHIACDAASNSEIV
YP_607257.1 GFYINRS-----EELVLGPFQ GQ-VACVRIP--FSRG CGAAAATR--TTQRVEDVHAFPGHIACDSASNSEIV
YP_001860169.1 GFYFALD-----GELVLGPFQ GK-PACVRIP--MGRG CGRAAQAL--ETQVVPDVDAFPGHIACDSASRSEIV
NP_465120.1 GFYLLEKETN-----QLVLGPFQ GL-PACIRIP--LCKG CGSAASDQ--KTYIVENVHDFPGHIACDAASNSEIV
NP_372242.1 GFYLL--EQN-----ELILGPFQ GH-PACVHIP--ICKG CGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
NP_374828.1 GFYLL--EQN-----ELILGPFQ GH-PACVHIP--ICKG CGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_001247140.1 GFYLL--EQN-----ELILGPFQ GH-PACVHIP--ICKG CGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV

YP_001316940.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFEGHIACDANSKSEIV
YP_001442294.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFEGHIACDANSKSEIV
YP_002314836.1 GFYIA--EGN-----ELVLGPFQGL-PACVRIP--FGKVCGTAAAEK--RTVRVADVHAFPGHIACDAASQSEIV
YP_001814701.1 GFYLLMDQEQN-----ELVLGPFQGL-PACVRIP--LGRVCGTAAQDQ--TTQLVMDVHAFPGHIACDAATNSEIV
NP_646478.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFEGHIACDANSKSEIV
YP_043764.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFEGHIACDANSKSEIV
YP_186602.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_494359.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_500333.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001332646.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001575592.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
ZP_02759966.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001487827.1 GFYLAKEG-----DELVLGPFQGL-PACVRIP--SGKVCGTAFATG--EVQRIADVHAFPGHIACDAASQSEIV
YP_417047.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFEGHIACDANSKSEIV
YP_041183.1 GFYLL--EQN-----ELILGPFQGH-PACVHIP--IGKVCGTAVSER--RTQVVDVHQFKGHIACDANSKSEIV
YP_176251.1 GFYFL--KEN-----ELVLGPFQGL-PACTRIA--IGKVCGTAVAEN--KTMRIDVHAFPGHIACDAASNAELV
YP_002349929.1 GFYLLLEKETN-----QLVLGPFQGL-PACIRIP--LQKVCGSAAADQ--KTYIVENVHDFPGHIACDAASNSEIV
YP_001376537.1 GFYVT--EGN-----ELVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_784647.1 GFYFFDG-----QELVLGPFQGK-PACVRIA--LARGVCGAAASLR--QTQVVPDVHAFPGHIACDAASRSEIV
YP_557760.1 GFYFHDG-----QELVVGPFQGK-PACVRIA--LQKVCGTAAQTR--QTQVVRDVHEFPGHIACDSASQSEIV
YP_275769.1 GFYLNLRN-----EELVLGPFQGQ-IACVRIP--FGRVCGAAAQSR--QTQRVQDVHEFPGHIACDSASNSELV
YP_001118914.1 GFYFFDG-----TELVVGPFQGK-PACVRIA--LQKVCGTAAQTR--QTQVVRDVHAFPGHIACDAASESEIV
YP_001782545.1 GFYLMKDG-----ELVLGPFQGM-PACDRIK--IGKVCGVAISSR--QIQRVDDVHKFEGHIACDSSTNSELV
NP_882192.1 GFYFHDG-----QELVLGPFQGK-PACVRIA--LSRVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
YP_001894915.1 GFYFHDG-----KELVVGPFQGK-PACVRIP--LQKVCGTAAQTR--QTQVVRDVREFPGHIACDSASQSEIV
YP_001647268.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_882465.1 GFYFHDG-----QELVLGPFQGK-PACVRIA--LSRVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
YP_002448219.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_834359.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_002369448.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_191278.1 GFYLVKEADR-----QLVLGPFQGR-LACTRIP--LGRVCGTVAQNR--ETLVVPDVHAFPGHIACDAASESEIV
NP_886654.1 GFYFHDG-----QELVLGPFQGK-PACVRIA--LSRVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
NP_847103.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_030799.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_981083.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_038702.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_002340711.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_085975.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--ATQLVADVHQFPGHIACDSASNSEIV
YP_002453717.1 GFYVT--EGN-----QLVLGPFQGM-PACVRIP--FGRVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_772921.1 GFYFFDG-----SELVVGPFQGK-PACVRIA--LQKVCGTAAQTR--ETQVVRDVHEFPGHIACDAASESEIV
YP_001843726.1 GFYRVQDG-----ELILGPFQGQ-VACVHIA--FGSVCGTAAQTG--KTQLVKDVHQFAGHIACDAASNSEIV
YP_942902.1 GFYLL--KDN-----QLVLGPFQGQ-PACIRIP--LNTGCVGAFSTN--TIQRVADVQQFVGHIACDANSQSEIV

YP_121651.1 GFYFAVG-----DELVVCPFFQK-PACVRIA--FGKGCCTAAATR--TTQLVVDVHAFPGHIACDADSRSEIV
YP_001633478.1 GFYFYDG-----RELVLCPFFQK-PACVRIA--LGRGCGTAASQR--QTQVVDVNAFPGHIACDVASRSEIV
YP_001392189.1 GFYLMKDG-----ELVLCPFFQGM-PACDRIK--IGKCGVAISSR--QIQRVDDVHKFEGHIACDSSSTNSEIV
YP_795394.1 GFYLYQPETD-----DLILCPFFQGN-VACVHIE--NNHCGTALATQ--TTQRIADVHQFPGHIACDASASNAEIV
YP_002284153.1 GFYFLRSE-----NELVLCPFFQGR-VACVRIA--VKGCGTAVAEA--RSVLVVDVHEFPGHIACDAAASRSEIV
YP_001255404.1 GFYLMKDG-----ELVLCPFFQGM-PACNRIK--IGKCGVAISSR--QIQRVNDVHKFEGHIACDSSSTNSEIV
YP_001385170.1 GFYLMKDG-----ELVLCPFFQGM-PACNRIK--IGKCGVAISSR--QIQRVNDVHKFEGHIACDSSSTNSEIV
YP_001388639.1 GFYLMKDG-----ELVLCPFFQGM-PACNRIK--IGKCGVAISSR--QIQRVNDVHKFEGHIACDSSSTNSEIV
YP_001347703.1 GFYLNRN-----EELLCPFFQK-VACVRIP--FAKCGAAARTR--QTQRVEDVHAFPGHIACDASARSEIV
NP_103622.1 GFYFLASD-----EELVLCPFFQK-PACVRIA--VKGCGTAVQLG--TSMLIKDVHDFPGHIACDADSRSEIV
YP_001807739.1 GFYFFDG-----SELVVCPPQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHEFPGHIACDAAESEIV
YP_001580337.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHAFPGHIACDAAESEIV
YP_002230176.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHDFPGHIACDAAESEIV
NP_251512.1 GFYLNRN-----EELVLCPFFQK-VACVRIP--FSKCGAAARTR--LTQRVDVHAFPGHIACDSSASSEIV
YP_459505.1 GFYRVVD-----GELVLCPFFQGR-PACIRIP--IGQCGAAASSG--ETQLVEDVHAFPGHIACDAAASQSEIV
YP_790351.1 GFYLNRN-----EELVLCPFFQK-VACVRIP--FSKCGAAARTR--LTQRVDVHAFPGHIACDSSASSEIV
YP_002439846.1 GFYLNRN-----EELVLCPFFQK-VACVRIP--FSKCGAAARTR--LTQRVDVHAFPGHIACDSSASSEIV
YP_001970489.1 GFYLYDG-----KELVVCPFFQL-PACVRIP--LDKCGAAASQR--VTQRVEDVDAFPGHIACDASARSEIV
YP_617023.1 GFYRFDG-----QELVLCPFFQK-AACIRIP--LDKCGAAARLR--ATQRVEDVHAFPGHIACDADSRSEIV
YP_620554.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHDFPGHIACDAAESEIV
YP_834795.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHDFPGHIACDAAESEIV
YP_148653.1 GFYLA--DGE-----ELVLCPFFQL-PACVRIP--FGKCGTAAAKR--RTVVVDVHQFPGHIACDAAASQSEIV
YP_442277.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASARSEIV
YP_334263.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_001067073.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_368503.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHDFPGHIACDAAESEIV
YP_109010.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_993758.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_001028775.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_002026852.1 GFYLYDG-----KELVVCPFFQL-PACVRIP--LDKCGAAASQR--VTQRVDVDAFPGHIACDASARSEIV
YP_001764423.1 GFYFFDG-----TELVVCPFFQK-PACVRIA--LKGCGTAAQTR--ETQVVDVHDFPGHIACDAAESEIV
NP_970249.1 GFYLLHNN-----ELLLSSPQL-PACTRIA--IGKCGTAAKTL--QTQLVADVDFPGHIVCDAAASKSEIV
YP_001059788.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTAAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_001081316.1 GFYFFDG-----RELVVCPFFQK-PACVRIP--LKGCGTTAQTR--GTQVVHDVHAFAGHIACDASASQSEIV
YP_001875000.1 GFYILKG-D-----KLVLCPPQK-PACVRIS--LKGCGVSARER--KTLVVDVHEFPGHIACDASASQSEIV
YP_278248.1 GFYLAEGE-----VLYLSSPQK-IACIQIP--FSRGGKAAAREQ--KTIVVDVHEFKDGHIACDASANSSEIV
YP_001309672.1 GFYLVKNN-----TLVLCPPQGM-PACTKIE--IGKCGKAAALEK--ETLLVVDVHNFEGHIACDAAANSSEIV
YP_001576818.1 GVYRFENG-----ELILCPFFQGM-PACVHIQ--MKGCGTAAQTE--KTQIVPNVHEFQGHIACDASANSSEIV
YP_193206.1 GVYRLENG-----ELILCPFFQGM-PACVHIP--VKGCGVAAKTQ--ITQIVPNVHEFAGHIACDASANSSEIV
YP_199405.1 GFYFYDG-----RELVVCPFFQL-PACVRIP--LDKCGAAASTR--QSQRIADVAFPGHIACDASARSEIV
YP_001915608.1 GFYFYDG-----RELVVCPFFQL-PACVRIP--LDKCGAAASTR--QSQRIADVAFPGHIACDASARSEIV
YP_002509654.1 GFYLMR-G-----GELVLCPFFQL-PACVRIK--VGRGCGSAVRDR--KTYLVDVHKFPGHIACDASARSEIV

YP_757341.1 GFYEVDPKAP-----TELVVCPYFGT-MGCLRIP--FGRGCGAAATG--ETQLVEDVHAFPGHIACD SRSNSEIV
YP_810624.1 TYYRLKNG-----ELILGPFQ GK-PACVHIP--LGRGCGTAAALK--KTVIVSNVHKFE GHIACDKDSNSEIV
YP_244708.1 GFYFYDG-----RELVVCPFQGL-PACVRIP--LDKGCGAAASTR--QTQRIADVDAFPGHIACD SASRSELV
NP_635978.1 GFYFYDG-----RELVVCPFQGL-PACVRIP--LDKGCGAAASTR--QTQRIADVDAFPGHIACD SASRSELV
XP_505859.1 GFYVIDPKKSE-----QLILGPFHQV-VACQEIP--LKGCGVAAKEQ--KTQLVKNVEEFP GHIACDGVTKSEIV
NP_964546.1 GVYRYENQ-----ELILGPFQ GK-PACVHIQ--IGKCGTAAKEK--QTQIVKNVHEFAGHIACD SESNSEIV
YP_001261873.1 GFYRAIG-----QELVLGPFQ GK-TACIRIP--FKGCGTAAADTR--TAQCVDVHEFP GHIACDAYSASEIV
YP_814311.1 GVYRYKNH-----ELILGPFQ GK-PACVHIA--IGKCGTTAKER--QTQIVKNVHEFAGHIACD SDSNSEIV
NP_298520.1 GFYFYDG-----RELVVCPFQGL-PACIRIP--LDKGCGTAAARTR--RTQRIDVDVTFPDHIACD SASRSELV
YP_001775202.1 GFYFYDG-----RELVVCPFQGL-PACIRIP--LDKGCGTAAARTQ--RTQRIDVDVTFPDHIACD SASRSELV
YP_001829224.1 GFYFYDG-----RELVVCPFQGL-PACIRIP--LDKGCGAAARTQ--RTQRIDVDVTFPDHIACD SASRSELV
NP_778732.1 GFYFYDG-----RELVVCPFQGL-PACIRIP--LDKGCGAAARTQ--RTQRIDVDVTFPDHIACD SASRSELV
NP_246070.1 GFYLVKDG-----VLKVGPFQ GK-VACSDIG--FKGCGYTWQTG--TTTVVDVHQFAGHIACD SASQSEV
XP_645637.1 GFYLVDTEN-----ELVLAPFQ GP-IACRIR--KGRGCGTAAWQQE--KTLIVPDVEKFP GHIACDSSLSKSEIV
XP_456263.1 GFYLTRENKK-----ELILGPFQ GK-VACQLIQ--FKGCGTAAASSQ--QTQLVPDVENFP GHIACDGETKSEIV
YP_001300883.1 GFYLVKDD-----QLILGPFQ GS-TACYRIH--KKGCGTAWAEA--RTLIVPNVEQFP GHIACDSSLSRSEIV
YP_812434.1 GVYRFKKG-----ELILGPFQ GM-PACVHIQ--LKGCGTAAALEK--KTQIVANVHDFP GHIACD SASNSEIV
NP_814796.1 GYVLYK-----DEELVLGPFQ GK-VSCTRIK--MKGCGESAQEK--ATLIVDNVKTHANYISCD SAAMSEIV
NP_358362.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
YP_816246.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
YP_001694316.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
YP_001835526.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
YP_002037500.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
NP_345351.1 GFYLFDF-----GKELVLGPFQ GK-VSCTIRIA--LKGCGEAAHFQ--ETVIVGDVTTYLNYISCD SLAKSEIV
XP_001385863.2 GFYVKNGEKNE-----LILGPFQ GK-VACQIIQ--IGKCGTAAHASK--ETQLVPDVEKFP GHIACDGETKSEIV
NP_985662.1 GFYITEKDDPN-----MLLILGPFQ GK-VACQRIQ--FGRGCGTAAALTQ--EVQVVPDVTVYP GHIACDGDTKSEMV
YP_759928.1 GFYLVDPKLP-----QELVVCPYFGT-LGCLRIP--FKGCGHVAATQ--APVIVPDVHEFP GHIACD SASNSEIV
YP_002123053.1 GFYLFDF-----GQELILGPFQ GK-VSCTVHIK--LKGCGEAAQSG--QTIIVDVVRKHANYISCD SAAMSEIV
YP_002131343.1 GFYVVDPAKG-----DELVVCPYFGT-LGCLRIA--FGRGCGAAATR--QTQLVEDVHAFPGHIACD SRSASEIV
XP_001683506.1 GFYLFQAPG-----LLALGPFQ GR-PACTEIR--VKGCGTVAESG--ESMVVQSVHEFP GHIACD SASKSEIA
NP_012854.1 GFYVTQASEEN-----TLILGPFQ GK-VACQMIQ--FKGCGTAAASTK--ETQIVPDVNKYP GHIACDGETKSEIV
XP_460154.1 GFYVLNEPRTD-----ELILGPFQ GK-VACQIIK--FKGCGNAASSA--KTQLVPNVNEYPGHIACD GETQSEIV
XP_001801028.1 GFYFTDPSNP---AK-----LILGPFQ GQ-VACQTIQ--FNRCGTAATQ--TTQLVADVDFP GHIACDGASKSEIV
YP_001201197.1 GFYLYD-----GSELILGPFQ GK-VSCTVHIA--LKGCGEVAAKR--QTILVDVRLHDNYISCD ATALSEIV
XP_001484046.1 GFYVINPEKQN-----ELILGPFQ GK-VACQTIQ--FKGCGNAASSK--NTQLVKNVNEYPGHIACD GDTQSEIV
NP_274882.1 GFYLVDTRS-----DELVLAPFQ GP-LACTRIP--FGRGCGQAWAKG--ETVVVDVNAHPDHIACD SSSLRSEIV
YP_001302571.1 GFYRVIDN-----QLVLAPFQ GP-IACRIR--YKGCCTAAWKET--RTIIVPDVDAFPGHIACD SSSARSEIV
YP_002000650.1 GFYLVDTRS-----DELVLAPFQ GP-LACTRIP--FGRGCGQAWAKG--ETVVVDVNAHPDHIACD SSSLRSEIV
YP_001682978.1 GFYVVDPLKD-----RELVVCPYFGT-LGCLRIA--YGRGCGAAAETG--QTQLVEDVHAFPGHIACD SRSQSEIV
YP_002342048.1 GFYLVDTRS-----DELVLAPFQ GP-LACTRIP--FGRGCGQAWAKG--GTVVVDVDAHPDHIACD SSSLRSEIV
YP_207199.1 GFYFVDTRS-----DELVLAPFQ GP-LACTRIP--FGRGCGQAWAKG--ETVVVDVNAHPDHIACD SSSLRSEIV
XP_844799.1 GFYLMHGPE-----LLALGPFQ GR-PACTEIK--MGRGCGTAAQQA--KTLVVDVHEFP GHIACD SASNSEIV

YP_001598496.1 GFYLVDRS-----DELVLAPFGP-LACTRIP--FGRGCGQAWAKG--GTVVVGVDADHPDHIACSSLSRSEIV
YP_974447.1 GFYLVDRS-----DELVLAPFGP-LACTRIP--FGRGCGQAWAKG--GTVVVGVDADHPDHIACSSLSRSEIV
XP_001465858.1 GFYLFQAPG-----LLALGPFQGR-PACTEIR--VKGKCGTVAESG--EPLVVQSVHEFPGHIACDSASKSEIA
YP_001280815.1 GFYRVDSAA-----NQLILGPFQGP-LACTRIP--HGKGCGEVWATG--TTQIVDVAHAHPNHIACSSLSQSEIV
NP_664843.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
NP_802083.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_280574.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_001128311.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_002286076.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
XP_390238.1 GFYVLDKSSKD--PQ-----LILGPFQGK-VACQTIK--FKGKCGTAAATH--ETQLVVDVEEFPGHIACDGDSKSEIV
XP_001565241.1 GFYLFQAPG-----LLALGPFQGR-PACTEIR--VKGKCGTVAESG--ESLVVSNVEEFPGHIACDSSSKSEVA
YP_596891.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMTEIV
YP_600772.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMTEIV
YP_001351938.1 GFYRVQK-----DELVLGPFQGP-LACTRIA--KKGKCGSAWQQA--ATIVVVDVDQFPGHIACSSLSRSEIV
XP_001910908.1 GFYVHDHTSPPTKPR-----LILGPFQGK-VACQTIK--FGRGCGTAAATQ--TTQLVVDVEQFPGHIACDSSSKSEIV
XP_446236.1 GFYIRNGE-KE-----QLLILGPFQGK-VACQTIK--FGRGCGIAASSQ--ETQLVVDVDKFPGHIACDGETKSEIV
NP_607473.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_060405.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_602707.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
NP_269471.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_282476.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_598792.1 GFYLFDF-----GQELILGPFQGR-VSCVHIK--LKGKCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_099090.1 GFYLVKQD-----QLVLGPFQGP-VACTRIR--KKGKCGTAWQES--ATLLVVDVEVFPGHIACSSLSRSEIV
NP_421082.1 GFYVVDPTKE-----RELVVCPYCGT-LGGLRIA--FGRGCGAAAATG--QTQLVVDVHAFPGHIACDGRSQSEIV
YP_001034962.1 GFYLFDF-----GKELILGPFQGG-VSCVHIT--LKGKCGESAQSR--QTIIVDVTQHANYISCDRAKSEIV
YP_211509.1 GFYLVKQD-----QLVLGPFQGP-VACTRIR--KKGKCGTAWQEG--ATLLVVDVEVFPGHIACSSLSRSEIV
XP_001940256.1 GFYFTDPAKE---TR-----LLLGPFQGQ-VACQSIK--FGRGCGTAAKEA--KTQLVVDVDKFPGHIACDGKSRSEIV
YP_001450196.1 GFYLFDF-----GSELILGPFQGG-VSCVHIT--LKGKCGESAQSR--QTIIVDVTQHANYISCDRAKSEIV
XP_680937.1 GFYIRQDQFPSKETAESTEAAARVKE---TLWLGPFGP-PACQEIR--FGRGCGAAAAR--ETVVVVDVLEFPGHIACDASSRSEIV
XP_001265696.1 GFYIRQDKFP-ALGSQNTSPSNSTNN---LLLGPFGP-PACQEIR--FGRGCGTAAEK--ETVIVVDVLSFPGHIACDASSRSEIV
XP_367405.1 GFYTLDPSSATR-QQ-----LILGPFQGK-VACQTIK--FGRGCGAAAQK--QTQLVVDVDKFPGHIACDADSRSEIV
YP_001033719.1 GFYLYN-----GDKLILGPFQGS-VSCVEIV--MKGKCGEAAQTR--QTMIVEDVKKHKNYISCDGHAMSEIV
YP_812051.1 GFYLYN-----GDKLILGPFQGS-VSCVEIV--MKGKCGEAAQTR--QTMIVEDVKKHKNYISCDGRAMSEIV
NP_810723.1 GFYLVVEGN-----ELVLAPFGP-IACTRIR--FGRGCGTAWKEA--QTLIVVDVEQFPGHIACSSDSKSEIV
YP_141537.1 GFYLMNDI-----KNELILGPFQGN-VSCVRIA--LKGKCGQSAAN--RTLIVQDVTKHANYIACDSAARSEIV
NP_721920.1 GFYLFDF-----GEELVLGPFQGG-VSCVRIA--LKGKCGQAAAK--QTLIVQDVTKHNDNYISCDSAAMSEIV
YP_820528.1 GFYLMNDI-----KNELILGPFQGN-VSCVRIA--LKGKCGQSAAN--RTLIVQDVTKHANYIACDSAARSEIV
XP_001246405.1 GFYVRKDKSPSSGVSTASDAKN-----ILLGPFQHGK-PACQLIP--FGRGCGTAAASKR--ETVIVSDVLEWEGHIACDADSRSEIV
YP_612063.1 GFYRVTEP-----EVLKICPYCG-HGCLVIP--FARGCGAAAARLN--ETQLVVDVDAFPGHIACSSSTRSEIV
YP_444835.1 GFYRAVS--D-----DELLVCPYCG-HGCLHID--VDRGCGAAAARTR--ETQLWVDVSKAPDHIACQSSTQSEIV
NP_943356.1 -----LFLDHSRGN-IACARIP--VGSKCGTAAVAEE--RILCINNVSFNHGHIACDTASNSEIV
YP_631535.1 GFYRVVTPG-----RLLRVCPYCGT-LGGLIIP--FKGKCGASAAK--ESVVVVDVHAFPGHIACDGRSASEIV

NP_687842.1 GFYLFDF-----GEELILGPFQGG-VSCVHIT--LKGKCGESAQTA--KTLIVD DVTKHANYISCD SKAMSEIV
 YP_329573.1 GFYLFDF-----GEELILGPFQGG-VSCVHIT--LKGKCGESAQTA--KTLIVD DVTKHANYISCD SKAMSEIV
 XP_001692446.1 GFYQAGPPGD-----DTLVLGPFQGH-MGCLRIP--YSRGCGAAARTK--TSQLVP DVSQFPGYTACASSTRSEV
 NP_390842.2 GFYFAK-EED-----GQLVLGPFQGL-PACVRIP--FGRGCGTAYANG--KVERIEDVNAFPGHIACDAASQSEIV
 YP_511991.1 GFYRVTEP-----DLLKIGPFQGG-HGCLQIP--FSKCGAAARTG--QTQLVP DVDAFPGHIACASSTRSEIV
 XP_719233.1 GFYLTVENKNEGQTSPT-----ELILGPFQGK-VACQLIK--FCHGCGTAAASKK--LTQLVP DVEKFPGHIACDGETKSEIV
 NP_735295.1 GFYLFDF-----GKELILGPFQGG-VSCVHIT--LKGKCGESAQTA--KTLIVD DVTKHANYISCD SKAMSEIV
 YP_168601.1 GFYRVTEP-----GVLKIGPFQGG-HGCLRIP--FDKCGAAASTG--EVQLVP DVDAFPGHIACASSTRSELV
 YP_001169098.1 GFYRVTEP-----EVLKIGPFQGS-HGCLVIP--FARGCGAAARTG--EIQLVP DVEAFPDHIACSSSTRSEIV
 XP_001216799.1 GFYIRQDKFP-VLSSTSQDTPKKDD---ILLGPFQGR-PACQEIR--FGRGCGTAAQKR--ETVVVP DVLNFPGHIACDASSRSEIV
 YP_437091.1 GFYRVVAH-----ELMKVGPYQGG-HGCLTIP--FSRCGAAAREQ--KVQLVSDVNEFPGHIACSSTTQSEIV
 XP_001390389.1 GFYIRQDKFP-VLGAPLEEDSKKKP---VLLGPFQGR-PACQEIR--FGRGCGTAAEKR--ETVVVP DVLNFPGHIACDASSRSEIV
 XP_001550500.1 -----LILGPFQGK-VACQTIA--FSRCGTAASTQ--TTQLVADVEKFPGHIACDAASQSEIV
 XP_001273080.1 GFYIRQDKFPQPASTTSSSTTEPPENLLLLLILGPFQGR-PACQEIR--FGRGCGAAAEQR--QTVLVP DVHAFPGHIACDASSRSEIV
 YP_001042044.1 GFYRVTGP-----ELLKIGPYGA-HGCLVIP--FSRCGAAARTG--RIQLVP DVEAFPDHIACSSSTRSEIV
 YP_916587.1 GFYRVTAP-----ELLKIGPYGG-HGCLVIP--FSRCGAAARTR--QVQLVP DVDAFPGHIACASSTRSELV
 XP_961529.1 -----GPFQGK-VACQTIA--FGRGCGAAATK--TTQLVP DVDAFPGHIACDGD SKSEIV
 YP_351544.1 GFYRVTGP-----ELLKIGPYGA-HGCLVIP--FSRCGAAARTG--RIQLVP DVEAFPDHIACSSSTRSEIV
 XP_001588803.1 -----LILGPFQGK-VACQTIA--FSRCGTAASTQ--TTQLVP DVEKFPGHIACDAASQSEIV
 XP_762195.1 GFYLLSPLL PSEIVSPKKRKHPTLLILGPE NGL-PACQLIVSVP GKG CADASAVLPPRVVRA DTDAYPGHIACD SLSKSEIV
 XP_001537789.1 GFYVRQDRFP-SLVSAGSNTKETPGTQ-ALLIDPEHGK-PACQLIQ--FGRGCGTAAEKQ--ETVLVP DVFNWEGHIACDAESKSEIV
 YP_001198994.1 GFYLYD-----GSELILGPFQGG-VSCVHIA--LKGKCGEVAAKR--QTILVHDVRLHDNYISCDATVP----
 NP_390842.1 GFYFAK-EED-----GQLVLGPFQGL-PACVRIP--FGRGCGTAYANG--KVERIEDVNAFPGHIACDAASQSEIV
 YP_001469780.1 GFYLTDFD-----EKNTLVLGPFVGEPTTEHV KIP--FGKIGCGAAAEK--TTFIVQ DVSKETNYLSCSPKVKSEIV
 NP_623333.1 GFYFMK-----DGELVIGPYIGRPTEHVRIK--VQGGVCGRAVAEK--STIIVD DVTKEDNYLAC SLETKSEIV
 YP_001567589.1 GFYLVDF-----ETEDNSLVLGPFVGEPTTEHTKIK--FGQIGCGAAASTK--TTFVVD DVSKEDNYLSCSPKTQSEIV
 NP_228329.1 GFYFVE-----DGKLLKLPFVGEPTTEHVEIP--FVGIGCGAAARE--ETFVVD DVSKETNYLSCSPKTKAEIV
 YP_001244004.1 GFYFVE-----DGKLLKLPFVGEPTTEHVEIP--FVGIGCGAAARE--ETFVVD DVSKETNYLSCSPKTKAEIV
 YP_001738457.1 GFYFVE-----DGKLLKLPFVGEPTTEHVEIP--FVGIGCGAAARE--ETFVVD DVSKETNYLSCSPKTKAEIV
 YP_591728.1 GFYMIEKSL-----PGVDPVVLGPFYVGAPTTHTTRIP--LNEGICGAAVTTG--DTVIVD DVNSDPRYLAC SLETKSEIV
 YP_001193798.1 GFYFANL-----ENKTLHLGPFYVGAETDHTVIP--FGKIGCGQVAESN--ANFVVP DVKAQDNYLAC S LTVKSEIV
 YP_001296902.1 GFYFANH-----ETKTLHLGPFYVGAETDHTVIP--FGKIGCGQVAESN--ENFVVP DVAAQDNYLAC SFTVKSEIV
 XP_001439866.1 GFYEVIDG-----NSLEVPYQST-ILATPRIE-KKGQCGQCWAEG--KVQIQE DVKVCQNYIACD NETQSEIV
 XP_001438779.1 GFYQVIDGI-----LIDSILDNHLEVPYQST-ILATPRIE-KKGQCGQCWAEG--KVQIQE DVKVCQNYIACD NETQSEIV
 YP_001319456.1 GFYMIE-----NGE---LVLGHYLGAHTDHTHIQ--IQGGICGQAAADRK--ETFIVD DVTKETNYLAC S FETASEIV
 XP_001022412.1 GFYVVVDELDS-----QGKNTGNKVLEVPYQST-ILATPRIA-IGKVCGTAWQEA--KTQVVNNVKECKNYIACD NETLSEIV
 YP_001306613.1 GFYMIND-----EGLLELYDFVGEPTTEHV KIP--ICEGICGQAAANLK--RTFIVQ DVSKETNYLSCSPKVKSEIV
 YP_002335308.1 GFYMINK-----DNVLELFEVGEPTTEHV KIN--ICEGICGQAAAMLK--DVFIVQ DVSKETNYLSCSPKVKSEIV
 NP_394308.1 GFYVLEH-----GKLLKLEAVGEKTDHVEIN--LCDGLCSLAVLKN--DIVNEY DVKSNPKYLACFPSTQSEIV
 YP_862969.1 GFYFKNG-----NKEELKLSVAGEPTDHEIIP--FGRGICGQVAESN--KNFVVP DVKAQNNYLAC S IHVKA EIV
 YP_023758.1 GVYVVNN-----GNLSLISYSKRTEHEIIN--LCSGLCSLAVTKR--MIINEG DVRSNSDYLACFPETNSEIV
 NP_111339.1 GVYVMEN-----RKLKLEAVGEKTEHEVIN--LCDGLCSLAVLKN--EVVNEY DVKSNPKYLACFPSTQSEIV

YP_001380982.1 YLYVLRG-----DVLELDAAGRETPHVRTIP--VQQLCGLSARTR--EVVLVDVAADPRYLACNLETRSEIV
 YP_806472.1 GFYLYNDKGTG-----QLDLGPFQGGK-VACMHIQ--PENGIVGTSYAQN--AVLRVNVHEFAGHICDSASNAEIV
 YP_001987415.1 GFYLYNDKGTG-----QLDLGPFQGGK-VACMHIQ--PENGIVGTSYAQN--AVLRVNVHEFTGHICDSASNAEIV
 NP_785802.1 GFYLFNEQGTG-----ELDLGPFQGGK-VACMHIK--VAGVVGTAFFETQ--TNQRVADVHQFPGHICDSASNSEIV
 YP_818384.1 GFYIYNDDAK-----ELDLGPFQGGK-VACMHIK--PGSGVGTAFANK--ESIVVNVVHDFAGHICDADSNSSELV
 YP_001621174.1 GFYMYKNG-----ILTLGPFQGGK-VACNLM--PKGVVGTSYSKN--EVIVVKNVHEIENHITCDPISMSSETV

| | 130 | | 140 | | 150 | | 160 | | 170 |
|----------------|----------|-------|-----|--------|--------|--------|-----------------|-----|-----|
| NP_288269.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002403065.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002397981.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002412850.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_001880631.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002391608.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_001730808.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| ZP_03003290.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_001743413.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002407226.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_002329476.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDADE | EQG-LRQLVAQLEK | VLA | |
| YP_002382399.1 | LPLVV-ED | ----- | QII | GVLDID | STQFGR | STEEDE | EQG-LRLLVTELEK | VLA | |
| ZP_03445135.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_688895.1 | LPLVV-KN | ----- | QII | GVLDID | STVFAR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| YP_669680.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| NP_416346.4 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| ZP_02903917.1 | LPLAV-KN | ----- | QII | GVLDID | STVFSR | STGEDE | EQG-LRQLVAQLEK | VLA | |
| ZP_03282954.1 | LPLVV-KG | ----- | QIM | GVLDID | STDFGR | STAEDE | EQG-LRELVANLENV | VLA | |
| YP_001452718.1 | LPLTV-DN | ----- | QII | GVLDID | STVFGR | STEEDE | EQG-LRELVARLET | VLA | |
| YP_001570142.1 | LPVAV-GK | ----- | KII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| NP_460803.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| YP_150310.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| YP_002141797.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| YP_002243298.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| NP_456352.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| NP_804849.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| YP_001587603.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| ZP_02666379.2 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| ZP_02832604.2 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| YP_001177122.1 | LPLSV-NN | ----- | QII | GVLDID | STAFGR | STDEDE | EQG-LRELVANLENV | VLN | |
| YP_002226291.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |
| NP_754135.1 | LPLVV-KN | ----- | QII | GVLDID | STVFGR | STDEDE | EQG-LRQLVAQLEK | VLA | |
| ZP_02347214.1 | LPVTV-GE | ----- | RII | GVLDID | STAFGR | STEEDE | EHG-LRTLVAQLET | VLA | |

| | | | | | |
|----------------|------------------|---------|--------|----------|---------------|
| ZP_02574893.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_02659310.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_02663512.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_02681881.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_02696584.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_002041101.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_002045892.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03075263.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_002114879.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_002146178.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03165709.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_002215294.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03221812.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03342093.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03347154.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03352578.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03357060.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03371453.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03379701.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03383790.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03412299.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| ZP_03217260.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLG |
| YP_002237781.1 | FPLRV-ND-----QII | GVLDIDS | STAYGR | STAEDQEQ | LRTLVEQLEKLI |
| YP_001336004.1 | FPLRV-NG-----QII | GVLDIDS | PAYGR | STAEDQEQ | LRTLVEHLEKLI |
| YP_001437512.1 | LPLRA-DG-----EII | GVLDIDS | TVFNRS | DADDEQEQ | LCTLVEQLEAVIA |
| ZP_03403067.1 | LPVTV-GE-----RII | GVLDIDS | STAFGR | STEEDEHG | LRTLVAQLETVLA |
| YP_001907453.1 | LPITV-NG-----HVI | GVLDIDS | VEYNRS | DEDEKDG | LKALTDGLCAVLS |
| YP_001006127.1 | LPITV-QG-----KVI | GVLDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| NP_669181.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| NP_993079.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_070893.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_651729.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_647856.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_001163103.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_01888376.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_001400647.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| YP_001607069.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_02221459.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_02227087.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_02229339.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_02237158.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |
| ZP_02307206.1 | LPITV-KG-----KVI | GILDIDS | SIVYDR | DEDEDELG | LISVVAGLCEHLE |

ZP_02314388.1 LPITV-KG-----KVI GILDIDSIVYDRFDEDDDELG-LISVVAGLCEHLE
 ZP_02318544.1 LPITV-KG-----KVI GILDIDSIVYDRFDEDDDELG-LISVVAGLCEHLE
 YP_001720525.1 LPITV-KG-----KVI GILDIDSIVYDRFDEDDDELG-LISVVAGLCEHLE
 YP_001872871.1 LPITV-KG-----KVI GILDIDSIVYDRFDEDDDELG-LISVVAGLCEHLE
 YP_002346706.1 LPITV-KG-----KVI GILDIDSIVYDRFDEDDDELG-LISVVAGLCEHLE
 YP_001478346.1 LPLNV-GG-----QII GVLDIDSTVYQRFDEQDERG-LEAVVAGLCAQLE
 YP_454994.1 LPITL-DG-----RVA GVLDIDSTIYQRFDQDDEEG-LKAVADYLCRQLG
 YP_001217052.1 IPFSI-NG-----KVA GVLDIDSPNIGRFSEIDEQG-LTYLMSEVEKLLN
 YP_002150764.1 LPLEV-NG-----QII GVLDIDSPIFGRFDKKDENG-LKALCDVLCHEHLK
 YP_856816.1 IPLRR-GG-----EII GVLDIDSPIFNRFDAQDRIG-LEETVQILESML-
 YP_001093925.1 IPLRR-DG-----QVI GVLDIDSPSFSRFDEADQLG-LESCVKSLEKALF
 YP_269525.1 VPVKI-AG-----KVI AILDIDSTVFDRFNLEDEQG-IVSIVQAFEQNIT
 NP_718189.1 IPVRA-DG-----KVI AILDIDSPIFDHFDDEEDQKG-LELLVKSFEKCLF
 YP_001366642.1 VPVCA-QG-----NVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLEF
 YP_001554990.1 VPVCA-QG-----NVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLEF
 YP_737796.1 IPVRA-DG-----KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLEF
 YP_869415.1 IPVRA-DG-----KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLEF
 NP_797987.1 IPFSI-KG-----KVA GVLDIDSPSIGRFNETDEEG-LTHFMSEVEKLLN
 YP_001501790.1 VPLRQ-NG-----TVI GVLDIDSPSLSRFSEIDQQG-VETLVKSLEICLI
 YP_927460.1 VPVRH-QG-----RVI AILDIDSPSFSRFDETDQAG-LETLVRTLEKALF
 YP_001474222.1 IPIMR-DH-----EVA VLDIDSPSLARFAETDQVG-LEKLVNSFEKALF
 YP_963195.1 VPVRA-NG-----KVI AILDIDSPSFDRFDEDDQKG-LELLVKSFENCLEF
 YP_001183724.1 VPVRA-NG-----KVI AILDIDSPSFDRFDEDDQKG-LELLVKSFENCLEF
 YP_733804.1 IPVRA-DG-----KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLEF
 YP_001050811.1 VPVCA-QG-----NVVAVLDIDSPIFDRFDEDDQKG-LELLVKSFENCLEF
 YP_750914.1 IPVRQ-GN-----EVA VLDIDSPLLNRFDQDDQIG-LENMVKSFETCLEF
 YP_002126297.1 VPLVL-ND-----SLV GVLDIDSPVFERFTEEDEKG-LVEIAQILMDSQK
 YP_130129.1 IPLTV-KG-----ELI GVLDIDSPSLSRFDQNDQIG-LVSFVKELQKSL-
 YP_001445537.1 IPFTI-NG-----EVA GVLDIDSPNIGRFNEIDEQG-LTYFMAEVEKLLN
 YP_661562.1 LPVYQ-KG-----QLL GVMIDLSPHTGRFDQDDMTG-LQNLVNILODAIA
 YP_001674584.1 VPLRK-NG-----SVI GVLDIDSPSLSRFSEIDQQG-VETLVKSLEFCLF
 YP_002417027.1 IPFSI-DG-----KIA GVLDIDSPNVGRFSQIDEDG-LTFFMAEVEKLLN
 NP_934438.1 IPFSI-NG-----QLA GVLDIDSPSIGRFSEIDEQG-LTFFMAEVEKLLN
 NP_761466.1 IPFSI-NG-----QLA GVLDIDSPSVGRFSEIDEQG-LTFFMAEVEKLLN
 YP_001760510.1 IPVFD-NN-----RLI GVLDIDSPILNRFTKTDQIG-LENLVKSFEFESALF
 YP_002312133.1 VPFRR-NG-----DVI GVLDIDSPSLSRFDEIDQQG-VETLVKSLENCLF
 YP_562727.1 IPVRR-GN-----EVI AILDIDSPSFDRFDSEDQLG-LEKLVKSLESCLF
 NP_693123.1 VPITI-NN-----QII GVLDIDSPINNRFDEEDQYF-LEKFLVHLTKHLQ
 YP_001699747.1 IPLMK-QG-----EVM GVLDIDSPTVNRFSKEDQAG-LEHFVQTLTLLHI-
 YP_001343242.1 IPVVV-NG-----QLI AVFDMSPKVGRFSDVDQAG-LESFVATFIEATD
 YP_001141806.1 IPVRR-GG-----EII GVLDIDSPIFNRFDEKDRIG-LEETVQILESML-
 YP_204661.1 IPFYR-NG-----ELA GVLDIDSPKFNRFSEIDEKG-LVDFMKEAEKRL-

| | | | | | | | | |
|----------------|---------------|-----|--------|--------|-----|------|------|----------------|
| YP_301135.1 | VPLRK-NN----- | QII | GVLDID | APITSR | FDV | DKN | G | LEQIVARIEKQIS |
| YP_002156075.1 | IPFYR-NG----- | KLA | GVLDID | SPKFN | R | SEI | DEK | LVDFMKEAEKRL- |
| NP_244074.1 | IPLFQ-NG----- | VLY | GVLDID | SPSLNR | F | SEEE | QAL | LESFVDVLSKNLK |
| YP_001559126.1 | IPMKK-NG----- | TIL | GVLDID | SPKFHR | F | DEVD | AEY | LEQIITLLVECCD |
| YP_080250.1 | VPLEV-DG----- | SII | GVLDID | SPIKNR | F | DET | DEYY | LKKFTEKLGERSL |
| YP_253122.1 | IPIHV-ND----- | EII | GVLDID | APITDR | F | TNE | DKEG | LEVIVKVIEKQLT |
| YP_261333.1 | VPLVK-DG----- | RLI | GVLDLD | SPKLAR | F | GVAD | QAG | IEQLAAIFLRLSD |
| YP_001187407.1 | VPLVK-EG----- | RLI | GVLDLD | SPSVGR | F | TEVD | QAG | IEALAAIFLAASD |
| YP_001422246.1 | LPIRV-NG----- | EVI | GVLDID | SPVKNR | F | DET | DETY | LAKFVETLEKRLG |
| YP_001748360.1 | IPLVK-EG----- | RLI | GVLDLD | SPKLAR | F | SEAD | QRG | LERLAAVFLELTD |
| YP_155660.1 | APIVV-NG----- | KTV | AVLDV | SPSVGR | F | SEED | AKA | IQAIADYCQTLNW |
| NP_744032.1 | IPLVK-DG----- | RLI | GVLDLD | SPTVGR | F | SEAD | QVG | LERLAAIFLALTD |
| YP_001269146.1 | IPLVK-DG----- | RLI | GVLDLD | SPTVGR | F | SEAD | QVG | LERLAAIFLALTD |
| YP_236680.1 | VPLIK-EG----- | RLI | GVLDLD | SPSVGR | F | NEED | QAG | IERLVAIFLASTD |
| YP_001667695.1 | IPLVK-EG----- | RLI | GVLDLD | SPKGR | F | GEAD | QVG | LERLAAIFLELTD |
| YP_188854.1 | IPLHQ-NG----- | KII | GVLDID | APIKNR | F | SQDD | QEG | LESIVKTLEKQLN |
| YP_002275225.1 | VPILD-GD----- | RLI | GVLDID | SPIRD | F | TPQD | RLV | LEQVAALLVPAPT |
| YP_804779.1 | LPLMV-GN----- | QLL | GVLDID | SPILNR | F | SEDE | BEAT | LIKFRDALVKHID |
| NP_470973.1 | LPIVK-NN----- | QLF | GVLDID | SPLFNR | F | DEVD | QLW | LEKIRDAITQEIN |
| YP_607257.1 | IPLVK-AG----- | RLI | GVLDLD | SPKLAR | F | SEAD | QAG | LERLAAIFLELTD |
| YP_001860169.1 | IPLQKANG----- | ELV | GVLDID | SPVLER | F | DD | DRRG | LEELAKIFVASLR |
| NP_465120.1 | LPIVK-NN----- | QLL | GVLDID | SPLFNR | F | NEVD | QLW | LEKIRDAIVQEIN |
| NP_372242.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| NP_374828.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001247140.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001316940.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001442294.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_002314836.1 | VPMMK-NG----- | VVI | GVLDID | SPIKNR | F | DEI | DEQY | LQRFVETLVQHLN |
| YP_001814701.1 | VPLIK-DG----- | KTI | GVLDID | SPELSR | F | DET | QAG | LEAFSEILLRHI- |
| NP_646478.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_043764.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_186602.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_494359.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_500333.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001332646.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001575592.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| ZP_02759966.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_001487827.1 | VPLKV-NG----- | EII | GVLDID | SPVKDR | F | SEV | DETY | LIQLTEVLQKALS |
| YP_417047.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_041183.1 | VPIFK-DD----- | KII | GVLDID | APITDR | F | DDND | KEH | LEAIVKIIIEKQLA |
| YP_176251.1 | IPIRH-EQ----- | NVI | AVLDID | SPSKGR | F | HEKE | AEA | LERLLKRVEPFLI |
| YP_002349929.1 | LPIVK-NN----- | QLL | GVLDID | SPLFNR | F | DEVD | QLW | LEKIRDAIVQEIN |

| | | | | | | | | | |
|----------------|---------------|-------|------|--------|--------|-----------|-------|-----------|------------|
| YP_001376537.1 | VPMIK-EG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QQY- | LEK | FVETLLKHM- |
| YP_784647.1 | VPLLR-DG----- | QLL | GVWD | VDS | SPTPGR | DDAD | REG- | MEAL | CQVFLHSLA |
| YP_557760.1 | VPLVAPDG----- | TLI | GVWD | VDS | SPVVAR | DEED | DAKG- | MEAL | CAVFIQAAL |
| YP_275769.1 | VPLIK-DG----- | RLI | GVLD | LDS | SPSVGR | NEVD | QAG- | IERL | AAIFLASTD |
| YP_001118914.1 | VPLVAADG----- | TLI | GVWD | VDS | SPVAAR | DDER | DRSG- | MEAL | CRVFVEHAW |
| YP_001782545.1 | VPIIK-EN----- | KIL | GVLD | LDS | SIEFGR | TELEEKY- | FEKF | VQIL | VENID |
| NP_882192.1 | VPLVH-KG----- | SLL | GVWD | VDS | SPLPGR | DDER | DRAG- | MEAL | CEVFLHSLG |
| YP_001894915.1 | VPLVAPDG----- | TLI | GVWD | VDS | SPSVAR | DEED | DAKG- | MEAL | CAVFLIEAAL |
| YP_001647268.1 | VPIMK-EG----- | NII | GVLD | IDS | SPEKNR | FDKVD | QQY- | LEK | FVETLLKHI- |
| NP_882465.1 | VPLLH-KG----- | SLL | GVWD | VDS | SPLPGR | DDER | DRAG- | MEAL | CEVFLHSLG |
| YP_002448219.1 | VPIVK-EG----- | AVI | GVLD | IDS | SPEKNR | FDKVD | QQY- | LEK | FVETLLKHM- |
| NP_834359.1 | VPIVK-EG----- | AVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_002369448.1 | VPIVK-EG----- | AVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_191278.1 | LPVVA-RG----- | ALV | GVLD | IDS | SPVKDR | SDADR | TR- | LKE | VVSLLVATLE |
| NP_886654.1 | VPLVH-QG----- | SLL | GVWD | VDS | SPLPGR | DDER | DRAG- | MEAL | CEVFLHSLG |
| NP_847103.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_030799.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| NP_981083.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_038702.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_002340711.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKRR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_085975.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKNR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_002453717.1 | VPIIK-DG----- | NVI | GVLD | IDS | SPEKRR | FDKVD | QRY- | LEK | FVETLLKHM- |
| YP_772921.1 | VPLVAADG----- | TLI | GVWD | VDS | SPVAAR | DDER | DRNG- | MEAL | CRVFVEHAW |
| YP_001843726.1 | VPIHQ-GD----- | KLW | GVLD | IDS | SPKLAR | DQED | QAG- | LER | LAPLFLPGA- |
| YP_942902.1 | LPISL-QG----- | APALL | LIDS | SPIKNR | FTQND | QDG- | LAYL | VEQIQOHL- | |
| YP_121651.1 | VPLVR-DD----- | TVI | GVLD | LDS | SPRPGR | DTVDR | DRG- | LEA | VARVFLDAL- |
| YP_001633478.1 | VPLVH-NG----- | ELI | GVWD | VDS | SEPGR | DEDD | DRQG- | MQAL | CAVFLASLA |
| YP_001392189.1 | VPIIK-ED----- | KIL | GGLD | LDS | SIEFGR | TELEEKY- | FKKF | VQIL | VENID |
| YP_795394.1 | VPLTI-GE----- | TKF | GVLD | IDS | SPTKDR | TAAD | QQV- | LEE | FARVLLSHID |
| YP_002284153.1 | IPLVK-DG----- | GIV | GVLD | LDS | SPTPHR | DAED | QAG- | FEAL | AAIYVAASA |
| YP_001255404.1 | VPIIK-ED----- | KVL | GVLD | LDS | SIEFGR | FTGLEEKY- | FKKF | VQIL | VENID |
| YP_001385170.1 | VPIIK-ED----- | KVL | GVLD | LDS | SIEFGR | FTGLEEKY- | FKKF | VQIL | VENID |
| YP_001388639.1 | VPIIK-ED----- | KVL | GVLD | LDS | SIEFGR | FTGLEEKY- | FKKF | VQIL | VENID |
| YP_001347703.1 | VPLLK-DG----- | RLI | GVLD | LDS | SPSEGR | FSARD | QAG- | IERL | VEIFLRLSD |
| NP_103622.1 | VLLED-DE----- | GVF | GVLD | LDS | SPLPGR | FDKAD | QAG- | IETL | AAIYAAASS |
| YP_001807739.1 | VPLVAADG----- | TLI | GVWD | VDS | SPVAAR | DDER | DRNG- | MEAL | CRVFVEHAW |
| YP_001580337.1 | VPLVADDG----- | TLI | GVWD | VDS | SPVAAR | DDER | DRRG- | MEAL | CRVFVEHAW |
| YP_002230176.1 | VPLVAGDG----- | TLI | GVWD | VDS | SPVAAR | DDER | DRQG- | MEAL | CRVFVEHAW |
| NP_251512.1 | VPLIK-DG----- | RLVA | VLDL | LDS | SPSVGR | FSARD | QAG- | IEGL | VEIFLRLTD |
| YP_459505.1 | VPVVR-DD----- | TVVA | VLDL | LDS | APTAKR | FSARAD | AEAG- | IEEL | ARRIAGRI- |
| YP_790351.1 | VPLLK-DG----- | RLVA | VLDL | LDS | SPSVGR | FSARD | QAG- | IEGL | VEIFLRLTD |
| YP_002439846.1 | VPLLK-DG----- | RLVA | VLDL | LDS | SPSVGR | FSARD | QAG- | IEGL | VEIFLRLTD |

YP_001970489.1 VPLLR-GD-----ELI G V F D I D S P K V G R F D A D D Q A G - L E A I A R V F V E A L G
 YP_617023.1 VPVIA-ND-----RLV G V L D L D S P L P A R F T A A D Q A G - A E A L V A R I A G A L A
 YP_620554.1 VPLVADDG-----TLI G V W D V D S P V A A R F D D E D R K G - M E A L C R V F V E H A W
 YP_834795.1 VPLVADDG-----TLI G V W D V D S P V A A R F D D E D R K G - M E A L C R V F V E H A W
 YP_148653.1 VPLIK-DD-----RVI G V L D I D S P V K N R F D D I D R R Y - L E Q F A S V L I S A --
 YP_442277.1 VPLIARDG-----ALI G V W D V D S P L I G R F D A E D A T G - M E A L C R V F V E V A W
 YP_334263.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A T G - M E A L C R V F V E V A W
 YP_001067073.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A T G - M E A L C R V F V E V A W
 YP_368503.1 VPLVAADG-----TLI G V W D V D S P V A A R F D D E D R K G - M E A L C R V F V E S A W
 YP_109010.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A I G - M E A L C R V F V E V A W
 YP_993758.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A I G - M E A L C R V F V E V A W
 YP_001028775.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A I G - M E A L C R V F V E V A W
 YP_002026852.1 VPLLR-GD-----ELI G V F D I D S P K V G R F D A D D Q A G - L E A I A R V F V E A L G
 YP_001764423.1 VPLVADDG-----TLI G V W D V D S P V A A R F D D E D R T G - M E A L C R V F V E H A W
 NP_970249.1 VPLIHNG-----KLL G V L D V D A P V L N R F D S E D Q K G - L Q G L V Q I L I S K T V
 YP_001059788.1 VPLVACDG-----ALI G V W D V D S P L V G R F D A E D A T G - M E A L C R V F V E V A W
 YP_001081316.1 VPLVARDG-----ALI G V W D V D S P L V G R F D A E D A I G - M E A L C R V F V E V A W
 YP_001875000.1 VPIIK-NG-----TLF G V L D I D S P I K N R F S D E D R I F - F E K V V G L F S K Y S D
 YP_278248.1 IPIIV-NS-----KLY G V L D L D S P L K S N F K D K E I V L T L E K I A T E L A N K I K
 YP_001309672.1 IPIIK-EG-----NLI G V L D L D S E E F E R F T E V E K T Y - L E K A V V I L S K Y I E
 YP_001576818.1 VPIFK-DD-----QLW G V F D F D S T K L D N F D E T D K K Y - L T E I A N V I F A ---
 YP_193206.1 IPIFK-NR-----KLW G V F D F D S T K L D D F D D M D K K Y - L T R I T K M V F A ---
 YP_199405.1 IPLVK-GD-----TLI G V L D L D S P E L D R F D A D D Q R G - L E A I A Q L F V D A L A
 YP_001915608.1 IPLVK-GD-----TLI G V L D L D S P E L D R F D A D D Q R G - L E A I A Q L F V D A L A
 YP_002509654.1 IPVEV-DG-----DII G V L D I D S P V K N R F D E I D K R Y - L E E F V K K I I N L S D
 YP_757341.1 VPVRDA-----AGALVA V L D I D S T E F G A F D E V D Q A G - L E A I C G E L -----
 YP_810624.1 VPIFS-KN-----EFW G V L D L D S T K F N Y F D K N D Q E Y - L E K I G N Y I F G I --
 YP_244708.1 IPLVR-GD-----TLI G V L D L D S P E L D R F D A D D Q R G - L E A I A Q V F V G A L T
 NP_635978.1 IPLVQ-GD-----TLI G V L D L D S P E L D R F D A D D Q R G - L E A I A Q V F V G A L T
 XP_505859.1 VPIV-LNG-----KLF G V I D I D C A S L E G F D E V D Q Q G - L E K I A L I L A K S C E
 NP_964546.1 VPIFKANG-----DLW G V F D F D S T K L A N F D E L D Q K Y - L E K I S K V F K F ---
 YP_001261873.1 VPIVH-EG-----RLI G V L D L D S P I K G R F D E Q D V V G - C A E L M A V L G P R I A
 YP_814311.1 IPIFKKNG-----ELW G V F D F D S T K I A N F D E L D Q K Y - L E A I S N I F K F ---
 NP_298520.1 VPLLK-DD-----TLI G V L D L D S P T C A R F D S D D Q T G - I E S L A R V F V S S L R
 YP_001775202.1 VPLLK-DD-----TLI G V L D L D S P T C A R F D S D D Q T G - I E S L A R V F V S S L R
 YP_001829224.1 VPLLK-DN-----TLI G V L D L D S P T C A R F D S D D Q T G - I E S L A R V F V S N L R
 NP_778732.1 VPLLK-DN-----TLI G V L D L D S P T C A R F D S D D Q T G - I E S L A R V F V S N L R
 NP_246070.1 VPIFQ-GG-----EMIAEL D V D S P R L A R F S A E D V A F - L E N C A A L I C I K E -
 XP_645637.1 LPLYK-----QGNII G V L D V D S D K L N S F D E I D E K Y - L T Q I L K L L D N ---
 XP_456263.1 VPIV-QNG-----ETV G V I D I D C L D Y N G F T K L D Q E F - L E K L A A S V S K T C V
 YP_001300883.1 VPILA-----NGQVK G V L D I D S N L L N F F D E T D R R Y - L E E V T A I V A K T --
 YP_812434.1 VPVFNRDG-----SLW G V Y D F D S I N L D N F D D L D Q K Y - L E Q I A A V A L G N --

NP_814796.1 VPMVKNHQ-----LVGVLDIDSGVTNSYDAVDQLY-LEKQVTLLEKSD
 NP_358362.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTT
 YP_816246.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTT
 YP_001694316.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTT
 YP_001835526.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTT
 YP_002037500.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTT
 NP_345351.1 VPMMKNGQ-----LLGVLDLDSSEIEDYDAMD RDY-LEQFVAILLEKTA
 XP_001385863.2 VPLV-KEG-----KILAVLDIDCLALDGFDES DKKF-LEELAVLIVRTCE
 NP_985662.1 VPIL-QDN-----RTVAIMLDCSEHNGFTDLDK KY-LEKLAQEISRTCK
 YP_759928.1 LPVFRA-----DGSLAAVLDVDS TQFDASDAVDQAG-LTAICESL-----
 YP_002123053.1 VPMLKAGR-----LVGVLDLDS SLVADYDQIDQDY-LEQFVALLLENTS
 YP_002131343.1 VPVLDA-----GGELIAVFDVDS ETKAAFDST DREW-LERILSET-----
 XP_001683506.1 VPVLIENG-----SVVAVIDVDS TELGHFSEEDREG-LERMAAVLAQHLD
 NP_012854.1 VPIISNDG-----KTLGVLDIDCLDYEGFDHVDKEF-LEKLAKLINKSCV
 XP_460154.1 VPIV-NKG-----QVVGVLIDIDCLTLEGFD DIDVKY-LEQLATLIGNTC D
 XP_001801028.1 VPIVQGG-----RTVAIIDVDC AELSGFDGEDQRA-LEELACMLGEACD
 YP_001201197.1 VPMVKNDQ-----LLGVLDLDS SRLVADYDTIDQDY-LEKQVALLVEKTD
 XP_001484046.1 VPIV-YEN-----KVVGVLDIDCLDLNGFDKVDQKY-LERLVEMISSSCK
 NP_274882.1 VPLFS-----DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
 YP_001302571.1 VPVIW-----ADKVIAVLDIDSDQPDSFDETDQIF-LEKIVELLPHQ--
 YP_002000650.1 VPLFS-----DGRCIGVLDADSEHLAQFDEADALY-LGELAKILERR--
 YP_001682978.1 VPVFDA-----AGTLLAVFDVDS DQPAASDAVDQHG-IEIILKAT-----
 YP_002342048.1 VPLFS-----DGRCIGVLDADSEYLAQFDETDALY-LGELAKILEKR--
 YP_207199.1 VPLFS-----DGRCIGVLDADSEHLAQFDEADALY-LGELAKILERR--
 XP_844799.1 VPIKSAAG-----EVVGIIDVDS TQLNFFDDVDAQG-LQEVAEIVARHVD
 YP_001598496.1 VPLFS-----DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
 YP_974447.1 VPLFS-----DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
 XP_001465858.1 VPVLIENG-----SVVAVIDVDS TELDHFSEEDKEG-LERMAAVLAQHLD
 YP_001280815.1 VPITDE-----AGEIIAVLDIDSDTLATFDEVDGEY-LQQLVDLLK----
 NP_664843.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 NP_802083.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 YP_280574.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 YP_001128311.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 YP_002286076.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 XP_390238.1 VPILVAQDDGT-----KKLVAIIDIDCAELNGFDEVDKAH-LEQLAALLATSCD
 XP_001565241.1 VPVLIENG-----SVVAVIDVDS TELGYFSEEDKEG-LERVAAVLEQHLN
 YP_596891.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 YP_600772.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LEAFVDLFLFKTT
 YP_001351938.1 VPIYFDR-----NGEVAQVLDVDS EYPDAFSAVDQLW-LEKIAALLS----
 XP_001910908.1 VPIMVEGK-----GVVGIIDIDCTVENGFDQIDKEY-LEKLADFLAKSSD
 XP_446236.1 VPIVSQSG-----KTLGVLDLDC LDFEGFTDVKKEY-LEKLALAITKSCD
 NP_607473.1 VPMVKEGH-----LIGVLDLDS SLVADYDEV DQEY-LETFVDLLEKTT

YP_060405.1 VPMVKEGH-----LI**G**VLDL**D**S**S**SLVADYDEVD**Q**EY-**L**ETFVDLLLEKTT
 YP_602707.1 VPMVKEGH-----LI**G**VLDL**D**S**S**SLVADYDEVD**Q**EY-**L**ETFVDLLLEKTT
 NP_269471.1 VPMVKEGH-----LI**G**VLDL**D**S**S**SLVADYDEVD**Q**EY-**L**EAFVDLFLEKTT
 YP_282476.1 VPMVKEGH-----LI**G**VLDL**D**S**S**SLVADYDEVD**Q**EY-**L**EAFVDLFLEKTT
 YP_598792.1 VPMVKEGH-----LI**G**VLDL**D**S**S**SLVADYDEVD**Q**EY-**L**EAFVDLFLEKTT
 YP_099090.1 VPLIK-----DGEVW**G**VLDL**D**S**S**DLLN**F****S**DETDRKY-**L**EEMCGYLSK---
 NP_421082.1 VPVFD**P**-----DGRLIA**V**F**D**V**S**DKPAA**F**DEIDQAW-**L**EKILKAT-----
 YP_001034962.1 VPMVKVGR-----LL**G**VLDL**D**S**S**SLTG**D**YDKV**D**Q**E**Y-**L**EKFVQILLEKTT
 YP_211509.1 VPLIK-----DGEVW**G**VLDL**D**S**S**DLLN**F****S**DETDRKY-**L**EEMCGYLSK---
 XP_001940256.1 VPVVQDG-----K**V**VAI**I**D**I**D**C**AELNG**S**TEEDQEA-**L**EELAELLSQSCD
 YP_001450196.1 VPMVKDGR-----LL**G**VLDL**D**S**S**ALT**G**DYDK**I**D**Q**EY-**L**EKFVQILLEKTI
 XP_680937.1 VPILAGG-----ET**V**AI**I**D**I**D**C**AEP**S**G**S**DEDEK**E**Y-**L**EELAALLSESCD
 XP_001265696.1 VPILVGG-----ET**V**AI**I**D**V**D**C**TE**P**AG**S**DEDEK**K**W-**L**EELASLLSECCD
 XP_367405.1 VPVLDSEG-----Q**L**RAI**I**D**V**D**C**AL**P**DG**S**DEVDR**V**W-**L**ESLAALVGRSCD
 YP_001033719.1 VPMVKDGN-----L**V****G**VLDL**D**S**S**SE**V**GFYDEID**Q**KY-**L**EEFASILCDMTD
 YP_812051.1 VPMVKDGN-----L**V****G**VLDL**D**S**S**SE**V**GFYDEID**Q**KY-**L**EEFASILCDMTD
 NP_810723.1 VPIIK-----E**G**N**V**I**G**VLDL**D**S**S**D**T**P**D**S**S**DE**T**D**A**R**Y**-**L**EEICTYIG----
 YP_141537.1 VPMEKGGK-----L**V****G**VLDL**D**S**S**SH**Q**VED**Y**N**Q**V**D**Q**D**Y-**L**EAFVKILLEKTD
 NP_721920.1 IPMVKDGQ-----L**I****G**VLDL**D**S**S**SH**L**V**G**D**Y**D**V**I**D**Q**G**Y-**L**EQFVTILLEKTN
 YP_820528.1 VPMEKGGK-----L**V****G**VLDL**D**S**S**SH**Q**VED**Y**N**Q**V**D**Q**D**Y-**L**EAFVKILLEKTD
 XP_001246405.1 VPIVVDG-----E**T**L**A**L**I**D**I**D**C**T**Q**P**S**G**S**DEID**K**K**G**-**L**EHLAKLLAENCD
 YP_612063.1 LPVRNR-----A**G**E**V**I**G**VLDL**D**S**S**N**Q**PAA**S**DRD**D**Q**D**A-**L**ETLLHAV-----
 YP_444835.1 VPVL-----**T****P**-----D**D**ALL**A**V**L**D**V**D**S**DT**L**G**A**S**D**P**T**D**R**E**H**-**L**EALCRDLG-----
 NP_943356.1 LPVRI-NN-----**T****V****V****G**A**L**D**I**D**S**PA**F**S**R****S**SE**D**E**L**-----
 YP_631535.1 VPVFGR-----N**R**E**L**L**A**V**L**D**I**D**S**E**H**K**N**A**S**DE**V**D**R**R**E**-**L**EELVRF-----
 NP_687842.1 VPMFKNGK-----L**L****G**VLDL**D**S**S**SLVADYDEID**Q**EY-**L**EKFVGILVEHTI
 YP_329573.1 VPMFKNGK-----L**L****G**VLDL**D**S**S**SLVADYDEID**Q**EY-**L**EKFVGILVEHTI
 XP_001692446.1 VPVVRSHISFNDGRTPSI-----P**E**K**L**V**A**V**L**D**V**D**S**D**H**PAA**S**TE**V**D**Q**E**C**-**L**EELCEWLG**D**HLQ
 NP_390842.2 LPIRV-DG-----**K**I**V****G**VLDL**D**S**S**SP**V**K**N**R**S**DEID**E**K**Y**-**L**TQFAETLEKALA
 YP_511991.1 LPVRDA-----S**G**A**I**I**A**V**L**D**I**D**S**N**Q**P**D**A**S**T**Q**D**D**A**T**-**L**ESILRAT-----
 XP_719233.1 VPIV-KNG-----E**T**K**G**V**I**D**L**D**C**L**D**L**E**G**S**D**Q**V**D**Q**K**Y-**L**EKLAELIAQSL-
 NP_735295.1 VPMFKNGK-----L**L****G**VLDL**D**S**S**SLVADYDEID**Q**EY-**L**EKFVGILVEHTI
 YP_168601.1 LPVFGR-----**T**G**A**L**I**A**V**F**D**I**D**S**S**D**L**P**D**A**S**T**V**A**D**A**E**A-**L**G**A**I**L**E**Q**V-----
 YP_001169098.1 LPVRNG-----E**G**R**L**L**G**VLDL**D**S**S**N**T**PAA**S**TEED**A**R**R**-**L**QALLDRT-----
 XP_001216799.1 VPILAGG-----E**T****V**A**I**I**D**I**D**C**T**E**P**D**G****S**DE**T**D**K**K**Y**-**L**EELAALLADSCD
 YP_437091.1 LPLFNA-----N**G**D**V**A**A**V**L**D**I**D**S**D**H**I**D**A**S**DA**V**D**E**R**F**-**L**A**E**L**L**T**D**L-----
 XP_001390389.1 VPILSKG-----E**T****V**A**I**I**D**I**D**C**A**E**P**D**G****S**DEDE**K**K**Y**-**L**EELAALLA**E**CCD
 XP_001550500.1 VPIVVDG-----**V****V**K**A**I**I**D**V**D**C**G**V**K**E**G**S**DE**V**D**R**K**W**-**L**EELAGV**L**A**E**G**C**D
 XP_001273080.1 VPVLVGG-----E**T****V**A**I**I**D**V**D**C**T**E**P**S**G****S**DE**V**D**R**K**Y**-**L**EELAALLA**E**CCD
 YP_001042044.1 LPVRNG-----E**G**R**L**L**G**VLDL**D**S**S**N**T**PAA**S**TEED**A**A**R**-**L**QALLDRT-----
 YP_916587.1 IPVIGG-----G**E**R**L**I**G**VLDL**D**S**S**D**Q**P**D**A**S**T**R**D**D**A**D**R-**L**Q**E**I**L**D**A**T-----
 XP_961529.1 VPIVVDD-----R**V**V**A**I**I**D**V**D**C**A**E**L**N**G**S**DE**V**D**K**V**W**-**L**EKLAD**L**V**G**R**S**C**D**

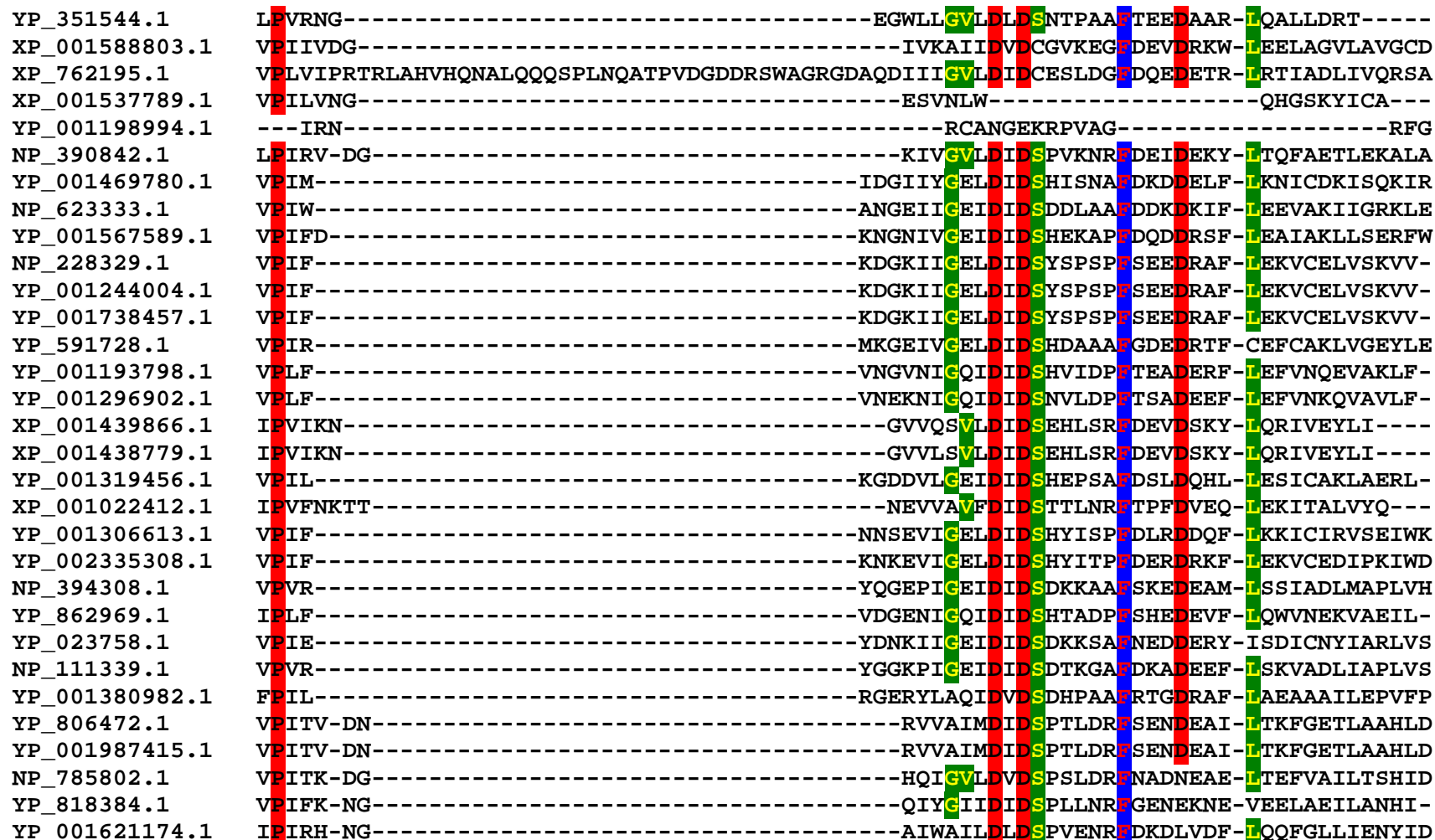


FIGURE SI: Sequence alignment of all putative fRMsr proteins.

The sequences of 358 putative fRMsr proteins were generated by PEDANT database (<http://pedant.gsf.de/>) and aligned with Bioedit software. The numbering of amino acid residues is based on the *E. coli* fRMsr sequence. For reason of clarity, only the alignment between residues 19 and 170 is represented. The residues conserved at least at 95 % are indicated in black on red boxes (such as Y66, C94, C118, E125, D141 and D143), those conserved at least at 90% in red on blue boxes (such as F78, and C84), and finally, in yellow on green boxes (for example W62) those conserved at 80%.

The name of each sequence correspond to NCBI reference:

C84/C94/C118 signature (333 sequences, 93%) :

NP_288269.1: *Escherichia coli* O157:H7 EDL933; YP_002403065.1: *Escherichia coli* 55989; YP_002397981.1 : *Escherichia coli* ED1a ; YP_002412850.1: *Escherichia coli* UMN026; YP_001880631.1: *Shigella boydii* CDC 3083-94; YP_002391608.1: *Escherichia coli* S88; YP_001730808.1: *Escherichia coli* str. K12 substr. DH10B; ZP_03003290.1 : *Escherichia coli* 53638; YP_001743413.1: *Escherichia coli* SMS-3-5; YP_002407226.1 : *Escherichia coli* IAI39; YP_002329476.1: *Escherichia coli* O127:H6 str. E2348/69; YP_002382399.1: *Escherichia fergusonii* ATCC 35469; ZP_03445135.1: *Escherichia coli* O157:H7 str. TW14588; YP_688895.1: *Shigella flexneri* 5 str. 8401; YP_669680.1: *Escherichia coli* 536; NP_416346.4: *Escherichia coli* str.K-12 substr. MG1655; ZP_02903917.1: *Escherichia albertii* TW07627; ZP_03282954.1: *Enterobacter cancerogenus* ATCC 35316; YP_001452718.1: *Citrobacter koseri* ATCC BAA-895; YP_001570142.1: *Salmonella enterica* subsp. *arizonae* serovar 62:z4,z23; NP_460803.1: *Salmonella typhimurium* LT2; YP_150310.1: *Salmonella enterica* subsp. *enterica* serovar *Paratyphi A* str. ATCC 9150; YP_002141797.1: *Salmonella enterica* subsp. *enterica* serovar *Paratyphi A* str. AKU_12601; YP_002243298.1: *Salmonella enterica* subsp. *enterica* serovar *Enteritidis* str. P125109; NP_456352.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. CT18; NP_804849.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. Ty2; YP_001587603.1: *Salmonella enterica* subsp. *enterica* serovar *Paratyphi B* str. SPB7; ZP_02666379.2: *Salmonella enterica* subsp. *enterica* serovar *Heidelberg* str. SL486; ZP_02832604.2: *Salmonella enterica* subsp. *enterica* serovar *Weltevreden* str. HI_N05-537; YP_001177122.1: *Enterobacter* sp. 638; YP_002226291.1: *Salmonella enterica* subsp. *enterica* serovar *Gallinarum* str. 287/91; YP_216830.1: *Salmonella enterica* subsp. *enterica* serovar *Choleraesuis* str. SC-B67; NP_754135.1: *Escherichia coli* CFT073; ZP_02347214.1: *Salmonella enterica* subsp. *enterica* serovar *Saintpaul* str. SARA29; ZP_02574893.1: *Salmonella enterica* subsp. *enterica* serovar 4,- str. CVM23701; ZP_02659310.1: *Salmonella enterica* subsp. *enterica* serovar *Kentucky* str. CDC 191; ZP_02663512.1: *Salmonella enterica* subsp. *enterica* serovar *Schwarzengrund* str. SL480; ZP_02681881.1: *Salmonella enterica* subsp. *enterica* serovar *Hadar* str. RI_05P066; ZP_02696584.1: *Salmonella enterica* subsp. *enterica* serovar *Newport* str. SL317; YP_002041101.1: *Salmonella enterica* subsp. *enterica* serovar *Newport* str. SL254; YP_002045892.1: *Salmonella enterica* subsp. *enterica* serovar *Heidelberg* str. SL476; ZP_03075263.1: *Salmonella enterica* subsp. *enterica* serovar *Kentucky* str. CVM29188; YP_002114879.1: *Salmonella enterica* subsp. *enterica* serovar *Schwarzengrund* str. CVM19633; YP_002146178.1: *Salmonella enterica* subsp. *enterica* serovar *Agona* str. SL483; ZP_03165709.1: *Salmonella enterica* subsp. *enterica* serovar *Saintpaul* str. SARA23; YP_002215294.1: *Salmonella enterica* subsp. *enterica* serovar *Dublin* str. CT_02021853; ZP_03221812.1: *Salmonella enterica* subsp. *enterica* serovar *Javiana* str. GA_MM04042433; ZP_03342093.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. 404ty; ZP_03347154.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. E00-7866; ZP_03352578.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. E01-6750; ZP_03357060.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. E02-1180; ZP_03371453.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. E98-2068; ZP_03379701.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. J185; ZP_03383790.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. M223; ZP_03412299.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. E98-3139; ZP_03217260.1: *Salmonella enterica* subsp. *enterica* serovar *Virchow* str. SL491; YP_002237781.1: *Klebsiella pneumoniae* 342; YP_001336004.1: *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578; YP_001437512.1: *Enterobacter sakazakii* ATCC BAA-894; ZP_03403067.1: *Salmonella enterica* subsp. *enterica* serovar *Typhi* str. AG3; YP_001907453.1: *Erwinia tasmaniensis* Et1/99; YP_001006127.1: *Yersinia enterocolitica* subsp. *enterocolitica* 8081; NP_669181.1: *Yersinia pestis* KIM; NP_993079.1: *Yersinia pestis* biovar *Microtus* str. 91001; YP_070893.1: *Yersinia pseudotuberculosis* IP 32953; YP_651729.1: *Yersinia pestis* Antiqua; YP_647856.1: *Yersinia pestis* Nepal516; YP_001163103.1: *Yersinia pestis* Pestoides F; ZP_01888376.1: *Yersinia pestis* CA88-4125; YP_001400647.1: *Yersinia pseudotuberculosis* IP 31758; YP_001607069.1: *Yersinia pestis* Angola; ZP_02221459.1: *Yersinia pestis*

biovar Orientalis str. F1991016; ZP_02227087.1: *Yersinia pestis* biovar Orientalis str. IP275; ZP_02229339.1: *Yersinia pestis* biovar Antiqua str. E1979001; ZP_02237158.1: *Yersinia pestis* biovar Antiqua str. B42003004; ZP_02307206.1: *Yersinia pestis* biovar Antiqua str. UG05-0454; ZP_02314388.1: *Yersinia pestis* biovar Orientalis str. MG05-1020; ZP_02318544.1: *Yersinia pestis* biovar Mediaevalis str. K1973002; YP_001720525.1: *Yersinia pseudotuberculosis* YPIII; YP_001872871.1: *Yersinia pseudotuberculosis* PB1/+; YP_002346706.1: *Yersinia pestis* CO92; YP_001478346.1: *Serratia proteamaculans* 568; YP_454994.1: *Sodalis glossinidius* str. 'morsitans'; YP_001217052.1: *Vibrio cholerae* O395; YP_002150764.1: *Proteus mirabilis* HI4320; YP_856816.1: *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966; YP_001093925.1: *Shewanella loihica* PV-4; YP_269525.1: *Colwellia psychrerythraea* 34H; NP_718189.1: *Shewanella oneidensis* MR-1; YP_001366642.1: *Shewanella baltica* OS185; YP_001554990.1: *Shewanella baltica* OS195; YP_737796.1: *Shewanella* sp. MR-7; YP_869415.1: *Shewanella* sp. ANA-3; NP_797987.1: *Vibrio parahaemolyticus* RIMD 2210633; YP_001501790.1: *Shewanella pealeana* ATCC 700345; YP_927460.1: *Shewanella amazonensis* SB2B; YP_001474222.1: *Shewanella sediminis* HAW-EB3; YP_963195.1: *Shewanella* sp. W3-18-1; YP_001183724.1: *Shewanella putrefaciens* CN-32; YP_733804.1: *Shewanella* sp. MR-4; YP_001050811.1: *Shewanella baltica* OS155; YP_750914.1: *Shewanella frigidimarina* NCIMB 400; YP_002126297.1: *Alteromonas macleodii* 'Deep ecotype'; YP_130129.1: *Photobacterium profundum* SS9; YP_001445537.1: *Vibrio harveyi* ATCC BAA-1116; YP_661562.1: *Pseudoalteromonas atlantica* T6c; YP_001674584.1: *Shewanella halifaxensis* HAW-EB4; YP_002417027.1: *Vibrio splendidus* LGP32; NP_934438.1: *Vibrio vulnificus* YJ016; NP_761466.1: *Vibrio vulnificus* CMCP6; YP_001760510.1: *Shewanella woodyi* ATCC 51908; YP_002312133.1: *Shewanella piezotolerans* WP3; YP_562727.1: *Shewanella denitrificans* OS217; NP_693123.1: *Oceanobacillus ihyensensis* HTE831; YP_001699747.1: *Lysinibacillus sphaericus* C3-41; YP_001343242.1: *Marinomonas* sp. MWYL1; YP_001141806.1: *Aeromonas salmonicida* subsp. *salmonicida* A449; YP_204661.1: *Vibrio fischeri* ES114; YP_301135.1: *Staphylococcus saprophyticus* subsp. *saprophyticus* ATCC 15305; YP_002156075.1: *Vibrio fischeri* MJ11; NP_244074.1: *Bacillus halodurans* C-125; YP_001559126.1: *Clostridium phytofermentans* ISDg; YP_080250.1: *Bacillus licheniformis* ATCC 14580; YP_253122.1: *Staphylococcus haemolyticus* JCSC1435; YP_261333.1: *Pseudomonas fluorescens* Pf-5; YP_001187407.1: *Pseudomonas mendocina* ymp; YP_001422246.1: *Bacillus amyloliquefaciens* FZB42; YP_001748360.1: *Pseudomonas putida* W619; YP_155660.1: *Idiomarina loihiensis* L2TR; NP_744032.1: *Pseudomonas putida* KT2440; YP_001269146.1: *Pseudomonas putida* F1; YP_236680.1: *Pseudomonas syringae* pv. *syringae* B728a; YP_001667695.1: *Pseudomonas putida* GB-1; YP_188854.1: *Staphylococcus epidermidis* RP62A; YP_002275225.1: *Gluconacetobacter diazotrophicus* PAI 5; YP_804779.1: *Pediococcus pentosaceus* ATCC 25745; NP_470973.1: *Listeria innocua* Clip11262; YP_607257.1: *Pseudomonas entomophila* L48; YP_001860169.1: *Burkholderia phymatum* STM815; NP_465120.1: *Listeria monocytogenes* EGD-e; NP_372242.1: *Staphylococcus aureus* subsp. *aureus* Mu50; NP_374828.1: *Staphylococcus aureus* subsp. *aureus* N315; YP_001247140.1: *Staphylococcus aureus* subsp. *aureus* JH9; YP_001316940.1: *Staphylococcus aureus* subsp. *aureus* JH1; YP_001442294.1: *Staphylococcus aureus* subsp. *aureus* Mu3; YP_002314836.1: *Anoxybacillus flavithermus* WK1; YP_001814701.1: *Exiguobacterium sibiricum* 255-15; NP_646478.1: *Staphylococcus aureus* subsp. *aureus* MW2; YP_043764.1: *Staphylococcus aureus* subsp. *aureus* MSSA476; YP_186602.1: *Staphylococcus aureus* subsp. *aureus* COL; YP_494359.1: *Staphylococcus aureus* subsp. *aureus* USA300; YP_500333.1: *Staphylococcus aureus* subsp. *aureus* NCTC 8325; YP_001332646.1: *Staphylococcus aureus* subsp. *aureus* str. Newman; YP_001575592.1: *Staphylococcus aureus* subsp. *aureus* USA300_TCH1516; ZP_02759966.1: *Staphylococcus aureus* subsp. *aureus* USA300_TCH1516; YP_001487827.1: *Bacillus pumilus* SAFR-032; YP_417047.1: *Staphylococcus aureus* RF122; YP_041183.1: *Staphylococcus aureus* subsp. *aureus* MRSA252; YP_176251.1: *Bacillus clausii* KSM-K16; YP_002349929.1: *Listeria monocytogenes* HCC23; YP_001376537.1: *Bacillus cereus* subsp. *cytotoxis* NVH 391-98; YP_784647.1: *Bordetella avium* 197N; YP_557760.1: *Burkholderia xenovorans* LB400; YP_275769.1: *Pseudomonas syringae* pv. *phaseolicola* 1448A; YP_001118914.1: *Burkholderia vietnamiensis* G4; YP_001782545.1: *Clostridium botulinum* B1 str. Okra; NP_882192.1: *Bordetella pertussis* Tohama I; YP_001894915.1: *Burkholderia phytofirmans* PsJN; YP_001647268.1: *Bacillus weihenstephanensis* KBAB4; NP_882465.1: *Bordetella parapertussis* 12822; YP_002448219.1: *Bacillus cereus* G9842; NP_834359.1: *Bacillus cereus* ATCC 14579; YP_002369448.1:

Bacillus cereus B4264; YP_191278.1: Gluconobacter oxydans 621H; NP_886654.1: Bordetella bronchiseptica RB50; NP_847103.1: Bacillus anthracis str. Ames; YP_030799.1: Bacillus anthracis str. Sterne; NP_981083.1: Bacillus cereus ATCC 10987; YP_038702.1: Bacillus thuringiensis serovar konkukian str. 97-27; YP_002340711.1: Bacillus cereus AH187; YP_085975.1: Bacillus cereus E33L; YP_002453717.1: Bacillus cereus AH820; YP_772921.1: Burkholderia ambifaria AMMD; YP_001843726.1: Lactobacillus fermentum IFO 3956; YP_942902.1: Psychromonas ingrahamii 37; YP_121651.1: Nocardia farcinica IFM 10152; YP_001633478.1: Bordetella petrii DSM 12804; YP_001392189.1: Clostridium botulinum F str. Langeland; YP_795394.1: Lactobacillus brevis ATCC 367; YP_002284153.1: Rhizobium leguminosarum bv. trifolii WSM2304; YP_001255404.1: Clostridium botulinum A str. ATCC 3502; YP_001385170.1: Clostridium botulinum A str. ATCC 19397; YP_001388639.1: Clostridium botulinum A str. Hall; YP_001347703.1: Pseudomonas aeruginosa PA7; NP_103622.1: Mesorhizobium loti MAFF303099; YP_001807739.1: Burkholderia ambifaria MC40-6; YP_001580337.1: Burkholderia multivorans ATCC 17616; YP_002230176.1: Burkholderia cenocepacia J2315; NP_251512.1: Pseudomonas aeruginosa PAO1; YP_459505.1: Erythrobacter litoralis HTCC2594; YP_790351.1: Pseudomonas aeruginosa UCBPP-PA14; YP_002439846.1: Pseudomonas aeruginosa LESB58; YP_001970489.1: Stenotrophomonas maltophilia K279a; YP_617023.1: Sphingopyxis alaskensis RB2256; YP_620554.1: Burkholderia cenocepacia AU 1054; YP_834795.1: Burkholderia cenocepacia HI2424; YP_148653.1: Geobacillus kaustophilus HTA426; YP_442277.1: Burkholderia thailandensis E264; YP_334263.1: Burkholderia pseudomallei 1710b; YP_001067073.1: Burkholderia pseudomallei 1106a; YP_368503.1: Burkholderia sp. 383; YP_109010.1: Burkholderia pseudomallei K96243; YP_993758.1: Burkholderia mallei SAVP1; YP_001028775.1: Burkholderia mallei NCTC 10229; YP_002026852.1: Stenotrophomonas maltophilia R551-3; YP_001764423.1: Burkholderia cenocepacia MC0-3; NP_970249.1: Bdellovibrio bacteriovorus HD100; YP_001059788.1: Burkholderia pseudomallei 668; YP_001081316.1: Burkholderia mallei NCTC 10247; YP_001875000.1: Elusimicrobium minutum Pei191; YP_278248.1: Mycoplasma synoviae 53; YP_001309672.1: Clostridium beijerinckii NCIMB 8052; YP_001576818.1: Lactobacillus helveticus DPC 4571; YP_193206.1: Lactobacillus acidophilus NCFM; YP_199405.1: Xanthomonas oryzae pv. oryzae KACC10331; YP_001915608.1: Xanthomonas oryzae pv. oryzae PXO99A; YP_002509654.1: Halothermothrix orenii H 168; YP_757341.1: Maricaulis maris MCS10; YP_810624.1: Oenococcus oeni PSU-1; YP_244708.1: Xanthomonas campestris pv. campestris str. 8004; NP_635978.1: Xanthomonas campestris pv. campestris str. ATCC 33913; XP_505859.1: Yarrowia lipolytica; NP_964546.1: Lactobacillus johnsonii NCC 533; YP_001261873.1: Sphingomonas wittichii RW1; YP_814311.1: Lactobacillus gasseri ATCC 33323; NP_298520.1: Xylella fastidiosa 9a5c; YP_001775202.1: Xylella fastidiosa M12; YP_001829224.1: Xylella fastidiosa M23; NP_778732.1: Xylella fastidiosa Temecula1; NP_246070.1: Pasteurella multocida subsp. multocida str. Pm70; YP_001997158.1: Chloroherpeton thalassium ATCC 35110; XP_645637.1: Dictyostelium discoideum AX4; XP_456263.1: Kluyveromyces lactis; YP_001300883.1: Bacteroides vulgatus ATCC 8482; YP_812434.1: Lactobacillus delbrueckii subsp. bulgaricus ATCC BAA-365; NP_814796.1: Enterococcus faecalis V583; NP_358362.1: Streptococcus pneumoniae R6; YP_816246.1: Streptococcus pneumoniae D39; YP_001694316.1: Streptococcus pneumoniae Hungary19A-6; YP_001835526.1: Streptococcus pneumoniae CGSP14; YP_002037500.1: Streptococcus pneumoniae G54; NP_345351.1: Streptococcus pneumoniae TIGR4; XP_001385863.2: Pichia stipitis CBS 6054; NP_985662.1: Ashbya gossypii ATCC 10895; YP_759928.1: Hyphomonas neptunium ATCC 15444; YP_002123053.1: Streptococcus equi subsp. zooepidemicus MGCS10565; YP_002131343.1: Phenylbacterium zucineum HLK1; XP_001683506.1: Leishmania major strain Friedlin; NP_012854.1: Saccharomyces cerevisiae; XP_460154.1: Debaryomyces hansenii CBS767; XP_001801028.1: Phaeosphaeria nodorum SN15; YP_001201197.1: Streptococcus suis 98HAH33; XP_001484046.1: Pichia guilliermondii ATCC 6260; NP_274882.1: Neisseria meningitidis MC58; YP_001302571.1: Parabacteroides distasonis ATCC 8503; YP_002000650.1: Neisseria gonorrhoeae NCCP11945; YP_001682978.1: Caulobacter sp. K31; YP_002342048.1: Neisseria meningitidis Z2491; YP_207199.1: Neisseria gonorrhoeae FA 1090; XP_844799.1: Trypanosoma brucei TREU927; YP_001598496.1: Neisseria meningitidis 053442; YP_974447.1: Neisseria meningitidis FAM18; XP_001465858.1: Leishmania infantum JPCM5; YP_001280815.1: Psychrobacter sp. PRwf-1; NP_664843.1: Streptococcus pyogenes MGAS315; NP_802083.1: Streptococcus pyogenes SSI-1;

YP_280574.1: *Streptococcus pyogenes* MGAS6180; YP_001128311.1: *Streptococcus pyogenes* str. Manfredo; YP_002286076.1: *Streptococcus pyogenes* NZ131; XP_390238.1: *Gibberella zeae* PH-1; XP_001565241.1: *Leishmania braziliensis* MHOM/BR/75/M2904; YP_596891.1: *Streptococcus pyogenes* MGAS9429; YP_600772.1: *Streptococcus pyogenes* MGAS2096; YP_001351938.1: *Janthinobacterium* sp. Marseille; XP_001910908.1: *Podospira anserina* DSM 980; XP_446236.1: *Candida glabrata* CBS 138; NP_607473.1: *Streptococcus pyogenes* MGAS8232; YP_060405.1: *Streptococcus pyogenes* MGAS10394; YP_602707.1: *Streptococcus pyogenes* MGAS10750; NP_269471.1: *Streptococcus pyogenes* M1 GAS; YP_282476.1: *Streptococcus pyogenes* MGAS5005; YP_598792.1: *Streptococcus pyogenes* MGAS10270; YP_099090.1: *Bacteroides fragilis* YCH46; NP_421082.1: *Caulobacter crescentus* CB15; YP_001034962.1: *Streptococcus sanguinis* SK36; YP_211509.1: *Bacteroides fragilis* NCTC 9343; XP_001940256.1: *Pyrenophora tritici-repentis* Pt-1C-BFP; YP_001450196.1: *Streptococcus gordonii* str. Challis substr. CH1; XP_680937.1: *Aspergillus nidulans* FGSC A4; XP_001265696.1: *Neosartorya fischeri* NRRL 181; XP_367405.1: *Magnaporthe grisea* 70-15; YP_001033719.1: *Lactococcus lactis* subsp. *cremoris* MG1363; YP_812051.1: *Lactococcus lactis* subsp. *cremoris* SK11; NP_810723.1: *Bacteroides thetaiotaomicron* VPI-5482; YP_141537.1: *Streptococcus thermophilus* CNRZ1066; NP_721920.1: *Streptococcus mutans* UA159; YP_820528.1: *Streptococcus thermophilus* LMD-9; XP_001246405.1: *Coccidioides immitis* RS; YP_612063.1: *Silicibacter* sp. TM1040; YP_001727726.1: *Leuconostoc citreum* KM20; YP_444835.1: *Salinibacter ruber* DSM 13855; NP_943356.1: *Klebsiella pneumoniae*; YP_631535.1: *Myxococcus xanthus* DK 1622; YP_809822.1: *Oenococcus oeni* PSU-1; NP_687842.1: *Streptococcus agalactiae* 2603V/R; YP_329573.1: *Streptococcus agalactiae* A909; XP_001692446.1: *Chlamydomonas reinhardtii*; YP_511991.1: *Jannaschia* sp. CCS1; XP_719233.1: *Candida albicans* SC5314; NP_735295.1: *Streptococcus agalactiae* NEM316; YP_168601.1: *Silicibacter pomeroyi* DSS-3; YP_001169098.1: *Rhodobacter sphaeroides* ATCC 17025; XP_001216799.1: *Aspergillus terreus* NIH2624; YP_437091.1: *Hahella chejuensis* KCTC 2396; XP_001390389.1: *Aspergillus niger* CBS 513.88; XP_001550500.1: *Botryotinia fuckeliana* B05.10; XP_001273080.1: *Aspergillus clavatus* NRRL 1; YP_001042044.1: *Rhodobacter sphaeroides* ATCC 17029; YP_916587.1: *Paracoccus denitrificans* PD1222; XP_961529.1: *Neurospora crassa* OR74A; YP_351544.1: *Rhodobacter sphaeroides* 2.4.1; XP_001588803.1: *Sclerotinia sclerotiorum* 1980; XP_762195.1: *Ustilago maydis* 521; XP_001537789.1: *Ajellomyces capsulatus* NAM1; YP_001198994.1: *Streptococcus suis* 05ZYH33; NP_390842.1: *Bacillus subtilis* subsp. *subtilis* str. 168.

C94/C118 signature (20 sequences, 5.6%):

NP_288269.1: *Escherichia coli* O157:H7 EDL933; YP_001469780.1: *Thermotoga lettingae* TMO; NP_623333.1: *Thermoanaerobacter tengcongensis* MB4; YP_001567589.1: *Petrogoga mobilis* SJ95; NP_228329.1: *Thermotoga maritima* MSB8; YP_001244004.1: *Thermotoga petrophila* RKU-1; YP_001738457.1: *Thermotoga* sp. RQ2; YP_591728.1: *Acidobacteria bacterium* Ellin345; YP_001193798.1: *Flavobacterium johnsoniae* UW101; YP_001296902.1: *Flavobacterium psychrophilum* JIP02/86; XP_001439866.1: *Paramecium tetraurelia* strain d4-2; XP_001438779.1: *Paramecium tetraurelia* strain d4-2; YP_001319456.1: *Alkaliphilus metalliredigens* QYMF; XP_001022412.1: *Tetrahymena thermophila* SB210; YP_001306613.1: *Thermosiphon melanesiensis* BI429; YP_002335308.1: *Thermosiphon africanus* TCF52B; NP_394308.1: *Thermoplasma acidophilum* DSM 1728; YP_862969.1: *Gramella forsetii* KT0803; YP_023758.1: *Picrophilus torridus* DSM 9790; NP_111339.1: *Thermoplasma volcanium* GSS1; YP_001380982.1: *Anaeromyxobacter* sp. Fw109-5.

C84/C118 signature (5 sequences, 1.4%):

YP_806472.1: *Lactobacillus casei* ATCC 334; YP_001987415.1: *Lactobacillus casei* BL23; NP_785802.1: *Lactobacillus plantarum* WCFS1; YP_818384.1: *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293; YP_001621174.1: *Acholeplasma laidlawii* PG-8A

LETTER

Substrate Binding in Free Methionine-*R*-sulfoxide Reductase

Recently, the group of Chi and co-workers published the crystal structures of the reduced, substrate-bound, and oxidized forms of fRMsR from *Staphylococcus aureus* (1). In parallel, our group proposed a catalytic mechanism of free methionine-*R*-sulfoxide (Met-*R*-O) reductase (fRMsR) from *Neisseria meningitidis* based on crystal structure and biochemical data (2). Superimposition of the crystal structures of the substrate-bound forms of *N. meningitidis* and *S. aureus* enzymes show significant differences; specifically, the sulfoxide in the *S. aureus* fRMsR occupies the position of the carboxylate in the *N. meningitidis* fRMsR and vice versa; in addition, the conformation of the flap, which contains the catalytic Cys¹¹⁸, is not superimposable (Fig. 1).

Four arguments favor a competent positioning of Met-*R*-O in the active site of the crystal structure of the *N. meningitidis* fRMsR. Firstly, the presence of an oxyanion hole rather excludes the productive binding of the sulfoxide function within the oxyanion hole as the sulfoxide reduction step requires protonation of the oxygen of the sulfurane transition state (3). Secondly, the *S. aureus* fRMsR complex was obtained with C84S fRMsR. Therefore, a single turnover catalytic event was expected to occur in the crystal, unless a noncompetent binary complex has been formed between fRMsR and Met-*R*-O. Thirdly, in the structure of the *N. meningitidis* fRMsR, the positioning of the catalytic cysteine is appropriate to allow a nucleophilic attack on the sulfur atom of the sulfoxide. Fourthly, all the residues of the *N. meningitidis* active site impose constraints geometrically on the L-Met-*R*-O, which is tightly bound.

Arnaud Gruez, Marouane Libiad, Sandrine Boschi-Muller, and Guy Branlant¹

AREMS, UMR CNRS-UHP 7214, Nancy Université, Faculté des Sciences et Technologies, Bld des Aiguillettes, BP 70239, 54506 Vandoeuvre-les-Nancy, France

1. Bong, S. M., Kwak, G. H., Moon, J. H., Lee, K. S., Kim, H. S., Kim, H. Y., and Chi, Y. M. (2010) *J. Biol. Chem.* **285**, 25044–25052

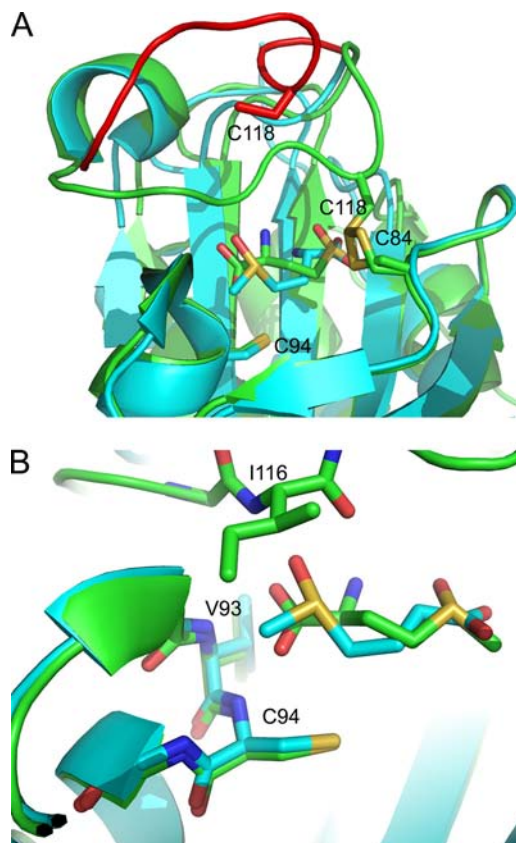


FIGURE 1. Ribbon representation of fRMsR from *N. meningitidis* (in green) and *S. aureus* (in cyan) in complex with L-Met-*R*-O. A, the conformation of a portion of the flap (residues 107–124, according to the numbering of the sequence of *Escherichia coli* fRMsR), between amino acids 113 and 120, which contains the catalytic Cys¹¹⁸, is not superimposable (in red for the structure of fRMsR of *S. aureus*). The side chains of Cys⁸⁴, Cys⁹⁴, and Cys¹¹⁸ are shown in stick representation. The L-Met-*R*-O is represented in stick with the oxygen atoms in red, nitrogen in blue, sulfur atoms in yellow, and carbon atoms in green (*N. meningitidis*) or in cyan (*S. aureus*). B, the main chain peptidic nitrogens of invariant residues Val⁹³, Cys⁹⁴, and Ile¹¹⁶ are shown in stick. The role of the oxyanion hole is to stabilize the binding of the carboxylate of the Met-*R*-O, whose negative charge is completely delocalized.

2. Gruez, A., Libiad, M., Boschi-Muller, S., and Branlant G. (2010) *J. Biol. Chem.* **285**, 25033–25043
3. Boschi-Muller, S., Gand, A., and Branlant, G. (2008) *Arch. Biochem. Biophys.* **474**, 266–273

DOI 10.1074/jbc.L110.103119

¹E-mail: Guy.Branlant@maem.uhp-nancy.fr

REFERENCES

BIBLIOGRAPHIQUES

Aachmann FL, Kwak GH, Del Conte R, Kim HY, Gladyshev VN, Dikiy A.

Structural and biochemical analysis of mammalian methionine sulfoxide reductase B2. *Proteins*, 2011, 79 : 3123-31.

Antoine M, Gand A, Boschi-Muller S, Branlant G.

Characterization of the amino acids from *Neisseria meningitidis* MsrA involved in the chemical catalysis of the methionine sulfoxide reduction step.

J Biol Chem., 2006, 281 : 39062-70.

Aravind L, Ponting CP.

The GAF domain: an evolutionary link between diverse phototransducing proteins.

Trends Biochem Sci., 1997, 22 : 458-9.

Asmus KD.

Stabilization of oxidized sulfur centers in organic sulfides. Radical cations and odd-electron-sulfur-sulfur bonds.

Acc Chem Res, 1979, 12 : 436-42.

Badger J, Sauder JM, Adams JM, Antonysamy S, Bain K, Bergseid MG, Buchanan SG, Buchanan MD, Batiyenko Y, Christopher JA, Emtage S, Eroshkina A, Feil I, Furlong EB, Gajiwala KS, Gao X, He D, Hendle J, Huber A, Hoda K, Kearins P, Kissinger C, Laubert B, Lewis HA, Lin J, Loomis K, Lorimer D, Louie G, Maletic M, Marsh CD, Miller I, Molinari J, Muller-Dieckmann HJ, Newman JM, Noland BW, Pagarigan B, Park F, Peat TS, Post KW, Radojicic S, Ramos A, Romero R, Rutter ME, Sanderson WE, Schwinn KD, Tresser J, Winhoven J, Wright TA, Wu L, Xu J, Harris TJ.

Structural analysis of a set of proteins resulting from a bacterial genomics project.

Proteins, 2005, 60 : 787-96.

Bamezai S, Banez MA, Breslow E.

Structural and functional changes associated with modification of the ubiquitin methionine *Biochemistry*, 1990, 29 : 5389-96.

Bar-Noy S, Moskovitz J.

Mouse methionine sulfoxide reductase B: effect of selenocysteine incorporation on its activity and expression of the seleno-containing enzyme in bacterial and mammalian cells.

Biochem. Biophys. Res. Commun., 2002, 297 : 956-61.

Bartlett RK, Bieber Urbauer RJ, Anbanandam A, Smallwood HS, Urbauer JL, Squier TC.

Oxidation of Met144 and Met145 in calmodulin blocks calmodulin dependent activation of the plasma membrane Ca-ATPase.

Biochemistry, 2003, 42 : 3231-8.

Bong SM, Kwak GH, Moon JH, Lee KS, Kim HS, Kim HY, Chi YM.

Structural and kinetic analysis of free methionine-R-sulfoxide reductase from *Staphylococcus aureus*: conformational changes during catalysis and implications for the catalytic and inhibitory mechanisms.

J. Biol. Chem., 2010, 285 : 25044-52.

Boschi-Muller S, Azza S, Sanglier-Cianferani S, Talfournier F, Van Dorsselear A, Branlant G.

A sulfenic acid enzyme intermediate is involved in the catalytic mechanism of peptide methionine sulfoxide reductase from *Escherichia coli*.

J. Biol. Chem., 2000, 275 : 35908-13

Boschi-Muller S, Gand A, Branlant G.

The methionine sulfoxide reductases: Catalysis and substrate specificities.

Arch. Biochem. Biophys., 2008, 474 : 266-73.

Boschi-Muller S, Olry A, Antoine M, Branlant G.

The enzymology and biochemistry of methionine sulfoxide reductases.

Biochim. Biophys. Acta., 2005, 1703 : 231-8

Breivik AS, Aachmann FL, Sal LS, Kim HY, Del Conte R, Gladyshev VN, Dikoy A.

¹H, ¹⁵N and ¹³C NMR assignments of mouse methionine sulfoxide reductase B2.

Biomol. NMR. Assign., 2008, 2 :199-201.

Brennan LA, Lee W, Giblin FJ, David LL, Kantorow M.

Methionine sulfoxide reductase A (MsrA) restores alpha-crystallin chaperone activity lost upon methionine oxidation.

Biochim. Biophys. Acta., 2009, 1790 : 1665-72.

Bruder S, Schultz A, Schultz JE.

Characterization of the tandem GAF domain of human phosphodiesterase 5 using a cyanobacterial adenylyl cyclase as a reporter enzyme.

J. Biol. Chem., 2006, 281 : 19969-76.

Cadenas E, Davie KJA.

Mitochondrial free radical production, oxidative stress, and aging.

Free. Radic. Biol., 2000, 29 : 222-30.

Carella M, Ohlenschläger O, Ramachandran R, Görlach M.

¹H, ¹³C and ¹⁵N resonance assignment of a zinc-binding methionine sulfoxide reductase type-B from the thermophilic *archaeobacterium* Methanothermobacter thermoautotrophicus.

Biomol. NMR. Assign., 2010, 4 : 93-5.

Chu JW, Trout BL.

On the mechanisms of oxidation of organic sulfides by H₂O₂ in aqueous solutions.

J. Am. Chem. Soc., 2004, 126 : 900-8.

Chu JW, Yin J, Wang DIC, Trout BL.

Molecular dynamics simulations and oxidation rates of methionine residues of granulocyte colony-stimulating factor at different pH values.

Biochemistry, 2004, 43 : 1019-29.

Ciorba MA, Heinemann SH, Weissbach H, Brot N, Hoshi T.

Modulation of potassium channel function by methionine oxidation and reduction.

Proc. Natl. Acad. Sci., 1997, 94 : 9932-37.

Dalle-Donne I, Rossi R, Giustarini D, Gagliano N, Di Simplicio P, Colombo R, Milzani

A.

Methionine oxidation as a major cause of the functional impairment of oxidized actin.

Free Radic. Biol. Med., 2002, 32 : 927-37.

Denkel LA, Horst SA, Rouf SF, Kitowski V, Böhm OM, Rhen M, Jäger T, Bange FC.

Methionine sulfoxide reductases are essential for virulence of *Salmonella typhimurium*.

PLoS One., 2011, 6 : e26974.

Douglas T, Daniel DS, Parida BK, Jagannath C, Dhandayuthapani S.

Methionine sulfoxide reductase A (MsrA) deficiency affects the survival of *Mycobacterium smegmatis* within macrophages.

J. Bacteriol., 2004, 186 : 3590-8.

Duescher RJ, Lawton MP, Philpot RM, Elfarra AA.

Flavin-containing monooxygenase (FMO)-dependent metabolism of methionine and evidence for FMO3 being the major FMO involved in methionine sulfoxidation in rabbit liver and kidney microsomes.

J Biol Chem., 1994, 269 : 17525-30.

Erickson JR, Joiner ML, Guan X, Kutschke W, Yang J, Oddis CV, Bartlett RK, Lowe JS, O'Donnell SE, Aykin-Burns N, Zimmerman MC, Zimmerman K, Ham AJ, Weiss RM, Spitz DR, Shea MA, Colbran RJ, Mohler PJ, Anderson ME.

A dynamic pathway for calcium-independent activation of CaMKII by methionine oxidation.

Cell, 2008, 13 : 462-74.

Etienne F, Spector D, Brot N, Weissbach H.

A methionine sulfoxide reductase in *Escherichia coli* that reduces the R enantiomer of methionine sulfoxide.

Biochem. Biophys. Res. Commun., 2003, 300 : 378-82

Ezraty B, Aussel L, Barras F.

Methionine sulfoxide reductases in prokaryotes.

Biochim. Biophys. Acta., 1703 : 221-29.

Ferrington DA, Sun H, Murray KK, Costa J, Williams TD, Bigelow DJ, Squier TC.

Selective degradation of oxidized calmodulin by the 20 S proteasome.
J. Biol. Chem., 2001, 276 : 937-43.

Fink TL, Francis SH, Beasley A, Grimes KA, Corbin JD.

Expression of an active, monomeric catalytic domain of the cGMP-binding cGMP-specific phosphodiesterase (PDE5).
J. Biol. Chem., 1999, 274 : 34613-20.

Galperin MY.

Structural classification of bacterial response regulators: diversity of output domains and domain combinations.
J. Bacteriol., 2006, 188 : 4169-82.

Gand A, Antoine M, Boschi-Muller S, Branlant G.

Characterization of the amino acids involved in substrate specificity of methionine sulfoxide reductase A.
J. Biol. Chem., 2007, 282 : 20484-91.

Gao, J, Yao, Y, et Squier, TC.

Oxidatively modified calmodulin binds to the plasma membrane Ca-ATPase in a nonproductive and conformationally disordered complex.
Biophys. J., 2001, 80 : 1791-801.

Gellman SH.

On the role of methionine residues in the sequence-independent recognition of nonpolar protein surfaces.
Biochemistry, 1991, 30 : 6633-6.

Gill SC, Von Hippel PH.

Calculation of protein extinction coefficients from amino acid sequence data.
Anal. Biochem., 1989, 182 : 319-26.

Glaser CB, Yamin G, Uversky VN, Fink AL.

Methionine oxidation, alpha-synuclein and Parkinson's disease.
Biochim. Biophys. Acta., 2005, 1703 : 157-69.

Granovsky AE, Natochin M, McEntaffer RL, Haik TL, Francis SH, Corbin JD, Artemyev NO.

Probing domain functions of chimeric PDE6alpha'/PDE5 cGMP-phosphodiesterase.
J. Biol. Chem., 1998, 273 : 24485-90.

Gross-Langenhoff M, Stenzl A, Altenberend F, Schultz A, Schultz JE.

The properties of phosphodiesterase 11A4 GAF domains are regulated by modifications in its N-terminal domain.
FEBS J., 2008, 275 : 1643-50.

Gruez A, Libiad M, Boschi-Muller S, Branlant G.

Structural and biochemical characterization of free methionine-R-sulfoxide reductase from *Neisseria meningitidis*.
J. Biol. Chem., 2010a, 285 : 25033-43.

Gruez A, Libiad M, Boschi-Muller S, Branlant G.

Substrate binding in free methionine-R-sulfoxide reductase.
Biol. Chem., 2010, 285.

Halliwel B.

Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life.
Plant. Physiol., 2006, 141 : 312-22.

Halliwel B, Gutteridge JMC.

Free Radicals in Biology and Medicine,
Oxford University Press., 1999, Ed 4.

Härndahl U, Kokke BP, Gustavsson N, Linse S, Berggren K, Tjerneld F, Boelens WC, Sundby C.

The chaperone-like activity of a small heat shock protein is lost after sulfoxidation of conserved methionines in a surface-exposed amphipathic alpha-helix.
Biochim. Biophys. Acta., 2001, 1545 : 227-37.

Heikaus CC, Stout JR, Sekharan MR, Eakin CM, Rajagopal P, Brzovic PS, Beavo JA, Klevit RE.

Solution structure of the cGMP binding GAF domain from phosphodiesterase 5: insights into nucleotide specificity, dimerization, and cGMP-dependent conformational change.
J. Biol. Chem., 2008, 283 : 22749-59.

Ho YS, Burden LM, Hurley JH.

Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor.
EMBO J., 2000, 19 : 5288-99.

Hou L, Kang I, Marchant RE, Zagorski MG.

Methionine 35 oxidation reduces fibril assembly of the amyloid abeta-(1-42) peptide of Alzheimer's disease.
J. Biol. Chem., 2002, 277 : 40173-6.

Huang D, Hinds TR, Martinez SE, Doneanu C, Beavo JA.

Molecular determinants of cGMP binding to chicken cone photoreceptor phosphodiesterase.
J. Biol. Chem., 2004, 279 : 48143-51.

Jacob C, Knight I, Winyard PG.

Aspects of the biological redox chemistry of cysteine: from simple redox responses to sophisticated signalling pathways.
Biol. Chem., 2006, 387 : 1385-97.

Kanacher T, Schultz A, Linder JU, Schultz JE.

A GAF-domain-regulated adenylyl cyclase from *Anabaena* is a self-activating cAMP switch.
EMBO J., 2002, 21 : 3672-80.

Kantorow M, Hawse JR, Cowell TL, Benhamed S, Pizarro GO, Reddy VN, Hejtmancik JF.

Methionine sulfoxide reductase A is important for lens cell viability and resistance to oxidative stress.
Proc. Natl. Acad. Sci., 2004, 101, 9654-9.

Katayama M, Ohmori M.

Isolation and characterization of multiple adenylyl cyclase genes from the *cyanobacterium* *Anabaena* sp. strain PCC 7120.
J. Bacteriol., 1997, 179 : 3588-93.

Kim HY, Gladyshev VN.

Characterization of mouse endoplasmic reticulum methionine-R-sulfoxide reductase.
Biochem. Biophys. Res. Commun., 2004a, 320, 1277-83.

Kim HY, Kim JR.

Thioredoxin as a reducing agent for mammalian methionine sulfoxide reductases B lacking resolving cysteine.
Biochem Biophys Res Commun., 2008, 371 : 490-4.

Kim MJ, Lee BC, Jeong J, Lee KJ, Hwang KY, Gladyshev VN, Kim HY.

Tandem use of selenocysteine: adaptation of a selenoprotein glutaredoxin for reduction of selenoprotein methionine sulfoxide reductase.
Mol. Microbiol., 2011, 79 : 1194-203.

Kwak GH, Kim MJ, Kim HY.

Cysteine-125 is the catalytic residue of *Saccharomyces cerevisiae* free methionine-R-sulfoxide reductase.
Biochem Biophys Res Commun., 2010, 395 : 412-5.

Kwak GH, Lim DH, Han JY, Lee YS, Kim HY.

Methionine sulfoxide reductase B3 protects from endoplasmic reticulum stress in *Drosophila* and in mammalian cells.
Biochem Biophys Res Commun., 2012, 420 : 130-5.

Le DT, Lee BC, Marino SM, Zhang Y, Fomenko DE, Kaya A, Hacıoglu E, Kwak GH, Koc A, Kim HY, Gladyshev VN.

Functional analysis of free methionine-R-sulfoxide reductase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 2009, 284 : 4354-64.

Lee TH, Kim HY.

An anaerobic bacterial MsrB model reveals catalytic mechanisms, advantages, and disadvantages provided by selenocysteine and cysteine in reduction of methionine-R-sulfoxide.

Arch. Biochem. Biophys., 2008, 478 : 175-80.

Levine RL, Berlett BS, Moskovitz J, Mosoni L, et Stadtman ER.

Methionine residues may protect proteins from critical oxidative damage.

Mech. Ageing Dev., 1999, 107, 323-32.

Levine RL, Moskovitz J, Stadtman ER.

Oxidation of methionine in proteins: roles in antioxidant defense and cellular regulation.

IUBMB Life., 2000, 50 : 301-7.

Levine RL, Mosoni L, Berlett BS, Stadtman ER.

Methionine residues as endogenous antioxidants in proteins.

Proc. Natl. Acad. Sci., 1996, 93 : 15036-40.

Lim DH, Han JY, Kim JR, Lee YS, Kim HY.

Methionine sulfoxide reductase B in the endoplasmic reticulum is critical for stress resistance and aging in *Drosophila*.

Biochem. Biophys. Res. Commun., 2012, 419 : 20-6.

Lim JC, You Z, Kim G, Levine RL.

Methionine sulfoxide reductase A is a stereospecific methionine oxidase.

Proc Natl Acad Sci U S A., 2011, 108 : 10472-7.

Lin Z, Johnson LC, Weissbach H, Brot N, Lively MO, Lowther WT.

Free methionine-(R)-sulfoxide reductase from *Escherichia coli* reveals a new GAF domain function.

Proc. Natl. Acad. Sci., 2007, 104 : 9597-9602.

Linder JU, Bruder S, Schultz A, Schultz JE.

Changes in purine specificity in tandem GAF chimeras from cyanobacterial cyaB1 adenylate cyclase and rat phosphodiesterase 2.

FEBS J., 2007, 274 : 1514-23.

Little R, Dixon R.

The amino-terminal GAF domain of *Azotobacter vinelandii* NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions.

J. Biol. Chem., 2003, 278 : 28711-8.

Lowther WT, Weissbach H, Etienne F, Brot N, Matthews BW.

The mirrored methionine sulfoxide reductases of *Neisseria gonorrhoeae* pilB.

Nat. Struct. Biol., 2002, 9 : 348-52.

Marchetti MA, Pizarro GO, Sagher D, Deamicis C, Brot N, Hejtmancik JF, Weissbach H, Kantorow M.

Methionine sulfoxide reductases B1, B2, and B3 are present in the human lens and confer oxidative stress resistance to lens cells.

Invest. Ophthalmol. Vis. Sci., 2005, 46 : 2107-12.

Martinez-Argudo I, Little R, Shearer N, Johnson P, Dixon R.

Nitrogen fixation: key genetic regulatory mechanisms.

Biochem. Soc. Trans., 2005, 33 : 152-6.

Martinez SE, Bruder S, Schultz A, Zheng N, Schultz JE, Beavo JA, Linder JU.

Crystal structure of the tandem GAF domains from a cyanobacterial adenylyl cyclase: modes of ligand binding and dimerization.

Proc. Natl. Acad. Sci., 2005, 102 : 3082-7.

Martinez SE, Wu AY, Glavas NA, Tang XB, Turley S, Hol WG, Beavo JA.

The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding.

Proc. Natl. Acad. Sci., 2002, 99 : 13260-65.

McAllister-Lucas LM, Haik TL, Colbran JL, Sonnenburg WK, Seger D, Turko IV, Beavo JA, Francis SH, Corbin JD.

An essential aspartic acid at each of two allosteric cGMP-binding sites of a cGMP-specific phosphodiesterase.

J. Biol. Chem., 1995, 270 : 30671-9.

Miles AM, Smith RL.

Functional methionines in the collagen/gelatin binding domain of plasma fibronectin: effects of chemical modification by chloramine T.

Biochemistry, 1993, 32 : 8168-78.

Miller BL, Kuczera K, Schoneich C.

One-electron photooxidation of N-methionyl peptides. Mechanism of sulfoxide and azasulfonium diastereomer formation through reaction of sulfide radical cation complexes with oxygen or superoxide:

J. Am. Chem. Soc., 1998, 120 : 3345-56.

Miller BL, Williams TD, Schöneich C.

Mechanism of Sulfoxide Formation through Reaction of Sulfur Radical Cation Complexes with Superoxide or Hydroxide Ion in Oxygenated Aqueous Solution.

J. Am. Chem. Soc., 1996, 118 : 11014-25.

Moskovitz J.

Methionine sulfoxide reductases: ubiquitous enzymes involved in antioxidant defense, protein regulation, and prevention of aging-associated diseases.

Biochim. Biophys. Acta., 1703 : 213-19.

Moskovitz J, Berlett BS, Poston JM, Stadtman ER.

The yeast peptide-methionine sulfoxide reductase functions as an antioxidant *in vivo*.

Proc. Natl. Acad. Sci., 1997, 94 : 9585-9.

Moskovitz J, Rahman MA, Strassman J, Yancey SO, Kushner SR, Brot N, Weissbach H.

Escherichia coli peptide methionine sulfoxide reductase gene: regulation of expression and role in protecting against oxidative damage.

J. Bacteriol., 1995, 177 : 502-7

Moskovitz J, Singh VK, Requena J, Wilkinson BJ, Jayaswal RK, Stadtman ER.

Purification and characterization of methionine sulfoxide reductases from mouse and *Staphylococcus aureus* and their substrate stereospecificity.

Biochem. Biophys. Res. Commun., 2002, 290 : 62-5.

Mössner E, Huber-Wunderlich M, Glockshuber R.

Characterization of *Escherichia coli* thioredoxin variants mimicking the active-sites of other thiol/disulfide oxidoreductases.

Protein Sci., 1998, 7 : 1233-44.

Mulrooney SB.

Application of a single-plasmid vector for mutagenesis and high-level expression of thioredoxin reductase and its use to examine flavin cofactor incorporation.

Protein Expr. Purif., 1997, 9 : 372-8.

Muradov H, Boyd KK, Artemyev NO.

Structural determinants of the PDE6 GAF A domain for binding the inhibitory gamma-subunit and noncatalytic cGMP.

Vision Res., 2004; 44 : 2437-44.

Neiers F.

Mécanisme, catalyse et spécificité structurale des méthionine sulfoxyde réductase de classe B et la protéine PilB de *Neisseria meningitidis*.

Manuscrit de thèse, Soutenance 2007.

Neiers F, Sonkaria S, Olry A, Boschi-Muller S, Branlant G.

Characterization of the amino acids from *Neisseria meningitidis* methionine sulfoxide reductase B involved in the chemical catalysis and substrate specificity of the reductase step.

J. Biol. Chem., 2007, 282 : 32397-405

Olry A, Boschi-Muller S, Branlant G.

Kinetic characterization of the catalytic mechanism of methionine sulfoxide reductase B from *Neisseria meningitidis*.
Biochemistry, 2004, 43 : 11616-22.

Olry A, Boschi-Muller S, Marraud M, Sanglier-Cianferani S, Van Dorsselear A, Branlant G.

Characterization of the methionine sulfoxide reductase activities of PILB, a probable virulence factor from *Neisseria meningitidis*.
J. Biol. Chem., 2002, 277 : 12016-22.

Pandit J, Forman MD, Fennell KF, Dillman KS, Menniti FS.

Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the X-ray structure of a near full-length construct.
Proc. Natl. Acad. Sci., 2009, 106 : 18225-30.

Pascual I, Larrayoz IM, Campos MM, Rodriguez IR.

Methionine sulfoxide reductase B2 is highly expressed in the retina and protects retinal pigmented epithelium cells from oxidative damage.
Exp. Eye Res., 2010, 90 : 420-8.

Perrin D, Koppenol WH.

The quantitative oxidation of methionine to methionine sulfoxide by peroxynitrite.
Arch. Biochem. Biophys., 2000, 377 : 266-72.

Premakumar R, Jacobson MR, Loveless TM, Bishop PE.

Characterization of transcripts expressed from nitrogenase-3 structural genes of *Azotobacter vinelandii*.
Can. J. Microbiol., 1992, 38 : 929-36.

Ranaivoson FM, Antoine M, Kauffmann B, Boschi-Muller S, Aubry A, Branlant G, Favier F.

A structural analysis of the catalytic mechanism of methionine sulfoxide reductase A from *Neisseria meningitidis*.
J. Mol. Biol., 2008, 377 : 268-80

Ripp SL, Itagaki K, Philpot RM, Elfarra AA.

Methionine S-oxidation in human and rabbit liver microsomes: evidence for a high-affinity methionine S-oxidase activity that is distinct from flavin-containing monooxygenase 3.
Arch. Biochem. Biophys., 1999, 367 : 322-32.

Robinet JJ, Dokainish HM, Paterson DJ, Gauld JW.

A sulfonium cation intermediate in the mechanism of methionine sulfoxide reductase B: a DFT study.

J. Phys. Chem. B., 2011, 115 : 9202-12.

Rouhier N, Kauffmann B, Tete-Favier F, Palladino P, Gans P, Branlant G, Jacquot JP, Boschi-Muller S.

Functional and structural aspects of poplar cytosolic and plastidial type a methionine sulfoxide reductases.

J. Biol. Chem., 2007, 282 : 3367-78

Ruan H, Tang XD, Chen ML, Joiner ML, Sun G, Brot N, Weissbach H, Heinemann SH, Iverson L, Wu CF, Hoshi T.

High-quality life extension by the enzyme peptide methionine sulfoxide reductase.

Proc. Natl. Acad. Sci., 2002, 99 : 2748-53.

Rybalkin SD, Rybalkina IG, Shimizu-Albergine M, Tang XB, Beavo JA.

PDE5 is converted to an activated state upon cGMP binding to the GAF A domain.

EMBO J., 2003, 22 : 469-78.

Sal LS, Aachmann FL, Kim HY, Gladyshev VN, Dikiy A.

NMR assignments of ¹H, ¹³C and ¹⁵N spectra of methionine sulfoxide reductase B1 from *Mus musculus*.

Biomol. NMR Assign., 2007, 1 :131-3.

Salmon AB, Pérez VI, Bokov A, Jernigan A, Kim G, Zhao H, Levine RL, Richardson A.

Lack of methionine sulfoxide reductase A in mice increases sensitivity to oxidative stress but does not diminish life span.

FASEB J., 2009, 23 : 3601-8.

Sarma BK, Mugesh G.

Redox regulation of protein tyrosine phosphatase 1B (PTP1B): a biomimetic study on the unexpected formation of a sulfenyl amide intermediate.

J. Am. Chem. Soc., 2007, 129 :8872-81.

Schoneich C.

Methionine oxidation by reactive oxygen species: reaction mechanisms and relevance to Alzheimer's disease.

Biochim. Biophys. Acta., 2005, 1703 : 111-19

Schoneich C, Aced A, Asmus KD.

Mechanism of Oxidation of Aliphatic Thioethers to Sulfoxides by Hydroxyl Radicals. The Importance of Molecular Oxygen.

J. Am. Chem. Soc., 1993, 115 : 11376-83.

Schoneich C, Bobrowski K.

Reaction of Hydroxysulfuranyl Radical with Molecular-Oxygen - Electron-Transfer Vs Addition.

J. Phys. Chem., 1994, 98 : 12613-20.

Schultz JE.

Structural and biochemical aspects of tandem GAF domains.

Handb. Exp. Pharmacol., 2009, 191 :93-109.

Shchedrina VA, Everley RA, Zhang Y, Gygi SP, Hatfield DL, Gladyshev VN.

Selenoprotein K binds multiprotein complexes and is involved in the regulation of endoplasmic reticulum homeostasis.

J. Biol. Chem., 2011, 286 :42937-48.

Singh VK, Moskovitz J.

Multiple methionine sulfoxide reductase genes in *Staphylococcus aureus*: expression of activity and roles in tolerance of oxidative stress.

Microbiology, 2003, 149 : 2739-47.

Skaar EP, Tobiason DM, Quick J, Judd RC, Weissbach H, Etienne F, Brot N, Seifert HS.

The outer membrane localization of the *Neisseria gonorrhoeae* MsrA/B is involved in survival against reactive oxygen species.

Proc. Natl. Acad. Sci., 2002, 99 : 10108-13.

Smetana JH, Zanchin NI.

Interaction analysis of the heterotrimer formed by the phosphatase 2A catalytic subunit, alpha4 and the mammalian ortholog of yeast Tip41 (TIPRL).

FEBS J., 2007, 274 :5891-904.

Sopory S, Balaji S, Srinivasan N, Visweswariah SS.

Modeling and mutational analysis of the GAF domain of the cGMP-binding, cGMP-specific phosphodiesterase, PDE5.

FEBS Lett., 2003, 539 :161-6.

Stadtman ER, Van Remmen H, Richardson A, Wehr NB, and R. Levine L.

Methionine oxidation and aging.

Biochim. Biophys. Acta., 2005, 1703 : 135-40.

Steggborn C, Litvin TN, Levin LR, Buck J, Wu H.

Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment.

Nat. Struct. Mol. Biol., 2005, 12 : 32-7.

Sunahara RK, Dessauer CW, Whisnant RE, Kleuss C, Gilman AG.

Interaction of Gs-alpha with the cytosolic domains of mammalian adenylyl cyclase.

J. Biol. Chem., 1997, 272 : 22265-71.

Swaim MW, Pizzo SV.

Methionine sulfoxide and the oxidative regulation of plasma proteinase inhibitors.

J. Leukoc. Biol., 1988, 43 : 365-79.

Tarrago L, Laugier E, Zaffagnini M, Marchand C, Le Maréchal P, Rouhier N, Lemaire SD, Rey P.

Regeneration mechanisms of *Arabidopsis thaliana* methionine sulfoxide reductases B by glutaredoxins and thioredoxins.

J. Biol. Chem., 2009, 284 :18963-71.

Taylor AB, Benglis DM, Jr., Dhandayuthapani S, Hart PJ.

Structure of *Mycobacterium tuberculosis* methionine sulfoxide reductase A in complex with protein-bound methionine.

J. Bacteriol., 2003, 185 : 4119-26.

Tete-Favier F, Cobessi D, Boschi-Muller S, Azza S, Branlant G, Aubry A.

Crystal structure of the *Escherichia coli* peptide methionine sulphoxide reductase at 1.9 Å resolution.

Structure Fold Des., 2000, 8 : 1167-78.

Tews I, Findeisen F, Sinning I, Schultz A, Schultz JE, Linder JU.

The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme.

Science, 2005, 308 :1020-3.

Turko IV, Haik TL, McAllister-Lucas LM, Burns F, Francis SH, Corbin JD.

Identification of key amino acids in a conserved cGMP-binding site of cGMP-binding phosphodiesterases. A putative NKXnD motif for cGMP binding.

J. Biol. Chem., 1996, 271 :22240-4.

Wardman P, Candeias LP.

Fenton chemistry: an introduction.

Radiat. Res., 1996, 145 : 523-31.

Waugh SI, Paulsen DM, Mylona PV, Maynard RH, Premakumar R, Bishop PE.

The genes encoding the delta subunits of dinitrogenases 2 and 3 are required for mo-independent diazotrophic growth by *Azotobacter vinelandii*.

J. Bacteriol., 1995, 177 :1505-10.

Weissbach H, Resnick L, Brot N.

Methionine sulfoxide reductases: history and cellular role in protecting against oxidative damage.

Biochim. Biophys. Acta., 2005, 1703 : 203-12.

Wells-Knecht MC, Lyons TJ, McCance DR, Thorpe SR, Baynes JW.

Age-dependent increase in ortho-tyrosine and methionine sulfoxide in human skin collagen is not accelerated in diabetes. Evidence against a generalized increase in oxidative stress in diabetes.

J. Clin. Invest., 1997, 100 : 839-46.

Wu AY, Tang XB, Martinez SE, Ikeda K, Beavo JA.

Molecular determinants for cyclic nucleotide binding to the regulatory domains of phosphodiesterase 2A.

J. Biol. Chem., 2004, 279 : 37928-38.

Vogt W, Hesse D.

Oxidants generated by the myeloperoxidase-halide system activate the fifth component of human complement, C5.

Immunobiology, 1994, 192 : 1-9.

Yao Y, Squier TC.

Variable conformation and dynamics of calmodulin complexed with peptides derived from the autoinhibitory domains of target proteins.

Biochemistry, 1996, 35 :6815-27.

Zhang G, Liu Y, Ruoho AE, Hurley JH.

Structure of the adenylyl cyclase catalytic core.

Nature, 1997, 386 :247-53.

Zoraghi R, Bessay EP, Corbin JD, Francis SH.

Structural and functional features in human PDE5A1 regulatory domain that provide for allosteric cGMP binding, dimerization, and regulation.

J. Biol. Chem., 2005, 280 : 12051-63.

Zoraghi R, Corbin JD, Francis SH.

Properties and functions of GAF domains in cyclic nucleotide phosphodiesterases and other proteins.

Mol. Pharmacol., 2004, 65 : 267-78.

Les Méthionine sulfoxyde réductases (Msr) catalysent la réduction spécifique des méthionine sulfoxydes (Met-O) en méthionines (Met). Elles sont impliquées dans la résistance des cellules à un stress oxydant et dans la virulence des bactéries pathogènes du genre *Neisseria*. Cette famille d'enzyme se compose de deux classes, les MsrA et B, structurellement distinctes, et présentent une stéréosélectivité respectivement pour l'isomère *S* et *R* de la fonction sulfoxyde du substrat. Une troisième classe, découverte récemment, et appelée fRMs, catalyse la réduction spécifique de l'isomère *R* de la fonction sulfoxyde de la Met-O libre. La fRMs appartient à la famille des domaines GAF, généralement impliqués dans les phénomènes de signalisation cellulaire, et elle représente le premier domaine GAF présentant une activité enzymatique.

Les études réalisées au cours de ma thèse sur la fRMs de *Neisseria meningitidis* ont permis de montrer que : 1) La fRMs présente un mécanisme de réduction de la fonction sulfoxyde similaire à celui des MsrA MsrB, qui passe par la chimie de l'acide sulfénique, et la formation d'au moins un intermédiaire pont disulfure intramoléculaire, entre la Cys118 et la Cys84, réduit par la Trx. 2) La Cys118 est le résidu catalytique sur lequel l'intermédiaire acide sulfénique doit se former ; 3) L'étape réductase est l'étape cinétiquement déterminante du mécanisme à deux étapes conduisant à la formation du pont disulfure Cys84-Cys118.

De plus, l'étude de la fRMs en termes de catalyse et de spécificité de substrat a permis de suggérer fortement 1) Un rôle de la fonction carboxylate du résidu Asp143 dans la catalyse de l'étape réductase ; 2) que le résidu Glu125 est impliqué dans la reconnaissance et/ou le positionnement du substrat Met-O probablement *via* la stabilisation du groupement NH_3^+ ; 3) Un rôle du résidu Asp141 dans le positionnement des résidus Asp143 et Glu125. Enfin, l'analyse de la structure tridimensionnelle de la fRMs de *N. meningitidis* en complexe avec le substrat a permis de montrer, l'existence d'un site de reconnaissance oxyanion impliqué dans la stabilisation de la fonction carboxylate.

Mots clés : Méthionine sulfoxyde réductase, fRMs, domaine GAF, mécanisme catalytique, catalyse, spécificité structurale.