



**HAL**  
open science

# Interactions des protéines du complexe de transduction du signal transmembranaire CnrYXH chez *Cupriavidus metallidurans* CH34

Widade Ziani

► **To cite this version:**

Widade Ziani. Interactions des protéines du complexe de transduction du signal transmembranaire CnrYXH chez *Cupriavidus metallidurans* CH34. Biologie structurale [q-bio.BM]. Université de Grenoble, 2014. Français. NNT : 2014GRENV059 . tel-01344080

**HAL Id: tel-01344080**

**<https://theses.hal.science/tel-01344080>**

Submitted on 11 Jul 2016

**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.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : Biologie structurale et nanobiologie

Arrêté ministériel : 7 août 2006

Présentée par

**Widade ZIANI**

Thèse dirigée par **Jacques Covès**

codirigée par **Antoine Maillard**

préparée au sein de l'**Institut de Biologie Structurale**

Equipe **Homéostasie des Métaux et Signalisation**

Groupe **Métalloprotéines**

dans l'**École Doctorale Chimie Sciences du Vivant**

## **Interactions des protéines du complexe de transduction du signal transmembranaire CnrYXH chez *Cupriavidus metallidurans* CH34**

Thèse soutenue publiquement le **18 décembre 2014**

devant le jury composé de :

**Prof. James STURGIS** (Université Aix-Marseille) **Rapporteur**

**Dr. Alexandra GRUSS** (INRA, Jouy-en-Josas) **Rapporteur**

**Prof. Christelle BRETON** (CERMAV, Grenoble) **Présidente**

**Dr. Virginie MOLLE** (CNRS, Montpellier) **Examinatrice**



# Table des matières

<b>Introduction bibliographique .....</b>	<b>9</b>
<b>I. La résistance aux stress chez les bactéries.....</b>	<b>10</b>
A. Définition du stress .....	10
B. Perception des stress chez la bactérie .....	11
C. Régulation de la réponse aux stress .....	12
1. Initiation de la transcription chez les bactéries.....	13
2. Systèmes à un composant.....	15
a. Un exemple de métallorégulateur : la protéine Fur (Ferric uptake regulator).....	15
3. Systèmes à deux composants.....	19
a. Un exemple représentatif des TCS : le système PhoQ/PhoP.....	22
b. Un exemple exceptionnel de TCS : le système PhyR:NepR: $\sigma^{ECF}$ .....	25
4. Systèmes dépendants des facteurs sigma ( $\sigma$ ) de type ECF .....	27
a. Généralités sur les facteurs $\sigma$ .....	27
b. Les facteurs sigma ECF ( $\sigma^{ECF}$ ).....	29
i. L'activation des facteurs $\sigma^{ECF}$ résulte de la protéolyse de son facteur anti- $\sigma$ .....	30
ii. L'activation du facteur $\sigma^{ECF}$ résulte du changement conformationnel de son facteur anti- $\sigma$ .....	34
iii. L'activation du facteur $\sigma^{ECF}$ résulte des interactions protéine:protéine entre les partenaires du système de transduction du signal.....	35
<b>II. Homéostasie des métaux lourds.....</b>	<b>40</b>
A. Définition de l'homéostasie des métaux lourds.....	40
B. Import des ions métalliques.....	41
1. Import des ions métalliques dans le périplasme .....	41
2. Import des ions métalliques dans le cytoplasme.....	42
C. Systèmes de résistance aux métaux .....	44
1. Efflux des ions métalliques.....	44
a. Systèmes d'efflux de la membrane interne.....	44
b. Systèmes d'efflux de la membrane externe.....	47
<b>III. <i>Cupriavidus metallidurans</i> CH34 : une bactérie hautement résistante aux métaux .....</b>	<b>51</b>

A.	Description générale de <i>C. metallidurans</i> CH34.....	51
B.	Les systèmes de résistances aux métaux.....	52
C.	Régulation de la résistance aux métaux.....	53
D.	Le déterminant Cnr de <i>Cupriavidus metallidurans</i> CH34.....	54
1.	Résistance au cobalt et au nickel.....	54
2.	Le système NccYXH de <i>C. metallidurans</i> 31A.....	55
3.	Régulation de l'expression de <i>cnrCBA</i> par <i>cnrYXH</i> .....	56
4.	Le système de régulation CnrYXH.....	57
a.	Description du système CnrYXH.....	57
b.	Induction du système par les métaux.....	59
c.	La protéine CnrX.....	60
i.	Caractérisation du domaine périplasmique de CnrX (CnrXs).....	62
	<b>Résultats.....</b>	<b>70</b>
I.	<b>Etude de l'interaction entre CnrY et CnrH dans le cytoplasme.....</b>	<b>71</b>
A.	Publication Maillard <i>et al.</i> , 2014.....	74
B.	CnrY est l'archétype d'une nouvelle classe d'ASD.....	104
II.	<b>Etude de l'interaction entre CnrX/NccX et CnrY/NccY dans la membrane.....</b>	<b>107</b>
A.	Homodimérisation de NccX.....	107
B.	Publication Ziani <i>et al.</i> , 2014.....	113
C.	Résultats complémentaires.....	144
D.	Homodimérisation de NccY.....	149
E.	Interaction entre NccX et NccY.....	151
III.	<b>Etude de l'interaction entre CnrX et CnrY dans le périplasmie.....</b>	<b>155</b>
A.	Mise en évidence de l'interaction entre les parties périplasmique de CnrX et CnrY....	155
B.	Caractérisation de l'interaction entre CnrYp et CnrXs.....	158
1.	Identification du déterminant de l'interaction de CnrYp.....	158
2.	Modulation de l'interaction entre CnrYp et CnrXs par les métaux.....	161
3.	Identification des régions de CnrXs interagissant avec CnrYp.....	164
a.	Identification par coprécipitation.....	164
b.	Identification par échange Hydrogène/Deutérium (HDX).....	166
ii.	Confirmation de l'interaction entre CnrXs et MYp par la méthode HDX-MS.....	168
iii.	Cartographie de CnrXs.....	169
iv.	Echanges H/D couplés à la MS.....	171
	<b>Conclusions et perspectives.....</b>	<b>180</b>
	<b>Matériels et méthodes.....</b>	<b>189</b>
I.	<b>Production et purification des protéines d'intérêt.....</b>	<b>190</b>

A.	Milieux de culture .....	190
B.	Souches bactériennes et plasmides.....	191
C.	Clonage.....	193
D.	Mutagenèse dirigée .....	193
1.	Mutagenèse par le protocole "QuikChange" .....	193
2.	Mutagenèse avec la polymérase Phusion .....	194
E.	Transformation bactérienne .....	195
F.	Culture et expression .....	195
G.	Purification de protéines .....	195
1.	Purification de CnrXs.....	195
a.	Lyse cellulaire.....	196
b.	Précipitation des protéines.....	196
c.	Chromatographie sur colonne échangeuse d'anions.....	196
d.	Filtration sur gel.....	197
2.	Purification de NccX.....	197
a.	Préparation de la fraction membranaire.....	197
b.	Filtration sur gel.....	198
c.	Chromatographie sur colonne échangeuse d'anions.....	198
3.	Purification de MBP-CnrYp.....	198
a.	Lyse cellulaire.....	199
b.	Chromatographie d'affinité : "Maltose Binding Protein" et résine d'amylose .....	199
c.	Chromatographie sur colonne échangeuse d'anions.....	199
d.	Filtration sur gel.....	199
4.	Purification de Strep-CnrH:CnrYc .....	200
a.	Chromatographie d'affinité (par gravité) : étiquette Strep:StrepTactine .....	200
b.	Chromatographie d'affinité (interactions hydrophobes) .....	200
c.	Filtration sur gel .....	201
<b>II.</b>	<b>Techniques biochimiques.....</b>	<b>201</b>
A.	Séparation des protéines par électrophorèse en conditions dénaturantes (SDS-PAGE) .....	201
B.	Révélations.....	201
1.	Coloration au bleu de Coomassie .....	201
2.	Western Blot.....	202
C.	Dosage des protéines .....	202
D.	Dosage colorimétrique des métaux avec du 4-(2-pyridylazo)resorcinol (PAR) .....	203
<b>III.</b>	<b>Tests d'interaction <i>in vivo</i> .....</b>	<b>204</b>
A.	Test TOXCAT .....	204
1.	Constructions utilisées.....	204

2.	Vérification de la topologie des protéines chimériques .....	204
a.	Test de complémentation <i>malE</i> .....	204
b.	Fractionnement cellulaire.....	205
3.	Culture et expression.....	205
4.	Mesure de l'activité CAT.....	206
B.	Test <i>in vivo</i> de l'inhibition de CnrH par la partie cytoplasmique de CnrY (CnrYc).....	207
<b>IV.</b>	<b>Test d'interaction <i>in vitro</i> .....</b>	<b>207</b>
A.	Expérience de coprécipitation.....	207
B.	Expérience d'échange hydrogène:deutérium couplé à la spectrométrie de masse (HDX-MS) .....	208
1.	Mise en évidence de l'interaction CnrXsΔNter:MBP-CnrYp.....	208
2.	Cartographie de CnrXsΔNter .....	209
3.	Identification des peptides de CnrXsΔNter impliqués dans l'interaction.....	209
	<b>Références bibliographiques .....</b>	<b>211</b>
	<b>Résumé .....</b>	<b>230</b>

# Remerciements

En premier lieu je tiens à exprimer mes plus sincères et profonds remerciements à Jacques Covès pour m'avoir accueillie au sein de son équipe ainsi que pour l'encadrement de ma thèse. Je tiens vivement à remercier Antoine Maillard qui a codirigé mon travail avec beaucoup de soin et d'attention. J'ai vraiment apprécié travailler avec vous deux dans une ambiance amicale et très constructive. J'ai eu un grand plaisir à évoluer à vos côtés durant ces trois années, dans une ambiance chaleureuse et très constructive grâce à nos riches discussions scientifiques où vous avez pu me transmettre vos connaissances techniques et transverses. Grâce à toi Antoine, la biomol n'a presque plus de secrets pour moi :) Merci également pour toute l'énergie que vous avez déployée pendant ma thèse et plus particulièrement pendant des périodes plus difficiles comme la rédaction... Vous avez veillé à ce que tout se passe le plus sereinement possible et cela m'a permis d'effectuer cette thèse dans le meilleur des cadres. Pour tout cela ainsi que pour votre disponibilité de tous les instants je vous adresse une nouvelle fois tous mes remerciements ainsi que ma plus grande reconnaissance.

Je tiens à exprimer ma reconnaissance aux autres membres du jury : James Sturgis et Alexandra Gruss pour l'honneur que vous m'avez fait en acceptant de juger ce travail, Christelle Breton pour avoir accepté de présider mon jury de thèse ainsi que Virginie Molle pour avoir accepté d'examiner mon travail mais aussi pour cette collaboration naissante qui, je l'espère, sera fructueuse.

Je voudrais ensuite remercier mes proches collaborateurs qui m'ont aidés tout au long de mes travaux de thèse : Serge Crouzy et Norbert Garnier pour toutes les expériences de simulations par dynamique moléculaire. Merci de m'avoir fait découvrir cet univers et d'avoir pris le temps de m'expliquer des notions qui m'étaient très floues. Un grand merci à Eric Forest pour toutes ces expériences H/D, ces longues discussions et explications notamment pour l'analyse des données. Il reste encore plusieurs choses à faire et je suis sûre que cela donnera de merveilleux résultats. Merci à Dominique Marion et Isabelle Ayala pour les expériences de RMN. Enfin, merci à Eric Girard et Anne Volbeda

pour la résolution, pas toujours facile, des nombreuses structures cristallographiques de protéines.

Je tiens aussi à remercier les membres de mon comité de suivi de thèse : Cécile Breyton, Agnès Rodrigues et Ludovic Pelosi. Merci pour votre présence et de vos conseils avisés. Ludovic, je te suis reconnaissante de m'avoir fait découvrir le monde fascinant de l'enseignement.

Durant ces trois années j'ai eu la chance de travailler avec des personnes formidables regroupées dans le laboratoire anciennement appelé LCCP et qui sont maintenant dispatchées dans les groupes Métallo et GSY sans oublier Kiki et Christophe mais aussi notre formidable et efficace secrétaire : Chantal Robesson. Merci à vous tous pour votre bonne humeur, votre soutien et aide à tout instant. J'ai vraiment passé de très bons moments à vos côtés au labo comme en dehors !! Vous allez tous me manquer... Je tiens particulièrement à remercier Patricia Amara. Les mots ne seront pas suffisants pour t'exprimer toute ma gratitude tant au niveau professionnel (surtout en info ;) ) que personnel. Tu as été une deuxième maman pour moi donc 1000 mercis :). Un énorme merci à Juan Fontecilla pour tes conseils de grand chef, ta disponibilité et ta bonne humeur quotidienne. Je suis peut être bon public mais tes blagues m'ont toujours fait rire. Et puis je sais maintenant comment faire disparaître un diamant sans laisser aucune trace ;). Une grosse pensée aussi à mes colocataires de bureau, Adrien et Roman, avec qui j'ai partagée de supers moments !!

Une énorme pensée à toutes les personnes qui étaient présentes à mes débuts au labo et qui s'en sont allées vers de nouvelles aventures. Vous avez été et êtes géniaux et c'est pour cela que vous êtes désormais mes amis : mes poulettes Marina et €llo (on aura fait les 400 coups ensemble et ça me manque) Yasser et Afi (mes petits parents de substitution), mon BG Hugo sans qui les soirées au labo n'auraient aucun sens et tu sais que c'est beaucoup plus (bonjour, merci et aurevoir!) mais aussi Tic et Tac, Sarah, Christine, Juliette, Myriam...

Il y a 3 ans, mon chemin a croisé celui d'autres thésards tout nouveaux aussi : la fameuse et fabuleuse "team Porquerolles" puis "groupe du H2 mercredi" (Louise alias ma super coloc, Gianluca, Mathieu, Yann HDK, Jules, Yann, Didier, Hicham, Vaness). Un énorme merci pour ces moments de qualité partagés avec vous, ces fous rires et nos conversations parfois gênantes surtout pour les autres :)

Merci aussi à toutes les personnes (IBS1 & 2) avec qui j'ai pu échanger, partager... Je pense particulièrement à Michel alias MichMich ou Macgyver. En plus d'avoir des doigts de fée et une imagination folle, tu as été d'une écoute, d'un soutien et d'un réconfort sans limite. Merci d'être une personne extraordinaire. Mes pensées vont également à Alain, Sylvie et Fabrice qui ont été d'une aide précieuse et ceci toujours avec la bonne humeur.



D'un point de vue plus personnel, je voudrais remercier Cyril qui a été d'un soutien sans faille pendant la dernière ligne droite mais aussi Yoan pour tes taquineries qui ont sues me redonner le sourire!!

Je remercie aussi tous mes amis de longue date qui ont fait le déplacement depuis de lointaines contrées : mon Moumou et ma So qui ne sont pas que des collègues de fac mais des amis pour la vie ;), Chadia et Régis qui ont toujours été là pour moi... Sans oublier ceux qui n'ont pas pu être présents mais qui le sont dans mon cœur : ma BFFE Ouided, Amel, Chloé, Julies, Delphine, Said...

Je finirai par remercier toute ma famille et plus particulièrement mes parents, ma sœur, Abdou et mes frères pour leurs encouragements et leur soutien inconditionnel. Merci d'avoir été présents et de croire toujours en moi. Je vous aime...

# Introduction bibliographique

L'équipe Homéostasie des Métaux et Signalisation (HMS) s'intéresse à la perception du stress métallique par les bactéries et aux mécanismes de propagation du signal perçu. Un des modèles d'étude est celui qui contribue à réguler la réponse au stress induit par le nickel et le cobalt chez la bactérie *Cupriavidus metallidurans* CH34. Il s'agit du complexe CnrYXH. Pour des raisons de commodité expérimentale, le complexe homologue NccYXH, issu de *Cupriavidus metallidurans* 31A, est aussi utilisé. Nous verrons plus loin que la description du système Cnr s'applique à celle de Ncc.

Une étude bibliographique décrivant l'état des connaissances sur le stress chez les bactéries est présentée dans la première partie de ce manuscrit. Nous essaierons de comprendre comment la bactérie gère les facteurs de stress qui l'entourent, de leur perception à la réponse aux stress causée par ces facteurs. Quelques exemples des différents systèmes de transduction du signal chez les bactéries seront détaillés, avec un accent sur des systèmes répondant à des stress métalliques. Dans une seconde partie, nous décrirons l'homéostasie des métaux. Nous détaillerons les systèmes d'import et d'efflux des métaux chez les bactéries notamment chez notre modèle d'étude, *Cupriavidus metallidurans* CH34. *C. metallidurans* CH34 est une bactérie multirésistante aux métaux. Sa description, ainsi que celle des nombreux systèmes de résistance au stress métallique que ce micro-organisme possède, fera l'objet d'un troisième point. Parmi ces systèmes, le déterminant Cnr sera présenté en détail.

# I. La résistance aux stress chez les bactéries

Les bactéries sont des micro-organismes unicellulaires, procaryotes, qui peuvent coloniser des niches écologiques très diverses et parfois considérées comme extrêmes (variation de pH, de température, de pression atmosphérique, teneur en oxygène, teneur en eau, radiations ionisantes, concentration en métaux...). Elles ont la capacité de prospérer dans des conditions changeantes et de résister au stress généré par des conditions défavorables. Cela suppose qu'elles puissent d'une part percevoir un facteur de stress, et d'autre part s'y adapter en mettant en œuvre une réponse assurant la résistance et/ou la survie dans les conditions de stress ressenties.

## A. Définition du stress

Un état de stress cellulaire est causé par tout changement du génome, du protéome, ou de l'environnement qui impose une croissance ou un potentiel de survie réduit. De tels changements sont susceptibles d'entraîner des réactions de survie de trois types différents : une migration vers des conditions plus favorables (tactisme), l'adaptation induite du métabolisme aux conditions nouvelles (résistance), ou la production de formes résistantes à l'issue d'une réaction de différenciation (mutations spontanées, sporulation). Toutes les étapes de l'expression génétique (transcription, traduction et modifications post-traductionnelles) sont considérées comme faisant partie du métabolisme. Aussi, les mutations peuvent générer un stress dans la cellule. Le phénotype résultant peut être amélioré par l'expression de gènes adaptés (Booth, 2002).

Les facteurs de stress mais aussi les mécanismes de réponse développés par la bactérie sont très divers (**Tableau 1**). Ces facteurs de stress sont majoritairement de nature physicochimique et affectent principalement l'enveloppe cellulaire et le cytoplasme des bactéries. Les réponses aux stress visent soit à rétablir l'homéostasie (expression de chaperons, synthèse de transporteurs...), soit à se protéger des effets du stress (modifications de l'enveloppe, formation de biofilm...).

**Tableau 1** : Diversité des facteurs de stress et des réponses aux stress chez les bactéries.

<b>Facteurs de stress environnementaux</b>	<b>Réponses cellulaires</b>
Température	Modification de l'enveloppe
pH	Synthèse de transporteurs / pompes d'efflux
Salinité	Expression de chaperons, protéases
Rayons UV	Modification métabolique
Oxygène	Formation de biofilm
Carences / Excès	Mort cellulaire programmée
	Entrée en phase stationnaire

## **B. Perception des stress chez la bactérie**

Les bactéries à Gram négatif possèdent une enveloppe cellulaire constituée d'une membrane externe (ME) et d'une membrane plasmique ou interne (MI). La ME est en contact direct avec le milieu environnant. Elle est constituée de façon asymétrique, avec un feuillet interne composé de phospholipides et un feuillet externe composé de lipopolysaccharides (LPS). La MI est composé d'une bicouche de phospholipides (Silhavy *et al.*, 2010). Ces deux membranes sont pourvues de nombreuses protéines et si elles disposent toutes deux de lipoprotéines, elles diffèrent par la nature de leurs protéines intrinsèques. En effet, les protéines intrinsèques de la MI traversent la bicouche de phospholipides grâce à des structures secondaires principalement en hélice  $\alpha$  (Dalbey *et al.*, 2011) alors que celles de la ME forment des structures en tonneau  $\beta$  (Fairman *et al.*, 2011) comme les porines chargées d'assurer le passage de petites molécules hydrophiles (voir section **II.B.**).

Ces deux membranes délimitent un compartiment intermembranaire appelé périplasma. Il contient une paroi fine composée de peptidoglycanes qui confère à la bactérie sa forme et lui permet de résister à la forte pression osmotique interne (Typas *et al.*, 2012). Le périplasma est un compartiment cellulaire métaboliquement actif (assemblage des peptidoglycanes, respiration bactérienne...) et est le lieu de transit de nombreuses molécules entre les deux membranes. La MI délimite aussi le compartiment intracellulaire, aussi appelé cytoplasme, où

est produite l'énergie et sont synthétisés les métabolites et les macromolécules nécessaires à la vie de la bactérie.

Ces membranes jouent un rôle essentiel dans la physiologie de la cellule. Elles sont le siège de nombreux processus vitaux pour les bactéries tels que la production d'énergie (MI seulement), les échanges de molécules *via* des transporteurs, la transduction de signaux... Ces membranes, situées à l'interface entre les milieux extra- et intracellulaire, sont les plus exposées aux agressions environnementales. Ainsi, ces agressions peuvent endommager la membrane externe et les protéines qu'elle contient, aboutissant à l'accumulation de protéines mal repliées dans le périplasma. Cette accumulation peut devenir toxique pour la cellule si elle ne répare pas ces dommages. La bactérie met alors en place un système de régulation, ce qui lui permet de répondre le plus rapidement à ces stress qui peuvent lui être fatals, comme par exemple la production de nouvelles protéines de la membrane, l'activation de synthèse de protéases, de chaperons ou de protéines de maturation. Ainsi, l'enveloppe cellulaire n'est pas une structure statique : elle est modifiée en réponse aux changements environnementaux (Silhavy, 2010).

Pour répondre à un stress, la bactérie doit dans un premier temps détecter les facteurs de stress correspondant. Les facteurs de stress peuvent être perçus directement par le senseur quand ce sont des objets physiques comme les métaux, les peptides... D'autres facteurs de stress ne sont pas perçus directement mais leurs effets le sont : la température, le pH et les radiations entraînent des dommages cellulaires qui seront ensuite détectés par le senseur. Par ailleurs, ces facteurs peuvent être perçus à différents niveaux de la cellule. Ainsi, on retrouve des senseurs à détection cytoplasmique et des senseurs à détection périplasmique. De part leur localisation, la transduction du signal initiée par ces senseurs sera différente. Cet aspect sera présenté dans la partie suivante à travers quelques exemples.

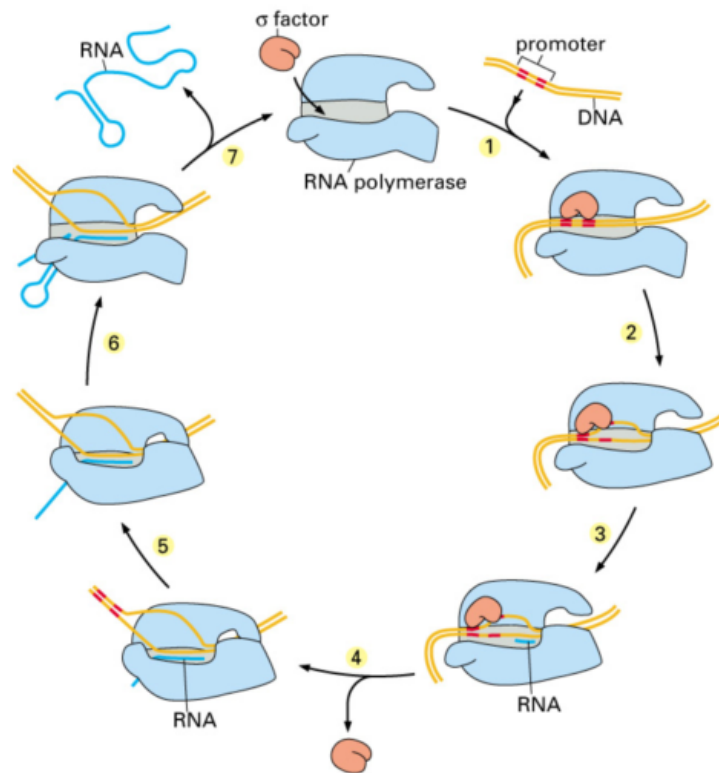
## **C. Régulation de la réponse aux stress**

Toute réponse adaptative repose sur le couplage d'un système de détection à la mise en place des effecteurs appropriés. Chez les bactéries, une réponse adaptative repose le plus souvent sur la modification du profil des gènes transcrits, c'est-à-dire sur la régulation de la transcription. La perception du facteur de stress par le senseur déclenche une cascade

d'événements qui transmet l'information du détecteur jusqu'à l'effecteur. Ce dernier contribue à mettre en place une réponse transcriptionnelle appropriée.

## 1. Initiation de la transcription chez les bactéries

L'enzyme qui catalyse la transcription est l'ARN polymérase (ARNP). Chez les bactéries, elle est composée de 5 sous-unités : 2  $\alpha$ ,  $\beta$ ,  $\beta'$  et  $\sigma$  dont la masse moléculaire totale est d'environ 500 kDa. Le complexe des 4 sous-unités  $\alpha_2\beta\beta'$  forme le cœur catalytique de l'enzyme et catalyse la synthèse de l'ARN mais ne peut initier la transcription (Darst *et al.*, 1989). Les fonctions de ses sous-unités ne sont pas encore clairement connues. La sous-unité  $\alpha$  semble intervenir dans l'assemblage du cœur de l'enzyme, dans la reconnaissance des promoteurs (voir ci-dessous), et dans l'interaction avec certains facteurs de régulation. Le site de fixation à l'ADN est sur  $\beta'$  et la sous-unité  $\beta$  se charge de la fixation des nucléosides triphosphates. La sous-unité  $\sigma$  ou facteur  $\sigma$  s'associe au cœur de l'enzyme et confère à l'holoenzyme ainsi formée ( $\alpha_2\beta\beta'\sigma$ ) la capacité de reconnaître spécifiquement une région particulière de l'ADN appelée promoteur, située en amont de la région codante d'un gène (Gruber et Gross, 2003; Helmann et Chamberlin, 1988). Le facteur  $\sigma$  reconnaît les boîtes -35 et -10 des promoteurs (35 et 10 nucléotides avant le site d'initiation de la transcription, respectivement) et permet la dissociation des deux brins d'ADN, indispensable à l'initiation de la transcription (Murakami *et al.*, 2003; Paget et Helmann, 2003) (**Figure 1**). Il existe plusieurs familles de facteurs  $\sigma$  chez les bactéries qui seront présentées par la suite (Partie **I.C.4.a.**).



**Figure 1 :** Représentation schématique de la transcription chez les bactéries. Le facteur  $\sigma$  (orange) en complexe avec l'ARN polymérase (bleu) va reconnaître spécifiquement les boîtes -35 et -10 du promoteur (tirés rouges) (**étape 1**) et dissocier les deux brins d'ADN (filaments jaunes) (**2**) pour initier la transcription. Dès que la synthèse de l'ARN commence, le facteur  $\sigma$  se dissocie du complexe enzyme:ADN (**3 et 4**). L'ARNP continue à synthétiser l'ARN (**5 et 6**) jusqu'à la terminaison de la transcription, accompagnée de la libération de l'ARN produit (**7**) (Hantke, 2001) (illustration tirée du livre "Molecular biology of the cell, Bruce Alberts, 2008).

L'expression des gènes est modulée principalement au niveau de l'initiation de la transcription par les facteurs de transcription (FT). Les FT sont des protéines qui se fixent sur une séquence spécifique de l'ADN et régulent l'activité du complexe de transcription suite à des signaux environnementaux ou intracellulaires. Ces stimuli peuvent soit interagir directement avec les FT soit être pris en charge par d'autres protéines qui vont transmettre le signal aux FT. Ces derniers agissent seuls ou en complexe avec d'autres protéines pour activer ou inhiber la fixation de l'ARNP sur l'ADN mais aussi pour activer ou inhiber la transcription en rendant l'ADN plus ou moins accessible.

Trois systèmes principaux de transduction du signal sont impliqués dans la régulation de la résistance aux stress. L'objectif de cette thèse est de caractériser les interactions entre les protéines du complexe de régulation de la résistance au nickel et au cobalt en vue de mieux

comprendre son mécanisme de transduction du signal. C'est pourquoi j'ai choisi de présenter par la suite l'état des connaissances sur les mécanismes d'activation des trois systèmes principaux en m'appuyant le plus souvent sur des exemples de systèmes répondant à des stress métalliques.

## **2. Systèmes à un composant**

Ce sont les formes les plus simples et les plus abondantes de signalisation chez les bactéries (Staroń et Mascher, 2010). Il s'agit de protéines possédant sur la même chaîne polypeptidique le domaine senseur et le domaine effecteur. Ces systèmes sont des protéines régulatrices qui détectent directement le signal, le plus souvent au niveau cytoplasmique. Sur le principe, le facteur perçu peut être un ligand dont la fixation modifie l'affinité de la protéine pour l'ADN. En se fixant sur la région promotrice des gènes, les protéines effectrices régulent la transcription de gènes cibles.

C'est le cas d'un certain nombre de métallorégulateurs. Ces derniers sont spécifiques pour un ion voire deux ions métalliques donnés. Par exemple, NikR, RcnR et Nur lient spécifiquement le nickel alors que NmtR et KmtR répondent au nickel et au cobalt (Iwig et Chivers, 2010). Cette sélectivité repose sur l'affinité d'un ion pour une protéine et sur la chimie de coordination du métal. La réponse transcriptionnelle repose le plus souvent sur changements allostériques dépendants de la coordination de l'ion métallique (Reyes-Caballero *et al.*, 2011).

### **a. Un exemple de métallorégulateur : la protéine Fur (Ferric uptake regulator)**

Fur est un FT soluble dans le cytosol qui contrôle l'expression d'une centaine de gènes impliqués dans de nombreux processus cellulaires (Fillat, 2014). Parmi ces gènes, on trouve ceux codant les protéines impliquées dans l'acquisition du fer (protéines de synthèse de sidérophores, systèmes d'import du fer...) (Hantke, 2001) mais aussi des gènes impliqués dans le métabolisme de la cellule et dans la réponse au stress oxydant (Massé et Arguin, 2005).



Fur est une métalloprotéine qui se comporte comme un senseur de fer. En complexe avec du fer ferreux ( $\text{Fe}^{2+}$ ), Fur présente une forte affinité pour l'ADN et se fixe sur la région du promoteur des gènes cibles (de Lorenzo *et al.*, 1987). L'ARNP ne pouvant plus se fixer sur l'ADN, il en résulte une inhibition directe de la transcription des gènes cibles.

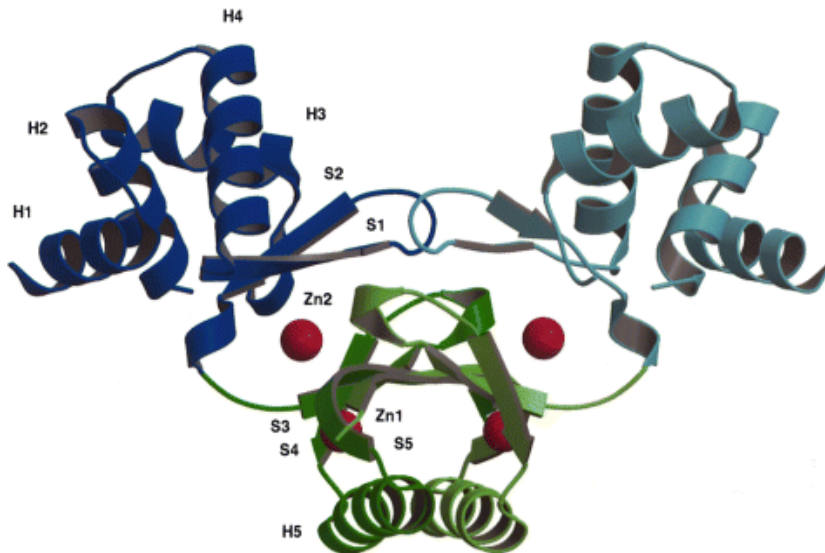
Fur a aussi donné son nom à la superfamille (FUR) de senseurs de métaux qui régulent la transcription des gènes impliqués dans l'homéostasie du fer, du zinc, du nickel et du manganèse et dans la réponse au stress oxydant. En plus de Fur, la famille de métallorégulateurs FUR est composée de Zur (Zinc uptake regulator), Mur (Manganese uptake regulator), Nur (Nickel uptake regulator), PerR (Peroxyde regulator) et Irr (Iron responsive regulator) (**Tableau 2**).

**Tableau 2** : Protéines appartenant à la superfamille FUR.

Réponse principale	protéine	métal régulateur	métal structural
<b>Disponibilité du métal</b>			
	Fur	$\text{Fe}^{2+}$	$\text{Zn}^{2+}$
	Zur	$\text{Zn}^{2+}$	$\text{Zn}^{2+}$
	Nur	$\text{Ni}^{2+}$	$\text{Zn}^{2+}$
	Mur	$\text{Mn}^{2+}$	-
<b>Stress peroxyde</b>	PerR	$\text{Fe}^{2+} / \text{Mn}^{2+}$	$\text{Zn}^{2+}$
<b>Disponibilité de l'hème</b>	Irr	$\text{Fe}^{2+}$ (hème)	-

Les protéines de la famille FUR sont similaires. En effet, ce sont des protéines de 120 résidus partageant un repliement structural commun dont la forme la plus commune semble être un homodimère. Elles possèdent un domaine N-terminal de liaison à l'ADN de type wHTH (winged-Helix-Turn-Helix) et un domaine C-terminal de dimérisation qui possède un site régulateur de liaison au métal.

Le nombre de sites métalliques et leur géométrie diffèrent d'une protéine à l'autre. Toutefois, il apparaît que seul un site par protomère correspond au site de régulation. C'est un site de haute affinité pour le métal, situé à la charnière entre le domaine régulateur et le domaine de dimérisation (site 2 ou Zn2, **Figure 2**). Le métal de ce site est pentacoordiné par 3 histidines, une cystéine et un résidu acide qui appartiennent aux deux domaines de Fur. Les autres sites ont des rôles structuraux importants notamment dans la stabilisation du dimère. Par exemple, les structures de Fur de *Pseudomonas aeruginosa* (Pohl, 2003) et *Vibrio cholerae* (Sheikh et Taylor, 2009) ainsi que la structure PerR de *B. subtilis* (Jacquamet *et al.*, 2009) possèdent 2 sites par protomères impliquant des azotes et des oxygènes comme ligands alors que celle de Fur d'*Helicobacter pylori* (Dian *et al.*, 2011) contient un troisième site impliquant 4 cystéines ligands.



**Figure 2** : Structure du dimère de Fur de *P. aeruginosa*. Les domaines de liaison à l'ADN des deux monomères sont colorés en bleu et cyan et le domaine de dimérisation en vert. Le zinc structural (site 1, Zn1) et le zinc coordonné dans le site de régulation (site 2, Zn2) sont représentés sous forme de sphères rouges (Pohl *et al.*, 2003).

Le mécanisme d'activation de la fixation des protéines FUR sur l'ADN est basé principalement sur les études effectuées sur PerR (Reyes-Caballero *et al.*, 2011). Chez *B. subtilis*, le manganèse qui se fixe sur le site régulateur est pentacoordiné par 3 histidines et 2 aspartates. Une des histidines appartenant au domaine N-terminal, est recrutée pour être ligand rapprochant le domaine de régulation du domaine de dimérisation. Cela permet à la

protéine d'adopter une conformation en forme d'étrier adéquate pour une interaction stable avec l'ADN (**Figure 3**).



**Figure 3:** Superposition des structures cristallographiques de la forme apo de PerR (PerR-Zn-apo, en orange) et de la forme métallée au niveau du site de régulation (PerR-Zn-Mn, en vert) (Jacquamet *et al.*, 2009). Les ions zinc structuraux sont en noirs et les manganèses régulateurs en jaune.

Ce résidu histidine est conservé chez toutes les protéines PerR et une majorité de protéines Fur. Sur la base du modèle défini avec PerR, il est proposé que les métallo-régulateurs de la famille Fur soient régulés de façon allostérique par le métal auquel elles sont sensibles.

La majorité des systèmes à un composant est activée par des modifications allostériques. Cependant, les modifications de conformation n'ont pas les mêmes effets selon les FT. Pour certains FT comme Fur, la perception des facteurs de stress induit une inhibition de la transcription par interaction des FT au niveau des promoteurs des gènes cibles. D'autres FT comme les membres de la famille des métallo-régulateurs LysR (Maddocks et Oyston, 2008) et TetR (Deng *et al.*, 2013) interagissent avec l'ADN dans des conditions où la cellule n'est pas stressée. Le changement de conformation de ces FT suite à la fixation des facteurs de stress induit soit une libération des FT (LysR) soit un changement de courbure de l'ADN permettant une activation de la transcription des gènes cibles (TetR).

Bien que la détection du stress pour les systèmes à un composant ait le plus souvent lieu dans le cytoplasme, il existe aussi quelques exemples de senseurs à détection périplasmique où le signal doit être transduit à travers la membrane comme la protéine ToxR de *Vibrio cholerae*.

La protéine ToxR a servi à développer un système permettant de détecter l'homodimérisation de protéines membranaires, le TOXCAT (Russ et Engelman, 1999), qui a été mis en œuvre dans le cadre de ce travail de thèse. Certaines caractéristiques de ToxR seront donc présentées dans le chapitre Résultats (partie **II**).

### **3. Systèmes à deux composants**

Les stimuli extracellulaires nécessitent une signalisation transmembranaire. En l'état actuel des connaissances, ces stimuli sont pris majoritairement en charge par les deux autres catégories de systèmes de transduction du signal chez les bactéries. Il s'agit des systèmes à deux composants (TCS pour *Two-Component System*) et des systèmes utilisant des facteurs sigma de type ECF (*ExtraCytoplasmic Function*).

Les TCS sont les systèmes les plus utilisés par les bactéries pour répondre aux facteurs de stress provenant du milieu extérieur (Stock *et al.*, 2000) et interviennent dans de nombreux processus biologiques (Gao et Stock, 2010). Chez *E. coli*, un grand nombre de TCS a été identifié (24 TCS). Ces derniers régulent divers processus biologiques tels que le métabolisme, la virulence, la différenciation, l'adaptation (**Tableau 3**) (Yamamoto *et al.*, 2005).

**Tableau 3** : Les 24 systèmes à deux composants identifiés chez *E. coli* (Yamamoto *et al.*, 2005).

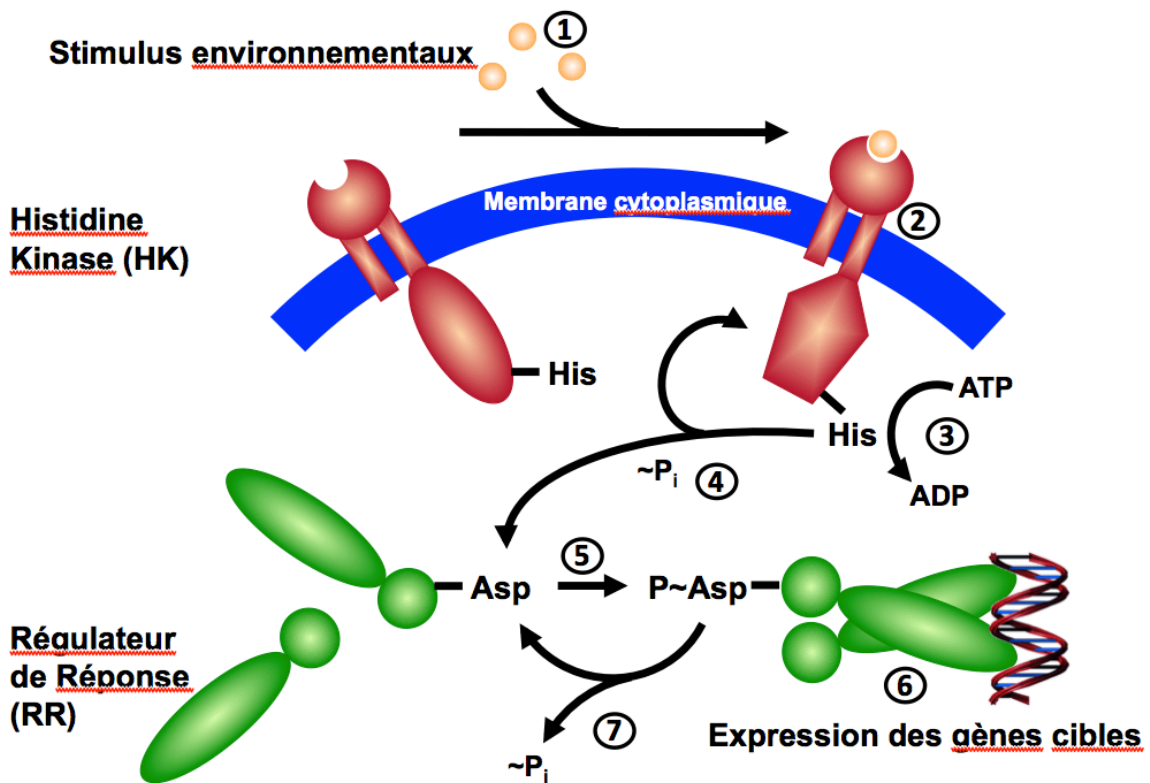
<b>Senseur Histidine Kinase</b>	<b>Régulateur de réponse</b>	<b>Rôle</b>
ArcB	ArcA	Respiration anaérobie
BaeS	BaeR	Réponse au stress d'enveloppe
BarA	UvrY	Régulation du stockage du carbone
BasS	BasR	Inconnu
CheA	CheB	Maturation du flagelle
CheA	CheY	Chemotactisme
CpxA	CpxR	résistance au cobalt, zinc et cadmium
CreC	CreB	Inconnu
CusS	CusR	Système d'efflux (cuivre/argent)
DcuS	DcuR	Respiration (à partir du fumarate)
EnvZ	OmpR	Synthèse protéines de la membrane externe
HydH	HydG	Système d'efflux (zinc/plomb)
KdpD	KdpE	Transport du potassium
NarX	NarL	Utilisation du nitrite et du nitrate
NarQ	NarP	Utilisation du nitrite et du nitrate
NtrB	NtrC	Assimilation du nitrite et du nitrate
PhoR	PhoB	Assimilation du phosphate
PhoQ	PhoP	Transport du magnésium
RstB	RstA	Inconnu
TorS	TorR	Respiration (à partir du triméthylamine N-oxyde)
UhpB	UhpA	Transport du glucose-6-phosphate
YedV	YedW	Inconnu
YehU	YehT	Inconnu
YfhK	YfhA	Inconnu
NtrB	NtrC	Assimilation de l'azote

A l'inverse des systèmes à un composant, les domaines senseur et effecteur des TCS sont présents sur deux protéines différentes. Ces systèmes se composent d'un senseur histidine kinase (HK) et d'un régulateur de réponse (RR). Le senseur HK est généralement constitué d'un détecteur N-terminal souvent périplasmique et d'un transmetteur C-terminal cytoplasmique. En présence d'agoniste, le domaine détecteur changerait de conformation, permettant l'autophosphorylation d'un résidu histidine conservé du domaine transmetteur.

Les RR sont eux aussi composés de deux domaines : un domaine receveur phosphorylé par le senseur au niveau d'un résidu aspartate très conservé et un domaine effecteur possédant un domaine de fixation à l'ADN.

Tous les TCS ne sont pas membranaires. Que leur senseur soit cytosolique ou membranaire, les TCS sont activés par le mécanisme de phosphorelais schématisé dans la **figure 4**. Ainsi, lorsque HK détecte un stimulus environnemental spécifique *via* son domaine senseur, souvent périplasmique (**Figure 4, étape 1**), cela mène à des changements de conformation (**étape 2**) résultant en une autophosphorylation dépendante de l'ATP du résidu histidine (His) conservé (**étape 3**). Cette histidine phosphorylée sert de donneur de phosphate au domaine receveur du RR (**étape 4**), résultant en la phosphorylation de l'aspartate (Asp) conservée ainsi qu'à sa dimérisation (**étape 5**). Cette dernière étape permet l'activation du domaine effecteur du RR qui se fixera sur l'ADN et ainsi pourra induire l'expression des gènes cibles (**étape 6**). Le retour à l'état basal est permis par la déphosphorylation de l'aspartate par le RR lui-même, par la HK ou par une phosphatase externe (**étape 7**) (Mascher *et al.*, 2006; Parkinson, 1993).

Par la suite, un exemple de TCS dont le mécanisme de fonctionnement est représentatif des autres TCS sera présenté. Ce système répond à certains métaux et est impliqué entre autre dans l'homéostasie du magnésium. Un second exemple de TCS, plus exceptionnel, sera également détaillé puisqu'il nous sera utile lors de l'analyse de nos résultats (voir **Résultats** partie **I**.)

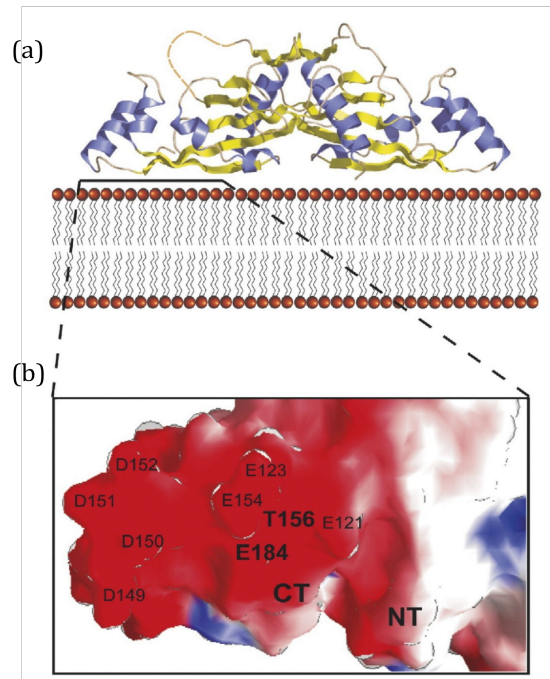


**Figure 4** : Mécanisme d'activation des systèmes à deux composants par phosphorelais (figure tirée du site [http://syntheticmicrobe.bio.lmu.de/daten\\_synbio/ressignaltr2.jpg](http://syntheticmicrobe.bio.lmu.de/daten_synbio/ressignaltr2.jpg))

### a. Un exemple représentatif des TCS : le système PhoQ/PhoP

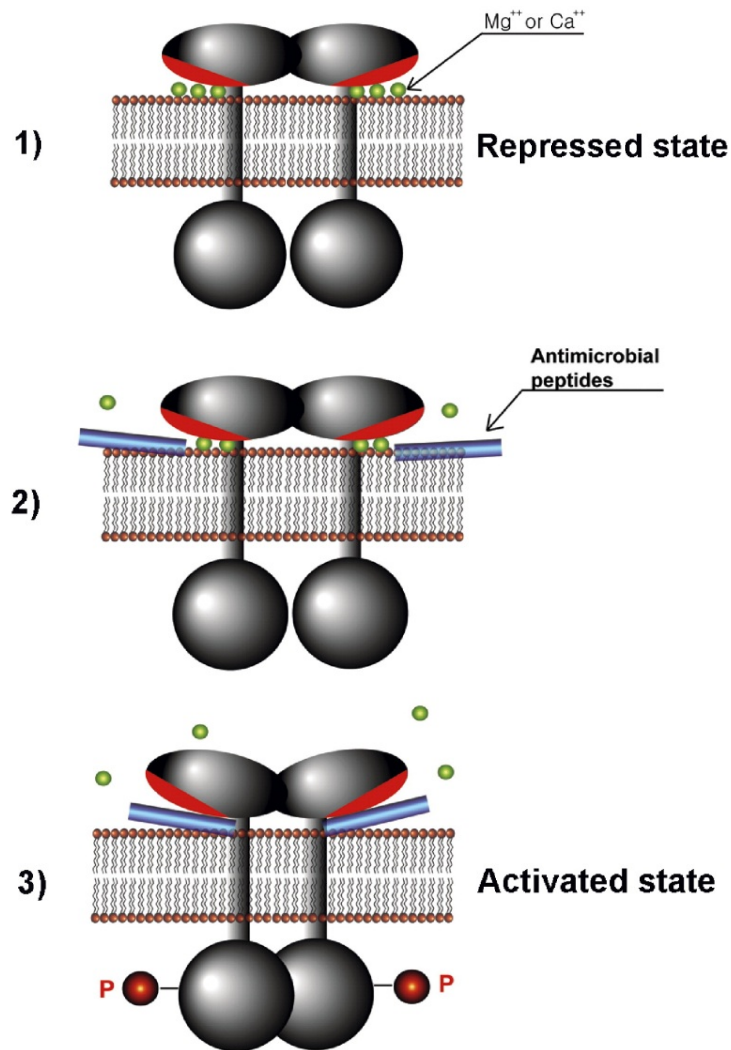
Identifié pour la première fois chez *Salmonella typhimurium* (Miller *et al.*, 1989), ce système est utilisé par la bactérie pour répondre aux changements environnementaux qui ont lieu durant une infection. Ce système est impliqué dans la virulence et permet la survie de la bactérie à l'intérieur des macrophages de l'hôte (Garvis *et al.*, 2001). Il est composé d'un senseur histidine kinase transmembranaire, PhoQ, et d'un régulateur de réponse soluble, PhoP. PhoQ détecte plusieurs signaux extracellulaires : elle est activée lorsque que le pH est acide (Alpuche Aranda *et al.*, 1992), lorsque la concentration en cations divalents est basse (García Véscovi *et al.*, 1996) mais aussi en présence de certains peptides antimicrobiens (Bader *et al.*, 2005). Le système PhoQ/PhoP régule l'expression d'une quarantaine de gènes impliqués dans différents processus cellulaires chez plusieurs bactéries à Gram négatif (Groisman, 2001). Il régule entre autres les gènes du transport de  $Mg^{2+}$  (Soncini *et al.*, 1996) et le système à deux composants PmrA/PmrB qui contrôle aussi les gènes impliqués dans la modification des LPS (Gunn et Miller, 1996) (Guo *et al.*, 1998).

La structure du domaine périplasmique de PhoQ (PhoQp) a été obtenue chez *S. typhimurium* (Cho *et al.*, 2006) sous forme dimérique (**Figure 5a**). Elle est composée d'une longue hélice N-terminale et d'un cœur formé par des brins  $\beta$  connectés par de petites hélices. Le dimère de PhoQp forme une surface plate chargée négativement face à la membrane (**Figure 5b**). En présence d'une forte concentration (de l'ordre du mM) en  $Mg^{2+}$  ou  $Ca^{2+}$ , la surface acide de PhoQp fixe ces ions. La conformation du domaine périplasmique de PhoQ est alors modifiée par ces cations divalents qui vont former un pont cationique entre la protéine et les phospholipides de la membrane, maintenant l'état inactif du senseur. La diminution de leur concentration, par exemple à la suite de leur déplacement par des peptides antimicrobiens, induit une répulsion entre la surface acide de PhoQp et la membrane permettant à PhoQ de retrouver une conformation active (Bader *et al.*, 2005) (**Figure 6**).



**Figure 5 :** (a) Structure du dimère de PhoQp en complexe avec du  $Ca^{2+}$  (code PDB : 1YAX). Chaque protomère possède une surface plane qui interagit avec la membrane. (b) Représentation de la surface électrostatique d'un protomère de PhoQp (vue de la membrane). La surface qui est chargée négativement (rouge), interagit avec la membrane (Cho *et al.*, 2006).





**Figure 6 :** Modèle de l'activation de PhoQ par les peptides antimicrobiens. Les cations divalents comme  $Ca^{2+}$  ou  $Mg^{2+}$  (sphères vertes) interagissent avec la surface acide (rouge) de PhoQ maintenant PhoQ dans une conformation inactive (1). Les peptides antimicrobiens cationiques (segment bleu) interagissent avec les phospholipides de la membrane et donc avec les sites de liaisons des cations divalents de PhoQ (2). Ces peptides déplacent donc ces cations divalents hors de la protéine. Cela induit un changement de conformation du senseur PhoQ qui entraîne la phosphorylation (sphères rouges) du domaine transmetteur de PhoQ (3) (Mascher *et al.*, 2006).

Ainsi, la première étape de transduction du signal à travers la membrane serait le déplacement des cations divalents qui est à l'origine d'une répulsion de charge entre PhoQ et la membrane. Cela mène à un changement d'orientation du senseur PhoQ par rapport à la membrane. PhoQ pourra alors s'autophosphoryler. Le groupement phosphate est ensuite transféré au régulateur de réponse, PhoP, dont la forme phosphorylée active la transcription des gènes cibles.

## **b. Un exemple exceptionnel de TCS : le système PhyR:NepR: $\sigma^{\text{EcfG}}$**

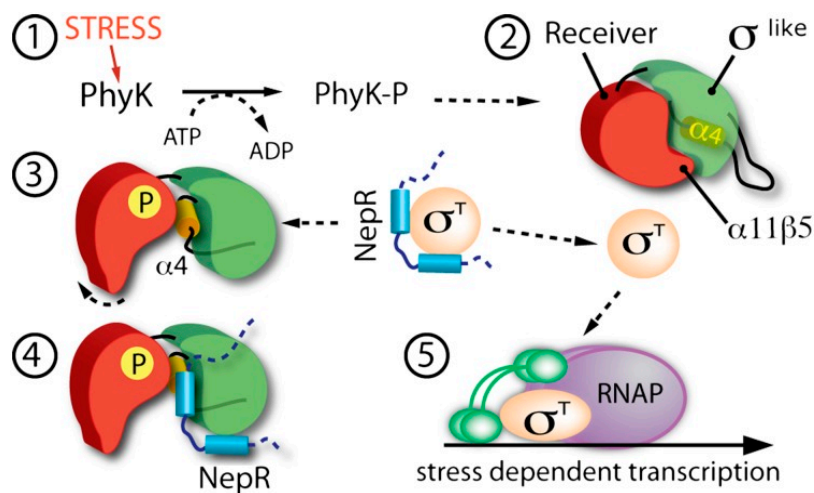
Dans leur majorité, les TCS régulent directement la transcription de gènes cibles comme le fait PhoP/PhoQ. Il existe cependant un cas, encore exceptionnel, où la signalisation par un TCS est médiée par un autre FT. C'est les cas du système PhyR:NepR: $\sigma^{\text{EcfG}}$  qui régule la réponse générale au stress chez les  $\alpha$ -protéobactéries (Francez-Charlot *et al.*, 2009). Le mécanisme a été établi à partir d'études effectuées chez *Methylobacterium extorquens* et *Caulobacter crescentus*. En l'absence de stress, le facteur  $\sigma^{\text{EcfG}}$  est séquestré par le facteur anti- $\sigma$  NepR, et PhyR est inactive. La perception d'un facteur de stress entraîne la phosphorylation de PhyR. Dans son état phosphorylé, PhyR interagit avec NepR, libérant ainsi le facteur  $\sigma^{\text{EcfG}}$  (*Sphingomonas sp.*) ou  $\sigma^{\text{T}}$  (*Caulobacter sp.*) et lui permet d'activer la transcription des gènes cibles. Les facteurs  $\sigma^{\text{EcfG}}$  et  $\sigma^{\text{T}}$  sont donc régulés par un mécanisme d'échange de partenaires.

Les dernières étapes de la transduction du signal aboutissant à la libération des facteurs  $\sigma^{\text{EcfG}}$  et  $\sigma^{\text{T}}$  ont été très étudiées, contrairement aux mécanismes de détection du facteur de stress et de phosphorylation de PhyR, qui restent mal définis. Toutefois, des histidines kinases potentielles ont été détectées par criblage génétique au voisinage du gène *phyR* de plusieurs  $\alpha$ -protéobactéries. Bien qu'elles ne soient pas conservées, ces kinases potentielles possèdent un domaine de dimérisation ainsi qu'un domaine "histidine phosphotransfer" qui déterminent la spécificité des kinases senseurs pour leurs régulateurs de réponse (Gao et Stock, 2010). Le rôle de la kinase potentielle, PhyK, de *Caulobacter crescentus* dans l'activation de la cascade PhyR:NepR: $\sigma^{\text{T}}$  a été étudié (Lourenço *et al.*, 2011). Lourenço et coll. ont démontré par des expériences de qRT-PCR que les gènes *sigT* et *sigU*, sous le contrôle de  $\sigma^{\text{T}}$ , ne sont plus exprimés lorsque le gène *phyK* est supprimé. Ce résultat indique que PhyK est nécessaire pour activer  $\sigma^{\text{T}}$ . PhyK est aussi nécessaire pour la phosphorylation de PhyR. Ainsi, PhyK est la kinase qui active la cascade PhyR:NepR: $\sigma^{\text{T}}$  via la phosphorylation de PhyR.

Les détails de la cascade de signalisation en aval de PhyR (PhyR:NepR: $\sigma^{\text{EcfG/T}}$ ) ont été révélés par des approches biophysiques. La structure cristallographique de PhyR de *Caulobacter crescentus* (Herrou *et al.*, 2010) a confirmé l'hypothèse selon laquelle, le domaine régulateur de PhyR ressemble à un facteur sigma ECF non fonctionnel (Francez-Charlot *et al.*, 2009). C'est donc par mimétisme moléculaire que PhyR régule l'activité d'un facteur sigma.

Les deux domaines de PhyR (domaine receveur, PhyR<sub>REC</sub> et domaine "sigma-like", PhyR<sub>SL</sub>) interagissent *via* une interface polaire qui masque une des faces du domaine PhyR<sub>SL</sub> et une région du domaine PhyR<sub>REC</sub> connue pour changer de conformation après phosphorylation chez d'autres régulateurs de réponse (Bourret et SilverSmith, 2010). La structure du complexe entre PhyR<sub>SL</sub> et NepR obtenue chez *C. crescentus* (Herrou *et al.*, 2012) et *Sphingomonas sp.* Fr1 (Campagne *et al.*, 2012) révèle que les deux hélices de NepR interagissent avec les deux domaines de PhyR<sub>SL</sub> comme une sorte d'agrafe. NepR interagit avec PhyR<sub>SL</sub> et  $\sigma^{\text{EcFG}}$  avec les mêmes déterminants d'interaction (Campagne *et al.*, 2012).

En résumé, en conditions de stress, PhyR est phosphorylée au niveau de l'histidine conservée de son domaine receveur qui transfère son groupement phosphate à l'aspartate dans le même domaine. Cela induit un changement de conformation de la protéine dans laquelle les deux domaines PhyR<sub>REC</sub> et PhyR<sub>SL</sub> se dissocient l'un de l'autre, exposant ainsi la surface de PhyR<sub>SL</sub>. L'affinité de NepR pour PhyR<sub>SL</sub> est plus importante que son affinité pour  $\sigma^{\text{EcFG}}$  et suffit pour que l'échange de partenaires ait lieu (Campagne *et al.*, 2012). Une fois libre,  $\sigma^{\text{EcFG}}$  peut alors s'associer à l'ARNP pour activer la transcription des gènes appropriés (Campagne *et al.*, 2012; Herrou *et al.*, 2012) (**Figure 7**).



**Figure 7** : Représentation schématique du modèle de régulation du facteur  $\sigma^{\text{EcFG/T}}$ . Le RR PhyR est composé du domaine receveur (PhyR<sub>REC</sub>) (rouge) et du domaine  $\sigma$ -like (PhyR<sub>SL</sub>) (vert) qui interagissent (2). En conditions de stress, le groupe phosphate (P) sur l'histidine de PhyK (1) est transféré au domaine PhyR<sub>REC</sub>. Les deux domaines PhyR<sub>REC</sub> et PhyR<sub>SL</sub> se dissocient, l'hélice  $\alpha$  (jaune) change de conformation stabilisant la conformation ouverte de PhyR (3). Dans cette conformation, le domaine PhyR<sub>SL</sub> peut interagir avec NepR (bleu) (4). Le facteur  $\sigma^{\text{T}}$  peut alors interagir avec l'ARNP (violet) (5) pour activer la transcription des gènes cibles (Herrou *et al.*, 2012).

Ce système bien qu'exceptionnel nous sera d'une grande utilité pour l'analyse des résultats obtenus pour l'interaction entre le domaine périplasmique de CnrY et CnrH (**Résultats partie I**).

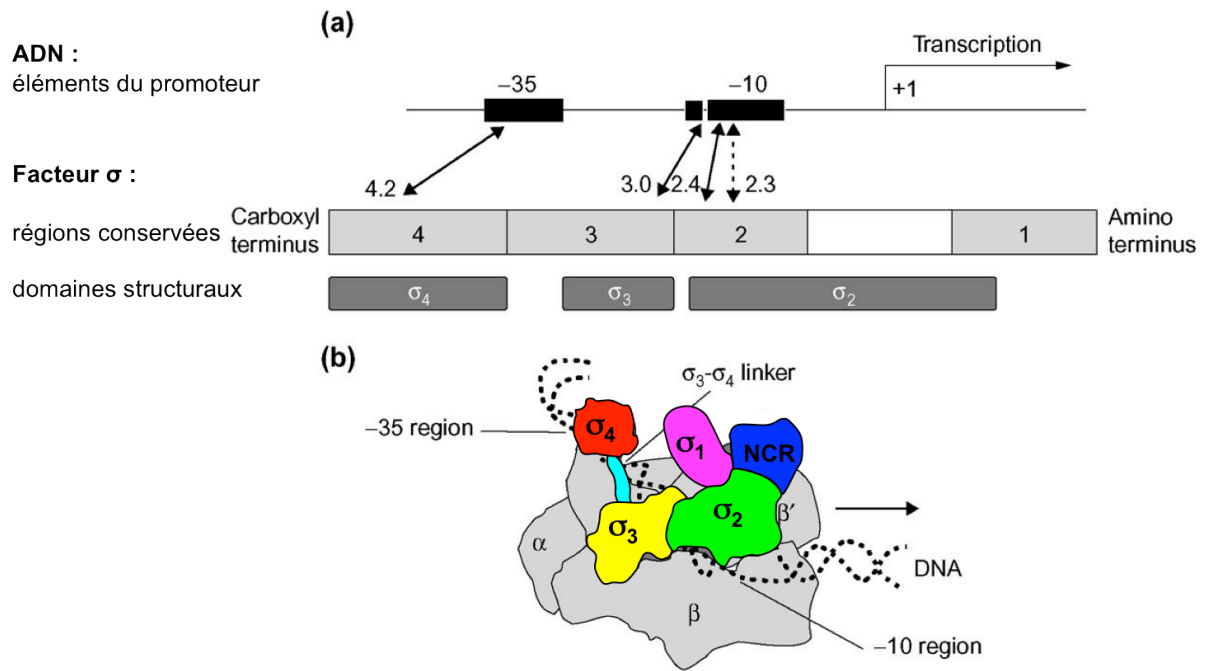
## 4. Systèmes dépendants des facteurs sigma ( $\sigma$ ) de type ECF

Le troisième mécanisme utilisé par la bactérie pour réagir à la présence de facteurs de stress est le système dépendant des facteurs  $\sigma$  de type ECF. Le système CnrYXH qui est l'objet d'étude de cette thèse fait partie de cette catégorie.

### a. Généralités sur les facteurs $\sigma$

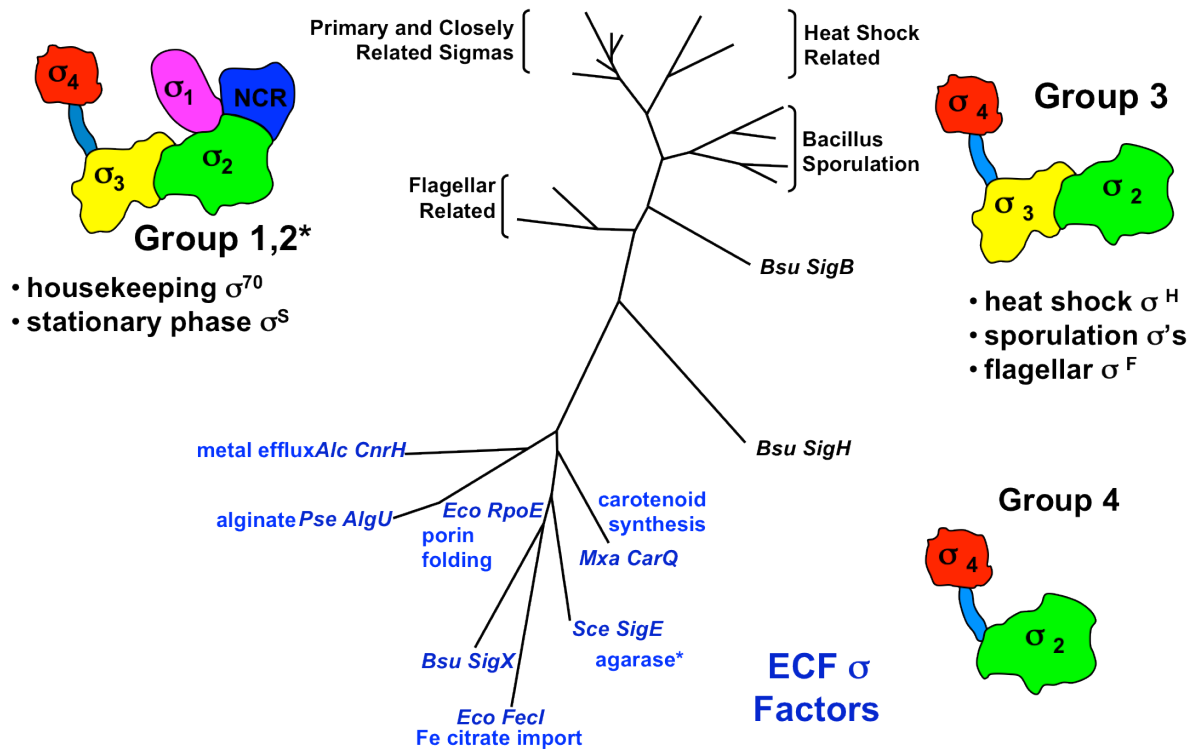
Comme décrit précédemment (partie **I.C.1.**), les facteurs  $\sigma$  sont des facteurs de transcription indispensables à l'initiation de la transcription. La majorité des facteurs  $\sigma$  appartient à la famille des facteurs  $\sigma^{70}$  qui est composée de quatre groupes (groupes 1 à 4, **Figure 9**) (Lonetto *et al.*, 1992; Paget et Helmann, 2003).

Les facteurs  $\sigma$  possèdent quatre domaines appelés  $\sigma 1$  à  $\sigma 4$ . Les analyses de séquences, la sélection de mutants et la détermination de la structure de certains facteurs  $\sigma$  ont permis d'identifier les régions conservées 1.1 (correspondant au domaine  $\sigma 1$ ), 1.2 et 2.1 à 2.4 (domaine  $\sigma 2$ ), 3.0 à 3.2 (domaine  $\sigma 3$ ) et 4.1 à 4.2 (domaine  $\sigma 4$ ), qui ne sont pas présents dans tous les facteurs  $\sigma$  (Campbell *et al.*, 2002, 2003; Lonetto *et al.*, 1992). Les domaines  $\sigma 2$  (région 2.4) et  $\sigma 4$  (région 4.2) reconnaissent les éléments -10 et -35 du promoteur, respectivement. Ces domaines, très conservés, sont suffisants pour que le facteur sigma soit fonctionnel (Gruber et Gross, 2003; Lonetto *et al.*, 1994). La région 2.3 du domaine  $\sigma 2$  interagit avec le brin codant de l'ADN dans le complexe ouvert. La région 1.1 empêcherait la fixation du facteur  $\sigma$  sur l'ADN. La région 3.0 du domaine  $\sigma 3$  interagit avec l'ADN en amont de l'élément -10 du promoteur ("extended-10") (**Figure 8**).



**Figure 8** : Caractéristiques structurales du facteur  $\sigma^{70}$  d'*E. coli*. **(a)** La séquence protéique a été divisée en quatre régions sur la base de la conservation des séquences des autres membres de la famille des facteurs  $\sigma^{70}$ . Les flèches indiquent des interactions entre l'ADN et le facteur  $\sigma$ . Les 3 domaines structuraux ( $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$ ), dont la structure cristallographique a été obtenue, sont représentés sous les régions conservées de  $\sigma^{70}$ . **(b)** Modèle de l'interaction entre la région promotrice de l'ADN (pointillés) et l'ARNP. Les différents domaines du facteur  $\sigma$  sont représentés :  $\sigma_1$ ,  $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$  en magenta, vert, jaune et rouge, respectivement. Le domaine de liaison entre les domaines  $\sigma_3$  et  $\sigma_4$  est en cyan et la région non conservée (NCR) présente que dans les groupes 1 et 2 des facteurs  $\sigma$  est en bleu (Paget et Helmann, 2003).

Les facteurs  $\sigma$  des groupes 1 et 2 possèdent une organisation structurale (domaines  $\sigma_1$  à  $\sigma_4$ ) similaire même si seuls les facteurs  $\sigma$  du groupe 1 sont essentiels pour la croissance cellulaire. Ceux du groupe 3 sont composés des domaines  $\sigma_2$  à  $\sigma_4$  et régulent les fonctions telles que le choc thermique ou la réponse au stress général, la sporulation et la biosynthèse de flagelles (Gruber et Gross, 2003). Les facteurs  $\sigma$  du groupe 4 sont composés des domaines  $\sigma_2$  et  $\sigma_4$  reliés par un connecteur. Ils sont spécialisés dans la réponse au stress oxydant ou extracellulaire, ce qui pourrait être la cause de leur grande diversité de séquences et de régulations (Helmann, 2002). Les facteurs  $\sigma$  qui régulent la réponse de la bactérie à un stimulus extracytoplasmique de la bactérie sont appelés ECF pour "ExtraCytoplasmic Function" ( $\sigma^{ECF}$ ) (Lonetto *et al.*, 1994) (**Figure 9**).



**Figure 9 :** Arbre phylogénétique sans racine des facteurs  $\sigma$ , appartenant à la famille des facteurs  $\sigma^{70}$ , sélectionnés dans différents organismes. Les facteurs  $\sigma$  sont classés en quatre groupes (1 à 4) selon leurs séquences. Ils sont composés des domaines  $\sigma_1$  à  $\sigma_4$  et d'un connecteur (bleu clair). Les groupes 1 et 2 possèdent en plus une région non conservée (NCR en bleu foncée) (Gruber et Gross, 2003; Lonetto *et al.*, 1994).

Toutes les bactéries possèdent au moins un facteur  $\sigma$  du groupe 1, qui contrôle l'expression des gènes de ménage, et des facteurs  $\sigma$  dits accessoires, qui contrôlent l'expression d'un nombre variable de gènes : de quelques uns à quelques centaines. Par le jeu des cascades d'activation, un facteur  $\sigma^{\text{ECF}}$  comme RpoE contrôle positivement l'expression des facteurs  $\sigma$  RpoH, RpoD et RpoN qui régulent à leur tour la transcription de dizaines de gènes (Bury-Moné *et al.*, 2009) alors que le facteur  $\sigma^{\text{ECF}}$  CnrH, appartenant au complexe CnrYXH, régule l'expression de trois gènes de régulation et trois gènes de résistance.

## b. Les facteurs sigma ECF ( $\sigma^{\text{ECF}}$ )

De récentes recherches sur les génomes microbiens ont démontré le rôle important des facteurs  $\sigma^{\text{ECF}}$  dans la transduction du signal chez les bactéries (Starón *et al.*, 2009) et notamment leur contribution dans la virulence chez un grand nombre de pathogènes (Rowley, 2006). Ceci est lié au fait que les facteurs  $\sigma^{\text{ECF}}$  coordonnent des réponses ciblées contre des

facteurs de stress extracellulaires et des facteurs de stress rédox. Or les défenses de l'hôte auxquelles les bactéries pathogènes doivent faire face sont précisément de ce type. L'activation d'un facteur  $\sigma^{\text{ECF}}$  en réponse à un stress extracellulaire résulte de l'interaction du facteur de stress avec un récepteur de la membrane cytoplasmique, ou dans quelques cas, de la membrane externe des bactéries à Gram négatif.

Au repos, chaque facteur  $\sigma^{\text{ECF}}$  est séquestré par un partenaire membranaire appelé facteur anti- $\sigma$ . Récemment, un motif de quatre hélices a été identifié dans les structures de deux facteurs anti-sigma (RseA de *E. coli* et RsiW de *B. subtilis*) (Campbell *et al.*, 2003, 2007). Ce motif long de 70 résidus, définit l'interface entre le facteur anti- $\sigma$  et  $\sigma^{\text{ECF}}$ . Ce domaine est appelé ASD pour "anti-sigma domain" (Campbell *et al.*, 2007). Le dispositif est complété par un senseur dont la conformation ou l'activité est sensible à un facteur de stress donné. En règle générale, les micro-organismes expriment d'autant plus de facteurs  $\sigma^{\text{ECF}}$  qu'ils sont soumis à diverses fluctuations environnementales (Paget et Helmann, 2003; Staroń *et al.*, 2009).

Environ 1700 à 3000  $\sigma^{\text{ECF}}$  provenant de 500 espèces ont été classés dans un des 43 groupes conservés de  $\sigma^{\text{ECF}}$  selon leur similarité de séquences ( $\sigma$ /anti- $\sigma$ ) (Staroń *et al.*, 2009). Une classification a été proposée selon les modes de transduction du signal (Mascher, 2013). Dans la suite de ce chapitre, nous présenterons les modes de transduction du signal les plus connus à ce jour. Nous détaillerons plus particulièrement les mécanismes d'activation des facteurs  $\sigma^{\text{ECF}}$  à travers des exemples de systèmes bien connus et qui sont impliqués dans l'homéostasie des métaux.

### **i. L'activation des facteurs $\sigma^{\text{ECF}}$ résulte de la protéolyse de son facteur anti- $\sigma$**

La transduction d'un signal à travers la membrane par protéolyse contrôlée est utilisée par des systèmes divers, notamment chez les eucaryotes, où elle intervient dans la régulation du facteur de transcription SREBP (Sterol regulatory element binding protein) (Espenshade et Hughes, 2007). Le mécanisme de protéolyse régulée intervient aussi dans la maturation des protéines TM neuronales APP (Amyloïde Precursor Protein). Ces protéines sont clivées séquentiellement par deux protéases (sécrétases) libérant les peptides  $\beta$ -amyloïde qui s'agrègent en présence de zinc et de cuivre. Une protéolyse altérée de la protéine APP est à

l'origine de la maladie d'Alzheimer (Faller *et al.*, 2014; Parkin *et al.*, 2002). La protéolyse est donc une solution générale pour un problème biologique commun à toutes les cellules.

Comment la protéolyse permet-elle d'activer les facteurs  $\sigma^{\text{ECF}}$  ?

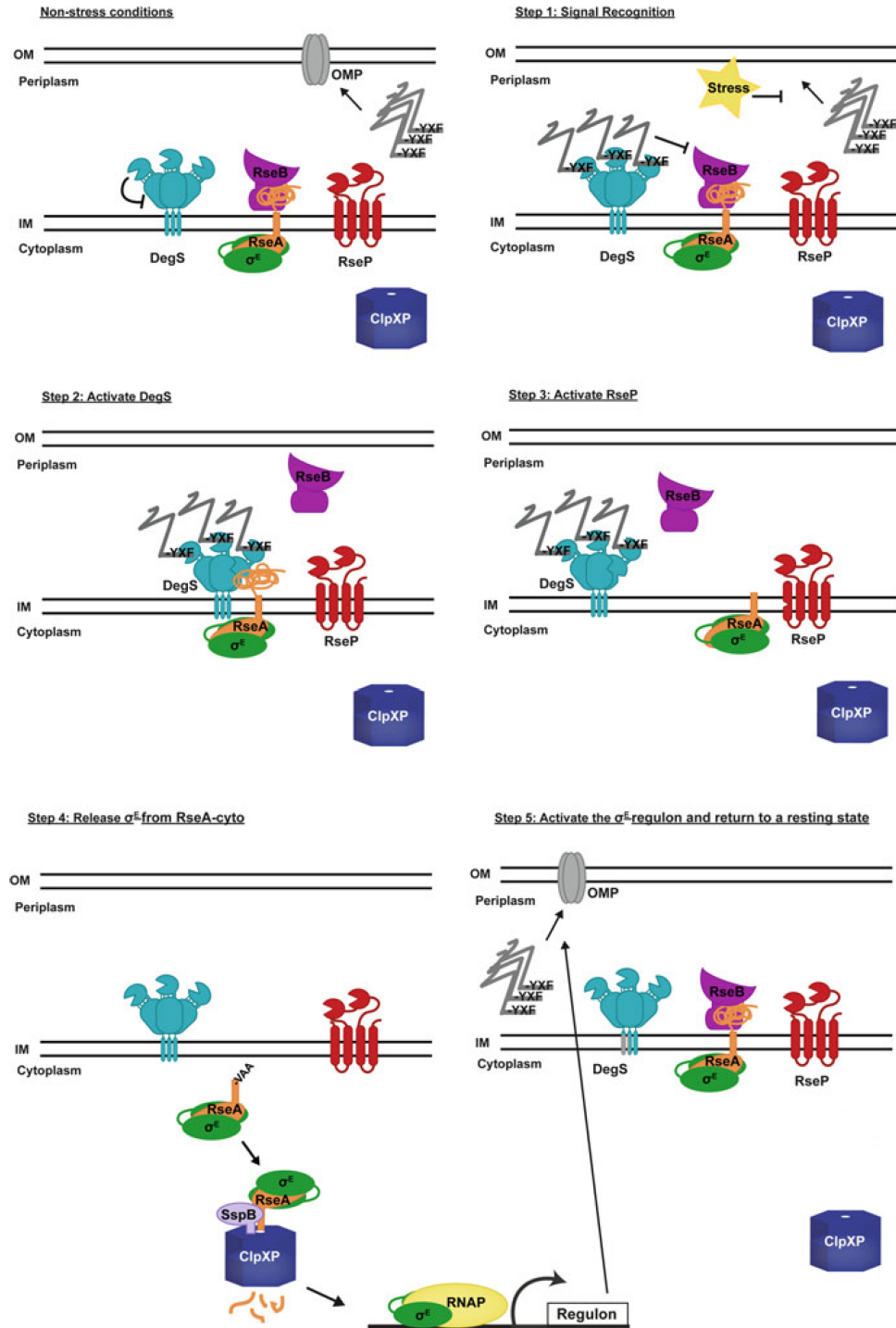
De manière simplifiée, le système protéolytique consiste en deux protéases qui clivent, séquentiellement, une protéine transmembranaire régulatrice (facteur anti- $\sigma$ ) de part et d'autre de la membrane. Cela résulte en la libération du facteur sigma dans le cytoplasme qui possède des fonctions biologiques distinctes (Brown *et al.*, 2000). Le signal est généralement perçu du côté extracytoplasmique de la membrane. La première protéase est souvent une protéase à sérine ou à aspartate (Brown *et al.*, 2000; Chen et Zhang, 2010). Pour la seconde protéase, plusieurs familles ont été identifiées permettant le clivage intramembranaire, notamment la protéase du site 2 ("site-2-protease" (S2P)) (Chen et Zhang, 2010).

### ❖ Le système RpoE ( $\sigma^{\text{E}}$ )-RseA chez *E. coli*

Le système permettant l'activation du facteur  $\sigma^{\text{ECF}}$  par protéolyse le mieux caractérisé est celui de la régulation de  $\sigma^{\text{E}}$  par RseA chez *E. coli*. Le facteur  $\sigma^{\text{E}}$  est impliqué dans la réponse au stress d'enveloppe (Hayden et Ades, 2008; De Las Peñas *et al.*, 1997) mais aussi dans la résistance au zinc, au cadmium et au cuivre (Egler *et al.*, 2005) chez *E. coli*.  $\sigma^{\text{E}}$  n'a pas été identifié dans un premier temps pour son implication dans la réponse aux stress d'enveloppe ou métalliques mais pour son rôle dans la transcription du gène codant un facteur sigma alternatif,  $\sigma^{32}$  en réponse à un choc thermique. Une connexion entre  $\sigma^{\text{E}}$  et le stress d'enveloppe a été trouvée par criblage génétique démontrant que l'activité de  $\sigma^{\text{E}}$  augmente en fonction de la surproduction des protéines de la membrane externe (OMP) (Mecsas *et al.*, 1993), de l'inactivation des gènes codant des chaperons et des protéases cytoplasmiques (Missiakas *et al.*, 1996), et de la délétion de gènes impliqués dans la biosynthèse de liposaccharides (Rhodius *et al.*, 2006). A ce jour, les systèmes  $\sigma^{\text{E}}$  ont été identifiés dans plus de 100 génomes bactériens séquencés et partagent un grand nombre de similarités (Staroń *et al.*, 2009). Malgré les différences de réponses à des stress spécifiques entre les espèces, il a été démontré, chez toutes les bactéries, que  $\sigma^{\text{E}}$  est impliqué dans les processus associés à l'enveloppe cellulaire (Rowley *et al.*, 2006; Staroń *et al.*, 2009). Le groupe majeur des gènes conservés dans le régulon  $\sigma^{\text{E}}$  code des protéases, des chaperons indispensables pour la synthèse, l'assemblage et le maintien des OMP et des lipopolysaccharides (Rhodius *et al.*, 2006).



Chez *E. coli*,  $\sigma^E$  est inhibé par RseA. Dans des conditions où la cellule n'est pas stressée, le facteur  $\sigma^E$  est séquestré à la membrane du côté cytoplasmique par le facteur anti- $\sigma$  transmembranaire, RseA. RseB, un régulateur négatif mineur de  $\sigma^E$ , est une protéine soluble dans le périplasme qui interagit avec RseA et la stabilise (Ades *et al.*, 1999). Les facteurs de stress qui endommagent l'enveloppe cellulaire dont le repliement des OMP, active la cascade protéolytique. Les résidus en C-terminal des OMP mal repliées (YxF), marqueurs du stress d'enveloppe, interagissent avec le senseur ancré à la membrane interne, DegS. DegS appartient à la famille des protéases à sérine, avec un ou plusieurs domaines PDZ en C-terminal (Clausen *et al.*, 2011). L'interaction entre YxF et le domaine PDZ a pour effet de stabiliser le site actif de DegS (Wilken *et al.*, 2004). Son activité protéolytique alors activée, elle pourra cliver le domaine périplasmique de RseA et ainsi inactiver RseB (Sohn *et al.*, 2007), créant un substrat pour la seconde protéase impliquée dans la transduction du signal, RseP. RseP est une métalloprotéase à zinc, transmembranaire, possédant deux domaines PDZ du côté périplasmique. Elle clive RseA dans la région cytoplasmique de l'hélice membranaire, par un processus de protéolyse intramembranaire régulée où les protéases interviennent séquentiellement (Li *et al.*, 2009). Par suite de la coupure réalisée par RseP,  $\sigma^E$  est libéré dans le cytoplasme en complexe avec un fragment cytoplasmique de RseA. Ce complexe est dégradé par ClpXP, une protéase dépendant de l'ATP, avec l'assistance d'une protéine adaptatrice, SspB (Campbell *et al.*, 2007; Flynn, 2004). ClpXP finalise la libération de  $\sigma^E$  qui peut alors conférer à la polymérase la possibilité d'exprimer les gènes cibles (**Figure 10**).



**Figure 10 :** Régulation de la réponse au stress d'enveloppe dépendant du facteur  $\sigma^E$ . En absence de stress, les OMP sont bien repliés et  $\sigma^E$  est inhibé à la membrane par RseA. En conditions de stress, l'interaction des domaines PDZ de DegS avec l'extrémité C-terminale exposée par les OMP mal repliés active DegS qui clive le domaine périplasmique de RseA levant l'inhibition de RseA par RseB. L'hélice TM de RseA est ensuite clivée par RseP dont le domaine PDZ reconnaît l'extrémité C-terminale générée par DegS. Ce clivage permet la libération, dans le cytoplasme, de  $\sigma^E$  en complexe avec le fragment cytosolique de RseA. Ce fragment de RseA est enfin protéolysé par ClpXP après avoir été pris en charge par SspB qui reconnaît son extrémité C-terminale.  $\sigma^E$  peut alors activer la transcription des gènes sous son contrôle (figure tirée du livre Bachinger et Ades, Subcellular Biochemistry, 2013).

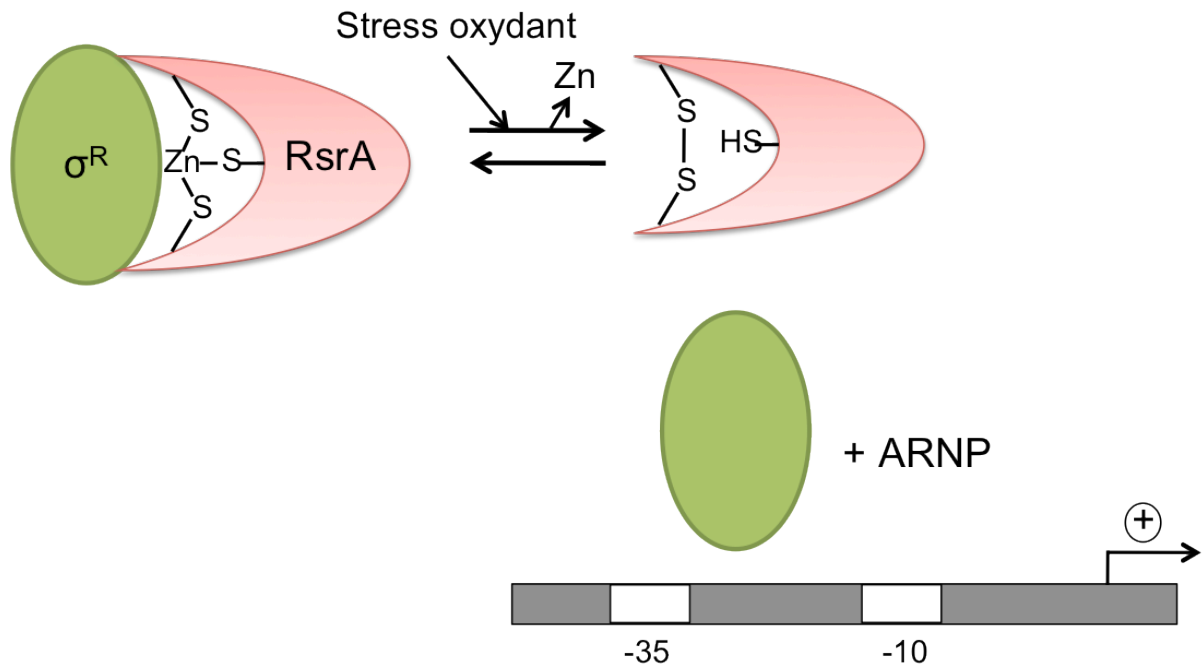
## ii. L'activation du facteur $\sigma^{\text{ECF}}$ résulte du changement conformationnel de son facteur anti- $\sigma$

Un nombre restreint de facteurs anti- $\sigma$  sont solubles dans le cytoplasme. Malgré la similarité de séquences qu'ils partagent avec les autres facteurs anti- $\sigma$  transmembranaires (Campbell *et al.*, 2007), le mécanisme de libération du facteur  $\sigma$  est très différent (**Figure 11**).

### ❖ Le système SigR ( $\sigma^{\text{R}}$ )-RsrA de *Streptomyces coelicolor*

Le système  $\sigma^{\text{R}}$ -RsrA de *Streptomyces coelicolor* (Kang *et al.*, 1999) illustre ce mode de transduction du signal. Le facteur anti- $\sigma$ , RsrA, cytoplasmique possède de nombreux résidus cystéine conservés qui sont les cibles potentielles du stress oxydant et qui coordinent un zinc. Ce facteur anti- $\sigma$  est un ASD appelé ZAS pour "zinc-binding anti-sigma" (Campbell *et al.*, 2007).

Dans des conditions où la bactérie n'est pas stressée, le facteur  $\sigma^{\text{R}}$  est maintenu à l'état inactif par liaison avec son facteur anti- $\sigma$ , RsrA. Cette interaction s'effectue *via* le domaine ZAS. Cet état est stabilisé par le zinc coordonné par les résidus cystéines de RsrA (Li *et al.*, 2003). Un changement du potentiel redox dans le cytoplasme, résultat d'un stress oxydant, induit la formation d'un pont disulfure entre les cystéines ligands du zinc. Cela aboutit à la libération du cofacteur zinc et au changement de conformation du facteur anti- $\sigma$  (Bae *et al.*, 2004). Le facteur  $\sigma^{\text{R}}$  est ainsi libéré et peut interagir avec l'ARNP pour activer la transcription de gènes de résistance au stress oxydant ainsi que les gènes de régulation (**Figure 11**). Ce mécanisme, contrairement au processus de protéolyses des facteurs anti- $\sigma$ , est réversible, ce qui permet de recycler le facteur anti- $\sigma$  (Mascher, 2013).



**Figure 11** : Mécanisme d'activation du facteur  $\sigma$  par un changement de conformation du facteur anti- $\sigma$ . En condition de non stress, le facteur anti- $\sigma$  RsrA (rouge), interagit avec le facteur  $\sigma^R$  (vert) *via* le ZAS. Lors d'un stress oxydant, les atomes de soufre (-S-) des cystéines ligands du zinc (Zn) forment un pont disulfure, aboutissant à la libération du facteur  $\sigma^R$  qui pourra interagir avec l'ARNP pour activer l'initiation de la transcription des gènes cibles (d'après Antelmann et Helmann, 2011).

### iii. L'activation du facteur $\sigma^{ECF}$ résulte des interactions protéine:protéine entre les partenaires du système de transduction du signal

#### ❖ Le système FecI-FecR de *E. coli*

Le fer sous sa forme oxydée  $Fe^{3+}$  précipite et est par conséquent très peu disponible dans l'environnement. Les bactéries produisent donc des molécules de faible poids moléculaire, appelées sidérophores qui chélatent l'ion ferrique. L'import du fer *via* des sidérophores s'opère par un transporteur de la membrane externe qui, dans certains cas, est également un détecteur capable de signaler la présence de fer assimilable à la cellule par l'intermédiaire d'un facteur  $\sigma^{ECF}$  (Mettrick et Lamont, 2009).

Deux des systèmes les plus étudiés sont les systèmes Fec d'*E. coli* et Fpv de *Pseudomonas aeruginosa* (Mahren *et al.*, 2005; Mettrick et Lamont, 2009). Ces systèmes sont d'autant plus étudiés qu'il a été démontré chez certaines bactéries pathogènes, qu'ils étaient liés à l'expression de la virulence (Ratledge et Dover, 2000). L'activation du facteur  $\sigma^{ECF}$  s'effectue

par des interactions entre les protéines de ces systèmes qui induisent des changements de conformation et des protéolyses. Toutefois, la régulation de la transcription des gènes d'acquisition des sidérophores est différente pour ces deux systèmes. En effet, chez *P. aeruginosa*, la régulation serait assurée par deux facteurs  $\sigma^{\text{ECF}}$  dissociés de leur facteur anti- $\sigma$  alors que chez *E. coli*, l'activité du facteur  $\sigma^{\text{ECF}}$  nécessiterait la partie cytoplasmique de son anti- $\sigma$  (Braun *et al.*, 2006; Llamas *et al.*, 2014).

Le régulon *fec* est composé de sept gènes contigus et deux promoteurs : en aval du promoteur *fecA* se trouvent les gènes *fecABCDE* qui interviennent directement dans l'import du fer à partir de l'environnement et, en aval du promoteur *fecI* se trouvent les gènes *fecI* et *fecR* impliqués dans la régulation du système (**Figure 12**).

FecA est une protéine de la membrane externe qui transporte le fer dicitrate ((Fe<sup>3+</sup> citrate)<sub>2</sub>) dans le périplasma par un mécanisme qui nécessite de l'énergie. Cette énergie est apportée par un complexe de trois protéines TM possédant un domaine périplasmique : TonB, ExbB et ExbD. FecBCDE est une perméase appartenant à la famille des transporteurs de type ABC (voir partie **II.B.2.**), permettant l'import du fer dicitrate dans le cytoplasme d'*E. coli* (Braun, 1997, 2002) (**Figure 12**).

FecA est aussi une protéine senseur qui en fonction de la concentration en fer dans l'environnement, peut déclencher une cascade de signalisation dont l'objectif est la régulation de l'opéron *fecABCDE* (Kim *et al.*, 1997). En plus de FecA, deux protéines sont nécessaires au mécanisme de signalisation : le facteur  $\sigma^{\text{ECF}}$  FecI et le facteur anti- $\sigma$  FecR qui est une protéine bitopique dont les domaines N-terminal et C-terminal sont localisés dans le cytoplasme et le périplasma, respectivement (Braun, 1997; Welz et Braun, 1998).

FecA est composée d'un tonneau de brins  $\beta$ , d'un domaine bouchon dans la lumière de ce tonneau et d'un domaine de signalisation N-terminal important pour la transcription des gènes *fec* (Mahren *et al.*, 2005). La fixation du fer dicitrate sur FecA induit de nombreux changements de conformation qui permettent aux brins  $\beta$  de se refermer sur le substrat et qui agissent sur la structure du domaine de signalisation (Ferguson *et al.*, 2002, 2007). Ce dernier, flexible, est situé dans le périplasma où il interagit avec la partie périplasmique de FecR (Enz *et al.*, 2000).

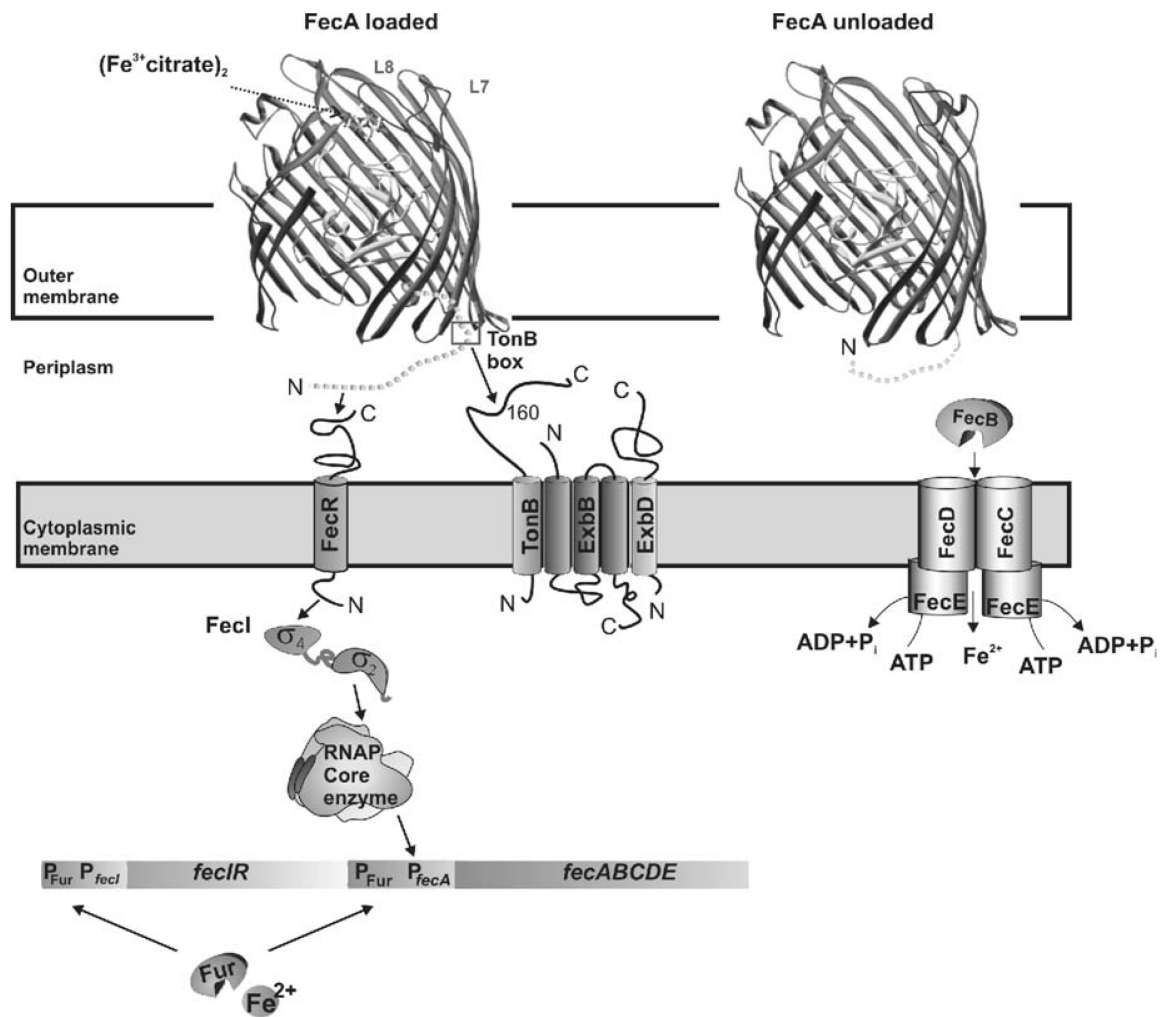
La manière dont FecR reçoit le signal et la façon dont il le transmet à FecI n'est pas encore clair (Braun *et al.*, 2006; Ferguson *et al.*, 2007). Toutefois, il semblerait que l'interaction entre le domaine de signalisation de FecA et l'extension périplasmique de FecR soit à l'origine de

changements de conformation de FecR (Ochs *et al.*, 1995) qui deviendrait sensible à la protéolyse. Cette protéolyse serait effectuée par la protéase RseP (décrite précédemment pour la protéolyse intramembranaire de RseA, partie **I.C.4.b.i.**) (Wriedt *et al.*, 1995). Il a été démontré que le fragment possédant les résidus 1 à 85 de FecR (FecR<sub>1-85</sub>) induit une transcription constitutive des gènes *fec* (Braun *et al.*, 2006; Ochs *et al.*, 1995) ce qui suggère que la libération de ce fragment par protéolyse serait le mécanisme par lequel le signal, généré par les changements de conformation de FecA et FecR, est transmis à FecI. De plus, FecR est nécessaire à l'activité de FecI. Ainsi, l'interaction de FecR<sub>1-85</sub> (qui mimerait la protéine complète FecR activée par le signal provenant de FecA en complexe avec du fer dicitrate) avec FecI induirait une conformation active de FecI ou protégerait FecI d'une dégradation. Le complexe FecR:FecI:ARNP initierait donc la transcription des gènes et, dans ce cas, FecR aurait une fonction définie comme "pro" sigma dans son état actif (Mettrick et Lamont, 2009; Ochs *et al.*, 1995).

Contrairement à d'autres groupes de gènes codant les facteurs  $\sigma$  et anti- $\sigma$  comme *RseA-rpoE* et *cnrY-cnrH*, les gènes *fecIR* ne sont pas autorégulés. En effet, ils sont soumis à l'influence d'un autre régulateur, le métallorégulateur Fur (voir partie **I.C.2.a.**) (Angerer et Braun, 1998). Lorsque le fer est en excès dans la cellule, Fur interagit avec Fe<sup>2+</sup> et inhibe la transcription de l'opéron *fec* : *fecIR* (Escolar *et al.*, 1999) et les gènes du système d'import du fer, *fecABCDE* (Braun *et al.*, 2006) (**Figure 12**).

Ainsi, par des changements de conformation et des événements de protéolyses induits par des interactions protéine:sidérophore et protéine:protéine, le facteur  $\sigma^{\text{ECF}}$  FecI est activé et permet de réguler l'homéostasie du fer.

Plusieurs homologues de FecI sont retrouvés chez d'autres organismes mais n'ont pas tous la même fonction. Par exemple, chez *Ralstonia solanacearum*, lorsque ces bactéries entrent en contact avec des cellules de plante, l'homologue de FecI, PrhI est activé. Il régule l'expression des gènes *hrp* codant un système de sécrétion de type III impliqué dans la colonisation et l'infection des cellules de plante (Aldon *et al.*, 2000). Il existe même des paralogues de FecI dans un même organisme comme *C. metallidurans* CH34 qui en compte au moins 6. Ces derniers seront évoqués dans le chapitre qui sera consacré à la présentation de *C. metallidurans* CH34 (partie **III.C.**).



**Figure 12 :** Représentation schématique du transport du fer dicitrate et son système de régulation. FecA aidée par le complexe TonB:ExbB:ExbD importent le fer dicitrate dans le périplasm où il est pris en charge par le transporteur ABC, FecBCDE. Une fois dans le cytoplasme, le fer pourra accomplir ses fonctions métaboliques. Si il est en excès, il interagira avec Fur qui réprimera la transcription de *fecIR* et *fecABCDE*. Par ailleurs, FecA, par des changements de conformation, interagit avec FecR. FecR est ensuite protéolysée au niveau de la membrane cytoplasmique par RseP (non illustré). Le domaine cytoplasmique de RseP forme un complexe avec FecI qui serait pris en charge par l'ARNP pour l'activation de la transcription de *fecABCDE* (Braun *et al.*, 2006).

Dans ce chapitre, nous avons vu que dans des conditions de stress, les bactéries couplent la détection des facteurs de stress avec l'activation des FT permettant d'ajuster leur métabolisme. L'homéostasie, qui est un processus commun à toutes les bactéries, regroupe les mécanismes décrits précédemment leur permettant de conserver un équilibre de fonctionnement. Parmi les nombreux facteurs de stress, le stress métallique est celui qui nous intéresse puisque que notre système d'étude répond aux métaux. L'homéostasie des métaux fera donc l'objet du deuxième chapitre de cette introduction.



## II. Homéostasie des métaux lourds

### A. Définition de l'homéostasie des métaux lourds

Depuis plusieurs décennies, le terme "métaux lourds" est largement utilisé par la communauté scientifique bien que cette notion soit un concept avant tout empirique (Duffus, 2002). Certains auteurs définissent les métaux lourds comme les éléments métalliques ayant une masse volumique supérieure à  $5\text{g/cm}^3$  (Nies, 1999). Pour d'autres, les métaux lourds sont des éléments qui ont des propriétés de forte toxicité. Plusieurs autres définitions existent sans qu'aucune ne soit universelle.

Les métaux, selon leur nature et/ou leur concentration, peuvent être toxiques pour la cellule. Certains métaux lourds tels que le cadmium, le mercure, l'arsenic, le palladium, l'or ou le plomb n'ont pas de rôle biologique connu (Nies, 1999) et sont purement toxiques pour la cellule car ils rentrent en compétition avec les métaux essentiels. D'autres métaux sont, quant à eux, des éléments essentiels à la vie puisqu'ils participent à de nombreuses réactions métaboliques dans toutes les cellules vivantes (Williams et Silva, 2005). Certains sont des cofacteurs indispensables pour le fonctionnement de ses enzymes. Par exemple, le zinc est un cofacteur des ADN polymérase (Springgate et Loeb, 1973), le nickel est nécessaire à certaines uréases (Farrugia *et al.*, 2013) et hydrogénases (Li *et al.*, 1987) et le cobalt est un cofacteur de certaines enzymes du métabolisme de la méthionine ou de la vitamine B12 (Kräutler, 2005). Cependant, ces métaux deviennent toxiques pour la cellule lorsque leur concentration est trop élevée. En effet, ils peuvent inhiber des enzymes en se fixant sur leurs sites actifs ou en rentrant en compétition avec le véritable cofacteur de l'enzyme (Chillappagari *et al.*, 2010). Ils peuvent aussi altérer la structure des acides nucléiques (Valko *et al.*, 2005) ou encore être responsables de la formation de radicaux libres (Imlay et Imlay, 1996).

L'homéostasie des métaux permet le maintien d'un subtil équilibre de la concentration intracellulaire en métaux pour éviter toute toxicité et carence. Les bactéries ont développé des systèmes de contrôle de la concentration en métaux permettant un apport suffisant de métaux

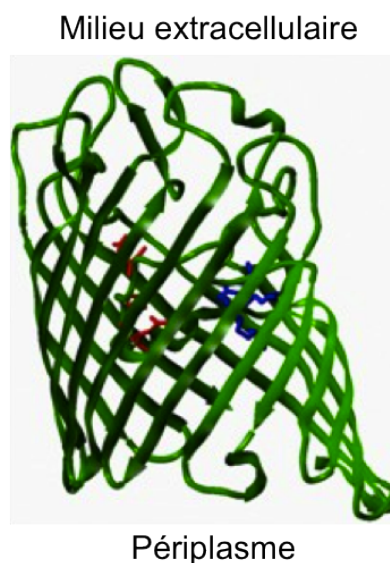
et une élimination efficace de l'excès pour maintenir l'intégrité de la cellule. Par la suite, ces systèmes seront présentés en essayant quand cela est possible de nous focaliser sur les systèmes présents chez *C. metallidurans* CH34.

## **B. Import des ions métalliques**

Chez les bactéries à Gram-négatif, les métaux doivent traverser la membrane externe et la membrane interne pour rejoindre le cytoplasme. L'accumulation des métaux dans le cytoplasme s'effectue par deux types de voies. La première, empruntée par une large gamme de substrats, est rapide, indépendante du métal, fait intervenir des protéines exprimées de manière constitutive et dépend uniquement d'un gradient chimiosmotique au travers de la membrane bactérienne. La seconde, plus spécifique, est inductible et requiert une consommation d'énergie souvent sous forme d'hydrolyse de l'ATP (Nies, 1999).

### **1. Import des ions métalliques dans le périplasme**

La majorité des ions métalliques entrent dans le périplasme par diffusion facilitée *via* des protéines appelées **porines** (Nikaido et Vaara, 1985). Les porines sont retrouvées sous la forme de trimères localisés dans la membrane externe. Chaque monomère est constitué de 16 brins  $\beta$  antiparallèles formant un tonneau  $\beta$  qui délimite un canal par lequel diffuse de petites molécules hydrophiles dont la masse moléculaire est inférieure à 600 Da (**Figure 13**). Les porines peuvent diffuser des ions ou d'autres petites molécules indépendamment de leur nature. Ce sont les porines non spécifiques telle que OmpF d'*E. coli*. D'autres porines transportent des substrats spécifiques comme LamB qui se charge de l'import du maltose et de la maltodextrine chez *E. coli* (Koebnik *et al.*, 2000).



**Figure 13** : Structure de la porine OmpF d'*E. coli* (Cowan *et al.*, 1992). La partie TM de OmpF est constituée par 16 feuillets  $\beta$  antiparallèles qui définissent le tonneau  $\beta$ . Les résidus colorés en rouge et bleu correspondent à des résidus acides et basiques, respectivement. Ces résidus localisés dans la région de constriction induisent un champ électrostatique qui impose des contraintes d'orientation aux solutés chargés qui empruntent le canal.

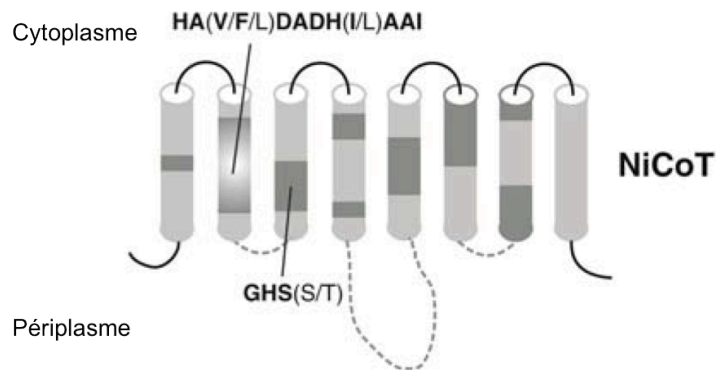
D'autres métaux, en plus faible concentration dans le milieu extracellulaire, sont pris en charge par des transporteurs actifs qui tirent l'énergie nécessaire à leur fonctionnement de la protéine TonB. La structure de ces protéines est similaire à celle des porines à l'exception de la présence d'une extension N-terminale qui se localise à l'intérieure du tonneau  $\beta$ , obstruant ainsi le canal (Koebnik *et al.*, 2000). Le transporteur FecA décrit précédemment (partie **I.C.4.b.iii.**) est un transporteur couplé à la protéine TonB.

## 2. Import des ions métalliques dans le cytoplasme

Une fois dans le périplasma, les métaux sont pris en charge par des systèmes tels que les perméases et les transporteurs de type ABC (ATP Binding Cassette).

**Les perméases** sont des transporteurs secondaires passifs permettant l'import des métaux dans le cytoplasme. Elles peuvent aussi transporter simultanément ou séquentiellement deux ions métalliques différents, soit dans le même sens (symport) soit dans des directions opposées (antiport). Les transporteurs NiCoT sont des perméases, composées de 8 domaines TM, qui importent le nickel et/ou le cobalt à travers la membrane interne (**Figure 14**). Elles ont des affinités pour leur substrat de l'ordre du nanomolaire. Certaines protéines de la famille

des NiCoT, comme HoxN de *Cupriavidus necator* H16 et NixA de *Helicobacter pylori*, sont spécifiques du nickel alors que NhfF de *Rhodococcus rhodochrous* J1 transporte les deux ions métalliques mais préfère le cobalt. L'obtention de la structure de ces transporteurs pourrait éclairer le mécanisme d'import encore inconnu (Eitinger *et al.*, 2005; Mulrooney et Hausinger, 2003).



**Figure 14** : Topologie du transporteur NiCoT. Les régions conservées des domaines TM sont colorées en gris foncé. Les boucles en pointillés désignent les séquences conservées. Les séquences en gras désignent des motifs très conservés parmi les membres de cette famille de transporteurs (Eitinger *et al.*, 2005).

**Les transporteurs de type ABC** utilisent l'énergie fournie par l'hydrolyse de l'ATP pour importer les métaux dans le cytoplasme (Higgins et Linton, 2001). Le transporteur le mieux caractérisé à ce jour est NikABCDE d'*E. coli* qui permet l'import du nickel à travers la membrane cytoplasmique (Navarro *et al.*, 1993). NikA est soluble dans le périplasma et est la seule protéine du complexe dont la structure est connue (Heddle *et al.*, 2003; Lebrette *et al.*, 2013). Elle lie le nickel et le délivre au complexe TM NikBC qui forme un pore. NikD et NikE sont deux ATPases cytoplasmiques qui hydrolysent l'ATP et fournissent cette énergie pour le transport du nickel (Mulrooney et Hausinger, 2003). FecABCDE appartient aussi à cette famille de transporteur. Ce complexe permet l'import du fer dans le cytoplasme. Le mécanisme de fonctionnement de ce transporteur est détaillé dans la section **I.C.4.b.iii.**

## C. Systèmes de résistance aux métaux

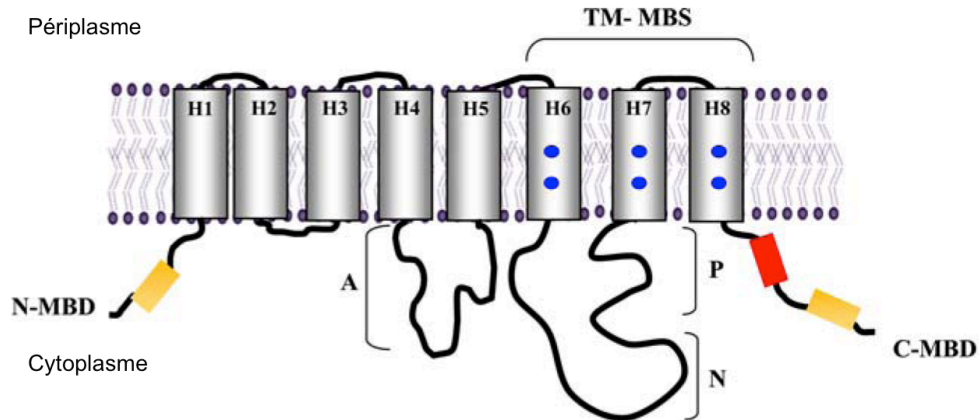
Pour résister à de trop fortes concentrations intracellulaires, les bactéries peuvent mettre en place des systèmes de résistance parmi lesquels la séquestration (Diels *et al.*, 1995), la conversion enzymatique (réduction et alkylation) (Silver et Phung, 1996) et l'efflux des ions métalliques. L'efflux apparait comme le mécanisme le plus utilisé par les bactéries. *C. metallidurans* CH34 possède un grand nombre de transporteurs (83 protéines dont la majorité sont des pompes d'efflux) ce qui pourrait expliquer sa capacité à résister à de fortes concentrations en ions métalliques (voir section III) (Rozycki et Nies, 2009).

### 1. Efflux des ions métalliques

#### a. Systèmes d'efflux de la membrane interne

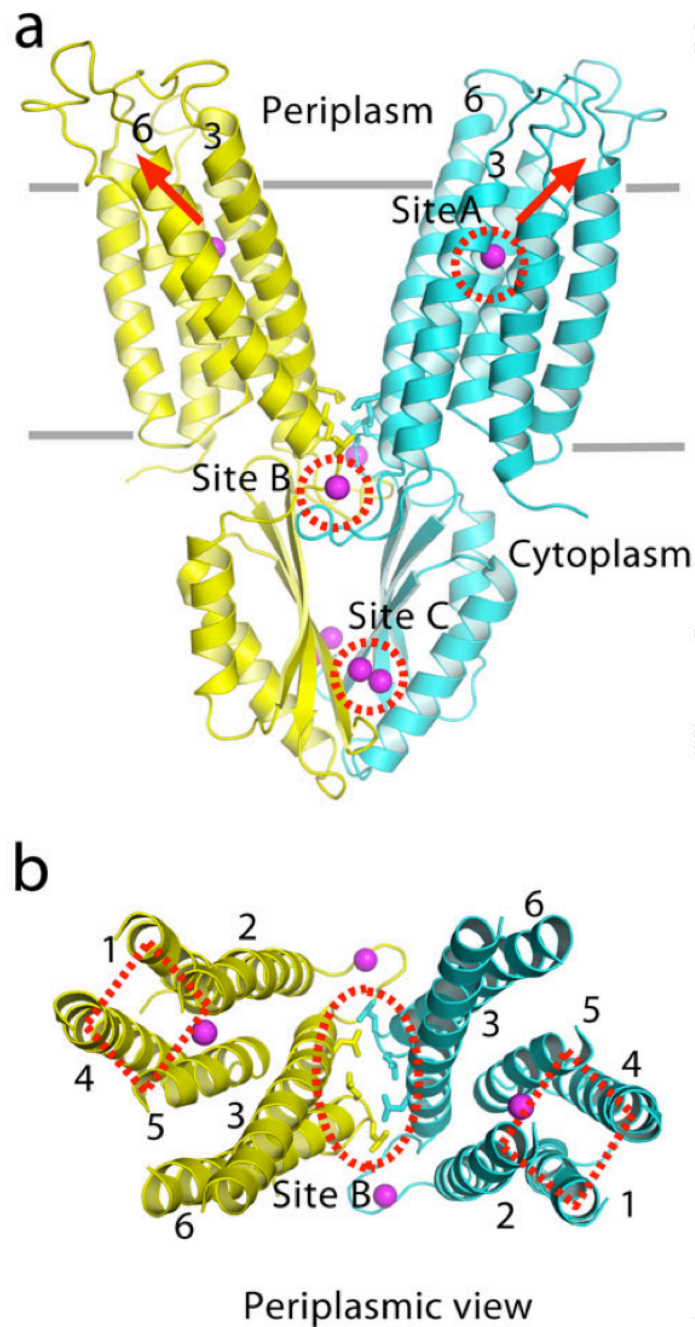
L'export des métaux du cytoplasme vers le périplasme est assuré par différents types de transporteurs :

- **Les ATPases de type P<sub>IB</sub>** constituent un sous-groupe de la superfamille des ATPases de type P dédié au transport des métaux à travers la membrane cytoplasmique (Argüello *et al.*, 2007; Nies, 2003). Ce transport nécessite de l'énergie qui est fournie par l'hydrolyse de l'ATP. Ces protéines se composent de 6 à 10 domaines TM, d'une boucle A (Actionneur) et de deux domaines de liaison à l'ATP (N pour "Nucleotide-binding" et P pour "Phosphorylation") cytosoliques (**Figure 15**). Le mécanisme d'action est similaire à celui des ATPases de type P (Palmgren et Nissen, 2011). La fixation de l'ATP sur le domaine N et de l'ion métallique sur la protéine entraîne la phosphorylation de l'aspartate conservé situé sur la boucle cytoplasmique reliant les hélices TM 6 et 7 (domaine P). Cette phosphorylation induit un changement conformationnel qui permet le transfert du métal vers le périplasme. Une des ATPases de type P<sub>IB</sub> la mieux caractérisée est la protéine CadA de *Staphylococcus aureus*. Elle permet la résistance aux ions cadmium mais aussi aux ions cobalt et zinc (Tsai *et al.*, 2002). Huit protéines de ce type sont présentes chez *C. metallidurans* CH34 contre environ 2 ou 3 chez les autres bactéries dont le génome a été séquencé (2 chez *E. coli*, 3 chez *B. subtilis*, *H. pylori*...) (Mergeay *et al.*, 2003).



**Figure 15** : Topologie du transporteur ATPases de type  $PI_B$ . Les domaines TM H1 à H8 sont indiqués. La position des domaines cytoplasmiques A, P et N est indiquée. Les résidus conservés (sphères bleues) dans les hélices H6, H7 et H8 forment le site métallique au niveau TM. Les domaines de liaison aux métaux au niveau des extrémités N-terminale et C-terminale sont représentés par des rectangles jaune et rouge (Argüello *et al.*, 2007).

- **Les protéines CDF** pour "Cation Diffusion Facilitators" forment une famille de transporteurs métalliques. Leurs substrats sont le zinc, le cobalt, le nickel, le cadmium et le fer. Leur transport s'effectue grâce à la force protomotrice de la membrane cytoplasmique (Nies, 2003). Parmi les pompes de ce type, CzcD de *C. metallidurans* CH34 permet de résister au zinc, cobalt et nickel en absence du système d'efflux, CzcCBA. La structure du transporteur de zinc YiiP d'*E. coli* (Lu *et al.*, 2009) a permis de mieux comprendre le mécanisme de fonctionnement des protéines de cette famille (**Figure 16**). YiiP est un dimère en forme de Y. Chaque monomère se compose de 6 segments TM, d'un domaine N-terminal et d'une longue partie C-terminale cytoplasmique riche en résidus histidines impliqués dans la reconnaissance de l'ion métallique. Chaque monomère possède 3 sites métalliques, deux mononucléaires (A et B) et un binucléaire (C). Le site C au niveau du domaine cytoplasmique détecte les ions  $Zn^{2+}$ , ce qui va entraîner un changement de conformation localisé. Celui-ci va être transmis aux hélices TM induisant une augmentation de l'affinité du site A pour le zinc. Ce site deviendra ainsi actif et permettra l'efflux du zinc dans le périplasmе (Coudray *et al.*, 2013; Gupta *et al.*, 2014). Cette régulation allostérique de la protéine effectrice permet de répondre à un stress très rapidement car elle ne nécessite pas la transcription de gènes codant de nouvelles pompes d'efflux.



**Figure 16** : Structure cristallographique du dimère de YiiP d'*E. coli* (code PDB : 2QFI) (Lu *et al.*, 2009). **(a)** Chaque monomère de YiiP est représenté en jaune et en cyan. Les ions zinc liés sont représentés par des sphères magenta entourées par des cercles en pointillés, rouge. Les sites métalliques A, B et C sont indiqués ainsi que les hélices 3 et 6. Les flèches indiquent la sortie des ions  $Zn^{2+}$  des sites actifs. **(b)** Vue de l'homodimère de YiiP depuis le périplasm. Le code couleur est identique à celui en **(a)**. Par souci de clarté, les boucles ont été supprimées. L'ellipse rouge marque les interactions hydrophobes à l'interface des hélices TM et les losanges rouges indiquent le faisceau des 4 hélices de chaque protomère. Les hélices sont numérotées.

- D'autres systèmes tels que les protéines de type **CHR**, **NreB** et **CnrT** permettent le transport des métaux depuis le cytoplasme (Nies, 2003).

Les membres de la famille des protéines CHR (CHromate Resistance) sont constitués de 10 domaines TM et sont probablement des pompes d'efflux du chromate qui utiliseraient le gradient chimiosmotique comme source d'énergie. Quatre protéines de ce type seraient présentes chez *C. metallidurans* CH34.

NreB, caractérisée chez *C. metallidurans* 31A et *E. coli*, confère une résistance au nickel.

Le gène *cnrT* est situé directement en aval des gènes *cnrCBA* codant pour la résistance au cobalt et au nickel chez *C. metallidurans* CH34. Le rôle de *cnrT* n'est pas très clair. Dans sa revue, Nies affirme que seul, le gène *cnrT* n'induit qu'une faible résistance au nickel sans avoir publié les données le prouvant (Nies, 2003).

NreB comme CnrT possèderaient 12 à 14 hélices TM et un domaine carboxyterminal riche en résidus histidines. Ces protéines ne possèdent pas de site de liaison à l'ATP potentiels et fonctionneraient grâce à la force protomotrice comme les protéines CDF.

## **b. Systèmes d'efflux de la membrane externe**

**La famille des protéines RND** pour "Resistance Nodulation et cell Division" permet d'ajuster la concentration en cations métalliques dans le périplasme. Le système RND est de type antiport cation-proton. Ce système est composé d'une pompe localisée dans la membrane interne (de la famille RND), d'une porine insérée dans la membrane externe (de la famille OMF pour "Outer Membrane Factor") et d'une protéine transpériplasmique (de la famille MFP pour "Membrane Fusion Protein") qui couple les deux protéines membranaires. La porine OMP et le transporteur sont des homotrimères d'environ 150 et 330 kDa respectivement alors que la protéine MFP est un hexamère d'environ 250 kDa. Le complexe de ces trois protéines forme un canal par lequel les cations métalliques en excès sont extrudés à l'extérieur de la cellule. Ce transport nécessite de l'énergie provenant du gradient de proton de la membrane interne (Nies, 2007).

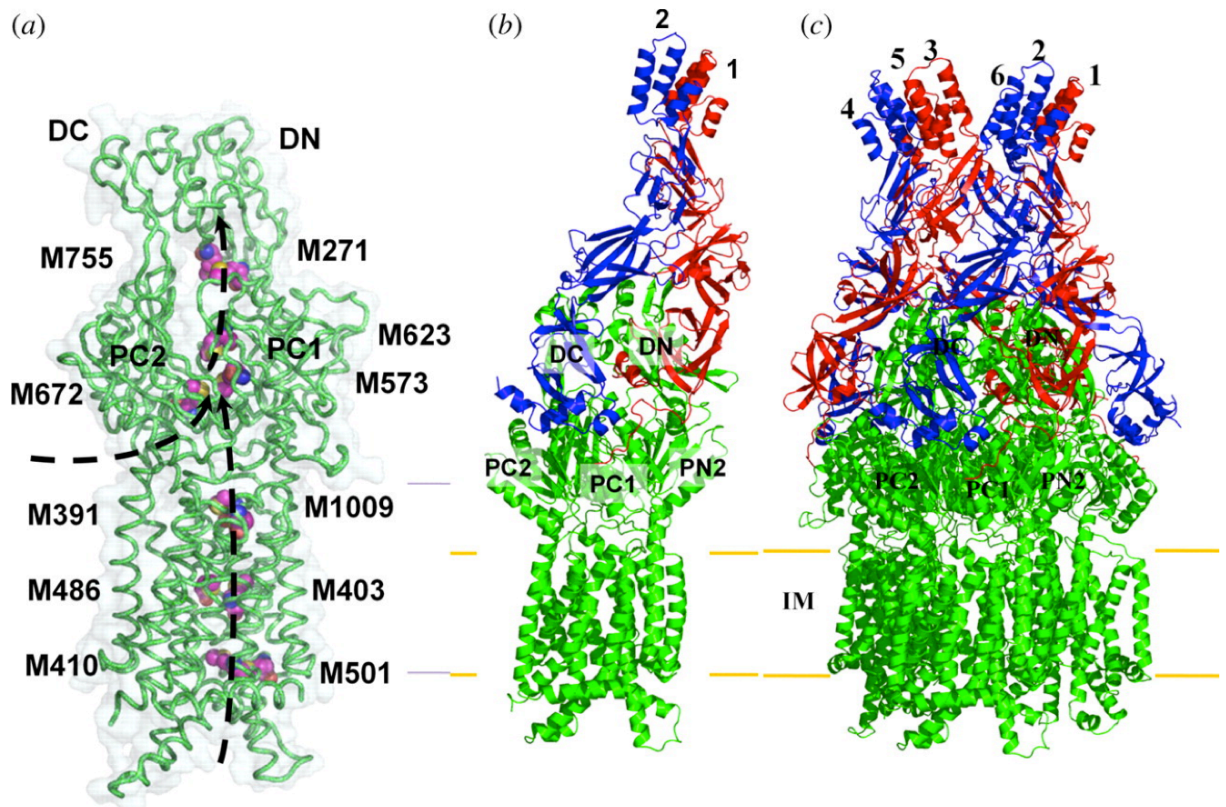
Un des systèmes bien caractérisé est le transporteur CusCBA d'*E. coli* responsable de l'export des ions cuivreux ( $\text{Cu}^+$ ) et argent ( $\text{Ag}^+$ ). CusC, CusB et CusA correspondent aux protéines appartenant aux familles OMP, MPF et RND, respectivement (Su *et al.*, 2011).



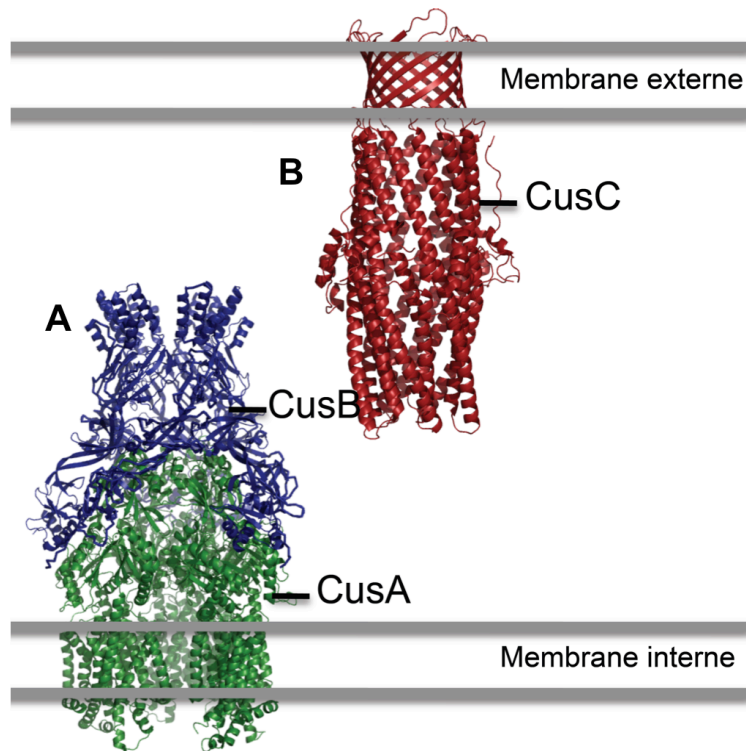
CusA (1047 résidus) est une protéine composée de 12 hélices TM et deux grandes boucles périplasmiques (Saier *et al.*, 1994).

C'est un homotrimère qui capte les métaux du cytoplasme et du périplasma. L'héxamère de CusB fixe les métaux du périplasma par son interaction avec la forme métallée de la protéine chaperon CusF et servirait d'intermédiaire entre CusA et CusC. CusC forme un canal en hélices  $\alpha$  dans le périplasma qui se prolonge dans la membrane externe par un tonneau  $\beta$  et permet l'export des ions métalliques dans le milieu extérieur (**Figure 18**).

Le mécanisme proposé pour l'export des ions métalliques est le suivant : CusA possède des paires de méthionines dans son domaine périplasmique et membranaire qui forment un chemin que peuvent emprunter les ions  $\text{Cu}^+$  et  $\text{Ag}^+$  (**Figure 17**). En l'absence de métal, le domaine périplasmique de CusA serait dans une conformation fermée. En présence de cuivre ou d'argent, ce domaine s'ouvrirait. Ces métaux se fixeraient au niveau d'un des trois sites métalliques formés par les méthionines au niveau périplasmique ou *via* la paire de méthionines dans le domaine TM. CusF transfère directement les ions métalliques aux sites métalliques périplasmiques de CusB qui les délivre aux paires de méthionines de CusA. Les métaux seraient acheminés jusqu'à la paire de méthionine M755/M271 puis seraient transmis à CusC qui les exportent jusqu'au milieu extracellulaire (Long *et al.*, 2012). Les structures cristallographiques de CusBA et CusC ont été déterminées séparément (Kulathila *et al.*, 2011; Su *et al.*, 2011) (**Figure 18**).



**Figure 17** : La voie de transport du cuivre et de l'argent proposée pour la pompe CusA et la structure cristallographique du complexe CusBA. **(a)** Les cinq paires de méthionines de la forme apo de CusA forment un chemin pour l'export des ions  $\text{Cu}^+$  et  $\text{Ag}^+$ . Ces méthionines sont représentées sous forme de sphères (les carbones en magenta, les oxygènes en rouge, les azotes en bleu et les soufres en orange). La paire M271-M755 est située au bout du canal périplasmique où les métaux pourraient être relargués pour l'extrusion finale. Le chemin emprunté par les métaux est représenté par une flèche noire. **(b)** Structure d'un protomère de CusA en vert et de deux protomères de CusB en rouge et bleu. DC, DN, PC1, PC2 et PN2 sont des sous domaines de CusA. **(c)** Structure du complexe CusBA. Le code couleur est similaire à celui de **(b)** (Long *et al.*, 2012).



**Figure 18** : Structure cristallographique de la pompe d'efflux CusCBA d'*E. coli*. **(A)** Structure de CusBA (code PDB : 3NE5) (Su *et al.*, 2011). **(B)** Structure de CusC (code PDB : 3PIK) (Kulathila *et al.*, 2011).

*C. metallidurans* CH34 possède au moins 12 opérons codant des systèmes d'efflux de métaux de type RND. Les systèmes les plus connus à l'heure actuelle sont CzcCBA pour l'efflux des cations cobalt, zinc et cadmium, CnrCBA pour le cobalt et le nickel et SilCBA pour l'argent (Nies, 2003). Comme décrit dans la section **I.C.**, les concentrations périplasmiques en cations métalliques peuvent aussi jouer le rôle de molécules signal dans les systèmes à deux composants ou dépendants des facteurs sigma de type ECF, induisant la régulation de l'expression des opérons codant les pompes d'efflux RND (van der Lelie *et al.*, 1997; Nies, 2004).

### III. *Cupriavidus metallidurans* CH34 : une bactérie hautement résistante aux métaux

#### A. Description générale de *C. metallidurans* CH34

*Cupriavidus metallidurans* CH34 (Vandamme et Coenye, 2004) est une bactérie à Gram négatif, du sol et des eaux, aérobie-anaérobie facultative appartenant à la classe des  $\beta$ -protéobactéries. D'abord connu sous le nom de *Alcaligenes eutrophus* (Mergeay *et al.*, 1985), *Ralstonia sp.*, *Ralstonia eutropha*, *Ralstonia metallidurans* (Goris *et al.*, 2001), et *Wautersia metallidurans* (Vanechoutte *et al.*, 2004), ce micro-organisme a été isolé pour la première fois en 1976 dans un bassin de décantation d'une usine produisant du zinc, près de Liège en Belgique (Mergeay *et al.*, 1978). Cette bactérie est un bacille de 0,4  $\mu\text{m}$  de diamètre et 1 à 2  $\mu\text{m}$  de longueur avec plusieurs flagelles. Elle s'observe sous une forme simple et isolée ou par paire, voire en courte chaîne.

La température de croissance optimale de *C. metallidurans* CH34 est de 30°C. Au delà de cette température, son phénotype devient hypermutateur (Taghavi *et al.*, 1997). Ce réarrangement génomique par mutagenèse, appelé TIMM ("Temperature-Induced Mutagenesis et Mortality"), révèle la plasticité génomique de *C. metallidurans* CH34 et aiderait la population bactérienne à s'adapter à des conditions environnementales extrêmes (Taghavi *et al.*, 1997). Toutefois, ces mutations peuvent aussi entraîner des dommages collatéraux. En effet, les bactéries peuvent, par exemple, perdre leur capacité à résister à certains métaux comme le cobalt et le zinc lorsque que l'opéron *czc* (pour "cobalt, zinc et cadmium") est muté (Mergeay, 2000).

La particularité de cette bactérie est d'être multirésistante aux métaux (Cadmium, Mercure, Plomb, Cuivre, Zinc, Nickel, Cobalt...) (Mergeay *et al.*, 1985). La concentration en métal suffisante pour inhiber *in vitro* la croissance des bactéries, appelée concentration minimale inhibitrice (CMI), a été mesurée chez *C. metallidurans* CH34 et *E. coli*. Les résultats indiquent que *C. metallidurans* CH34 est généralement plus résistante aux métaux,

notamment aux cations  $\text{Co}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$  et  $\text{Cd}^{2+}$  (**Tableau 4**) (Mergeay *et al.*, 1985; Nies, 1999, 2003).

**Tableau 4** : Concentration minimale inhibitrice (CMI) de nickel, cobalt, zinc et cadmium déterminée chez *E. coli* et *C. metallidurans* CH34, dans un milieu minimum contenant des sels minéraux.

Métal	CMI (mM) chez <i>E. coli</i> (Nies, 1999)	CMI (mM) chez <i>C. metallidurans</i> CH34 (Mergeay, 1985)
Nickel	1	2,5
Cobalt	1	25
Zinc	1	12
Cadmium	0,5	2,5

*C. metallidurans* CH34 représente ainsi une souche modèle pour l'étude de la résistance aux métaux lourds chez les bactéries et permet son utilisation dans le développement de nouveaux outils biotechnologiques. Elle peut servir de capteurs sensibles aux doses biodisponibles (biosenseurs) (van der Lelie *et al.*, 1994; Tibazarwa *et al.*, 2001). Elle est aussi utilisée pour la détoxification d'environnements pollués (Diels *et al.*, 2009; Nies, 2000; Valls et de Lorenzo, 2002; Wei *et al.*, 2014b), ou pour récupérer des métaux précieux comme l'or *via* des processus de biofiltration (Reith *et al.*, 2009).

## B. Les systèmes de résistances aux métaux

Le génome de *C. metallidurans* CH34 a été totalement séquencé (Janssen *et al.*, 2010). Il est constitué de quatre réplicons : un chromosome de 3,9 Mb, un mégaplasmide de 2,6 Mb et deux plasmides géants pMOL28 et pMOL30 de 171 kb et 234 kb, respectivement. Ce séquençage a permis de répertorier les déterminants de résistance aux métaux lourds qui sont principalement localisés sur les deux plasmides pMOL28 et pMOL30 (Mergeay *et al.*, 1985; Monchy *et al.*, 2007; Rozycki et Nies, 2009). Ces déterminants codent principalement pour des pompes d'efflux (**Tableau 5**). Le plasmide pMOL30 contient également le déterminant *ncc* impliqué dans la résistance aux nickel, cobalt et cadmium mais *ncc* est silencieux dans la

souche CH34 : il ne possède pas de gènes de régulation (*nccYXH*) et aucune expression n'a été mesurée CH34 (Nies *et al.*, 2006).

**Tableau 5** : Exemples de déterminants de résistance aux métaux chez *C. metallidurans* CH34

Emplacement	Opérons	Métaux	Mécanisme de résistance	Références
<b>Chromosome</b>	<i>cup</i>	cuivre	efflux	(Monchy <i>et al.</i> , 2006)
<b>Mégaplasmide</b>	<i>chr</i>	chromate	efflux	(Juhnke <i>et al.</i> , 2002)
	<i>znt</i>	zinc	efflux	(Monchy <i>et al.</i> , 2007)
	<i>cop</i>	cuivre	efflux	(Monchy <i>et al.</i> , 2007)
<b>pMOL28</b>	<i>cnr</i>	nickel/cobalt	efflux	(Liesegang <i>et al.</i> , 1993)
	<i>mer</i>	mercure	réduction	(Misra, 1992)
	<i>chr</i>	chromate	efflux	(Juhnke <i>et al.</i> , 2002)
<b>pMOL30</b>	<i>czc</i>	cobalt/zinc/cadmium	efflux	(Nies, 1995)
	<i>cop</i>	cuivre	efflux	(Monchy <i>et al.</i> , 2007)
	<i>pbr</i>	plomb	séquestration/efflux	(Borremans <i>et al.</i> , 2001)
	<i>sil</i>	argent	efflux	(Monchy <i>et al.</i> , 2007)

## C. Régulation de la résistance aux métaux

L'expression des systèmes de résistance est majoritairement régulée par les systèmes à deux composants (TCS), dont le mécanisme est détaillé dans la section **I.C.3**. *C. metallidurans* CH34 possède 42 kinases et 67 protéines régulatrices identifiées, contre respectivement 29 et 37 chez *E. coli*. Ces TCS sont souvent associés aux gènes RND comme le système CzsRS de *C. metallidurans* CH34 qui régule l'expression de la pompe d'efflux CzcCBA (Grosse *et al.*, 1999). *C. metallidurans* CH34 compte au moins sept autres systèmes RND dont cinq opérons de type *czc* (pompes RND régulées par un TCS).

La transcription des gènes de résistance est aussi régulée par les systèmes dépendants des facteurs  $\sigma^{\text{ECF}}$ . Parmi les 17 facteurs sigma annotés dans le génome de *C. metallidurans* CH34, 11 sont du type ECF (Nies, 2004). Ces 11 facteurs  $\sigma^{\text{ECF}}$  ont été répertoriés dans 4 sous-groupes (**Tableau 6**), parmi lesquels un groupe de  $\sigma^{\text{ECF}}$  typique (Lonetto *et al.*, 1994) et deux groupes d'ECF de type FecI distingués sur la base des similarités de séquence. RpoI a été le

premier facteur  $\sigma$  identifié pour son rôle dans l'assimilation du fer chez *C. metallidurans* CH34. Il porte ce nom pour indiquer une fonction similaire à FecI d'*E. coli*. Parmi ces facteurs  $\sigma^{\text{ECF}}$ , seuls ceux appartenant au sous-groupe ECF:FecI 1, CnrH et RpoE semblent importants pour la résistance aux métaux (Grosse *et al.*, 2007).

**Tableau 6** : Répertoire des 11 facteurs  $\sigma^{\text{ECF}}$  de *C. metallidurans* CH34 (Grosse *et al.*, 2007).

Sous-groupes	Facteurs $\sigma^{\text{ECF}}$
ECF:FecI 1	RpoJ, RpoK et RpoI
ECF:FecI 2	RpoL, RpoM et RpoO
ECF:RpoE	RpoP, RpoE et CnrH
'(ECF)'	RpoQ et RpoR

## D. Le déterminant Cnr de *Cupriavidus metallidurans* CH34

### 1. Résistance au cobalt et au nickel

La résistance au nickel et au cobalt est principalement conférée par les déterminants portés par les plasmides pMOL28 et/ ou pMOL30. En effet, la souche *C. metallidurans* CH34 sans ses deux plasmides (AE104) voit sa CMI diminuée de plus de 4 fois pour le nickel et plus de 350 fois pour le cobalt comparée à la souche sauvage CH34. La présence de pMOL28 dans AE104 (AE126) ou seulement du déterminant *cnr* (AE104 + *cnr*) augmente de 70 et de 4 la CMI du cobalt et du nickel, respectivement. La présence de pMOL30, qui contient le déterminant *czc*, dans AE104 (AE128) augmente environ 300 fois la CMI du cobalt. Les valeurs des CMI pour le nickel et le cobalt sont reportées dans le **tableau 7** (Liesegang *et al.*, 1993; Mergeay *et al.*, 1985).

**Tableau 7 :** Concentration minimale inhibitrice de nickel et de cobalt pour différentes souches de *C. metallidurans* CH34. La souche AE104 correspond à celle de CH34 sans plasmides. AE126 est la souche qui ne comporte pas le plasmide pMOL30 et AE128 ne comporte pas le plasmide pMOL28 (Liesegang *et al.*, 1993; Mergeay *et al.*, 1985).

Souche	CMI Nickel (mM)	CMI Cobalt (mM)
<b>CH34</b>	2,5	25
<b>AE104</b>	0,6	0,07
<b>AE126</b>	2,5	5
<b>AE128</b>	0,6	20
<b>AE104 + <i>cnr</i></b>	2,5	5

Ainsi la résistance au nickel n'est apportée que par le plasmide pMOL28 et plus particulièrement par le déterminant *cnr* (pour "cobalt nickel resistance") alors que la résistance au cobalt est apportée majoritairement par pMOL30 qui contient l'opéron *czc* mais aussi partiellement par pMOL28.

## 2. Le système NccYXH de *C. metallidurans* 31A

Un autre déterminant de résistance au nickel et au cobalt similaire à *cnr*, a été identifié sur le plasmide pTOM9 de *C. metallidurans* 31A. Comme l'opéron *cnr*, l'opéron *ncc* permet à la bactérie de résister à de fortes concentrations de nickel et de cobalt. Il se compose des gènes *nccYXHCBAN*. A l'exception de *nccN*, les produits de ces gènes partagent de fortes similarités de séquence avec les protéines codées par *cnrYXHCBA* (Schmidt et Schlegel, 1994). En effet, NccX, NccY et NccH partagent 76%, 59% et 65% d'identité de séquence avec CnrX, CnrY et CnrH, respectivement. Cela laisse supposer que ces deux systèmes ont un mode de fonctionnement similaire. De plus, NccX partage les mêmes propriétés spectroscopiques que CnrX (Trepreau *et al.*, 2011a).

Lors de notre étude, nous avons tiré avantage de la forte homologie que partage le système Ncc avec Cnr pour basculer d'un système à l'autre lorsque cela a été nécessaire.



### 3. Régulation de l'expression de *cnrCBA* par *cnrYXH*

L'opéron *cnr* se compose des gènes *cnrYXHCBA*. La transcription de *cnr* est initiée au niveau du promoteur *CnrYp* en amont des gènes de régulation, *cnrYXH* et *CnrCp* en amont des gènes de résistance au cobalt et au nickel, *cnrCBA*. Ces deux promoteurs possèdent des séquences conservées aux positions -10 (CCGTATA) et -35 (CRAGGGGRAG) (Grass *et al.*, 2000).

Les gènes *cnrCBA* codent un antiport cation/proton, composé de trois protéines appartenant à la famille des RND (partie **II.C.1.b**). Son expression est dépendante de la concentration de certains métaux dans l'environnement. Grass et coll. ont mesuré l'activité  $\beta$ -galactosidase d'une souche de CH34 recombinante contenant la fusion *cnrCBA-lacZ* en présence de différents métaux. Cette activité rend compte du taux d'expression de la pompe CnrCBA. En présence de zinc, l'activité  $\beta$ -galactosidase est similaire à celle mesurée lorsque les bactéries sont cultivées sans métaux (activité basale). En présence de 0,5 mM de cobalt, l'activité augmente d'un facteur 2 alors qu'elle augmente d'un facteur 6 lorsque les bactéries sont cultivées avec 0,5 mM de nickel. L'expression de la pompe CnrCBA est donc plus induite avec du nickel qu'avec du cobalt contrairement au zinc qui n'a aucun effet (Grass *et al.*, 2000).

Dans le même système de gène rapporteur (*cnrYXH-cnrCBA-lacZ*), lorsque les gènes *cnrYXH* sont délétés ( $\Delta$ *cnrYXH-cnrCBA-lacZ*), l'activité  $\beta$ -galactosidase est deux fois plus faible, même en présence de nickel. L'expression de la pompe n'est donc pas constitutive, elle dépend du complexe CnrYXH dont l'expression est induite par le nickel. L'expression individuelle ou couplée de *cnrX* et de *cnrY* dans la souche  $\Delta$ *cnrYXH*, en absence ou en présence de nickel, ne modifie pas l'activité  $\beta$ -galactosidase. Au contraire, *cnrH* seul ou en présence de *cnrX* ou *cnrY* conduit à une forte augmentation de l'activité mais cette activité n'est pas augmentée en présence de nickel. CnrH est ainsi un activateur constitutif de l'expression de la pompe (Grass *et al.*, 2000). La régulation de l'expression de la pompe d'efflux CBA dépend donc de l'expression des trois protéines, CnrY, CnrX et CnrH.

## 4. Le système de régulation CnrYXH

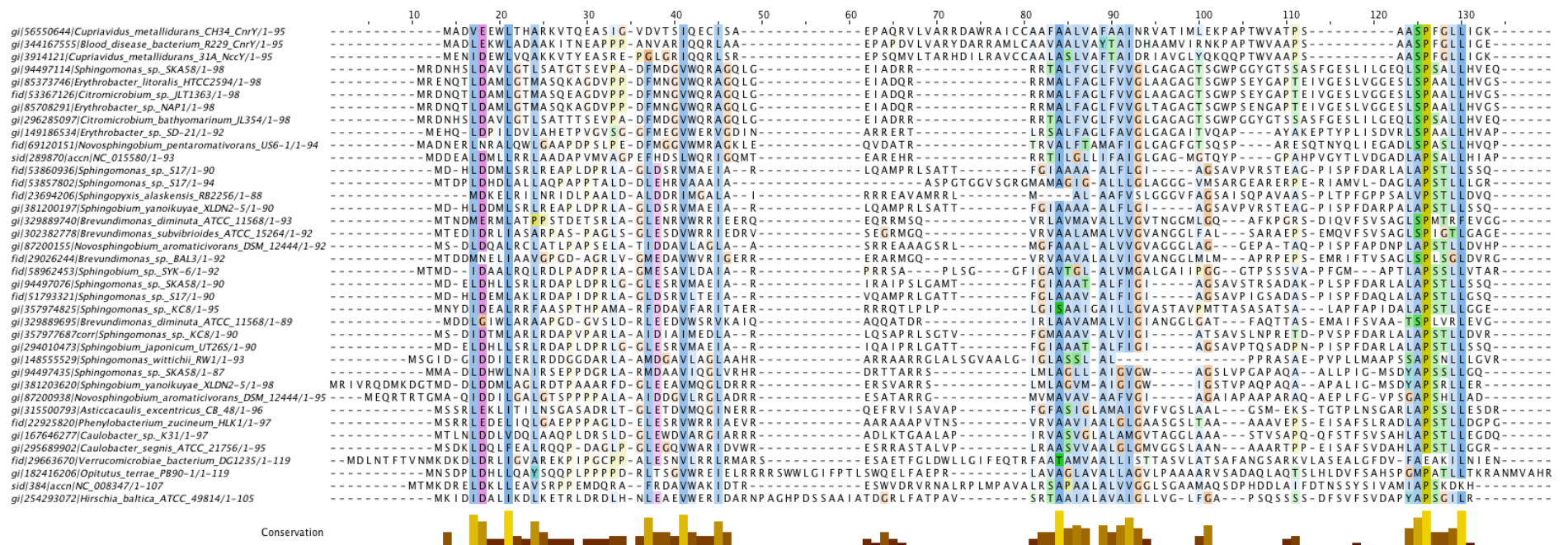
### a. Description du système CnrYXH

CnrYXH est un complexe entre trois partenaires protéiques de petite taille : CnrX, CnrY et CnrH.

- **CnrH** est une protéine de 21 kDa (191 résidus), soluble dans le cytosol. Elle fait partie de la famille des facteurs  $\sigma^{\text{ECF}}$  qui ne possèdent que deux des quatre domaines  $\sigma$  ( $\sigma_2$  et  $\sigma_4$ ) (voir partie **I.C.4.a.**), suffisants pour une interaction avec l'ARNP et les promoteurs cibles. CnrH peut se lier aux promoteurs *CnrYp* et *CnrCp* et initier la transcription des gènes sous leur contrôle (Grass *et al.*, 2005).

- **CnrY** est une protéine bitopique de 10 kDa (95 résidus). Son domaine N-terminal, long de 45 résidus se localise dans le cytoplasme alors que son domaine C-terminal est périplasmique (30 résidus) (Grass *et al.*, 2000). Dans les bases de données UniProtKB, très peu d'homologues à CnrY sont recensés. Toutefois, au voisinage de gènes homologues à *cnrX*, 37 gènes codant des homologues à CnrY, dont NccY de *C. metallidurans* 31A, ont été trouvés (**Figure 19**).

- **CnrX** est une protéine de 13 kDa (148 résidus). D'après les outils de prédiction de topologie, les 30 premiers acides aminés de CnrX serviraient d'ancrage membranaire à la protéine et les 118 suivants pointerait dans le périplasma (Grass *et al.*, 2000) pour former le domaine senseur de métal. Cette protéine est la plus étudiée du complexe. Une présentation détaillée de cette protéine sera effectuée dans les prochaines sections.

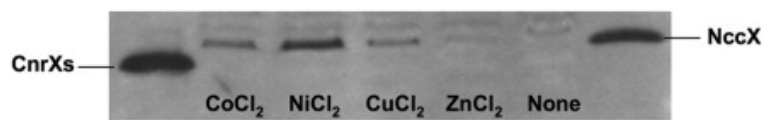


**Figure 19 :** Alignement de séquences des 38 protéines apparentées à CnrY (dont CnrY) dont les gènes sont voisins des homologues de CnrX. L'alignement est mis en forme avec le logiciel Jalview 2.8.1. Le code couleur utilisé est celui de Clustalx qui tient compte des propriétés physicochimiques des résidus. L'histogramme en dessous des séquences décrit le degré de conservation des résidus alignés dans chaque colonne.

## b. Induction du système par les métaux

Monchy et coll. (Monchy *et al.*, 2007) ont montré par des analyses de transcriptomique *via* l'utilisation de puces à ADN que les gènes de *cnrYXHCBA* sont surexprimés en présence de 0,5 mM de  $\text{Cd}^{2+}$ , de 0,1 mM de  $\text{Cu}^{2+}$  et de 0,6 mM de  $\text{Ni}^{2+}$  dans le milieu de culture. Le zinc (0,8 mM de  $\text{Zn}^{2+}$ ) et le mercure (5  $\mu\text{M}$  de  $\text{Hg}^{2+}$ ) n'induisent aucune surexpression alors que le cobalt (2 mM de  $\text{Co}^{2+}$ ) permet seulement de surexprimer les gènes *cnrB* et *cnrA*. Cependant, Grosse et coll. (Grosse *et al.*, 2007), par des expériences de RT-PCR, montrent que 0,3 mM de  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Cd}^{2+}$  mais aussi de  $\text{Co}^{2+}$  induisent une surexpression de *cnrH*. Le nickel, le cadmium, le cuivre et le cobalt seraient donc des métaux inducteurs du système contrairement au zinc.

De plus, une étude sur la production de CnrX dans des conditions où *C. metallidurans* CH34 est cultivée en présence de nickel, de cobalt, de cuivre ou de zinc (concentrations identiques à celles utilisées par Monchy *et al.*, 2007) a été effectuée afin de confirmer les données de transcriptomique décrites précédemment. Les résultats sont présentés dans la **figure 20**. CnrX a été détectée par western-blot anti-CnrXs (partie périplasmique de CnrX) lorsque la bactérie a été stressée avec du nickel, du cobalt et du cuivre. A l'inverse, CnrX n'est pas détectée lorsque les bactéries CH34 sont incubées avec du zinc et lorsqu'aucun métal n'est ajouté au milieu de culture. Ces résultats confirment que lorsque le système *cnr* est induit par le nickel, le cuivre et le cobalt, la protéine CnrX est produite. La production de CnrX est plus induite par du nickel que du cobalt qui serait lui-même plus inducteur que le cuivre (Trepreau *et al.*, 2011b).



**Figure 20** : Détection de la production de CnrX dans *C. metallidurans* CH34 par western-blot anti-CnrXs. Les cellules ont été incubées pendant 2 heures avec 2 mM de  $\text{CoCl}_2$ , 0,6 mM de  $\text{NiCl}_2$ , 0,1 mM de  $\text{CuCl}_2$  ou 0,8 mM de  $\text{ZnCl}_2$ . Une culture dans laquelle aucun ion métallique n'a été ajouté a servi de contrôle (None). Les protéines pures CnrXs et NccX (protéine homologue de CnrX) ont été utilisées comme contrôles de taille (Trepreau *et al.*, 2011b).

### c. La protéine CnrX

42 protéines homologues à CnrX, dont NccX de *C. metallidurans* 31A, ont été trouvées dans les bases de données UniProtKB. Pour la majorité, ces protéines sont impliquées dans la résistance aux métaux. Les autres ne sont pas caractérisées. Cependant, toutes ces protéines sont issues de protéobactéries (**Figure 21**).

Sur ces 43 séquences, 16 résidus sont très conservés (plus de 97% d'identité) (**Figure 22**). Parmi ces résidus, on trouve 3 histidines (H42, H46 et H119) ainsi qu'un glutamate (E63) situés dans le domaine périplasmique de CnrX. Ces histidines ainsi que les 3 autres de la protéine (H32, H38 et H131) ont été testées par des expériences de mutagénèse couplées à une étude *in vivo* via le système rapporteur *cnrCBA-lacZ*. Les résultats indiquent que seules 4 histidines dont les 3 conservées sont essentielles pour l'induction de *cnr* (Grass *et al.*, 2005).

Par ailleurs, on retrouve sur la séquence de CnrX le motif HExxH. Ce motif est caractéristique d'un type de protéase à zinc (Jongeneel *et al.*, 1989). Ces protéases sont impliquées dans différents mécanismes de régulation parmi lesquels la protéolyse du domaine intramembranaire des facteurs anti- $\sigma$ , permettant de libérer le facteur  $\sigma$  dans le cytoplasme (Heinrich et Wiegert, 2009). La protéine possède aussi dans sa partie TM, la séquence consensus SxxxGxxxGxxxA. SxxxG, GxxxG et GxxxA sont connus pour être des motifs de dimérisation des segments TM (Lemmon *et al.*, 1992). Ces motifs ne sont pas conservés. Cependant, on retrouve le motif GxxxG dans 4 autres séquences dont NccX de *C. metallidurans* 31A et dans 11 autres mais décalé de quelques résidus (**Figure 21**).

CnrX et plus particulièrement son domaine périplasmique a fait l'objet de nombreuses recherches dans l'équipe qui seront décrites dans la partie suivante.



1                    10                    20                    30                    40                    50  
 MMKSRTRRLS LSTLFGALLG VSVAAAWLYY SHRNEAGHGD LHEILHEAVP  
  
                   60                    70                    80                    90                    100  
 LDANEREILE LKEDAFQRR REIETRLRAA NGKLADAIK NPAWSPEVEA  
  
                   110                    120                    130                    140                    148  
 ATQEVERAAG DLQRATLVHV FEMRAGLKPE HRPAYDRVLI DALRRGSQ

**Figure 22** : Séquence de CnrX. La séquence soulignée correspond au segment TM prédit (logiciel THMM). Les résidus en rouge ont un très fort degré de conservation dans les 45 séquences homologues à CnrX (plus de 97% d'identité), ceux en bleus partagent plus de 90% d'identité de séquence et ceux en vert ont un fort degré de conservation.

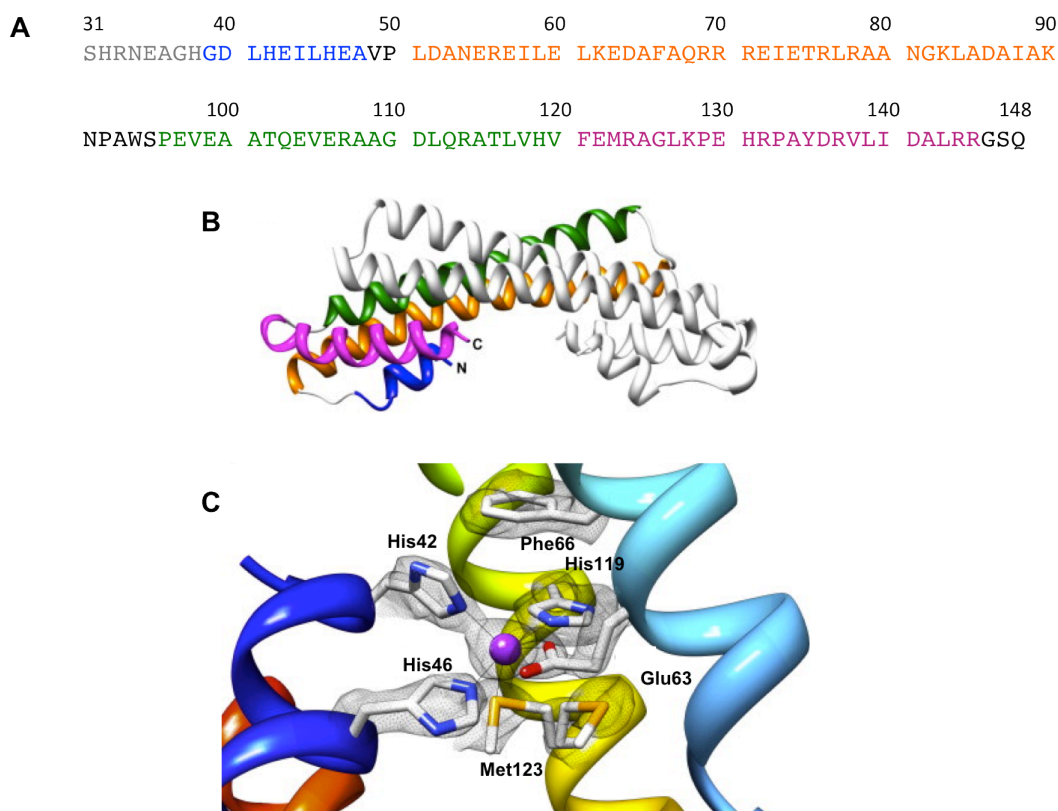
### i. Caractérisation du domaine périplasmique de CnrX (CnrXs)

Précédemment dans l'équipe, le domaine périplasmique de CnrX (résidus 31 à 148, **CnrXs**, **Figure 23A**) a été produit chez *E. coli*, purifié et cristallisé. La structure tridimensionnelle, à 1,74 Å de résolution, a été obtenue par diffraction aux rayons X (Pompidor *et al.*, 2009) (**Figure 23B**). La protéine est dimérique et entièrement constituée d'hélices  $\alpha$  formant plusieurs épingles à cheveux. Les préparations de CnrXs étaient contaminées par une petite quantité de cuivre qui a permis d'obtenir la structure de la forme apo de CnrXs (dépourvue de métal, CnrXs-apo) et sa forme en complexe avec du cuivre (CnrXs-Cu).

Cette structure a permis de caractériser un site métallique composé des atomes d'azote de la chaîne latérale des 3 résidus histidines conservés et décrits comme indispensables à l'induction de *cnr* (H42, H46 et H119) et les deux atomes d'oxygène de la chaîne latérale du résidu glutamate conservé (E63). Ces résidus sont accompagnés par un sixième ligand axial inattendu : le soufre de la liaison thioéther de la seule méthionine de la séquence (M123) (**Figure 23C**). Contrairement à la majorité des enzymes et des métallo-régulateurs, la séquence de CnrX ne possède pas de cystéine mais une méthionine strictement conservée (M123). Cette conservation pourrait indiquer que CnrX est l'archétype d'une nouvelle famille de senseur de métaux (Pompidor *et al.*, 2008).

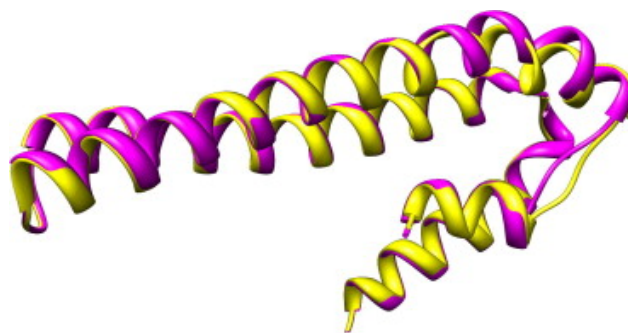
Le site semble préformé dans la forme apo et de légères modifications conformationnelles sont associées à la métallation de CnrXs par le cuivre (**Figure 24**). Une autre forme apo de la

protéine a été cristallisée après l'avoir purifiée en présence d'EDTA mais cette forme était contaminée par du zinc (Trepreau *et al.*, 2011b). Trepreau et coll. ont donc muté un résidu ligand du métal afin de détruire le site métallique (CnrXs-E63Q). La structure de ce mutant ne contient aucun métal. Cette structure sous forme apo servira de référence pour la comparaison avec les autres structures en complexe avec des métaux.



**Figure 23** : Structure du dimère de CnrXs (code pdb : 3epv). **(A)** Séquence de CnrXs. La protéine a été exprimée à partir de la sérine 31. Les résidus colorés en bleu forment l'hélice H1, ceux en orange l'hélice H2, ceux en vert l'hélice H3 et ceux en magenta l'hélice H4. Les résidus en gris n'apparaissent pas sur la structure. **(B)** Le dimère de CnrXs est entièrement en hélices. Les hélices d'un des monomères sont colorées avec le même code couleur que dans (A). **(C)** Zoom sur le site métallique de CnrXs ayant lié un ion cuivre (sphère violette). Trois des hélices (H1, H2 et H3) d'un monomère contribuent à fournir les ligands du métal. Ces ligands sont annotés et les hélices sont colorées avec le même code couleur que dans (A) (Pompidor *et al.*, 2008).

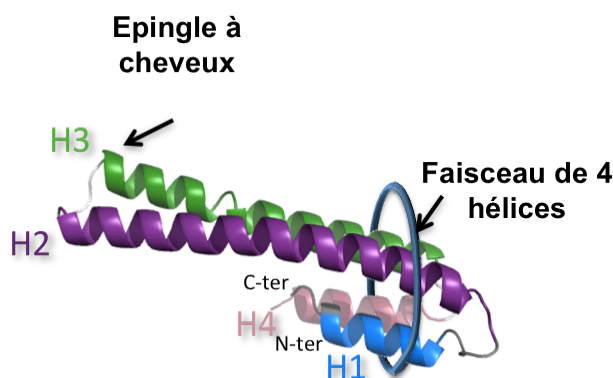




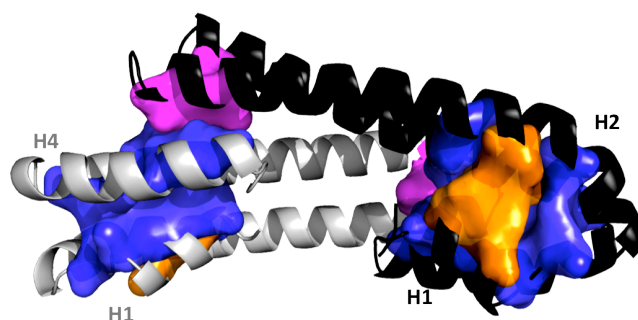
**Figure 24** : Superposition de CnrXs-apo en jaune et CnrXs-Cu en magenta (Pompidor *et al.*, 2008).

Le mécanisme de perception du stress métallique ainsi que la nature du signal transmis ont fait l'objet d'une étude approfondie lors de la thèse réalisée précédemment dans l'équipe par Juliette Trepreau.

L'étude structurale de CnrXs a été complétée par l'obtention des structures cristallographiques à haute résolution (entre 2,5 et 1,4 Å de résolution) de CnrXs en complexe avec du zinc (CnrXs-Zn), du cobalt (CnrXs-Co) et du nickel (CnrXs-Ni) (Trepreau *et al.*, 2011b). Ces structures ont la même architecture globale, similaire à celle de CnrXs-apo et CnrXs-Cu (**Figure 23B**). Chaque monomère est constitué de 4 hélices  $\alpha$  repliées en épingle à cheveux. Les deux courtes hélices H1 et H4, la partie N-terminale de H2 et la partie C-terminale de H3 forment un faisceau de 4 hélices (**Figure 25**). Ce repliement des 4 hélices empilées est stabilisé par des interactions hydrophobes entre des résidus conservés. Ces résidus forment un cœur hydrophobe qui englobe les 4 hélices et les stabilise (**Figure 26**). Les deux monomères s'associent pour former un dimère qui est l'unité biologique (Pompidor *et al.*, 2009). Ce dimère est stabilisé principalement par le cœur hydrophobe qui s'étend à l'interface dimérique (**Figure 26**) (Trepreau *et al.*, 2011b).



**Figure 25** : Structure d'un monomère de CnrXs. Les hélices H1, H2, H3 et H4 sont colorées en bleu, violet, vert, et rose, respectivement. Le faisceau de 4 hélices est indiqué par un cercle bleu (Trepreau *et al.*, 2011b).



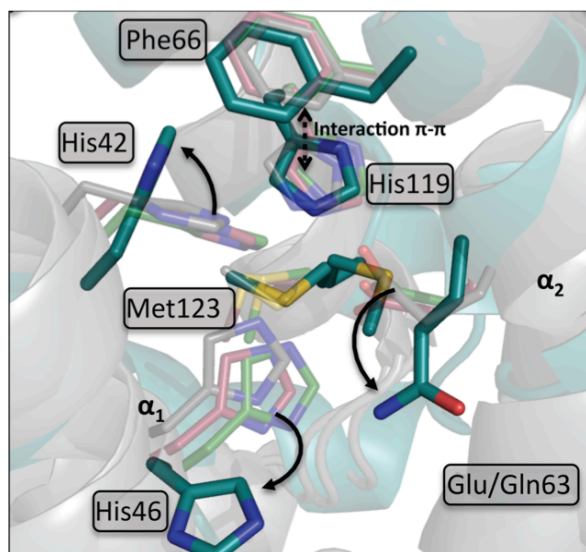
**Figure 26** : Interactions hydrophobes intra et inter protomère. Les deux monomères du dimère de CnrXs sont en noir et gris. La surface moléculaire des résidus formant le cœur hydrophobe du faisceau de 4 hélices est colorée en bleu, celle des résidus du site métallique en orange. La surface moléculaire colorée en magenta correspond aux résidus impliqués dans la stabilisation du dimère (résidus hydrophobes d'un monomère interagissant avec le cœur hydrophobe de l'autre monomère) (Trepreau *et al.*, 2011b).

Le dimère de CnrXs possède deux sites métalliques, un par monomère. Les sites métalliques de CnrXs-Ni et CnrXs-Co sont similaires entre eux et à celui de CnrXs-Cu. Les ligands du métal (H42, H46, H119, E63 et M123) appartenant aux hélices H1, H2 et H3 sont presque superposables (**Figure 27**). Le nickel, le cobalt comme le cuivre sont liés par une géométrie octaédrique 3N2O1S.

La structure de CnrXs en complexe avec du zinc (CnrXs-Zn) est différente des autres. Trois atomes de zinc sont présents dans chaque dimère, 2 dans le monomère A et 1 dans le monomère B. Dans ce dernier, le zinc est pentacoordiné dans une sphère 3N2O (N de H42,

H46 et H119 et les deux O de E63). Les deux ions zinc du monomère A sont en coordination tétraédrique. Le premier zinc (ZnA1) est lié par les résidus H38, E63 et H119 et le deuxième (ZnA2) par H42, H46 et E63. Toutefois, ce site est un artéfact induit par la construction de la protéine et les conditions de cristallisation (Trepreau *et al.*, 2011a).

Lorsque les sites de liaison au métal dans la forme apo (CnrXs-E63Q) sont comparés à ceux dans les formes métallées de CnrXs, la position des carbones  $\alpha$  des résidus ligands E63 (ou Q63 pour la forme apo), H119 et M123 est similaire dans la forme apo et dans les formes métallées (CnrXs-Ni, -Co et -Zn). Cependant, seule H119 a la même conformation dans les 4 formes. L'imidazole de H119 est stabilisé par une interaction  $\pi$ - $\pi$  avec F66. Les chaînes latérales des autres résidus ont une position différente : ceux des formes métallées sont plus proches du métal que ceux de la forme apo (**Figure 27**). Cela suggère que H119 fonctionne comme un piège à cations et que les autres résidus du site de liaison replient leurs chaînes latérales pour compléter la sphère de coordination.

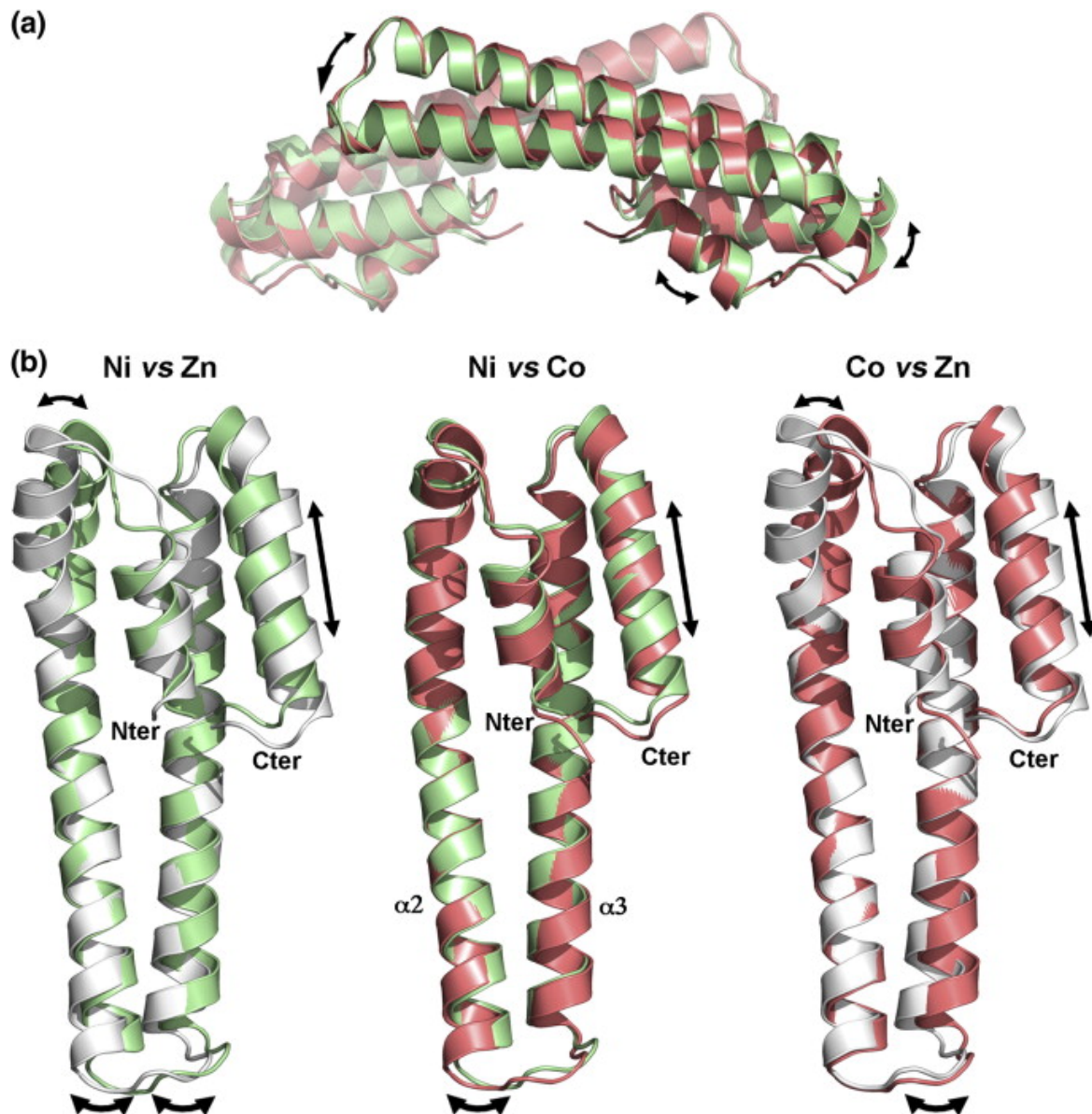


**Figure 20** : Superposition des résidus du site métallique de CnrXs-Ni (vert), CnrXs-Co (rouge) et CnrXs-Zn (gris) et des mêmes résidus de la forme apo (CnrXs-E63Q) (bleu) (Trepreau *et al.*, 2011b).

Le recrutement du soufre de M123 dans la sphère de coordination du nickel, du cobalt et du cuivre marque le passage de la forme inactive de la protéine (CnrXs-Zn) à la forme active (CnrXs-Ni, CnrXs-Co, CnrXs-Cu). Ce recrutement est à l'origine de changements de

conformation localisés (rapprochement des hélices H2 et H3, translation verticale des hélices H1 et H4) qui vont s'étendre au monomère puis au dimère (**Figure 28**).

De plus, la différence essentielle entre les trois métaux activateurs de la protéine est la distance du métal au soufre de M123. Cette distance est plus longue pour l'ion cuivre (S-Cu : 2,77 Å) que pour l'ion cobalt (S-Co : 2,54 Å) qui est plus longue que l'ion nickel (S-Cu : 2,45 Å). Ce petit écart est propagé par un réseau d'interactions et va affecter la structure globale des dimères. Dans la structure CnrXs-Ni : la courbure de la partie N-terminale de H2 et de la partie C-terminale de H3 est plus accentuée dans la forme nickel, les hélices H1 et H4 sont légèrement plus repliées vers la face concave de la protéine, l'extrémité N-terminale du monomère A et l'extrémité C-terminale du monomère B sont plus proches et les hélices H2 et H3 sont légèrement plus resserrées. Les deux monomères sont plus proches. Cet ensemble de légères modifications engendre une conformation de la protéine CnrXs-Ni plus compacte que la forme contenant du cobalt (**Figure 28**).



**Figure 21** : Changement de conformation de CnrXs lié à la métallation. **(a)** Superposition des structures dimériques de CnrXs-Ni (vert) et CnrXs-Co (rouge). Les doubles flèches indiquent les différences de positionnement entre les deux formes. **(b)** Superposition des monomères B de CnrXs-Ni (vert), CnrXs-Co (rouge) et CnrXs-Zn (blanc). Les doubles flèches du haut montrent le mouvement de translation de l'hélice H4 (un même mouvement est effectué par H1), celles du bas illustrent les modifications conformationnelles des épingles à cheveux (Trepreau *et al.*, 2011b).

En résumé, CnrXs-Zn est la forme inactive de la protéine. Lorsqu'ils sont en concentrations suffisantes, les ions  $\text{Ni}^{2+}$  ou  $\text{Co}^{2+}$  qui sont plus affins pour le site métallique (Trepreau *et al.*, 2011b), se fixent à la protéine. Ils recrutent le soufre de la Met123 comme sixième ligand, induisant un changement dans la géométrie du site métallique (de bipyramidal à octaédrique). Ce changement entraîne des modifications allostériques (rapprochement d'une

partie de H2 vers H3 et distorsion latérale de H1 et H4) conduisant à un dimère plus compact qui peut être la conformation active. Le recrutement de M123 dans la sphère de coordination est donc l'élément déclencheur de la transduction du signal à travers la membrane. Sa mutation en alanine (M123A) abolit l'inductibilité du système Cnr bien que la structure cristallographique de CnrXs-M123A-Ni montre que la structure du site métallique n'est pas affectée par la mutation (Trepreau, 2014). Les événements qui suivent la perception du métal par CnrXs sont toujours inconnus. Afin de documenter ces événements, mon projet de thèse vise à identifier et caractériser les interactions entre les protéines du complexe CnrYXH.

# Résultats

Chez *Cupriavidus metallidurans* CH34, le complexe CnrYXH contribue à réguler l'expression des gènes de résistance au cobalt et au nickel en fonction de la concentration de ces cations dans le périplasma, qui est représentative des concentrations environnementales. La fixation de cobalt ou de nickel sur le domaine périplasmique senseur de CnrX est à l'origine de la transduction d'un signal transmembranaire conduisant à rendre le facteur sigma CnrH disponible pour sa liaison à l'ARN polymérase (ARNP). Le complexe CnrH:ARNP se fixe alors spécifiquement sur les promoteurs des gènes *cnrY* et *cnrC* pour initier la transcription des gènes de régulation (*cnrYXH*) et de résistance (*cnrCBAT*). Les travaux précédents mon arrivée ont permis d'établir que la liaison des ions métalliques inducteurs sur CnrX induit des modifications conformationnelles pouvant être à l'origine de la transduction du signal de la concentration environnementale de ces ions (Trepreau *et al.*, 2011b). Le mécanisme de signalisation transmembranaire par CnrYXH est inconnu. Dans le but de déterminer la nature de ce signal, mon projet de thèse visait à cartographier les interactions protéine:protéine au sein du complexe CnrYXH dans les trois compartiments concernés : le périplasma, la membrane plasmique et le cytoplasme. La documentation des déterminants d'interaction entre CnrX, CnrY et CnrH a permis d'élaborer un modèle structural et fonctionnel pour le complexe CnrYXH et ses homologues, certes incomplet mais soulevant des hypothèses nouvelles sur le fonctionnement du système Cnr.

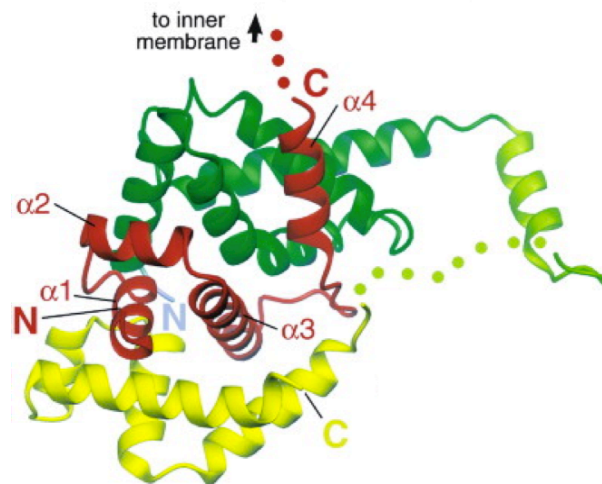
# I. Etude de l'interaction entre CnrY et CnrH dans le cytoplasme

Chez les bactéries, les facteurs sigma ( $\sigma$ ) sont des acteurs clés de l'initiation de la transcription. Les facteurs  $\sigma$  sont une sous-unité de l'ARNP bactérienne en charge de la reconnaissance des promoteurs et de l'ouverture du double brin d'ADN à ce niveau. La plupart des facteurs  $\sigma$  appartiennent à la famille des facteurs  $\sigma^{70}$  et partagent une organisation similaire. Cette famille est divisée en quatre groupes qui sont présentés dans l'introduction (partie **I.C.4.a.**). Les facteurs  $\sigma$  de type ECF ( $\sigma^{\text{ECF}}$ ) forment le groupe 4. Ces derniers se distinguent par l'absence des domaines  $\sigma 1$  et  $\sigma 3$ . Ce sont les plus divers en termes de fonction métabolique et de régulation (Helmann, 2002). Dans leur majorité, ces facteurs sont régulés par interaction directe avec des protéines inhibitrices spécifiques, les facteurs anti- $\sigma$ . Ces derniers sont divers de par leurs séquences, leurs structures et les mécanismes de leur régulation (Brown et Hughes, 1995).

Récemment, un domaine N-terminal constitué de quatre hélices a été caractérisé dans les facteurs anti- $\sigma$  RseA d'*E. coli* et ChrR de *R. sphaeroides* (Campbell *et al.*, 2003, 2007). Ce domaine appelé ASD (anti-sigma domain), long d'environ 70 résidus, interagit avec le facteur  $\sigma$  et est suffisant pour inhiber son activité. Une analyse fondée sur les similarités de séquence et les prédictions de structure secondaire a permis de prédire que l'ASD serait présent dans 30% des facteurs anti- $\sigma^{\text{ECF}}$  (Campbell *et al.*, 2007).

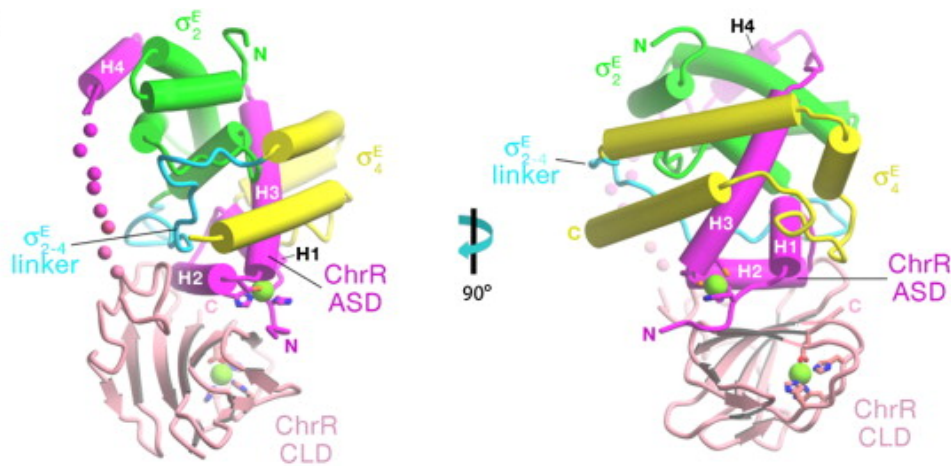
L'obtention de plusieurs structures de complexes  $\sigma$ :anti- $\sigma$  a permis de mieux comprendre le mode d'inhibition des facteurs  $\sigma$  par les anti- $\sigma$ . Chez *E. coli*, le facteur  $\sigma^{\text{E}}$  impliqué dans la réponse à un choc thermique, est régulé négativement par RseB et le facteur anti- $\sigma$  RseA (le système est présenté dans l'introduction, section **I.C.4.b.i.**). La structure cristallographique du complexe entre la partie cytoplasmique de RseA et  $\sigma^{\text{E}}$  révèle comment RseA inhibe  $\sigma^{\text{E}}$ . Le domaine ASD de RseA interagit avec les deux domaines  $\sigma 2$  et  $\sigma 4$  et masque tous les déterminants de l'interaction de  $\sigma^{\text{E}}$  avec l'ARNP et l'ADN (**Figure 1**) (Campbell *et al.*, 2003, 2008).





**Figure 1** : Structure cristallographique du complexe  $\sigma^E$ :RseA d'*E. coli*. Les domaines  $\sigma 2$  et  $\sigma 4$  de  $\sigma^E$  sont en vert et jaune, respectivement. La séquence de liaison entre les deux domaines est en vert clair. Entre cette séquence de liaison et le domaine  $\sigma 4$ , un fragment de 8 résidus n'est pas ordonné et est représenté sous forme de sphères vert clair. Le facteur anti- $\sigma$  RseA est en rouge (les hélices  $\alpha 1$  à  $\alpha 4$  sont annotées), il interagit à l'interface des deux domaines  $\sigma$  de  $\sigma^E$ . L'extrémité C-terminale de RseA se prolonge par une ancre TM (Campbell *et al.*, 2003).

Chez *Rhodobacter sphaeroides*, la réponse transcriptionnelle à l'oxygène singulet est contrôlée par le facteur  $\sigma^E$  et le facteur anti- $\sigma$ , ChrR. Dans la structure cristallographique du complexe, ChrR se compose de deux domaines, une partie N-terminale (ASD) qui lie un ion zinc et interagit avec le facteur sigma. Ce domaine interagit avec les déterminants de l'interaction avec l'ARNP situés à l'interface des deux domaines,  $\sigma 2$  et  $\sigma 4$  de  $\sigma^E$ , inhibant ainsi l'activité de  $\sigma^E$ . ChrR se compose aussi d'un domaine C-terminal (CLD pour Cupin-Like Domain) qui peut coordonner un zinc additionnel et est nécessaire pour la réponse transcriptionnelle au stress oxydant (**Figure 2**) (Campbell *et al.*, 2007).



**Figure 2 :** Structure cristallographique du complexe  $\sigma^E$ :ChrR de *R. sphaeroides*. Les hélices sont représentées sous forme de cylindres. Les domaines  $\sigma_2$  et  $\sigma_4$  de  $\sigma^E$  sont en vert et jaune, respectivement. La séquence de liaison entre les deux domaines est en cyan. Le domaine ASD (les 4 hélices sont annotées H1 à H4) de ChrR est coloré en magenta et le domaine CLD en rose. Les ions zinc (un sur chaque région de ChrR) sont représentés par une sphère verte. La séquence de liaison entre les deux domaines ChrR est désordonnée et représentée sous forme de sphères magenta (Campbell *et al.*, 2007).

Dans le système CnrYXH, CnrY est le facteur anti- $\sigma$  qui séquestre le facteur  $\sigma$ , CnrH, à la membrane plasmique. Cette inhibition est effectuée plus précisément par le domaine cytosolique de CnrY (CnrYc) composé de 45 résidus. Cela en fait un des plus petits domaines anti- $\sigma$  identifiés. L'étude de l'interaction entre CnrH et CnrYc a été un enjeu important pour une meilleure compréhension du mode d'inhibition de CnrH par CnrY. C'est cette problématique qui est abordée dans le premier article présenté dans ce manuscrit. Ce projet a bénéficié d'une approche structurale par cristallographie et du développement d'un test fonctionnel qui a permis la caractérisation fine du site d'interaction entre les différents partenaires du complexe. C'est à cette deuxième partie que j'ai le plus contribué.

## **A. Publication Maillard *et al.*, 2014**

### **The crystal structure of the anti- $\sigma$ factor CnrY in complex with the $\sigma$ factor CnrH shows a new structural class of anti- $\sigma$ factors targeting ExtraCytoplasmic-Function $\sigma$ factors**

Antoine P. Maillard<sup>1-4,\*</sup>, Eric Girard<sup>1-5</sup>, Widade Ziani<sup>1-5</sup>, Isabelle Petit-Härtlein<sup>1-4</sup>, Richard Kahn<sup>1-4</sup> and Jacques Covès<sup>1-4</sup>

<sup>1</sup> Université Grenoble Alpes, Grenoble, France

<sup>2</sup> Commissariat à l'Energie Atomique and aux Energies Alternatives (CEA), Grenoble, France

<sup>3</sup> Centre National de la Recherche Scientifique (CNRS), Grenoble, France

<sup>4</sup> Institut de Biologie Structurale, F-38000 Grenoble, France

<sup>5</sup> These authors contributed equally to this work.

Running title:

**Molecular basis for  $\sigma$  factor regulation**

\* Correspondence to Antoine Maillard: [antoine.maillard@ibs.fr](mailto:antoine.maillard@ibs.fr)

Article publié dans *J. Mol. Biol.* en 2014 (Maillard and *al.*, 2014)

## Abstract

Gene expression in bacteria is regulated at the level of transcription initiation, a process driven by  $\sigma$  factors. The regulation of  $\sigma$  factor activity proceeds from the regulation of their cytoplasmic availability, which relies on specific inhibitory proteins called anti- $\sigma$  factors. With anti- $\sigma$  factors regulating their availability according to diverse cues, ExtraCytoplasmic Function  $\sigma$  factors ( $\sigma^{\text{ECF}}$ ) form a major signal transduction system in bacteria. Here, structure: function relationships have been characterized in an emerging class of minimal-size, transmembrane anti- $\sigma$  factors, using CnrY from *Cupriavidus metallidurans* CH34 as a model. This study reports the 1.75 Å resolution structure of CnrY cytosolic domain in complex with CnrH, its cognate  $\sigma^{\text{ECF}}$  and identifies a small hydrophobic knob in CnrY as the major determinant of this interaction *in vivo*. Unsuspected structural similarity with the molecular switch regulating the general stress response in  $\alpha$ -proteobacteria unravels a new class of anti- $\sigma$  factors targeting  $\sigma^{\text{ECF}}$ . Members of this class carry out their function *via* a 30-residue stretch that displays helical propensity but no canonical structure on its own.

## Keywords

Protein:protein interaction

Sigma factor

Bacterial transcription

Transmembrane signaling

Metal resistance

## Introduction

The bacterial RNA polymerase (RP) uses one of several dissociable subunits called  $\sigma$  factors to recognize gene promoters and to melt the double-stranded DNA at the transcription initiation site<sup>1</sup>. Most  $\sigma$  factors belong to the  $\sigma^{70}$  family: they show related transcription initiation mechanisms<sup>2</sup> and display two to four of the typical domains named  $\sigma_1$  through  $\sigma_4$ <sup>3</sup>. The simplest of  $\sigma^{70}$  factors only display domains  $\sigma_2$  and  $\sigma_4$  connected by a linker unrelated to  $\sigma_3$ : this is the signature of a large, diversified subfamily of  $\sigma$  factors collectively called ExtraCytoplasmic Function  $\sigma$  factors ( $\sigma^{\text{ECF}}$ ) because most of them respond to extracellular stress<sup>4,5</sup>.

Bacteria possess several  $\sigma$  factors that control the transcription of different sets of genes<sup>6</sup> and compete to bind the  $\alpha_2\beta\beta'\omega$  core of the RP<sup>7</sup>. Sigma factor availability is regulated by anti- $\sigma$  factors (anti- $\sigma$ ) and proteins that build regulatory networks around them<sup>8</sup>. For most  $\sigma^{\text{ECF}}$ , regulatory networks reach across the cytoplasmic membrane, allowing their cytoplasmic availability to reflect a particular condition sensed in the envelope. Consistently,  $\sigma^{\text{ECF}}$  play a prominent role in bacterial signal transduction<sup>9</sup>. Sequence profiling has allowed sorting  $\sigma^{\text{ECF}}$  in an open classification of more than forty subgroups which reflect stimulus and promoter specificities<sup>10</sup>. In half the subgroups, the corresponding anti- $\sigma$ s display a similar 70-residue long structural motif, called ASD for anti $\sigma$  domain, in combination with diverse sensory modules<sup>11</sup>.

A co-founder of the  $\sigma^{\text{ECF}}$  subfamily<sup>5</sup>, CnrH controls cobalt and nickel resistance in *Cupriavidus metallidurans* CH34. CnrH is regulated by a complex of two transmembrane proteins: the periplasmic sensor CnrX and the anti- $\sigma$  CnrY<sup>12; 13; 14; 15</sup>. At rest, CnrH is sequestered by CnrY whose 45 residue long cytosolic domain is one of the shortest anti- $\sigma$  domains. Upon Ni(II) or Co(II) ions detection by CnrX in the periplasm, CnrH is released by an unknown mechanism (**Figure 3a**). We report the crystal structure at 1.75 Å resolution of a complex between CnrH and the cytosolic domain of CnrY (CnrY<sub>C</sub>) as well as structure: function relationships in CnrY. The CnrH:CnrY<sub>C</sub> complex displays an unexpected structural similarity to the anti- $\sigma$  NepR in complex with its antagonist PhyR<sup>16; 17</sup> whereas NepR shares no sequence similarity with CnrY. Based on structural similarity alone, they form a new class of intrinsically unfolded anti- $\sigma^{\text{ECF}}$  with a distinct mode of inhibitory binding

that involves a closed conformation of the  $\sigma^{\text{ECF}}$ . In keeping with the nomenclature in use, the linear anti- $\sigma^{\text{ECF}}$  motif was named a class II ASD.

## Results

### Protein production, purification, crystallization and structure determination

Strep-tagged CnrH (strepCnrH) and CnrY<sub>C</sub> were coproduced in *E. coli* and purified with three chromatographic steps. The strepCnrH:CnrY<sub>C</sub> complex behaved like a globular protein of ~30 kDa in gel filtration, which is compatible with a 1:1 stoichiometry. Crystallogenesis occurred only after uncontrolled proteolysis of strepCnrH:CnrY<sub>C</sub> had taken place (Materials and Methods): crystals belonged to the I222 space group with one CnrH:CnrY<sub>C</sub> complex per asymmetric unit and 25% solvent content. A model of the complex was first determined at medium resolution by single wavelength anomalous dispersion, using a Gd:DO3A derivative<sup>18</sup>. This model was then used as a molecular replacement search model to determine the structure from native data at 1.75 Å resolution (**Figure 3b**). Iterative rounds of refinement and model building converged to R<sub>work</sub>/R<sub>free</sub> 0.174/0.201 (**Table 1**). Three polypeptide chains were resolved in the asymmetric unit: the  $\sigma_2$  and  $\sigma_4$  domains, encompassing CnrH residues 5 to 91 and 123 to 189 (chain A) and CnrY<sub>C</sub> residues 2 to 30 (chain B) (**Figure 3c**).

### Structure of the CnrH:CnrY<sub>C</sub> complex

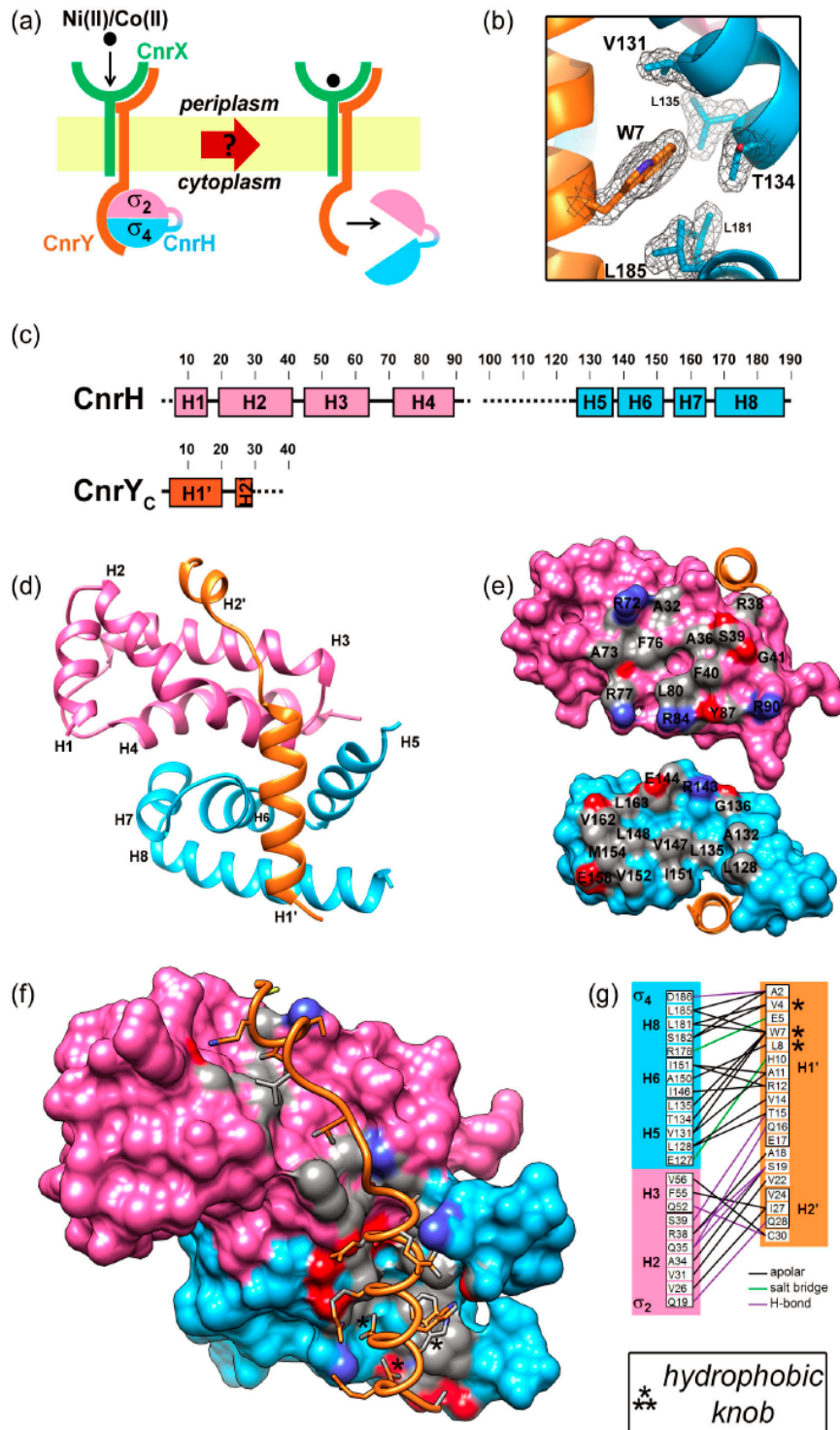
In CnrH:CnrY<sub>C</sub>, a L-shaped CnrY<sub>C</sub> binds a globular CnrH. Domains  $\sigma_2$  and  $\sigma_4$  of CnrH contact each other through a flat interface that buries 1494 Å<sup>2</sup> of solvent accessible surface. Most contact pairs between  $\sigma_2$  and  $\sigma_4$  involve apolar moieties carried by helices H2 and H5, H2 and H6, H4 and H7 (**Figure 3d,e**). Tight packing within CnrH did not alter the fold integrity of  $\sigma_2$  and  $\sigma_4$  domains: domain-to-domain comparison between CnrH and SigA in the 2.6 Å-resolution structure of *T. thermophilus* RP holoenzyme<sup>19</sup> showed that the topology of CnrH domains was consistent with that of SigA  $\sigma_2$  and  $\sigma_4$  domains in a functional RP (**Table 2**). At the surface of CnrH, a remarkable groove is delineated by helices H2 and H3 of  $\sigma_2$  and extends into a hydrophobic pocket shaped by H5, H6 and H8 of  $\sigma_4$  (**Figure 3f**). CnrY<sub>C</sub> fits well in this 40 Å-long hydrophobic groove, as 40% (1245 Å<sup>2</sup>) of its molecular surface is buried.

In the complex, CnrY<sub>C</sub> residues 3-19 are folded as a well-defined  $\alpha$ -helix (H1') that fits into  $\sigma_4$  hydrophobic pocket. The peptide further extends along the hydrophobic groove of  $\sigma_2$  with no canonical structure except for a short helical turn (H2') spanning residues 24-28 (**Figure 3g**). Depending on the domain of CnrH they interact with, CnrY<sub>C</sub> residues part in two groups with little overlap: the area of solvent accessible surface buried is 570 Å<sup>2</sup> on  $\sigma_2$  and 660 Å<sup>2</sup> on  $\sigma_4$ . A hydrophobic knob made of V4, W7 and L8 side-chains protrudes into  $\sigma_4$  hydrophobic pocket and contributes 320 Å<sup>2</sup> to the interface. In contrast, the contact surface on  $\sigma_2$  is patchy with residues in and upstream H2' (V22, V24, I27) making hydrophobic contacts along the groove. Finally, the complex is also stabilized by polar interactions (**Figure 3g**).

**Table 1:** Data collection, phasing and refinement statistics

	Native	SAD data set
<b>Data collection</b>		
Space group	I 2 2 2	I 2 2 2
Cell dimensions		
<i>a, b, c</i> (Å)	61.01, 79.12, 80.54	60.43, 78.71, 81.59
Wavelength (Å)	0.8726	1.7102
Resolution (Å)*	39.56-1.75 (1.84-1.75)	47.93-2.50 (2.63-2.49)
<i>R</i> <sub>sym</sub> *	0.038 (0.489)	0.078 (0.351)
<i>I</i> / $\sigma I$ *	20.4 (2.5)	20.8 (8.6)
Completeness (%)*	99.2 (99.2)	98.8 (94.6)
Redundancy*	3.7 (3.6)	7.6 (7.7)
<b>Refinement</b>		
Resolution (Å)	39.56-1.75	
No. reflections	19830	
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.174 / 0.201	
No. atoms		
CnrH Protein	1242	
CnrY Protein	224	
Ligand/ion	21	
Water	144	
<i>B</i> -factors (Å <sup>2</sup> )		
CnrH Protein	11.55	
CnrY Protein	25.65	
Ligand/ion	47.51	
Water	16.83	
R.m.s deviations		
Bond lengths (Å)	0.007	
Bond angles (°)	1.071	

\*Values in parentheses are for the highest-resolution shell.



**Figure 3: CnrY embraces the closed conformation of CnrH.** (a) CnrYXH signaling cascade involves Ni(II) or Co(II) binding to the sensor CnrX (green), which induces the anti- $\sigma$  factor CnrY (orange) to release CnrH ( $\sigma_2$  in pink,  $\sigma_4$  in blue). CnrH drives the transcription of cobalt and nickel resistance genes. The color code in use here applies to all panels. (b)  $2F_o - F_c$  electron density map near CnrY residue W7, contoured at  $1 \sigma$ . (c) Diagram representation of CnrH and CnrY<sub>C</sub> fragments present in the crystal with segments unresolved in the electron density pictured as dashed lines. See text for details. (d) Ribbon representation of CnrY<sub>C</sub> in



complex with CnrH domains  $\sigma_2$  and  $\sigma_4$ . Helices were tagged near their N-terminus. **(e)** Open book display of  $\sigma_2$  and  $\sigma_4$  molecular surfaces with atoms of the  $\sigma_2$ : $\sigma_4$  interface in CPK colors. CnrY<sub>C</sub> H1' (orange) is shown for orientation. Residues of  $\sigma_2$  in contact with  $\sigma_4$  (top) make a central hydrophobic patch surrounded by a rim of arginyl residues contributing apolar contacts (R38, R72, R77) and interdomain salt bridges (R72:E158, R84:E144). Likewise, most  $\sigma_4$  residues in contact with  $\sigma_2$  are hydrophobic (bottom). **(f)** CnrH:CnrY<sub>C</sub> interface atoms. Intermolecular atomic contacts are CPK colored on CnrH molecular surface and CnrY<sub>C</sub> stick representations. V4, W7 and L8 side chains (stars) form a hydrophobic knob that protrudes into  $\sigma_4$  hydrophobic pocket. **(g)** CnrH:CnrY<sub>C</sub> interface residues. CnrH domains  $\sigma_2$ ,  $\sigma_4$  and CnrY<sub>C</sub> appear as pink, blue and orange boxes respectively. Only residues with atomic contacts shorter than 4.0 Å are displayed. Line colors depend on the chemical nature of the interaction, as specified in the picture.

### CnrH functional surfaces are occluded in the CnrH:CnrY<sub>C</sub> complex

To achieve functionality,  $\sigma$  factors bind simultaneously the RP core and promoter DNA in the transcription initiation complex (TIC). The TIC loaded with CnrH as a  $\sigma$  factor was modeled in order to shed light on CnrH inhibition by CnrY. To this end, the high resolution crystal structures of *T. aquaticus* SigA domains  $\sigma_2$  and  $\sigma_4$  in complex with mimics of appropriate promoter regions<sup>20, 21</sup> and the 6 Å-resolution crystal structure of the *T. aquaticus* TIC<sup>22</sup> were used as templates (**Table 2**). As expected from the high degree of conservation within the  $\sigma$ <sup>70</sup> family<sup>23</sup>, only minor steric clashes could be observed after  $\sigma_2$  and  $\sigma_4$  domains were superimposed. Consequently, the models were used to assess steric hindrances that impede CnrH function when bound to CnrY.

The functional organization of CnrH domains (**Figure 4a**) was inferred from a refined structural alignment of  $\sigma$  factors<sup>24</sup>. Domain  $\sigma_2$  binds the promoter -10 element *via* conserved regions 2.3 and 2.4<sup>21</sup> and binds the so-called  $\beta'$  coiled-coil domain, a major site of interaction with  $\sigma$  factors<sup>25</sup> *via* regions 2.1 and 2.2. In CnrH:CnrY<sub>C</sub>,  $\sigma_2$  cannot take part in any of these interactions since CnrY residues 21-30 compete with the  $\beta'$  coiled-coil domain and CnrH  $\sigma_4$  conceals the DNA-binding surface of  $\sigma_2$  (Figure 2b,c). Domain  $\sigma_4$  binds the promoter -35 element *via* region 4.2<sup>20</sup> which is freely accessible in CnrH:CnrY<sub>C</sub> and it docks the  $\beta$  flap-tip helix at the same hydrophobic pocket where CnrY<sub>C</sub> binds (Figure 2b,d). Although CnrY<sub>C</sub> H1' and the  $\beta$ -flap tip helix run antiparallel to each other, it is remarkable that CnrY<sub>C</sub> residues forming the aforementioned hydrophobic knob, V4, W7 and L8 superimpose with conserved residues of the  $\beta$ -flap tip helix (**Figure 4e**) that mediate  $\sigma_4$  function<sup>26</sup>.

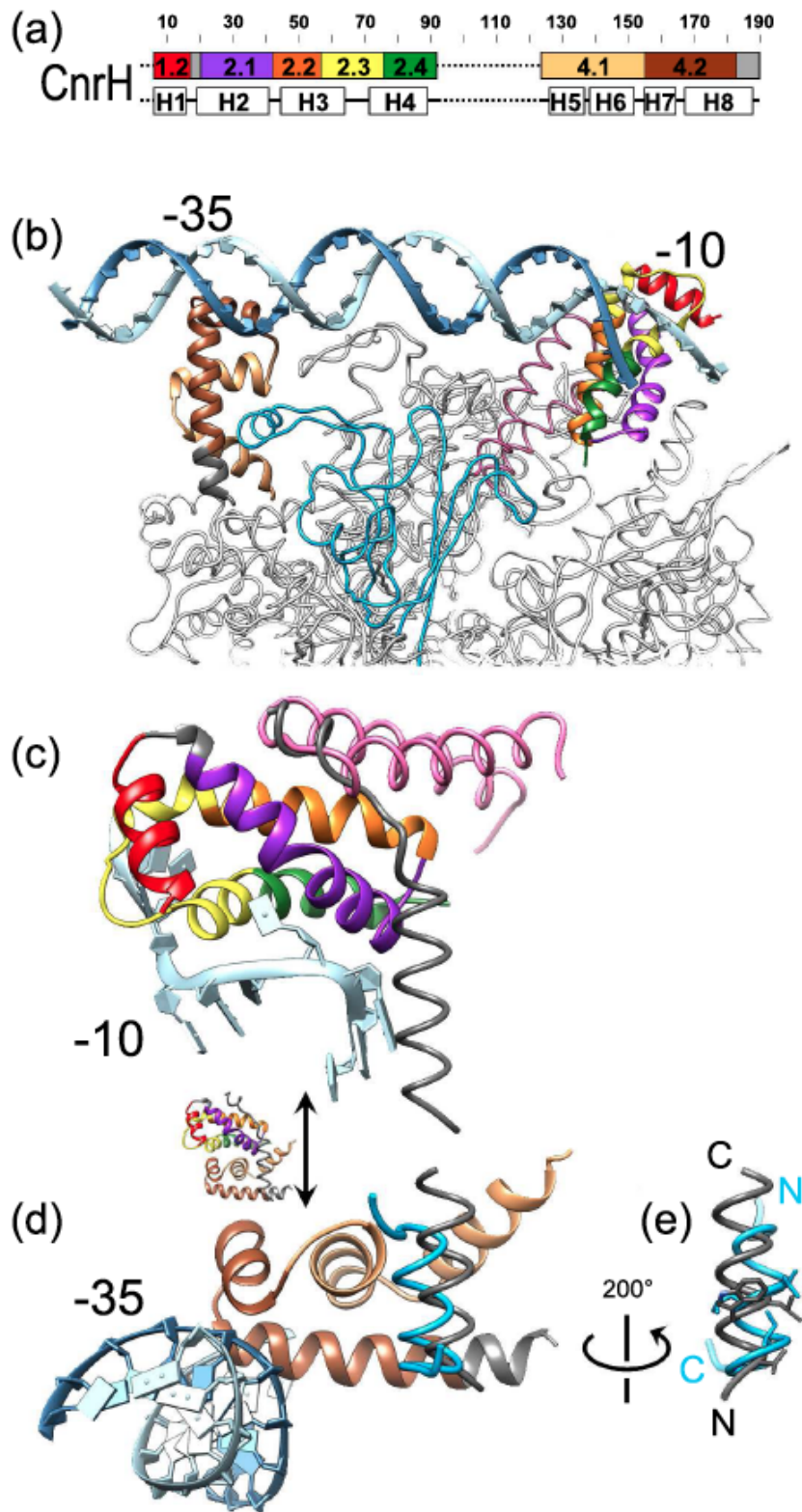
**Table 2:** Structural superimpositions indicators.<sup>a</sup>

Shown	Domain	Target b	Template b	Rmsd (Å)	Number of Ca pairs
Fig. 4b	$\sigma_2$	CnrH	1L9Z	2.19	71
Fig. 4b	$\sigma_4$	CnrH	1L9Z	1.53	65
Fig. 4c	$\sigma_2$	1IW7	CnrH	1.76	74
Fig. 4c	$\sigma_3$	3UGO	1IW7	1.63	159
Fig. 4c	$\sigma_2$	3UGO	CnrH	1.76	74
Fig. 4d,e	$\sigma_4$	1IW7	CnrH	1.88	60
Fig. 4d,e	$\sigma_4$	1KU7	1IW7	1.68	62
Fig. 4d,e	$\sigma_4$	1KU7	CnrH	1.81	60
Fig. 5b	$\sigma_2$ & $\sigma_4$	CnrH	2LFW	2.45 <sup>c</sup>	121
Fig. 5b	–	CnrY	2LFW	2.57 <sup>c</sup>	24
Fig. 9a	$\sigma_1$	1OR7	CnrH	1.19	56
Fig. 9a	$\sigma_1$	1RP3	CnrH	1.48	58
Fig. 9b	$\sigma_1$	CnrH	1IW7	1.76	74
Fig. 9b	$\sigma_3$	1OR7	1IW7	2.57	61
Fig. 9b	$\sigma_3$	1RP3	1IW7	1.68	78
Fig. 9b	$\sigma_1$	CnrH	1IW7	1.88	60
Fig. 9b	$\sigma_1$	1OR7	1IW7	1.99	56
Fig. 7b	$\sigma_1$	1RP3	1IW7	1.98	59
Fig. 7b	$\sigma_1$	2P7V	1IW7	1.80	46

<sup>a</sup> Using UCSF Chimera, structural superimpositions have been performed with equal weights on secondary structure matching and local sequence similarity (Smith-Waterman algorithm with the BLOSUM45 similarity matrix).

<sup>b</sup> Target and template models are designated by the PDB ID of the complex they belong to. 1L9Z, *T. aquaticus* transcription initiation complex at 6.5 Å resolution <sup>21</sup>. 1IW7, *T. thermophilus* RNA polymerase at 2.6 Å resolution <sup>18</sup>. 3UGO, *T. aquaticus* SigA  $\sigma_2$ :DNA complex at 2.1 Å resolution <sup>20</sup>. 1KU7, *T. aquaticus* SigA  $\sigma_4$ :DNA complex at 2.4 Å resolution. 1OR7, *E. coli* RpoE:RseA<sub>C</sub> complex at 2.0 Å resolution <sup>23</sup>. 1RP3, *A. aeolicus* FliA:FlgM complex at 2.3 Å resolution <sup>33</sup>. 2P7V, *E. coli* RpoD  $\sigma_4$ :Rsd complex at 2.6 Å resolution <sup>35</sup>. See text and figures' legends for details.

<sup>c</sup> Each of the 15 models of PhyR (2LFW) was superimposed onto CnrH and the mean of the 15 root mean square deviations (rmsd) was calculated separately for PhyR:CnrH and for NepR:CnrY<sub>C</sub>.

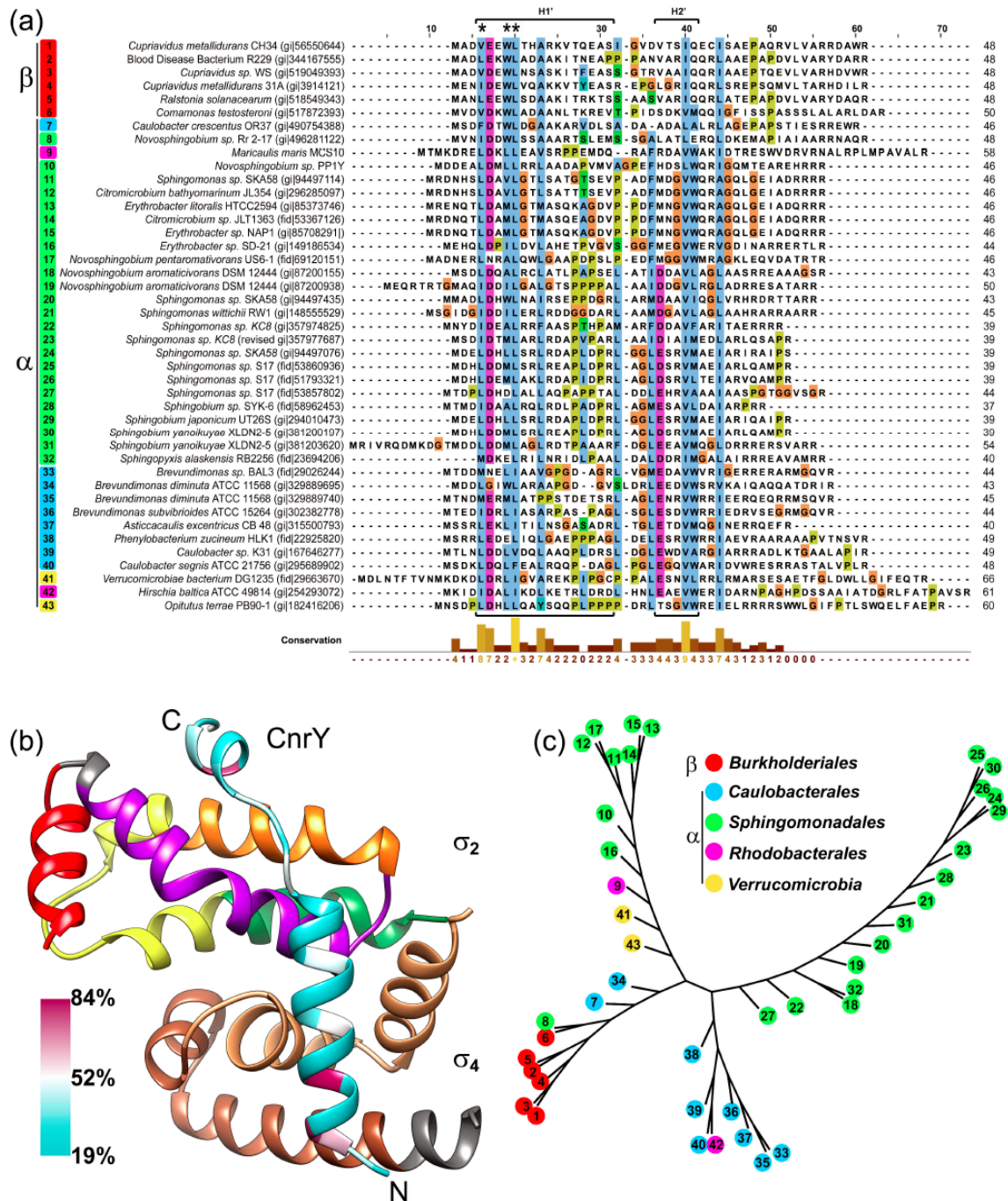


**Figure 4: Molecular basis of CnrH inhibition by CnrY<sub>c</sub>.** (a) Diagram representations of CnrH show conserved  $\sigma^{70}$  regions as color-coded boxes (top),  $\alpha$ -helices as bricks (bottom) and portions not determined in the electron density as dashed lines. (b) Model of CnrH within the transcription initiation complex. The RNA polymerase core (coil) contacts CnrH (ribbons) *via* the  $\beta'$  coiled-coil domain (pink) and the  $\beta$  flap tip helix (blue). The tips of the unresolved

segment connecting CnrH domains stand 55 Å apart, which is compatible with the length of a 31-residue linker. **(c)** A close-up view at CnrH  $\sigma_2$  shows that CnrY<sub>C</sub> helix H2' (grey) competes with the  $\beta'$  coiled-coil domain (pink) while CnrH  $\sigma_i$  domain competes with the promoter -10 element. The latter is suggested by the double-arrow between panels (c) and (d) while the inset pictures CnrH:CnrY<sub>C</sub>. **(d)** A close-up view at CnrH  $\sigma_i$  shows that CnrY<sub>C</sub> (grey) prevents  $\sigma_4$  binding to the  $\beta$ -flap tip helix (blue) but not to the promoter -35 element. **(e)** CnrY<sub>C</sub> residues V4, W7 and L8 superimpose with conserved residues I777, L774 and L773 of the  $\beta$ -flap tip helix, respectively.

### **CnrY proteins: an emerging group with limited sequence conservation**

Similarity searches with *C. metallidurans* CH34 CnrY protein sequence returned five homologs. These are 95 residue-long transmembrane proteins performing their anti- $\sigma$  function *via* a 45 residue-long cytosolic N-terminal domain. More CnrY-like sequences were collected using gene association within *cnrYXH*-like operons. This suggests that CnrY is the prototype of an emerging group of proteins mediating metal-dependent transmembrane signaling in connection with a  $\sigma^{\text{ECF}}$  and a CnrX-like periplasmic metal sensor<sup>27</sup>. CnrY<sub>C</sub> sequences were readily aligned even though similarity was confined to portions of the cytosolic domain (**Figure 5a**): four-to-five residues from the N-terminal half of H1', two hydrophobic residues in and downstream H2'. These residues are those that occlude the RP binding determinants of CnrH (**Figure 5b**). Residue conservation directly downstream H2' suggests that H2' might be longer than observed in CnrY<sub>C</sub>. The phylogenetic analysis of CnrY-like proteins also revealed that sequences from  $\beta$ - and  $\alpha$ -proteobacteria cluster together (Figure 3c): this and the observation that the *cnr* determinant displays a distinct G+C content in *C. metallidurans* indicate that *cnrYXH* may have been acquired by horizontal gene transfer along with the resistance genes that they control<sup>28; 29</sup>.

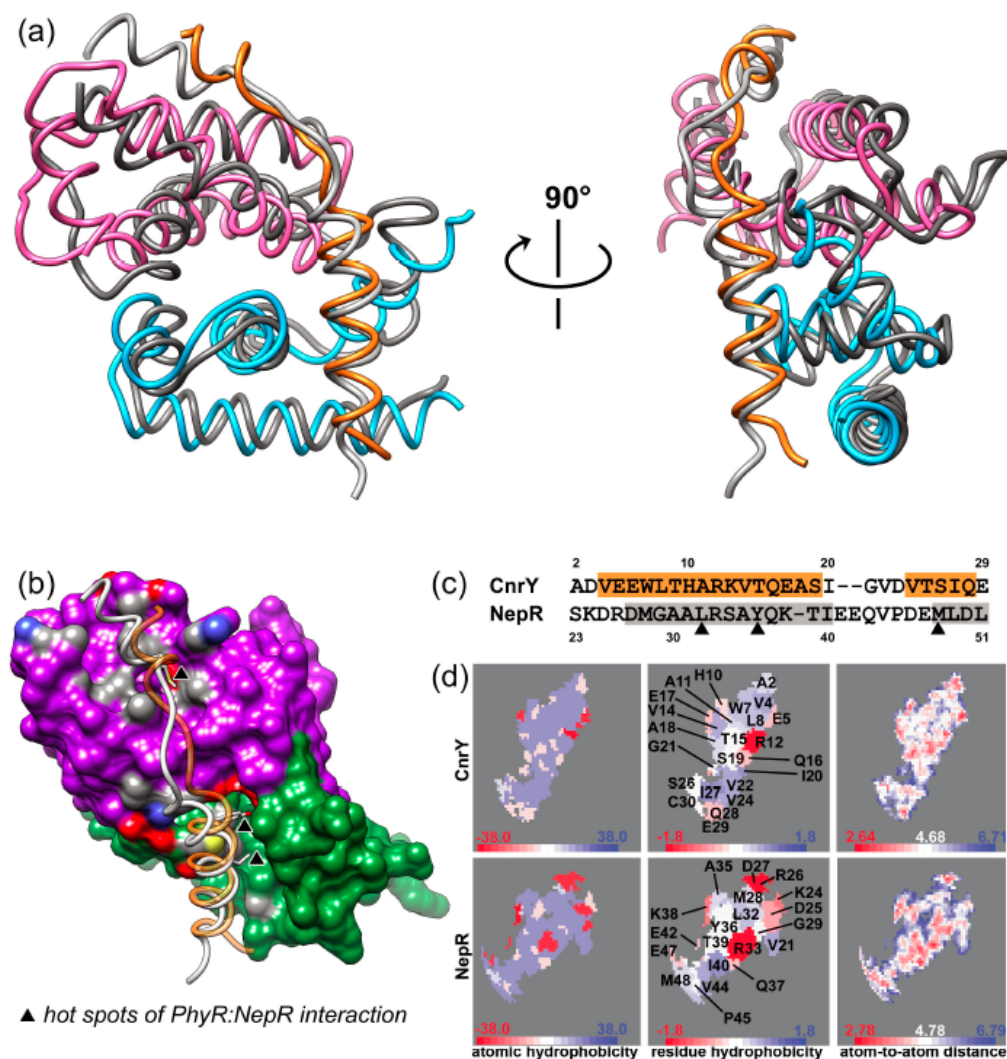


**Figure 5: Sequence conservation analysis of CnrY-like proteins.** (a) Multiple sequence alignment of cytoplasmic sequences from the CnrY-like proteins identified in this work. Brackets span the helices H1' and H2' seen in CnrY<sub>C</sub> and stars highlight the residues of the hydrophobic knob. Each sequence is identified by the name of the source organism and a specific ID from NCBI (gi) or PATRIC (fid) database. Transmembrane and periplasmic domains are not shown. The alignment was edited with Jalview and rendered with the ClustalX color display which takes physicochemical properties (coloring scheme) and conservation (shading) into account. Sequences were number-coded from top to bottom and color-coded based on their phylogenetic lineage: red, Burkholderiales; blue, Caulobacteriales; green, Spingomonadales; magenta, Rhodobacterales; gold, Verrucomicrobia. Brackets and single  $\alpha$  and  $\beta$  letters designate groups of sequences from  $\alpha$ - and  $\beta$ -proteobacteria. (b) Mapping of CnrY sequence conservation on a ribbon representation of CnrH:CnrY<sub>C</sub>.

CnrY<sub>C</sub> residues are colored from cyan to maroon as conservation increases (see color key) and CnrH conserved regions are colored as in Fig 2. CnrY<sub>C</sub> partially occludes conserved regions 2.1 (purple), 2.2 (orange), 4.1 (sand) and 4.2 (brown) while the closed conformation of CnrH partially occludes regions 2.1, 2.3 (yellow), 2.4 (green), 4.1 and 4.2. (c) Unrooted phylogenetic tree inferred from the alignment shown in panel (a). The distribution of CnrY-like proteins in  $\beta$ -proteobacteria (Burkholderiales, red) may be the result of horizontal gene transfer from  $\alpha$ -proteobacteria (Caulobacteriales, Sphingomonadales, Rhodobacterales, Verrucomicrobia). See text for details.

### Unexpected structural similarity with a known molecular switch

CnrH:CnrY<sub>C</sub> was found to be structurally superimposable to PhyR:NepR (**Figure 6a**). The response regulator of a two-component system that controls the general stress response in  $\alpha$ -proteobacteria, PhyR has an effector domain that mimicks a  $\sigma^{\text{ECF}}$  and titrates the anti- $\sigma$  NepR away from its  $\sigma$  factor. Upon phosphorylation of PhyR, NepR switches from the  $\sigma$  factor EcfG to the  $\sigma$ -like domain of PhyR<sup>16; 17; 30</sup>. The  $\sigma$ -like domain of *Sphingomonas sp.* Fr1 PhyR<sup>16</sup> superimposed to CnrH with a 2.45 Å rmsd over 121 C $\alpha$  pairs. Consistently, CnrY<sub>C</sub> and NepR aligned over 24 C $\alpha$  pairs with an rmsd of 2.57 Å as a result of this superimposition, thus indicating that both complexes are similar (**Figure 6a-c**). Both interfaces were of comparable size (~1200 Å<sup>2</sup>) and hydrophobic character, although the identity and distribution of the contact points were specific (**Figure 6d**).

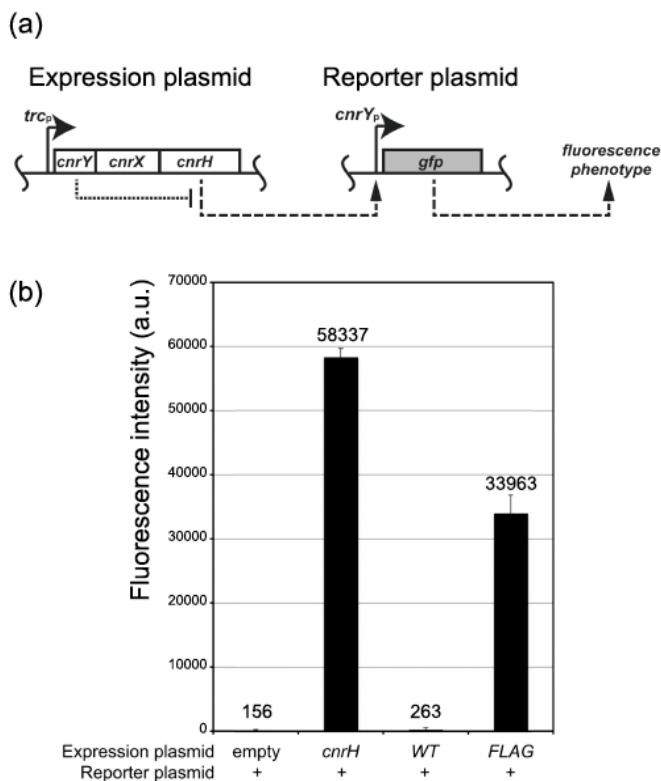


**Figure 6: CnrY and NepR define a new structural class of anti- $\sigma$ s.** (a) *Sphingomonas sp.* Fr1 PhyR (grey) and CnrH ( $\sigma_2$  in pink and  $\sigma_4$  in blue) have been superimposed: not only do CnrH and PhyR effector domain adopt similar closed conformations, but they also build similar complexes with CnrY<sub>C</sub> (orange) and NepR (lighter grey), respectively. (b) The molecular surface of PhyR  $\sigma_2$ -like (purple) and  $\sigma_4$ -like (green) domains is shown with the backbone traces of CnrY<sub>C</sub> (transparent orange) and NepR (grey) running as coils. The side chains of three major determinants of PhyR:NepR interaction, NepR residues L32, Y36 and M48 are displayed as sticks<sup>16</sup> and highlighted with triangles. Intermolecular atomic contacts are CPK colored on PhyR surface and on NepR-borne sticks. (c) Structure-based alignment of CnrY and NepR sequences. Residues of H1' and H2' are shaded. The three residues of NepR shown to be important for PhyR binding<sup>16</sup> are highlighted with triangles. (d) The contact surfaces on CnrY<sub>C</sub> (upper panels) and NepR (lower panels) are similar in shape and hydrophobic character (left), but they show specificity in residue composition (middle) and in the distribution of contact points with their target (right). Color keys for hydrophobicity or atom-to-atom distances are displayed. Interface stamps were produced with MolSurfer<sup>61</sup>.



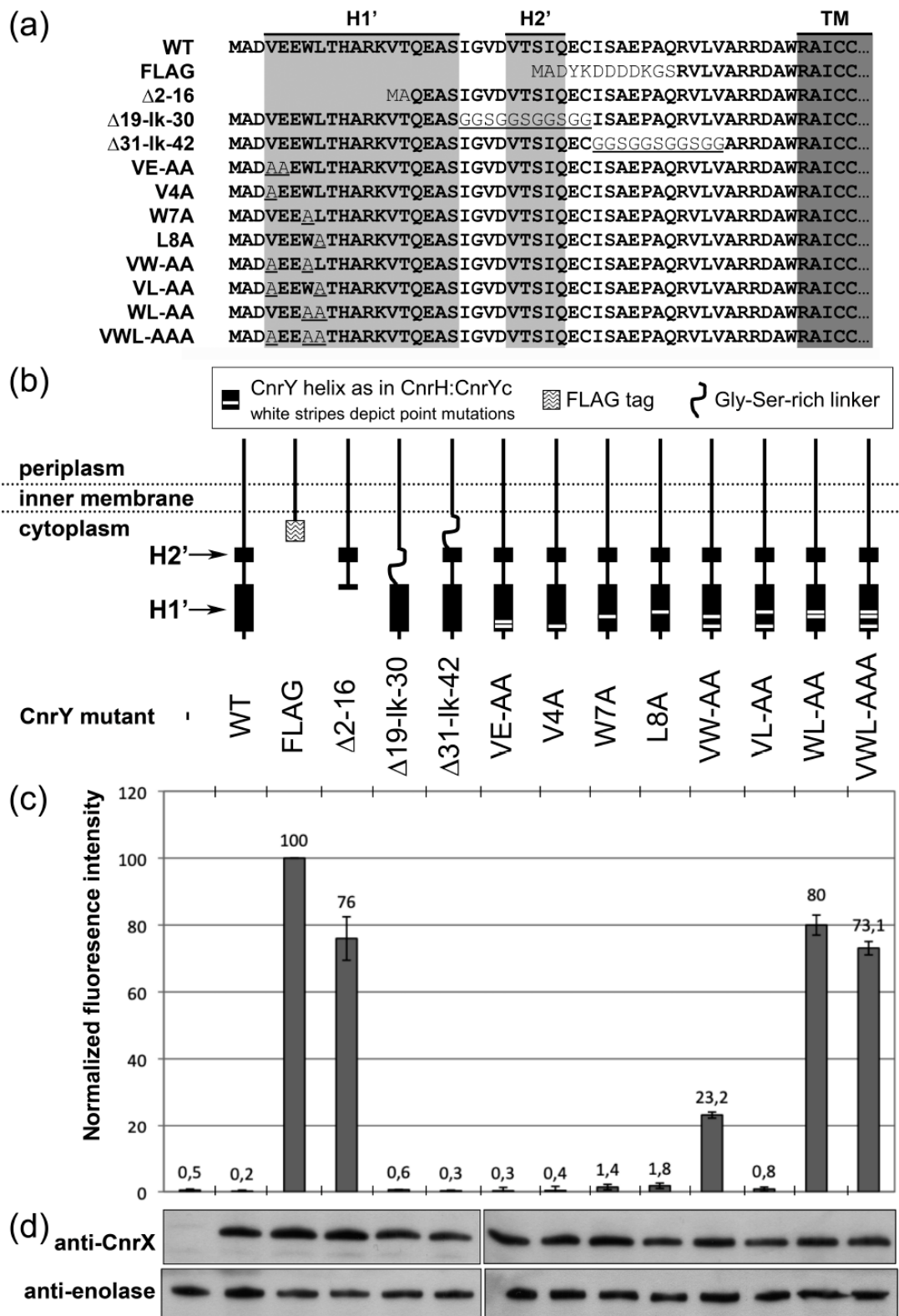
## A two-residue hotspot mediates inhibition of CnrH by CnrY *in vivo*

The structural traits of CnrH:CnrY<sub>C</sub> interaction were assessed for function. Because *C. metallidurans* CH34 exhibits slow growth, has a hypermutator phenotype and a limited number of genetic tools are available, we took advantage of CnrH activity in *E. coli* (**Figure 7**). A dual vector assay was set up, in which the production of Green Fluorescent Protein (GFP) was used to assess how changes made to CnrY sequence affect CnrH availability. Coproducing CnrH with the regulatory complex CnrYX lowered fluorescence intensity to background, whereas replacing CnrY residues 4 to 37 with a FLAG tag restored fluorescence, hence CnrH availability, to a level that was used as a reference to assess mutations in CnrY.



**Figure 7: CnrY-dependent activity of CnrH in *E. coli*.** (a) Design principle of the *in vivo* functional assay. Compatible *E. coli* vectors carry the *cnrYXH* operon downstream an IPTG-inducible promoter and the *gfp* reporter gene downstream a CnrH-specific promoter. GFP fluorescence depends on CnrH functionality and it reflects the failure of CnrY to sequester CnrH when CnrY is present. (b) CnrH activity in the heterologous host *E. coli*. Overexpression of *cnrH* triggered high fluorescence intensity in *E. coli* cells thus demonstrating CnrH functionality in this host (counting from left, compare bars 1 and 2: - and *cnrH*). CnrH activity was efficiently quenched when *cnrH* was expressed from the *cnrYXH* operon (compare bars 2 and 3: *cnrH* and WT). The positive control of CnrH activity when produced from *cnrYXH* was obtained by replacing CnrY cytosolic domain by an irrelevant sequence (FLAG tag) in *cnrYXH* (compare bars 3 and 4: WT and FLAG). Error bars depict standard deviation about the mean.

First, CnrY mutants were made to assess the role of the different parts outlined by the secondary structure of CnrY<sub>C</sub> in complex with CnrH (**Figure 8**). The deletion of H1' (residues 2 to 16) resulted in 76% fluorescence being restored, whereas the replacement of H2' (residues 19 to 30) or residues downstream (31 to 42) by a flexible, polar linker failed to produce a detectable phenotype (**Figure 8c**). Then site-directed mutagenesis was performed to replace conserved or interface residues by alanyl residues: this preserved the helicity<sup>31</sup> and presumably the stability of the constructs. When conserved dipeptides V4-E5 (VE-AA) or W7-L8 (WL-AA) were mutated, only the WL-AA mutant yielded fluorescence, with 80% signal being restored. An additional V4-to-A mutation (VWL-AAA) had no effect. Alternatively, when V4-W7 (VW-AA) or V4-L8 (VL-AA) were substituted, only the VW-AA mutant yielded fluorescence, albeit poorly. Finally, the substitution of either V4 or W7 or L8 to Ala had no phenotype. Because fluorescence intensity primarily reflects the quantity of CnrH available, a possible polar effect of *cnrY* mutations downstream on *cnrXH* expression was investigated. Protein production from *cnrX* was assessed by western analysis of the same OD-normalized cell suspension that was used to measure fluorescence. Since CnrX production was similar in all constructs (**Figure 8d**) fluorescence intensity was comforted as a quantitative measure of CnrY mutants' inability to sequester CnrH. Altogether, these results showed that the W7-L8 dyad is a hotspot mediating the inhibitory binding of CnrY to CnrH *in vivo*, while the rest of CnrY sequence contributes dispensable, diffuse determinants.



**Figure 8: *In vivo* investigation of CnrY function pinpoints part of the hydrophobic knob as a hotspot in CnrH inhibitory binding.** (a) Partial sequence of the CnrY mutated proteins tested. Shaded are sequence portions aligned with helices H1' and H2' (light grey) and the cytosolic end of the transmembrane anchor (dark grey). As discussed in the text, the intrinsic stability of H1' and H2' is probably low and they are stabilized upon CnrY binding to CnrH. The sequence of wild type CnrY is in bold type whereas mutations are not; substitutions are

underlined. **(b)** Schematic representation of the CnrY mutants tested. These are (left to right) wild type CnrY (*WT*); a CnrY derivative with a FLAG tag replacing the cytosolic domain (*FLAG*); a CnrY deletion mutant with residues 2 to 16 missing ( $\Delta 2-16$ ); CnrY mutants where a linker was substituted to residues 19 to 30 ( $\Delta 19-1k-30$ ) or residues 31 to 42 ( $\Delta 31-1k-42$ ); point mutants of CnrY where V4, E5, W7, L8 have been changed to alanyl residues either alone (*V4A*, *W7A*, *L8A*) or in combination (all others). **(c)** Fluorescence intensity was measured on live cells expressing either nothing (-, empty vector) or the *cnrYXH* operon encoding the CnrY construct specified above. The impact on CnrY function was maximized with the *WL-AA* double mutation; see text for details. Error bars depict standard deviation about the mean. **(d)** The quantitative comparison of fluorescence measurements was comforted by ruling out a polar effect of *cnrY* mutations on downstream expression of *cnrXH*. Using antibodies directed against CnrX, the amount of recombinant proteins produced from each construct was assessed by western blotting. *E. coli* enolase was used as an internal control. The same washed-cell suspension was used to perform western analyses and fluorescence spectroscopy (panel c).

## Discussion

The complex between CnrH, a  $\sigma^{\text{ECF}}$  from *C. metallidurans* CH34, and the cytosolic portion of its minimal-size anti- $\sigma$  CnrY has been characterized. This work unravels structure: function relationships in a new group of anti- $\sigma$ s and delivers several new insights regarding the regulation of  $\sigma^{70}$ -family  $\sigma$  factors by anti- $\sigma$ s.

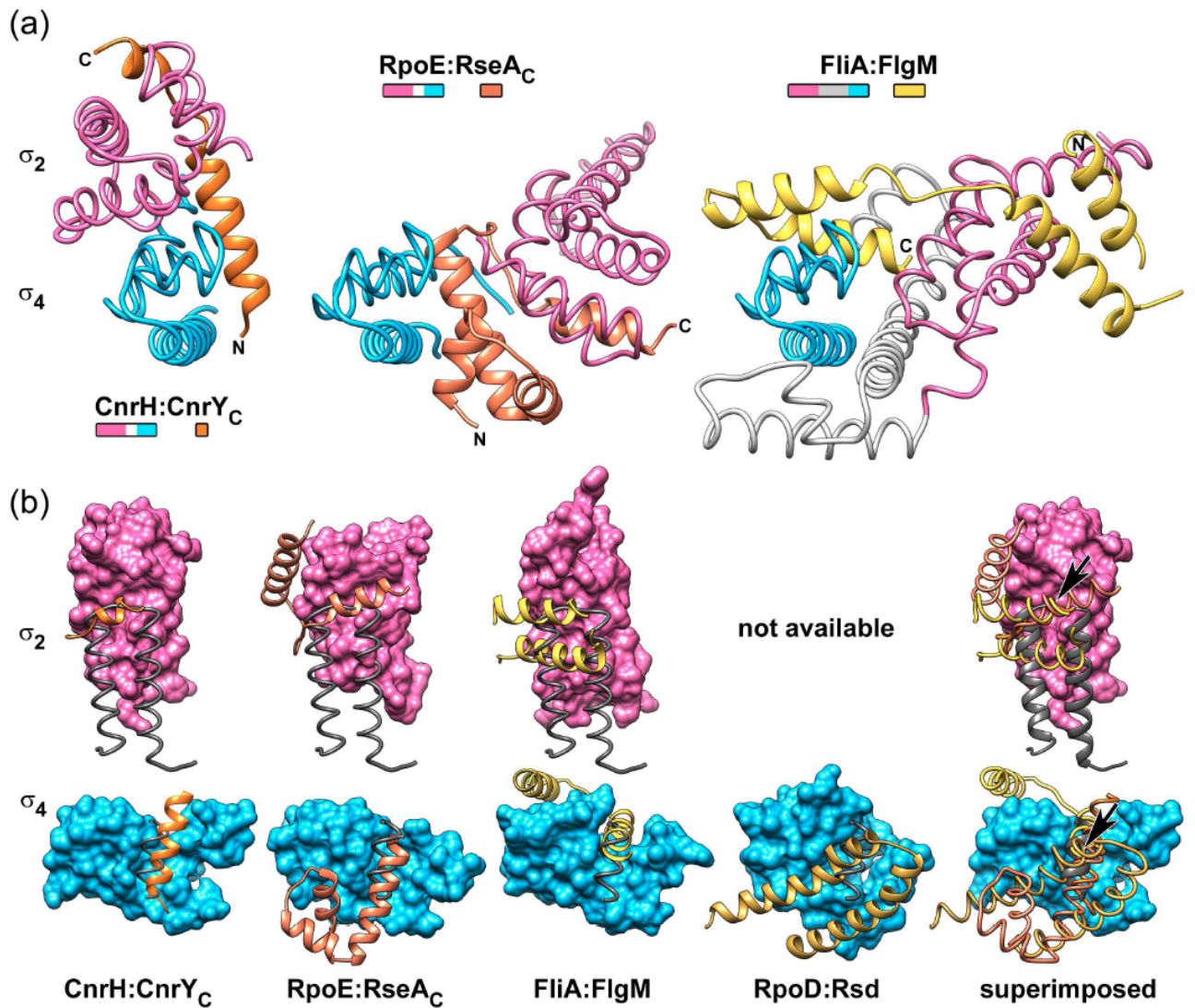
### A new structural class of anti- $\sigma^{\text{ECF}}$

The structural classification of anti- $\sigma^{\text{ECF}}$  was pioneered by Darst and colleagues following the discovery of the ASD fold, a four-helix motif, in *E. coli* RseA and *R. sphaeroides* ChrR<sup>11</sup>. Unlike RseA and ChrR, CnrY<sub>C</sub> and NepR display no sequence similarity to each other (Figure 4c). This and a contrasted topology around H1' and H2' (upstream a membrane-anchor in CnrY; in-between two disordered termini in NepR) makes the prediction of this motif outside CnrY and NepR-like proteins very tentative. The high structural similarity of CnrY<sub>C</sub> and NepR in complex may result from convergent selection *via* binding to similar targets: it defines a new structural class that we refer to as a class II ASD, in keeping with the ASD nomenclature suggested for anti- $\sigma^{\text{ECF}}$ <sup>11</sup>. While the distribution of the class II ASD is presently limited to CnrY and NepR proteins, the rather poor predictive power of the corresponding motif (see above) leaves this class open to additions in the future.

The class II ASD is considerably shorter than other structurally characterized anti- $\sigma$ s, e.g. RseA (70 residues), FlgM (80) and Rsd (150)<sup>32</sup>. The structure of CnrH:CnrY<sub>C</sub> shows how the minimal-size class II ASD achieves function: as CnrH domains show typical folds, inhibition proceeds from competition, not remodeling<sup>32</sup>. CnrY<sub>C</sub> stabilizes a conformation of CnrH that cannot bind the promoter -10 element and it occludes the RP core-binding determinants on  $\sigma_2$  and  $\sigma_4$ . Size-wise efficiency thus relies on binding the closed conformation of a  $\sigma^{\text{ECF}}$ , *i.e.* a  $\sigma$  factor of the simplest form (**Figure 9a**), and tolerance of promoter -35 element binding (**Figure 4d**).

### Closed and open conformations of $\sigma$ as alternative targets

CnrY<sub>C</sub> acts similarly as longer anti- $\sigma$ s<sup>32</sup>, which suggests that these are minimum requirements for  $\sigma$  factor inhibition (**Figure 9b**). These requirements are satisfied in a variety of ways but it is remarkable that CnrY, NepR and FlgM share similar properties: (i) the lack of a hydrophobic core in CnrY<sub>C</sub> suggests that the cytosolic domain of CnrY is intrinsically unfolded, like NepR and the C-terminal region of *S. typhimurium* FlgM<sup>16; 33</sup>; (ii) CnrY and FlgM bind a closed conformation of  $\sigma$  that restrains promoter binding<sup>34</sup>. This is in contrast with RseA or ChrR whose ASD is sandwiched between the two domains of the  $\sigma^{\text{ECF}}$ <sup>11</sup>, and with Rsd possibly doing the same<sup>35; 36; 37</sup>: these anti- $\sigma$ s compete directly for the DNA-binding and the core-binding determinants as well<sup>32</sup>. Thus, anti- $\sigma$ s may be parted in two groups: those with a folded anti- $\sigma$  domain that wedges between  $\sigma$  domains (RseA, Rsd) and those whose intrinsically unfolded anti- $\sigma$  domain embraces a closed conformation of  $\sigma$  (NepR, FlgM). The diversity of anti- $\sigma^{\text{ECF}}$  is such that they are found in both. While the correlation between intrinsic disorder in the anti- $\sigma$  and sequestration of a closed  $\sigma$  factor is only emerging, it is worth noting that folding-upon-binding can compensate for low-affinity (reversible) interactions in signaling events by promoting specificity and speed<sup>38; 39</sup>: these may contribute to efficient-yet-labile  $\sigma$  factor sequestration.



**Figure 9: Comparison of CnrH:CnrY<sub>C</sub> to relevant  $\sigma$ :anti- $\sigma$  complexes.** (a) CnrY<sub>C</sub> is a minimal size anti- $\sigma$ . Left: the L-shaped CnrY<sub>C</sub> sequesters a compact conformation of the  $\sigma^{\text{ECF}}$  CnrH. Middle: the three-helix bundle of *E. coli* RseA cytosolic domain (RseA<sub>C</sub>, a typical ASD) is wedged in-between  $\sigma_2$  and  $\sigma_4$  domains of the  $\sigma^{\text{ECF}}$  RpoE<sup>24</sup>. Right: *A. aeolicus* FlgM sequesters a compact conformation of FliA which, unlike  $\sigma^{\text{ECF}}$ , also possesses a  $\sigma_3$  domain<sup>34</sup>. To help with comparison, a similar orientation was devised for the domains  $\sigma_4$  (blue) of all complexes by means of structural superimposition. Colored boxes depicting protein domains were drawn to scale. (b) Anti- $\sigma$ s constantly occlude the RNA polymerase binding determinants of domains  $\sigma_2$  and  $\sigma_4$  of their respective  $\sigma$  factor. The domains  $\sigma_2$  (top) and  $\sigma_4$  (bottom) of *C. metallidurans* CH34 CnrH, *E. coli* RpoE, *A. aeolicus* FliA and *E. coli* RpoD<sup>36</sup> have been superimposed onto relevant SigA domains from the *T. thermophilus* holoenzyme<sup>19</sup>. The  $\beta'$  coiled-coil (top) and the  $\beta$ -flap tip helix (bottom) appear in grey while the anti- $\sigma$ s appear in shades of orange. Overall superimposition allowed pinpointing a focal point in anti- $\sigma$ s' superimposed paths onto  $\sigma_2$  and  $\sigma_4$  domains (arrowheads). See Table 2 for relevant rmsd's.

## The CnrYXH signaling pathway

Transmembrane signaling *via*  $\sigma^{\text{ECF}}$  factors has been shown to involve regulated anti- $\sigma$  proteolysis in all systems investigated to date<sup>40; 41; 42; 43</sup>. All but *B. subtilis* RsiV display the ASD<sup>10; 11</sup> and are no model for CnrY. On the other hand, similarity with NepR and FlgM supports new hypotheses regarding CnrYXH function: the cytosolic release of CnrH from CnrY may be the consequence of (i) a partner-mediated switch, (ii) a conformational change, (iii) post-translational modification, (iv) regulated proteolysis or a combination thereof. A partner-mediated switch like FlgM's relocation to the periplasm<sup>8</sup> is irrelevant because of CnrY membrane anchorage. A titration switch like PhyR-mediated regulation of NepR, is unlikely since all  $\sigma^{\text{ECF}}$  detected in CH34 genome are functional<sup>44; 45</sup> but the binding versatility of intrinsically unfolded proteins allows suggesting that CnrY might be titrated by a partner lacking  $\sigma^{\text{ECF}}$  homology<sup>38; 39</sup>. The structure of CnrH:CnrY<sub>C</sub> shows that CnrY<sub>C</sub> has no hydrophobic core to mediate a conformational change. It also shows that CnrH-bound CnrY<sub>C</sub> is accessible to the solvent from one end to the other, a property that is favorable to post-translational modification (two phosphorylation sites were predicted at Ser26 and Ser32<sup>46</sup>) or proteolysis. Although this was not investigated *per se*, the existence of specific factor(s) supporting Cnr operation in *C. metallidurans* CH34 is suggested by the observation that expressing the whole *cnr* determinant in *E. coli* failed to confer cobalt resistance<sup>47</sup> whereas CnrH is functional in *E. coli* (this study). Therefore the CnrYXH pathway may be incompletely described and examples from the literature suggest that additional factor(s) be a kinase and/or proteinases<sup>9; 48</sup>.

## Conclusion

The class II ASD is a structural module found in anti- $\sigma^{\text{ECF}}$  that consists in a stretch of 30 residues with a propensity to form two helices upon  $\sigma^{\text{ECF}}$  binding. No clear aminoacid sequence signature is expected for this module because of target diversity and relaxed tertiary structural constraints. Like class I ASD, class II ASD is present in proteins with different topologies and contributes to cytosolic or transmembrane signaling pathways. Thus the functional modularity stressed in anti- $\sigma^{\text{ECF}}$  with a class I ASD<sup>11</sup> also applies to anti- $\sigma^{\text{ECF}}$  with a class II ASD. The identification of class II ASD highlights essential schemes underlying anti- $\sigma$ -mediated inhibition of  $\sigma$  factors: (i) CnrY<sub>C</sub> minimal size makes sharper the observation that

all anti- $\sigma$ s structurally documented so far target the same binding sites; (ii) closed conformation of the  $\sigma$  factor and conformational disorder in the anti- $\sigma$  appear to correlate, suggesting that for class II ASD at least, conformational disorder plays a role in regulating  $\sigma$ :anti- $\sigma$  interaction; (iii) structural convergence of CnrH:CnrY<sub>C</sub> and PhyR:NepR questions the prevalence of proteolysis in transmembrane  $\sigma^{\text{ECF}}$  activation since the latter operates a reversible titration switch. These points shall be clarified by the characterization of select  $\sigma$ :anti- $\sigma$  complexes; they allow to suggest that although  $\sigma^{\text{ECF}}$  regulation is diverse, it is supported by a limited set of schemes.

## Materials and Methods

**Plasmid construction.** The expression plasmid pET-strepH-Yc was made in two steps. The DNA encoding the cytosolic portion of CnrY (CnrY<sub>C</sub>) was PCR-amplified from *C. metallidurans* CH34 total DNA and cloned into pET-Duet-1 (Novagen) between NdeI and XhoI restriction sites, while *cnrH* was PCR-amplified and cloned into pASK7+ (IBA). The open reading frame with a StrepTagII upstream of CnrH was then amplified from the recombinant pASK7+ and inserted between the XbaI and EcoRI restriction sites of the pET-Duet-1 *cnrY*<sub>C</sub> derivative.

The *in vivo* functional assay was based on two plasmids. The operon *cnrYXH* was PCR amplified and ligated into pTrc99a using the NcoI and SalI restriction sites. The original GTG initiation codon of *cnrH* was mutated to ATG in pTrc-*cnrYXH*, which resulted in higher CnrH production upon IPTG induction of *cnrYXH* expression, while no leaky expression was detected. The reporter construct pACYC-*cnrYp-gfp* was based on the vector pACYC-Duet-1 (Novagen) in which the gene encoding enhanced-GFP was inserted downstream the *cnrY* promoter (Grass and *al.*, 2000), such as to replace both T7 promoters by the DNA of interest in a tripartite ligation involving EcoNI and NdeI from the plasmid and PagI to bridge both inserts. Mutant constructs of pTrc-*cnrYXaH* were obtained by site-directed mutagenesis with the QuikChange™ protocol (Agilent). The *cnrYXH* operon of each mutant was fully sequenced.

**Protein production and purification.** CnrY<sub>C</sub> and strep-tagged CnrH were coproduced in Rosetta (DE3) cells freshly transformed with the plasmid pET-strepH-Yc. Cultures were grown to 0.5-0.6 optical density (600 nm) in LB and protein production was induced for



4 hours at 30°C with 0.4 mM IPTG. The cells were harvested, resuspended in buffer A (Tris 100 mM pH 8.0, NaCl 200 mM, KCl 100 mM, glycerol 5 %) and disrupted by sonication after supplementation with an antiprotease cocktail (Chymostatin 1  $\mu$ g/mL, Leupeptin 1  $\mu$ g/mL, Antipain 1  $\mu$ g/mL, Pepstatin 1  $\mu$ g/mL, Aprotinin 5  $\mu$ g/mL, Sigma-Aldrich). The CnrH:CnrY<sub>C</sub> complex was obtained pure from the clarified extract (30000 rpm, Ti70, Beckman) after a three-step purification procedure. First, CnrH was purified by gravity-flow StrepTactin affinity (10 mL, IBA). The column loaded with tagged CnrH was equilibrated in buffer B (Tris 20 mM pH 8.0, NaCl 40 mM, KCl 20 mM, glycerol 5 %) prior elution with 2.5 mM desthiobiotine in buffer B. The selected fractions were pooled, pH-adjusted to 7.0 with Hepes 1 M and KCl concentration was set to 1 M. The sample was loaded onto a 1 mL Phenyl-Sepharose Hi-Trap column (GE Healthcare) equilibrated in buffer C (Hepes 50 mM pH 7.0, KCl 1 M, DTT 2 mM). The strepCnrH:CnrY<sub>C</sub> complex was present in the flow-through, pure from contaminants and excess strepCnrH, and it was concentrated by ultrafiltration (5000 MW cut-off, Amicon, Millipore). The purification of strepCnrH:CnrY<sub>C</sub> was finalized by size-exclusion chromatography, yielding a single peak on a Superdex75 column (GE Healthcare) equilibrated in buffer D (Hepes 50 mM pH 8.0, NaCl 200 mM, KCl 20 mM, DTT 2 mM).

**Protein crystallization and structure determination.** Crystals of strepCnrH:CnrY<sub>C</sub> were grown by vapor diffusion at 293 K by mixing 1  $\mu$ L of protein solution at 15 mg/mL with 1  $\mu$ L of crystallization solution (2.0 to 2.2 M ammonium sulfate, 100 mM HEPES pH 7.5 and 2 % PEG 400). Crystals appeared about one month after set-up but were reproduced overnight with a one-month old protein sample kept at 4°C. Mass spectrometry analyses identified products matching the CnrH domains and a portion of CnrY<sub>C</sub>: peptides of 10785.14 Da, 9533.82 Da and 4065.94 Da molecular mass were respectively identified as strepCnrH<sub>14-108</sub> (CnrH residues -2 to 93 – calculated molecular mass: 10785.16 Da), strepCnrH<sub>113-206</sub> (CnrH residues 98 to 191 – 9533.83 Da) and CnrY<sub>C</sub> (residues 2 to 38 – 4066.01 Da). The inverse correlation between extent of degradation and delay in crystal nucleation suggested that proteolysis is a prerequisite for crystallogenesis. Crystals were soaked for 15 minutes in a solution containing 1.95 M ammonium sulfate, 40 mM HEPES pH 7.5 and 120 mM Gd:DO3A complex (Girard and *al.*, 2003) to get an anomalous derivative. Both native and derivatized crystals were cryocooled using paratone as cryoprotectant. Crystals belonged to the I222 spacegroup with one CnrH:CnrY<sub>C</sub> complex per asymmetric unit and a solvent

content of 25 %. Diffraction data were integrated with XDS<sup>49</sup> and integrated intensities were scaled and merged using SCALA and TRUNCATE from the CCP4 suite of programs<sup>50</sup>.

The structure was determined in two steps combining single-wavelength diffraction (SAD) phasing and molecular replacement. The initial structure was determined by SAD phasing using the Gd:DO3A complex. SAD data were collected on ID23-EH1 at ESRF (Grenoble) at the Gd<sup>3+</sup> fluorescence peak ( $L_{III}$  absorption edge). Gd positions were determined within the asymmetric unit using SHELXD<sup>51</sup>. Heavy atom refinement, phasing and density modification calculations using SHARP<sup>52</sup> yielded a 2.49 Å resolution interpretable electron density map (Figures of merit of 0.257 and 0.854 after SHARP and density modification, respectively). Automatic model building was performed with BUCCANEER<sup>53</sup> leading to an initial model consisting in 115 over the 240 expected residues. This initial model was then used as search model for molecular replacement with PHASER<sup>54</sup> using native data collected on the fixed-wavelength microfocus beamline ID23-EH2 at ESRF and leading to a unique solution (TFZ=19.3; LLG=420). Hendrickson-Lattman coefficients from PHASER were improved by solvent flattening and histogram matching using DM<sup>50</sup> prior to automatic model building with BUCCANEER. The resulting model (144 residues) was manually improved in COOT<sup>50</sup>, refined with PHENIX<sup>55</sup> and optimized through iterative rounds of refinement and model building. At the end stages of refinement, TLS was used with TLS-groups determined with the TLSMD server<sup>56</sup>. The final 1.75 Å resolution model consists of residues 5 to 91 and 123 to 189 of CnrH, and in residues 2 to 30 of CnrY – all residues are in allowed regions of the Ramachandran plot (99.3 % in preferred regions). Structure analysis and representation were performed with UCSF Chimera<sup>57</sup>.

### **Modeling and structure superimpositions.**

The CnrH  $\sigma$  factor in complex with CnrY<sub>C</sub> was compared to  $\sigma$  factors in *T. thermophilus* RP holoenzyme, *T. aquaticus* transcription initiation complex (TIC) or to  $\sigma$  factors in complex with their anti- $\sigma$ . A model of the CnrH-loaded holoenzyme was obtained by superimposing high resolution structures of *T. aquaticus* SigA  $\sigma_2$  and  $\sigma_4$  domains in complex with DNA onto the same domains within *T. aquaticus* TIC and by superimposing CnrH  $\sigma_2$  and  $\sigma_4$  domains on high resolution structures of SigA relevant domains. Indicators of structural superimpositions are provided in Table 2.

***In vivo* assay of CnrH inhibition by CnrY.** CnrY, CnrX and CnrH were coproduced in TOP10 *E. coli* carrying either wild-type or a mutant derivative of pTrc-*cnrYXH* (test plasmid)

and pACYC-*cnrYp-gfp* (reporter plasmid). The cells, grown on LB supplemented with ampicillin and chloramphenicol, were induced overnight at 30°C with 0.4 mM IPTG, then harvested and washed by resuspension in cold PBS three times. The washed suspension was diluted to 0.2 optical density (600 nm) and aliquots were assessed for CnrH activity and CnrYXH protein production as follows. CnrH activity: aliquots of 100  $\mu$ L were loaded on 96-well plaques-for-fluorimetry (Greiner) and GFP fluorescence (absorption at 490 nm, emission at 520 nm) was measured on a FluoStar Optima (BMGLabtech). All constructs were assessed in three independent experiments at least and, for each experiment, the fluorescence of the strain producing a FLAG derivative of CnrY was used to normalize the data. CnrYXH production: in parallel, OD-normalized aliquots of the washed cell suspension were analyzed by western blotting after SDS-PAGE. Because the *cnrXH* portion of the various *cnrYXH* constructs was never altered and a highly specific polyclonal antibody raised against CnrX periplasmic domain is available<sup>14</sup>, CnrX detection was used to monitor the possible polar effect of each *cnrY* mutation on protein production. In order to monitor the consistency of all samples in each experiment, the endogenous *E. coli* enolase was detected, using a polyclonal antibody raised against *P. aeruginosa* enolase (gift from A.-M. di Guilmi).

**Collection and analysis of CnrY-like sequences.** PSI-BLAST searches were carried against the database of non-redundant protein sequences accessible *via* the NCBI web site (<http://blast.ncbi.nlm.nih.gov/>) using the BLOSUM 45 similarity matrix and default parameters. Because the sequence of CnrY<sub>C</sub> returned only five homologs, we took advantage of the physical linkage with *cnrX*-like genes within *cnrYXH* operons to detect *cnrY*-like genes. Proteins homologous to *C. metallidurans* CH34 CnrX were identified *via* PSI-BLAST and the genetic context of the corresponding genes was examined through the KEGG (<http://www.genome.jp/kegg>), PATRIC (<http://www.patricbrc.org>) and GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) databases. Thirty-eight unique sequences translated from *cnrX*-neighboring genes shared the same topology as CnrY (~100 residues, single pass transmembrane protein). Multiple sequence alignments have been performed with MAFFT-L-INS-i<sup>58</sup> using the BLOSUM 45 similarity matrix. Minor editing and presentation were carried out with Jalview<sup>59</sup>. From this alignment, the most likely phylogenetic tree was inferred using PhyML 3.0<sup>60</sup> with default algorithms for tree topology searching, 1000 bootstrapped datasets and LG as the aminoacid substitution model.

## Accession Number

The PDB accession code for the coordinates and structure factors of the CnrH:CnrY<sub>C</sub> complex structure is 4BGS.

## Acknowledgements

This work is dedicated to the memory of R.K. We thank A.-M. di Guilmi for the anti-enolase antibody and J.-C. Fontecilla-Camps, A. Kolb and S. Lacour for helpful comments on the manuscript. We thank the IBS-PG people for fluorimeter access, J.-P. Andrieux for Edman sequencing, L. Signor for mass spectrometry analysis, the staff of the EMBL HTX platform and staffs of ESRF beamlines BM30a and ID23. We thank BRACCO Imaging Spa., Milan, for kindly providing the Gd:DO3A complex. This work used platforms of the Grenoble Partnership for Structural Biology (PSB) and Instruct centre (ISBG; UMS 3518) supported by FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) and was supported by the CNRS (PEPS program) and Région Rhône-Alpes (CIBLE program). Molecular graphics and analyses were performed with UCSF Chimera (NIGMS P41-GM103311).

## References

1. Mooney, R. A., Darst, S. A. & Landick, R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell* **20**, 335-45.
2. Campagne, S., Marsh, M. E., Capitani, G., Vorholt, J. A. & Allain, F. H. (2014). Structural basis for -10 promoter element melting by environmentally induced sigma factors. *Nat Struct Mol Biol*.
3. Lonetto, M., Gribskov, M. & Gross, C. A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* **174**, 3843-9.
4. Helmann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* **46**, 47-110.
5. Lonetto, M. A., Brown, K. L., Rudd, K. E. & Buttner, M. J. (1994). Analysis of the *Streptomyces coelicolor* sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *Proc Natl Acad Sci U S A* **91**, 7573-7.
6. Gruber, T. M. & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**, 441-66.

7. Murakami, K. S. & Darst, S. A. (2003). Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* **13**, 31-9.
8. Hughes, K. T. & Mathee, K. (1998). The anti-sigma factors. *Annu Rev Microbiol* **52**, 231-86.
9. Mascher, T. (2013). Signaling diversity and evolution of extracytoplasmic function (ECF) sigma factors. *Curr Opin Microbiol*.
10. Staroń, A., Sofia, H. J., Dietrich, S., Ulrich, L. E., Liesegang, H. & Mascher, T. (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* **74**, 557-81.
11. Campbell, E. A., Greenwell, R., Anthony, J. R., Wang, S., Lim, L., Das, K., Sofia, H. J., Donohue, T. J. & Darst, S. A. (2007). A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol Cell* **27**, 793-805.
12. Grass, G., Grosse, C. & Nies, D. H. (2000). Regulation of the *cnr* cobalt and nickel resistance determinant from *Ralstonia* sp. strain CH34. *J Bacteriol* **182**, 1390-8.
13. Tibazarwa, C., Wuertz, S., Mergeay, M., Wyns, L. & van Der Lelie, D. (2000). Regulation of the *cnr* cobalt and nickel resistance determinant of *Ralstonia eutropha* (*Alcaligenes eutrophus*) CH34. *J Bacteriol* **182**, 1399-409.
14. Trepreau, J., Girard, E., Maillard, A. P., de Rosny, E., Petit-Haertlein, I., Kahn, R. & Coves, J. (2011). Structural basis for metal sensing by CnrX. *J Mol Biol* **408**, 766-79.
15. Trepreau, J., Grosse, C., Mouesca, J. M., Sarret, G., Girard, E., Petit-Haertlein, I., Kuennemann, S., Desbourdes, C., de Rosny, E., Maillard, A. P., Nies, D. H. & Coves, J. (2013). Metal sensing and signal transduction by CnrX from *Cupriavidus metallidurans* CH34: role of the only methionine assessed by a functional, spectroscopic, and theoretical study. *Metallomics*.
16. Campagne, S., Damberger, F. F., Kaczmarczyk, A., Francez-Charlot, A., Allain, F. H. & Vorholt, J. A. (2012). Structural basis for sigma factor mimicry in the general stress response of Alphaproteobacteria. *Proc Natl Acad Sci U S A* **109**, E1405-14.
17. Herrou, J., Rotskoff, G., Luo, Y., Roux, B. & Crosson, S. (2012). Structural basis of a protein partner switch that regulates the general stress response of alpha-proteobacteria. *Proc Natl Acad Sci U S A* **109**, E1415-23.
18. Girard, E., Stelter, M., Vicat, J. & Kahn, R. (2003). A new class of lanthanide complexes to obtain high-phasing-power heavy-atom derivatives for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **59**, 1914-22.
19. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S. & Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712-9.
20. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L. & Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* **9**, 527-39.

21. Feklistov, A. & Darst, S. A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **147**, 1257-69.
22. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285-90.
23. Sharp, M. M., Chan, C. L., Lu, C. Z., Marr, M. T., Nechaev, S., Merritt, E. W., Severinov, K., Roberts, J. W. & Gross, C. A. (1999). The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized. *Genes Dev* **13**, 3015-26.
24. Campbell, E. A., Tupy, J. L., Gruber, T. M., Wang, S., Sharp, M. M., Gross, C. A. & Darst, S. A. (2003). Crystal structure of Escherichia coli sigmaE with the cytoplasmic domain of its anti-sigma RseA. *Mol Cell* **11**, 1067-78.
25. Arthur, T. M., Anthony, L. C. & Burgess, R. R. (2000). Mutational analysis of beta '260-309, a sigma 70 binding site located on Escherichia coli core RNA polymerase. *J Biol Chem* **275**, 23113-9.
26. Geszvain, K., Gruber, T. M., Mooney, R. A., Gross, C. A. & Landick, R. (2004). A hydrophobic patch on the flap-tip helix of E.coli RNA polymerase mediates sigma(70) region 4 function. *J Mol Biol* **343**, 569-87.
27. Pompidor, G., Maillard, A. P., Girard, E., Gambarelli, S., Kahn, R. & Coves, J. (2008). X-ray structure of the metal-sensor CnrX in both the apo- and copper-bound forms. *FEBS Lett* **582**, 3954-8.
28. Monchy, S., Benotmane, M. A., Janssen, P., Vallaeys, T., Taghavi, S., van der Lelie, D. & Mergeay, M. (2007). Plasmids pMOL28 and pMOL30 of Cupriavidus metallidurans are specialized in the maximal viable response to heavy metals. *J Bacteriol* **189**, 7417-25.
29. von Rozycki, T. & Nies, D. H. (2009). Cupriavidus metallidurans: evolution of a metal-resistant bacterium. *Antonie Van Leeuwenhoek* **96**, 115-39.
30. Francez-Charlot, A., Frunzke, J., Reichen, C., Ebneter, J. Z., Gourion, B. & Vorholt, J. A. (2009). Sigma factor mimicry involved in regulation of general stress response. *Proc Natl Acad Sci U S A* **106**, 3467-72.
31. Cole, C., Barber, J. D. & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197-201.
32. Campbell, E. A., Westblade, L. F. & Darst, S. A. (2008). Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr Opin Microbiol* **11**, 121-7.
33. Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T. & Dahlquist, F. W. (1997). The C-terminal half of the anti-sigma factor FlgM, becomes structured when bound to its target, sigma 28. *Nat Struct Biol* **4**, 285-91.
34. Sorenson, M. K., Ray, S. S. & Darst, S. A. (2004). Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. *Mol Cell* **14**, 127-38.

35. Westblade, L. F., Ilag, L. L., Powell, A. K., Kolb, A., Robinson, C. V. & Busby, S. J. (2004). Studies of the Escherichia coli Rsd-sigma70 complex. *J Mol Biol* **335**, 685-92.
36. Patikoglou, G. A., Westblade, L. F., Campbell, E. A., Lamour, V., Lane, W. J. & Darst, S. A. (2007). Crystal structure of the Escherichia coli regulator of sigma70, Rsd, in complex with sigma70 domain 4. *J Mol Biol* **372**, 649-59.
37. Yuan, A. H., Gregory, B. D., Sharp, J. S., McCleary, K. D., Dove, S. L. & Hochschild, A. (2008). Rsd family proteins make simultaneous interactions with regions 2 and 4 of the primary sigma factor. *Mol Microbiol* **70**, 1136-51.
38. Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208.
39. Uversky, V. N. & Dunker, A. K. (2010). Understanding protein non-folding. *Biochim Biophys Acta* **1804**, 1231-64.
40. Brooks, B. E. & Buchanan, S. K. (2008). Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Biochim Biophys Acta* **1778**, 1930-45.
41. Draper, R. C., Martin, L. W., Beare, P. A. & Lamont, I. L. (2011). Differential proteolysis of sigma regulators controls cell-surface signalling in Pseudomonas aeruginosa. *Mol Microbiol* **82**, 1444-53.
42. Hastie, J. L., Williams, K. B. & Ellermeier, C. D. (2013). The activity of sigmaV, an extracytoplasmic function sigma factor of Bacillus subtilis, is controlled by regulated proteolysis of the anti-sigma factor RsiV. *J Bacteriol* **195**, 3135-44.
43. Heinrich, J. & Wiegert, T. (2009). Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Res Microbiol* **160**, 696-703.
44. Grosse, C., Friedrich, S. & Nies, D. H. (2007). Contribution of extracytoplasmic function sigma factors to transition metal homeostasis in Cupriavidus metallidurans strain CH34. *J Mol Microbiol Biotechnol* **12**, 227-40.
45. Janssen, P. J., Van Houdt, R., Moors, H., Monsieurs, P., Morin, N., Michaux, A., Benotmane, M. A., Leys, N., Vallaeys, T., Lapidus, A., Monchy, S., Medigue, C., Taghavi, S., McCorkle, S., Dunn, J., van der Lelie, D. & Mergeay, M. (2010). The complete genome sequence of Cupriavidus metallidurans strain CH34, a master survivalist in harsh and anthropogenic environments. *PLoS One* **5**, e10433.
46. Miller, M. L., Soufi, B., Jers, C., Blom, N., Macek, B. & Mijakovic, I. (2009). NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* **9**, 116-25.
47. Nies, A., Nies, D. H. & Silver, S. (1989). Cloning and expression of plasmid genes encoding resistances to chromate and cobalt in Alcaligenes eutrophus. *J Bacteriol* **171**, 5065-70.
48. Barik, S., Sureka, K., Mukherjee, P., Basu, J. & Kundu, M. (2010). RseA, the SigE specific anti-sigma factor of Mycobacterium tuberculosis, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol Microbiol* **75**, 592-606.
49. Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32.

50. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42.
51. Sheldrick, G. M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* **66**, 479-85.
52. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215-30.
53. Cowtan, K. (2012). Completion of autobuilt protein models using a database of protein fragments. *Acta Crystallogr D Biol Crystallogr* **68**, 328-35.
54. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.
55. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J. & Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-9.
56. Painter, J. & Merritt, E. A. (2006). TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39**, 109-11.
57. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12.
58. Katoh, K. & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9**, 286-98.
59. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91.
60. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-21.
61. Gabdoulline, R. R., Wade, R. C. & Walther, D. (2003). MolSurfer: A macromolecular interface navigator. *Nucleic Acids Res* **31**, 3349-51.

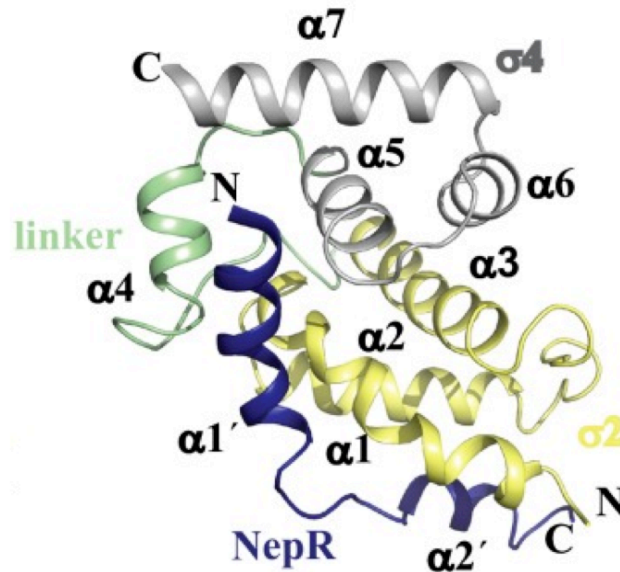


## B. CnrY est l'archétype d'une nouvelle classe d'ASD

Rappelons que pour être fonctionnel, le facteur sigma CnrH doit se fixer simultanément sur le cœur de l'ARN polymérase et sur l'ADN au niveau du promoteur. La structure du complexe a montré que CnrYc maintient, comme une sorte d'agrafe, les deux domaines  $\sigma 2$  et  $\sigma 4$  de CnrH. Les déterminants de l'interaction avec l'ADN se localisant à l'interface des deux domaines  $\sigma$ , ils ne sont plus accessibles à l'ADN. De plus, les déterminants de l'interaction avec l'ARNP sont masqués par CnrY. CnrY inhibe ainsi la fonction de CnrH.

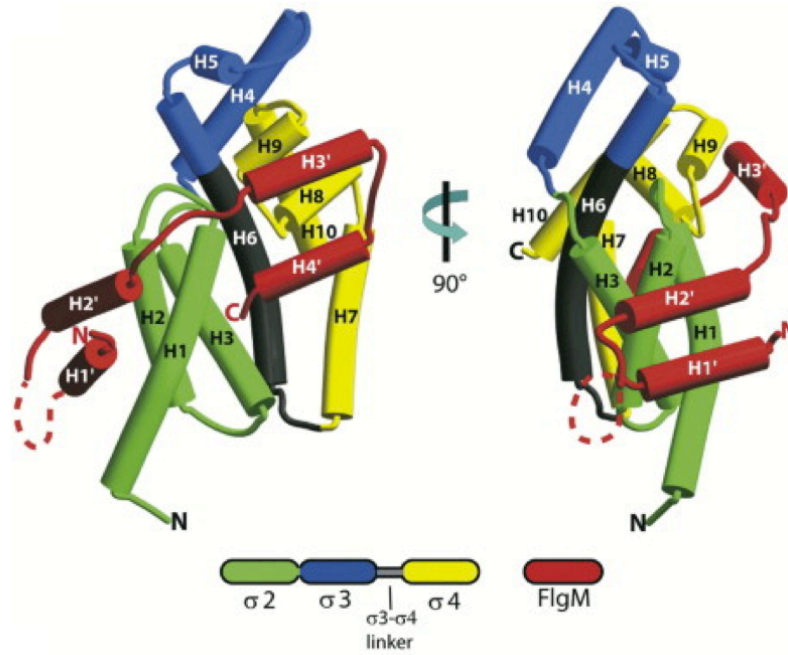
Par ailleurs, une étude fonctionnelle de détection *in vivo* de la libération de CnrH, basée sur la transcription d'un gène rapporteur sous contrôle de CnrH a permis d'identifier le dipeptide  $W_7L_8$  de CnrY comme déterminant majeur de l'interaction entre CnrY et CnrH et donc de l'inhibition de CnrH par CnrY.

De manière surprenante, la structure de CnrH:CnrYc est superposable à celle de PhyR<sub>SL</sub>:NepR (**Figure 10**), décrite dans l'introduction (section **I.C.3.a.**) (Campagne *et al.*, 2012; Herrou *et al.*, 2012), alors que NepR ne partage aucune similarité de séquence avec CnrY et que PhyR est similaire aux facteurs  $\sigma^{ECF}$  mais n'est pas fonctionnelle. Ces deux facteurs anti- $\sigma$  forment une nouvelle classe structurale de facteurs anti- $\sigma^{ECF}$ . Leur mode d'inhibition (inhibition du facteur  $\sigma$  par interaction avec le facteur anti- $\sigma$ ) est différent de celui décrit pour les autres facteurs anti- $\sigma$ . En effet, les facteurs anti- $\sigma$  de cette nouvelle classe maintiennent les facteurs  $\sigma$  dans une conformation fermée contrairement aux autres facteurs anti- $\sigma$  qui s'insèrent entre les deux domaines des facteurs  $\sigma$ .



**Figure 10** : Structure du complexe PhyR<sub>SL</sub>:NepR de *C. crescentus* déterminée par RMN. Le domaine SL (sigma like) de PhyR est composé de deux domaines  $\sigma$ ,  $\sigma 2$  et  $\sigma 4$  en jaune et gris, respectivement. La séquence de liaison entre ces deux domaines est en vert. Le facteur anti- $\sigma$  NepR est en bleu. Il maintient PhyR dans une conformation fermée (Campagne *et al.*, 2012).

Ceci n'est pas sans précédent chez les facteurs  $\sigma$ . Le facteur anti- $\sigma$  FlgM d'*Aquiflex aeolicus* régule le facteur  $\sigma^{28}$  qui appartient au groupe 3 des facteurs  $\sigma^{70}$ . Ce dernier est codé par le gène *fliA* et impliqué dans l'expression des gènes flagellaires tardifs (Kutsukake et Iino, 1994). Dans la structure cristallographique du complexe  $\sigma^{28}$ :FlgM (**Figure 11**), FlgM interagit avec les domaines  $\sigma 2$ ,  $\sigma 3$  et  $\sigma 4$  de  $\sigma^{28}$  (Sorenson *et al.*, 2004). Des expériences de pontage covalent réalisées en présence ou en absence de FlgM, ont montré que la conformation compacte observée pour FliA au sein du complexe FliA:FlgM préexiste en solution (Sorenson et Darst, 2006). FlgM régule l'activité de FliA en stabilisant cette conformation inactive. Dans cette conformation, les déterminants d'interaction des domaines de liaison à l'ADN ( $\sigma 2$  et  $\sigma 4$ ) ne sont plus accessibles à l'ADN. De plus, les déterminants de l'interaction avec l'ARNP sont sur la surface de  $\sigma$  mais masqués par FlgM. FlgM inhibe ainsi l'activité de  $\sigma^{28}$ .



**Figure 4** : Structure cristallographique du complexe  $\sigma^{28}$ :FlgM de *A. aeolicus*. Les hélices sont représentées sous forme de cylindres. Le code couleur est indiqué en dessous de la structure : les domaines  $\sigma 2$ ,  $\sigma 3$  et  $\sigma 4$  de  $\sigma^{28}$  sont en vert, bleu et jaune, respectivement. La séquence de liaison entre les domaines  $\sigma 3$  et  $\sigma 4$  est en noir. Le facteur anti- $\sigma$  FlgM est en rouge. La boucle non structurée entre les hélices H1' et H2' de FlgM est représentée par des pointillés (Sorenson *et al.*, 2004).

## II. Etude de l'interaction entre CnrX/NccX et CnrY/NccY dans la membrane

La conformation de CnrX est sensible à la présence de certains ions métalliques dans le périplasma. L'interaction de CnrX avec le nickel ou le cobalt déclenche une cascade de signalisation de nature inconnue qui fait passer un message à travers la membrane jusqu'au cytoplasme. Les protéines du complexe CnrYXH possédant une ancre TM sont CnrX et CnrY. L'étude du complexe formé par ces deux protéines transmembranaires a été engagée dans le but de découvrir l'organisation du lien physique que tendent CnrX et CnrY entre le périplasma et le cytoplasme et ainsi mieux comprendre le rôle des segments TM dans la transduction du signal.

L'un des objectifs atteints par l'équipe HMS avant mon arrivée fut d'obtenir la structure du senseur après solubilisation en détergent. Les premières tentatives de solubilisation de CnrX et NccX ont montré qu'il était possible de purifier NccX dès l'étape de solubilisation, contrairement à CnrX. De plus, les séquences et probablement les structures de CnrX et de NccX sont similaires (76% d'identité de séquence) ainsi que les signaux spectroscopiques en UV-visible (Trepreau *et al.*, 2011a). C'est donc le modèle Ncc de *Cupriavidus metallidurans* 31A que nous avons utilisé pour caractériser l'interaction entre NccX et NccY dans la membrane, à l'aide du test TOXCAT réalisé *in vivo*. Nous avons d'abord observé l'organisation oligomérique de NccX lorsqu'elle est produite seule dans la membrane, puis celle de NccY, avant de mettre en évidence l'interaction entre NccX et NccY lorsqu'elles sont coproduites.

### A. Homodimérisation de NccX

Le domaine périplasmique de la protéine CnrX (CnrXs) a été purifié et cristallisé sous forme dimérique (Pompidor *et al.*, 2008; Trepreau *et al.*, 2011b, 2014). Chaque protomère est formé par 4 hélices repliées en épingle à cheveux. Les deux protomères s'associent pour

former un dimère qui correspond à l'unité biologique. La description détaillée de la structure de CnrXs est présentée dans l'introduction (section **III.D.4.c.i.**). Cependant, cette structure ne présente que la partie périplasmique de la protéine. La détermination de la structure de NccX avant mon arrivée m'a permis d'étudier l'organisation d'une protéine complète homologue de CnrX. Ces travaux sont le fruit d'une collaboration. Isabelle Petit-Haertlein, membre de l'équipe HMS, a effectué la purification et la cristallisation de la protéine NccX. Eric Girard (groupe ELMA, IBS) a congelé les cristaux de protéines, collecté les données de diffraction aux rayons X et résolu la structure de NccX avec l'aide de Pierre Legrand (Synchrotron Soleil, Gif sur Yvette).

Dans l'article ci-après, la structure du dimère de NccX obtenue à 3,12 Å de résolution est comparée à un modèle de son homologue CnrX, basé sur la structure de CnrXs. Bien que la structure de chaque protomère soit similaire à celle de CnrXs, la structure du dimère de NccX diffère significativement du modèle. L'analyse de la structure obtenue indique que le détergent dodécylphosphocholine (DPC), utilisé pour solubiliser les membranes et maintenir la protéine en solution lors des étapes de purification et de cristallisation, a redistribué les interactions hydrophobes situées à l'interface des domaines périplasmiques des deux protomères, menant à une conformation aberrante du dimère. Cette structure artefactuelle n'est pas compatible avec l'insertion des segments TM dans la membrane.

Les données structurales ont été réconciliées avec le modèle physiologique du dimère de NccX en utilisant des simulations par dynamique moléculaire couplées à des expériences *in vivo* de TOXCAT. Les simulations utilisant la structure cristalline ont été réalisées par Serge Crouzy (CEA, Grenoble). Celles qui ont concernées l'interaction des segments transmembranaires isolés sont le fruit du travail de Norbert Garnier (Université d'Orléans). Ma contribution a consisté à interpréter les résultats de simulation, à mettre au point et à réaliser les tests d'oligomérisation *in vivo*.

Avant de présenter l'article, je vais détailler le fonctionnement du système TOXCAT, développé par Russ et Engelman (Russ et Engelman, 1999). Ce test est basé sur la dimérisation des domaines activateurs de la transcription de la protéine ToxR (Kolmar, 1995).

La protéine ToxR est une protéine bitopique d'environ 32 kDa (294 résidus) ancrée dans la membrane interne par un simple segment TM (**Figures 12 et 13**). Cette protéine est issue de la bactérie à Gram négatif *Vibrio cholerae* connue pour causer la maladie du choléra (Miller *et al.*, 1987). Son domaine périplasmique carboxyterminal jouerait le rôle de senseur en

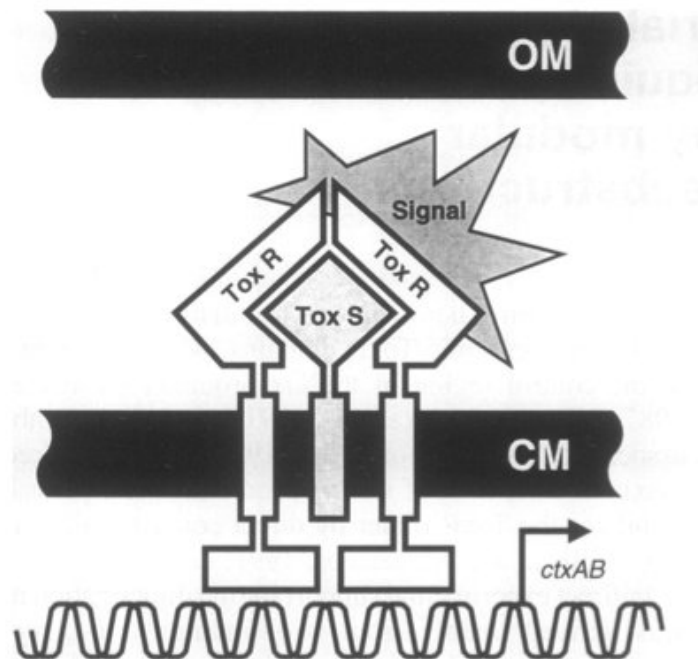
permettant la perception de signaux extracellulaire (température, pH, osmolarité) (Faruque *et al.*, 2000). En réponse à ces stimuli, les domaines périplasmiques de ToxR dimérisent. Cela conduit au rapprochement de deux segments TM et de deux domaines cytoplasmiques de ToxR (régulateur), ce qui permet d'activer la transcription des gènes de la toxine du choléra, *ctxA* et *ctxB* et d'autres gènes en relation avec la virulence placés sous le contrôle du promoteur *ctx* (DiRita, 1992) (**Figure 13**). D'autre part, il a été démontré que l'activité maximale de ToxR est conférée par une protéine régulatrice, ToxS (19 kDa) qui possède une ancre TM et un domaine périplasmique. Le domaine périplasmique de ToxS interagit avec le dimère de ToxR et le stabilise (DiRita et Mekalanos, 1991; Kolmar *et al.*, 1995) (**Figure 13**).

```

1           10           20           30           40           50           60
MFGLGHNSKE ISMSHIGTKF ILAEKFTFDP LSNTLIDKED SEEIIRLGSN ESRILWLLAQ
           70           80           90           100          110          120
RPNEVISRND LHDFVWREQG FEVDDSSLTQ AISTLRKMLK DSTKSPQYV KTVPKRGYQLI
           130          140          150          160          170          180
ARVETVEEEM ARESEAAHDI SQPESVNEYA ESSSVSSAT VVNTQPANV VTNKSAPNLG
           190          200          210          220          230          240
NRLLILIAVL LPLAVLLLTN PSQTSFKPLT VVDGVAVNMP NNHPDLSNWL PSIELCVKKY
           250          260          270          280          290          294
NEKHTGGLKP IEVIATGGQN NQLTLNYIHS PEVSGENITL RIVANPNDAI KVCE

```

**Figure 12** : Séquence protéique de ToxR. Son segment TM prédit est en gras et souligné (Kolmar *et al.*, 1995).

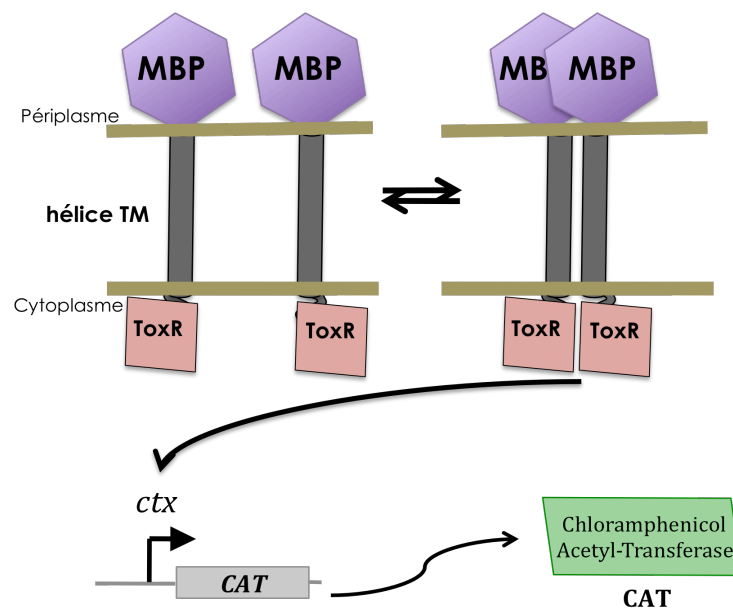


**Figure 13** : Activation de la transcription à partir du promoteur *ctx* dans *Vibrio cholerae*. L'interaction entre les domaines périplasmiques de ToxR et ToxS induit la dimérisation passive des domaines cytoplasmiques de ToxR qui permet l'activation de la transcription des gènes *ctxAB*. La flèche indique la direction de la transcription. OM et CM désignent la membrane externe et interne, respectivement (Kolmar *et al.*, 1995).

Le segment TM prédit de ToxR (résidus 83-198, **Figure 12**) est un signal suffisant pour assurer l'insertion dans la membrane mais il n'assure pas un ancrage membranaire stable. Pour que le segment TM de ToxR soit ancré dans la membrane efficacement, son extrémité C-terminale doit être fusionnée à une protéine périplasmique telle que la Maltose Binding Protein (MBP, produit du gène *malE*). Le segment TM de ToxR est donc nécessaire mais pas suffisant pour une localisation dans la membrane et n'a pas de fonction spéciale dans la transduction du signal si ce n'est que de faire le lien entre le périplasmique et le cytoplasme (Kolmar *et al.*, 1995). Kolmar et collaborateurs ont aussi montré que la substitution du segment TM de ToxR par un autre segment TM préserve la fonction de ToxR. De plus, la protéine chimérique est fonctionnelle dans un hôte hétérologue tel que *E. coli*, ce qui peut en faire un indicateur génétique pour les interactions protéine:protéine (Kolmar *et al.*, 1995).

Sur la base de ces travaux, le groupe de Donald Engelman a mis au point le test TOXCAT qui permet de mesurer les associations entre domaines TM dans la membrane biologique d'*E. coli* (Russ et Engelman, 1999). Le principe de ce test consiste à fusionner le

segment TM d'intérêt avec la MBP du côté périplasmique (extrémité C-terminale) et le domaine N-terminal de la protéine ToxR du côté cytoplasmique (extrémité N-terminale). La dimérisation de ToxR dépend de l'interaction des segments TM et induit l'activation de la transcription du gène codant l'enzyme Chloramphénicol AcétylTransférase (CAT), placé sous le contrôle du promoteur *ctx* (**Figure 14**).



**Figure 14 :** Représentation schématique du mode de fonctionnement du système TOXCAT (d'après Russ et Engelman, 1999).

Le fonctionnement de ce test a été validé (Russ et Engelman, 1999) en utilisant le modèle de dimérisation du segment TM de la glycopherine A (GpA WT) qui dimérise par son motif GxxxG. Alors que GpA WT sert de contrôle positif dans les études TOXCAT publiées, le mutant GpA G83I, dans lequel un résidu glycine du motif GxxxG a été remplacé par un résidu Ile et dont la dimérisation est lourdement affectée sert de contrôle négatif.

Dans le travail présenté ci-dessous, un modèle natif de la protéine NccX a été proposé à partir de données cristallographiques concernant NccX et CnrXs, de la reconstitution *in silico* de l'insertion de NccX dans une bicouche lipidique et d'expériences de TOXCAT qui ont permis d'identifier les déterminants de l'homodimérisation de NccX. Néanmoins, le message mis en avant concerne davantage la biochimie des protéines membranaires, en



particulier celles qui sont bitopiques, qui est le cas de NccX. En effet, les données obtenues ont permis de montrer que le détergent utilisé a induit un réarrangement conformationnel au sein du domaine périplasmique.

## B. Publication Ziani *et al.*, 2014

**The X-ray structure of NccX from *Cupriavidus metallidurans* 31A illustrates potential dangers of detergent solubilization when generating and interpreting crystal structures of membrane proteins.**

**Widade Ziani<sup>1°</sup>, Antoine P. Maillard<sup>1°</sup>, Isabelle Petit-Hartlein<sup>1</sup>, Norbert Garnier<sup>2</sup>,  
Serge Crouzy<sup>3</sup>, Eric Girard<sup>1</sup>, and Jacques Covès<sup>1\*</sup>**

<sup>1</sup> Institut de Biologie Structurale, UMR 5075 CNRS-CEA-Université Grenoble-Alpes, 71, Avenue des Martyrs, 38044 Grenoble Cedex 9, France

<sup>2</sup> Centre de Biophysique Moléculaire, CNRS UPR4301 affiliated to the University of Orléans, Rue Charles Sadron, 45071 Orléans cedex 2, France.

<sup>3</sup> Laboratoire de Chimie and Biologie des Métaux, UMR 5249, CEA-Grenoble, 17, Avenue des martyrs, 38054 Grenoble Cedex 09, France.

Running title: *Structure of detergent-solubilized NccX*

° WZ and APM contributed equally to this work

\*To whom correspondence should be addressed: Dr. Jacques Covès, Institut de Biologie Structurale, Campus EPN, CS 10090, 71, Avenue des Martyrs, 38044 Grenoble Cedex 9, France, Tel.: +33-4-57-42-85-16; Fax: 33-4-76-50-18-90, e-mail: jacques.coves@ibs.fr

<sup>1</sup> Institut de Biologie Structurale, UMR 5075 CNRS-CEA-Université Grenoble-Alpes, 71, Avenue des Martyrs, 38044 Grenoble Cedex 9, France

<sup>2</sup> Centre de Biophysique Moléculaire, CNRS UPR4301 affiliated to the University of Orléans, Rue Charles Sadron, 45071 Orléans cedex 2, France.

<sup>3</sup> Laboratoire de Chimie and Biologie des Métaux, UMR 5249, CEA-Grenoble, 17 rue des martyrs, 38054 Grenoble Cedex 09, France

<sup>4</sup> The abbreviations used are: CAT, chloramphenicol acetyl transferase; CG-MD, coarse-grained molecular dynamics; CMC, critical micellar concentration; DPC, dodecylphosphocholine; DTNB, dithiobis-nitrobenzoate; GpA; glycoporphine A; MBP, maltose binding protein; TM, transmembrane; POPC, 1-palmitoyl-2-oleoylphosphatidylcholine

**Keywords:** NccX; dodecylphosphocholine; X-ray structure; molecular dynamic simulations

Article publié dans *JBC* en 2014 (Ziani and *al.*, 2014)

**Background:** NccX is the membrane-anchored periplasmic metal-sensor of the NccYXH transmembrane signal transduction complex.

**Results:** Detergent-induced redistribution of the hydrophobic interactions led to a non-native X-ray structure.

**Conclusion:** Molecular dynamics simulations along with *in vivo* assays reconciled the structural data with a physiological model of NccX dimer.

**Significance:** Complementary investigations are needed to rationalize three-dimensional structures obtained in non-native conditions.

## ABSTRACT

We have determined the X-ray structure of NccX from *Cupriavidus metallidurans* 31A, a type II transmembrane metal sensor, at a resolution of 3.12 Å. This was achieved after solubilization by dodecylphosphocholine and purification in the presence of the detergent. NccX crystal structure did not match the model based on the extensively characterized periplasmic domain of its closest homologue CnrX. Instead, the periplasmic domains of NccX appeared collapsed against the hydrophobic transmembrane segments, leading to an aberrant topology, incompatible with membrane insertion. This was explained by a detergent-induced redistribution of the hydrophobic interactions among the transmembrane helices and a pair of hydrophobic patches keeping the periplasmic domains together in the native dimer. Molecular dynamics simulations performed with the full-length protein or with the transmembrane segments were used along with *in vivo* homodimerization assays (TOXCAT) to evaluate the determinants of the interactions between NccX protomers. Taken as a whole, computational and experimental results are in agreement with the structural model of CnrX where a cradle-shaped periplasmic metal-sensor domain is anchored into the inner membrane by two N-terminal helices. In addition, they show that the main determinant of NccX dimerization is the periplasmic soluble domain while the interaction between transmembrane segments is highly dynamic. The present work introduces a new crystal structure for a transmembrane protein and, in line with previous studies, substantiates the use of complementary theoretical and *in vivo* investigations to rationalize a three-dimensional structure obtained in non-native conditions.

*Cupriavidus metallidurans* is a Gram-negative bacillus best known for its outstanding ability to grow in the millimolar concentration range of many toxic heavy metals (1-4). To respond to environmental stress due to excess of these cations, *C. metallidurans* must sense extracellular changes and pass this information to the cytoplasm where transcription factors will be specifically activated to set up the appropriate efflux systems. The surplus of metal ions in the periplasm is commonly pumped back to the outside, by redundant RND (Resistance, Nodulation and cell Division) heavy-metal efflux pumps (3). For instance, pumping out the surplus of Co and Ni cations essentially involves the RND efflux pumps CnrCBA in *C. metallidurans* CH34 or NccCBA in *C. metallidurans* 31A (3,5). The transcription of the genes *cnrCBA* or *nccCBA* encoding the efflux pumps is under the control of the products of three additional genes, *cnrYXH* or *nccYXH*, also organized in operons (3,6,7). While the Cnr and Ncc systems parallel each other in many respects, Cnr from *C. metallidurans* CH34 has emerged as the experimental model for their characterization (6-10). Metal sensing relies on CnrX, a dimeric type II transmembrane (TM) protein that displays a 13 kDa-large C-terminal metal-sensor domain in the periplasm (9-12). Upon binding of Ni or Co to the sensor domain of CnrX, CnrY-mediated inhibition of the extra-cytoplasmic function sigma factor CnrH (13) is relieved, and CnrH enables the RNA-polymerase to transcribe *cnrCBA* resistance genes (6,7).

To characterize the structural basis of metal sensing by CnrX, we have previously produced a soluble form of the periplasmic domain spanning residues 31-148. This soluble metal-sensor domain is referred to as CnrX<sub>s</sub> in the following. The high-resolution structures of CnrXs under the Ni-, Co-, and Zn-bound forms as well as in the *apo*-form (10, 12) have shown that CnrXs is a cradle-shaped dimer that contains one physiologically relevant metal-binding site per monomer. They also pointed out the crucial role of the only methionine (Met123) (10,12) in metal selectivity and affinity. M123 is a key player of the allosteric switch corresponding to the first step of signal transduction (10, 12). Our understanding of subsequent signal propagation is limited by the lack of information on the TM domain of the metal sensor and on its relation with the bitopic protein CnrY that couples metal binding to CnrX in the periplasm to CnrH release in the cytoplasm. However, it has not been possible to purify CnrX as yet, in spite of very good overproduction (10). In contrast with CnrX, we have successfully solubilized and purified NccX (10). NccX displays 76% overall sequence identity with CnrX, with 80% identity between the soluble domains and 63% sequence

identity between the TM anchors (**Figure 15**). A detailed analysis of the spectroscopic data obtained with the Co-bound forms of CnrXs or NccX led to the conclusion that CnrXs is a relevant model of the NccX metal-sensor domain (10,11). We have thus used NccX for a structural investigation of the full-length sensor protein of the CnrYXH/NccYXH complex. In the native state, CnrX/NccX was hypothesized to be flexibly anchored to the inner membrane with the two TM segments running perpendicular with respect to the periplasmic domain (Fig. 2, panel A). In this paper we present the X-ray structure at a resolution of 3.12 Å of NccX purified and crystallized in the presence of foscholine-12 (dodecylphosphocholine or DPC in the following). The structure of NccX differs markedly from the prediction. The structural data were reconciled with a physiological model of NccX dimer using molecular dynamics (MD) simulations and *in vivo* dimerization (TOXCAT) assays (14) as complementary approaches. It appeared that the detergent not only solubilized the membrane bilayer but also destabilized the hydrophobic dimer interface, thus leading to the redistribution of the hydrophobic interactions that maintained the native structure of NccX. The two main conclusions of this work are that first, we present one of the rare complete structures of a bitopic, type II membrane protein and second, we illustrate the need to use complementary approaches to extract meaningful biological information from a potentially artifactual crystal structure.

## EXPERIMENTAL PROCEDURES

*Cloning, expression and purification of full-length NccX* - These steps were already described (10). Briefly, NccX was overproduced in *Escherichia coli* cells and the membrane fraction was recovered by centrifugation. For solubilization of NccX, the membrane preparation was 10-fold diluted in a buffer composed of 50 mM Hepes pH 7.5, 500 mM NaCl, 10 mM EDTA, 0.1% w/v (2 CMC) DPC (Anagrade from Anatrace) prior to Dounce homogenization and passage through a microfluidizer M110-P (Microfluidics). After an incubation of 3 hours on ice, the solubilized material was purified in two chromatographic steps consisting of a gel filtration on Superdex-200 (GE-Healthcare) followed by an anion exchange on a Q Hi-Trap column (GE Healthcare). The fractions considered electrophoretically pure on SDS-PAGE were pooled and NccX was concentrated at 13 mg.mL<sup>-1</sup> in the following buffer: 50 mM Hepes pH 7.5, 100 mM NaCl, 0.1% w/v DPC.

After concentration the protein was aliquoted and stored at  $-80^{\circ}\text{C}$  for further use. Protein concentrations were determined using the Bradford protein assay with bovine serum albumin as standard (15).

*Crystallization, data collection, structure determination and refinement* - Crystals were obtained by mixing  $1\ \mu\text{L}$  of NccX at  $13\ \text{mg/mL}$  with  $1\ \mu\text{L}$  of reservoir solution consisting in 14% PEG4000, 0.1 M Na-acetate pH 4.5 and 0.1 M Zn-acetate. Crystals grew after 2 months at  $20^{\circ}\text{C}$ . Prior to data collection, crystals were cryocooled using paratone-N as cryoprotectant.

Diffraction data were collected to a resolution of  $3.12\ \text{\AA}$  (Table 1) on the Proxima1 beamline at SOLEIL (Saint Aubin, France) taking advantage of the Pilatus-6M detector. The beam energy was set to the zinc K-absorption edge. Data integration and scaling were done using the XDS package (16). The structure of NccX was determined by the single-wavelength anomalous diffraction (SAD) method. The SHELXD program (17) was used to locate the zinc substructure. Initial phases were obtained with PHASER (18) and improved with PARROT (19) by density modification and by taking advantage of the two-fold non-crystallographic symmetry (NCS).

Using the phased translation function of the MOLREP program (20) and using one chain of the homologous structure, CnrXs (PDB entry: 2y3d) (10) as template, two molecular replacement solutions were obtained. From that initial model and using SAD experimental electron density maps improved by density modification, iterative cycles of manual model building with Coot and refinement with autoBuster-TNT (21, 22) were carried out. Temperature factors were refined with the translation-libration-screw (TLS) approach with a single TLS group for the NCS dimer. Zinc atom positions were cross-checked by an anomalous Fourier synthesis using the final model. Coordinates and structure factors for the structure of NccX have been deposited in the Protein Data Bank with accession number as 4clv. Structure figures were generated with Chimera (<http://www.cgl.ucsf.edu/chimera>).

*Model building and simulation parameters for NccX* - Two models of NccX were constructed. The first one is based on the structure of DPC-solubilized NccX (this work), and was referred to as NccX<sub>DPC</sub>. The X-ray structure of NccX was oriented with VMD (25) so that its principal axes are superimposed with the coordinate system. All missing hydrogens were built with CHARMM and the CHARMM19 force field (26,27). The structure was energy-minimized down to a gradient of  $0.1\ \text{kcal/mol/\AA}$  subject to harmonic restraints of 5

kcal/mol/Å on heavy atoms. Simulations from NccX<sub>DPC</sub> include residues Thr6 to Leu143 for protomer 1 and Thr6 to Lys144 for protomer 2. The second model, called NccX<sub>pred</sub>, was conceived such as to represent the structure predicted for NccX/CnrX proteins, based on the characterization of the periplasmic domain of CnrX and was used for control experiments. We started from the X-ray structure of the soluble sensor domain of CnrX in the Ni-bound form (PDB entry: 2y39) defined from Gly39 to Gln148 (10). The CnrXs dimer was built using the symmetry operations given in the PDB file. Thanks to the high sequence similarity between NccX and CnrX (**Figure 15**), we could build a model of the soluble domain of NccX by mapping the residues between the two proteins one by one and determining the missing coordinates for the soluble part of NccX (Thr39 to Gln148) from internal coordinates with CHARMM. In the PDB entry 2y39, the coordinates of the N-terminal residues of CnrX, from Met1 to His38, are absent (**Figure 16A**). To obtain NccX<sub>pred</sub>, these N-terminal missing residues, including the membrane-anchoring sequence, were initially built from internal coordinates only. To have a valid structure for the amino-acid residues tethering the TM segment to the soluble part of the protein, we fitted the backbone atom coordinates of Pro37 to Leu41 in the structure of NccX<sub>DPC</sub> to those calculated for NccX<sub>pred</sub>. Then after applying the same coordinate transformation to all atoms of NccX<sub>DPC</sub> from Thr6 to Glu36, we inserted these coordinates in the initial structure of NccX<sub>pred</sub>. Missing residues Met1 to Arg5 and all missing hydrogens were built with CHARMM so that simulations from NccX<sub>pred</sub> included the full-length sequence for both protomers, *i.e.* residues Met1 to Gln148. The structure was energy-minimized as described for NccX<sub>DPC</sub>. Both all-atom systems, NccX<sub>DPC</sub> and NccX<sub>pred</sub> were coarse-grained (CG) (*i.e.* each amino acid was simplified to a couple of interacting spheres) with a script developed for GROMACS (28) before running simulations.

The CG NccX<sub>DPC</sub> protein contained 296 particles for chain A and 299 particles for chain B while the CG NccX<sub>pred</sub> protein contained 638 particles. The structures were embedded either in CG POPC (1-palmitoyl-2-oleoylphosphatidylcholine) or in CG DPC, and solvated with CG water molecules. The protein bears a total charge of -13, so that 13 Na<sup>+</sup> were added to the system to ensure electric neutrality. Coarse-grained molecular dynamics (CG-MD) simulations were performed using the GROMACS software package (version 4.5.4) (29,30). The proteins, lipids, ions and water molecules were described with the MARTINI force field version 2.1 (31,32). The system was simulated with periodic boundary conditions using a time step of 20 fs. During the simulations, the pressure was maintained at 1 atm using a Berendsen barostat with a coupling constant of 2 ps and a compressibility of 3.10<sup>-4</sup> bar<sup>-1</sup> while the

temperature was maintained at 323.15 K with a Berendsen thermostat using a coupling constant of 1 ps. The pressure coupling was either isotropic, leaving all dimensions of the simulation box free to vary or semi-isotropic in which case only the Z-dimension of the box could vary.

During the coarse-graining process, helical structures are maintained as restraints in the GROMACS MD runs so that varying this number of restraints may have an incidence on the structure finally reached. As the helix H1 in NccX<sub>DPC</sub> propagated through the flexible linker in-between the TM and periplasmic domains, the secondary structure status of residues 32 to 44 in the model has been relaxed to allow flexibility to this sequence. Conditions for the simulations are given in the legend of the corresponding figures.

All atom coordinates corresponding to the last frame of the simulations using NccX<sub>DPC</sub> in DPC or POPC were converted to united-atom resolution using the tool "backward" (33), after adding map files for DPC. The back-mapping was performed using a wrapper which relaxes the resulting structure through energy minimization and short MD-based relaxation. The size of the united-atom systems increased by a factor varying from 8 to 12 with respect to the CG systems reaching 326,184 and 89,102 atoms for the simulations in DPC and POPC, respectively. Two nanoseconds MD simulations in the NVT (constant Number of particles, Volume and Temperature) ensemble followed by 5 to 100 ns MD simulations at constant pressure and temperature (NPT) were subsequently run with a time step of 2 fs.

*Self-assembly simulations of the NccX TM segments* - To study the self association of NccX TM segments, these were coarse-grained as described previously for the whole protein. The sequence spanning Thr6 to Ser40, thus encompassing the predicted TM segment (Leu9 to Tyr29) of NccX, was used for simulations and called NccX-TMsim (Fig. 1). The structure of NccX in DPC (PDB entry: 4clv, this work) provided the initial atomistic structure of TMsim monomers. All protein, lipids and solvent CG particles were described with the MARTINI force field. Monomers of CG NccX-TMsim were inserted parallel to each other in a preformed POPC bilayer containing 492 lipids and 6,000 water particles, at a distance of 6 nm from each other. The X, Y, Z dimensions of the bilayer were 13.6, 12.2, 8.7 nm. The Z axis represents the normal to the bilayer plane (X, Y). Prior to the insertion, the NccX-TMsim monomers were aligned along the Z axis and were randomly rotated about the Z axis. The ionic strength of the system was set with 100 mM NaCl, which, in addition, contributed to neutralize the charges of the NccX-TMsim monomers.



The system was energy-minimized for 10,000 steps using the steepest descent integrator. Then MD simulations were performed at T=300 K with a coupling time of 1 ps using a v-rescale algorithm. The pressure was coupled using a semi-isotropic scheme at P=1 bar with a coupling time of 5 ps and a compressibility of  $3 \times 10^{-4} \text{ bar}^{-1}$ . A time step of 40 fs was used during the production period. Varying the initial orientation of the two monomers relative to each other as well as initial velocity, two distinct simulations of 60  $\mu\text{s}$  were performed in POPC, resulting in a total simulation time of 120  $\mu\text{s}$ .

*TOXCAT assays* - The homodimerization of wild-type and mutant NccX, full-length or truncated, was assessed using the TOXCAT assay (14). The expression vectors pccKAN, pccGpA-wt, pccGpA-G83I, and the *malE* *E. coli* strain MM39 were kindly provided by Donald Engelman (Yale University, New Haven, CT). As in any TOXCAT experiment, the expression of the constructs produced chimeras where the peptide or protein of interest was fused N-terminally with the dimerization-dependent transcriptional activator domain of ToxR from *Vibrio cholerae* and with the *E. coli* maltose binding protein (MBP) at its C-terminus. We inserted the NccX-TMsim segment used for simulation experiments and the full-length sequence of NccX (NccX-FL) between ToxR and MBP. The GGII double mutation in NccX-TMsim or NccX-FL was generated using the Quik-Change mutagenesis kit (Agilent). The sequence of all the constructs was confirmed by DNA sequencing. The resulting plasmids were transformed into *E. coli* MM39 for further use. Oligomerization of the inserted sequences turn on the transcription of the chloramphenicol-acetyl-transferase (CAT) reporter gene proportionally to the strength of the interaction between TM segments. Insertion efficiency and proper orientation were tested on the basis of the ability to grow on M9 minimal medium plates with 0.4% maltose as the only carbon source. Membrane localization of NccX-FL was checked by cell fractionation and subsequent detection of the construct in the membrane fraction by Western blot with either CnrXs antibodies (10) or MBP antibodies (anti-MBP monoclonal antibody, HRP conjugated, New England Biolabs). Briefly, the MM39 *E. coli* cells bearing either the ToxR-NccX-FL-MBP or the ToxR-GpA-MBP constructs and used for the TOXCAT assays were lysed by sonic oscillation in 25 mM Tris-HCl pH 8.0, 2 mM EDTA. The resulting lysates were centrifuged at 4,500 g for 20 min (Thermo Scientific). The pellet consisting of cell debris, unbroken cells and inclusion bodies was washed in sonication buffer containing 0.5% Triton X-100 and centrifuged again at 4,500 g for 20 min to obtain the inclusion bodies. The supernatant of the first centrifugation was subjected to a centrifugation of 118,000 g for 60 min (Beckmann) to pellet the membranes. The resulting

supernatant was used for electrophoretic and Western blot analysis while the pellet was suspended in sonication buffer and centrifuge again the same way. The new supernatant and the membrane fraction were used for further analysis. CAT activity was assayed spectrophotometrically at room temperature by measuring the rates of chloramphenicol acetylation in the presence of dithiobis-nitrobenzoate (DTNB) and acetyl-coenzymeA (34). The rate of CoA formation was monitored by recording the linear increase of the absorbance at 412 nm due to its reaction with DTNB and thionitrobenzoate dianion release. The background rate of acetyl-CoA hydrolysis in the absence of chloramphenicol was recorded and subtracted before calculating the specific activities, expressed as  $\Delta A_{412} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  of protein from total extracts. Each experiment was performed at least in triplicate. The wild type and mutant forms of the glycoporphin A (GpA) TM domain, GpA-wt and GpA-G83I, were used as positive and negative controls in the TOXCAT assays, respectively. Normalization for the relative expression level of each construct was carried out using Western blotting with MBP antibodies and densitometry with ImageJ (<http://imagej.nih.gov/ij>).

## RESULTS

*X-ray structure of NccX* - The X-ray structure of DPC-solubilized full-length NccX, crystallized in the presence of an excess of zinc, was determined at a resolution of 3.12 Å (**Figure 16B**). It shows a dimer with an all- $\alpha$ -helix secondary structure, like CnrXs (**Figure 16A**) (9, 10,12). Only few residues are missing in the 148-residue long NccX since each protomer spans residues Ser4 to Arg145 (sequence shown in **Figure 15**). Both NccX and CnrXs are characterized by a long hairpin with the NccX helices H2 and H3 coinciding with CnrXs helices H2' and H3' (**Figure 16C**). However, while CnrXs protomer is best described as a hooked hairpin with the short helices H1' and H4' (N- and C-terminal helices, respectively) folded back on the H2'-H3' hairpin, full-length NccX is a three-helix bundle with the short C-terminal helix H4 sticking out at an angle of about 50° to avoid a steric clash with the H3 helix of the opposite protomer (**Figure 16**). NccX helix H1 (residues 3 to 48) is a long helix whose C-terminal end corresponds to CnrXs helix H1' and that extends upstream

to include the TM domain and the linker in-between. Noteworthy, a kink upstream Pro37 physically delineates the C-terminal part of H1 that coincides with CnrXs H1' (residues 37 to 48). The NccX protomers A and B actually differ slightly in the angle of this kink.

The NccX structure was stabilized by numerous zinc ions, consistent with the presence of 100 mM Zn-acetate in the crystallization conditions of NccX. The anomalous diffraction properties of Zn(II) allowed precisely locating ten Zn ions in the asymmetric unit. Two Zn(II) ions were bound by each NccX protomer with the same trigonal bipyramidal coordination geometry at the same metal-binding sites (**Figure 16C and D**) than those were argued to be an artifact in protomer A of Zn-bound CnrXs (Zn<sub>A</sub>-CnrXs) (10). Accordingly, NccX protomers were best superimposable to Zn<sub>A</sub>-CnrXs (rmsd = 0.93 Å for 78 Cα-atom pairs) which is loaded with two Zn(II) ions (**Figure 16C**), than to Zn<sub>B</sub>-CnrXs (rmsd = 0.99 Å for 48 Cα-atom pairs) which is loaded with a single Zn(II) ion. Extra Zn(II) ions were also located outside the metal binding site (**Figure 16B**), at the interface between two monomers, either at the dimer interface or at crystal contacts. Two Zn ions of particular importance for the conformation of the NccX dimer link a nitrogen of His32 and Glu36 from H1 of one protomer with an oxygen of Glu106 from H3 of the opposite protomer. The coordination sphere of these extra Zn(II) ions is supposed to be completed with counter-anions or ligands from the solvent, undefined at this resolution.

```

1          10          20          30          40          50
iiiiiiiIMM MMMMMMMMMM MMMMMMMMMM oooooooooo oooooooooo
NccX  MMKSRTFRLS VSTLVGALVG VLMAIVGVYV THSTEEPHTS LHEMLHDAVP
CnrX  MMKSRTRRLS LSTLFGALLG VSVAAAWLYY SHRNEAGHGD LHEILHEAVP

          60          70          80          90          100
oooooooooooo ooooooooooooo ooooooooooooo ooooooooooooo ooooooooooooo
NccX  LDSNEREILE LKEEFTARR REIESRLRAA NGKLAESIAK NPQWSPEVEE
CnrX  LDANEREILE LKEDAFARRR REIETRLRAA NGKLADAIK  NPAWSPEVEA

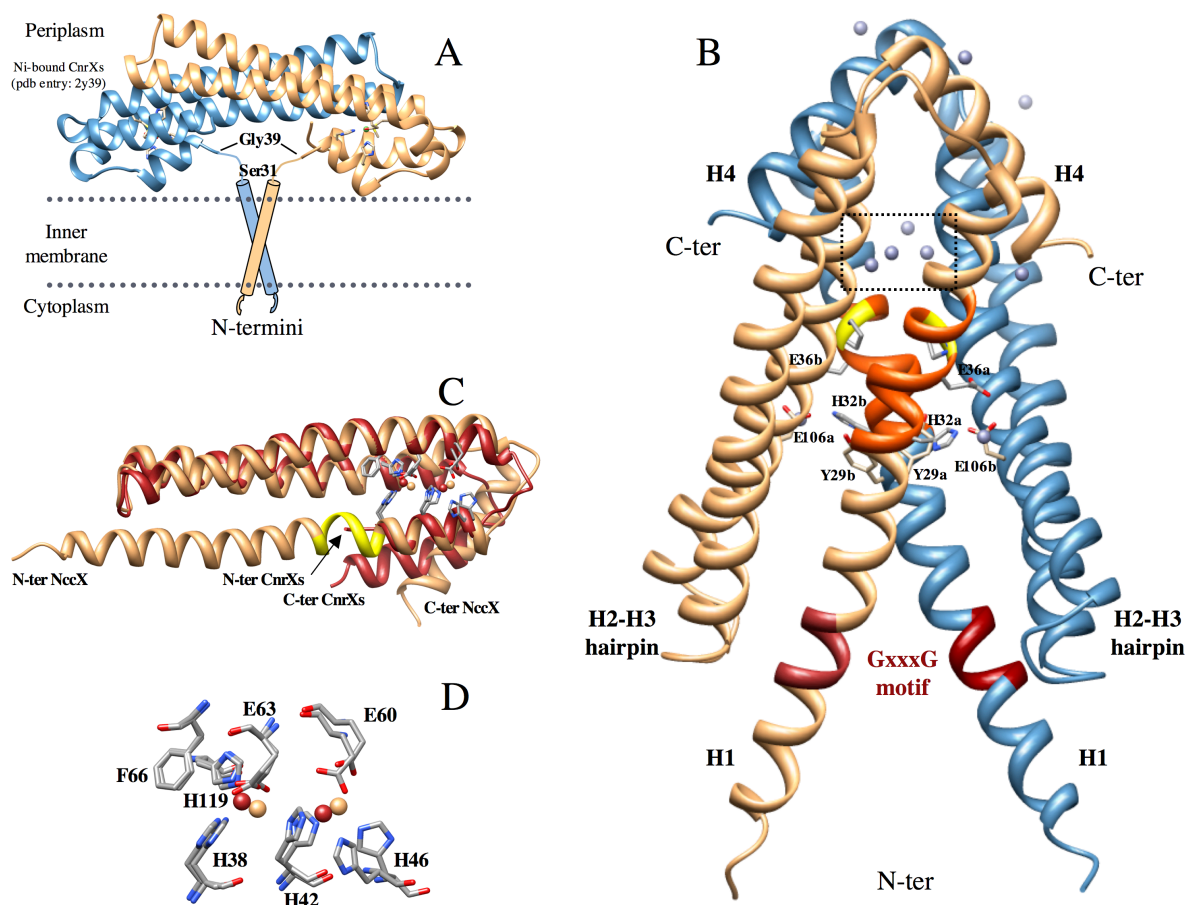
          110         120         130         140         148
oooooooooooo ooooooooooooo ooooooooooooo ooooooooooooo oooooooooo
NccX  ATREVERAAA DLQRATLVHV FEMRAGLKPE HRAAYDRVLV DALKRGSQ
CnrX  ATQEVERAAG DLQRATLVHV FEMRAGLKPE HRPAYDRVLI DALRRGSQ

```

**Figure 15:** CnrX and NccX sequence alignment, and topology prediction generated by the OCTOPUS program (available online at <http://octopus.cbr.su.se/>). The differences are highlighted in grey. The sequence of NccX-TMsim is underlined and includes the transmembrane segment shown in bold. Letter code: i, inside (cytoplasm); M, membrane; o, outside (periplasm)

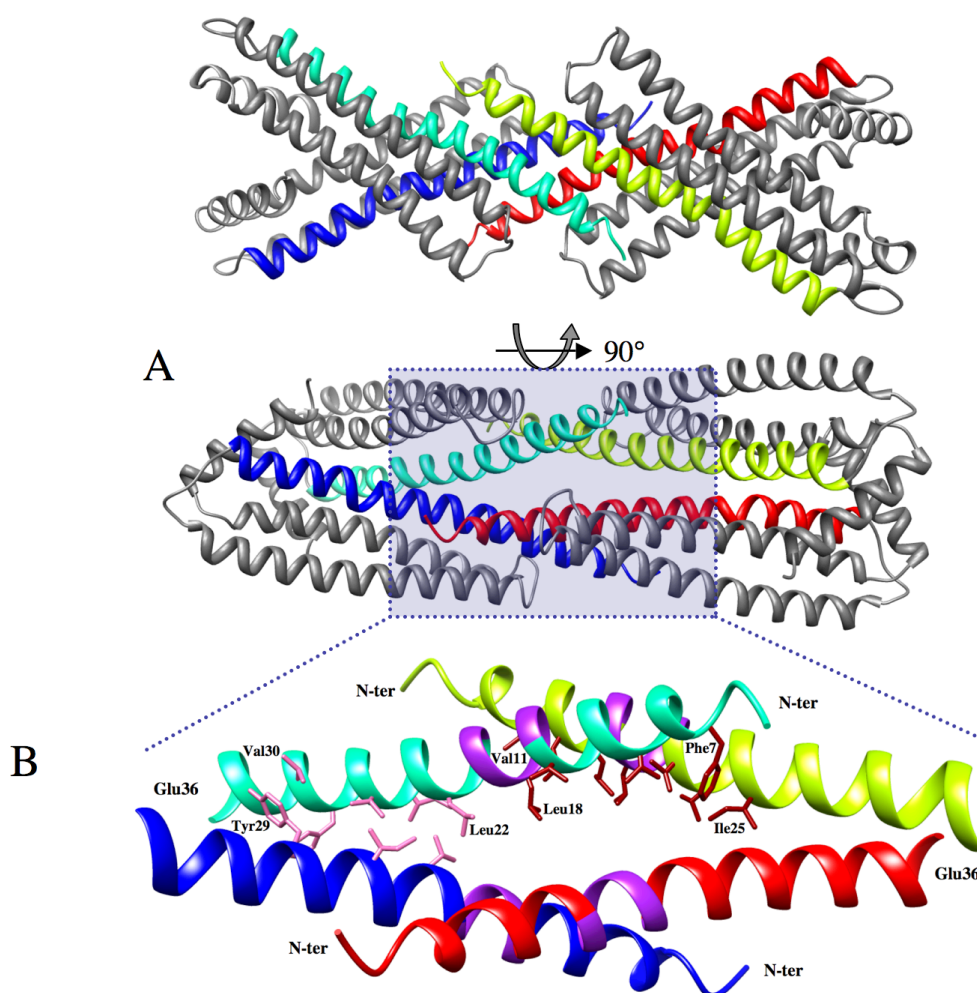
The TM domains cross each other near their Tyr29 residue at an angle of about 55° (**Figure 16B**), which brings both N-terminal Ser3 approximately 34 Å apart. Moreover, the putative GxxxG dimerization motifs in the middle of the TM segments are distant by about 18 Å whereas they are supposed to interact *in vivo*. Each GxxxG motif makes contacts with the C-terminal tip of H2 helix of the same protomer on one side, and with the TM domain of a symmetry-related dimer in the crystal on the other side (**Figure 17**). In this way, two NccX dimers interlock by the fork formed by their TM domains, which act as hydrophobic clamps.

This crystal contact is contributed by the N-terminal halves of the TM domains that run antiparallel to each other (from Phe7 to Leu18). Nevertheless parallel docking of H1 helices is still visible in the C-terminal half of the TM (from Leu22 to Tyr29), a portion that is delimited in-between the kink at Pro37 and a second kink, between Val21 and Leu22. The packing of H1 helices at the C-terminal half of the TM is likely a remnant of the physiological arrangement, as suggested by *in silico* and *in vivo* experiments (see below).



**Figure 16: X-ray structure of DPC-solubilized full-length NccX dimer.** *Panel A*, putative model of full-length NccX inserted in the inner membrane adapted from the model of CnrX (16). The two TM segments are supposed to anchor the periplasmic sensor domain. In the CnrXs construct that starts with Ser31, Gly39 was the first N-terminal residue to be structurally defined. *Panel B*, X-ray structure of NccX determined in this study (pdb entry: 4clv). Protomer a is colored sandy brown and protomer b is colored steel blue. The side chains of some amino acids stressed in the text are shown as sticks. These amino acids are identified by their one-letter code, numbering in the sequence, and a lowercase letter corresponding to the protomer they belong to. The elements of secondary structure mentioned in the text are annotated (e.g. H2-H3 hairpin) and the location of the Zn ions is shown. The red portion in each H1 helix corresponds to the GxxxG motif. The orange portion corresponds

to a stretch of about 10 residues (AA31 to 40) that were present in the sequence of CnrXs (panel A) but not structurally defined (16, 19). This stretch includes Pro37 depicted as sticks and highlighted in yellow. Ten Zn ions were located in the asymmetric unit (grey spheres). Most of them have an incomplete coordination sphere at this resolution. The four Zn ions in the dotted square are those of the metal-binding sites detailed in panels C and D. The ligands of two Zn ions involved in interprotomer interactions are developed as sticks: these are H1 residues Glu36 and His32, respectively contributing one oxygen and one nitrogen and Glu106 from H3 of the opposite protomer, that contributes one oxygen. **Panel C**, superimposition of NccX and Zn<sub>A</sub>-CnrXs protomers. NccX is colored sandy brown and Zn<sub>A</sub>-CnrXs red. The portion of H1 in NccX colored yellow corresponds to the kink upstream Pro37 that makes a break in the helix as mentioned in panel B. This kink marks the putative junction between the sensor domain and the TM segment (see discussion). Note that the C-terminal helix H4 of NccX sticks out of the helix bundle. **Panel D**, close-up view of the superimposition of the metal-binding sites (color code for the zinc ions as for proteins in panel C). Phe66 is not a metal ligand but it interacts by  $\pi$ - $\pi$  stacking with His119, thus providing an interaction between H2 and H3.



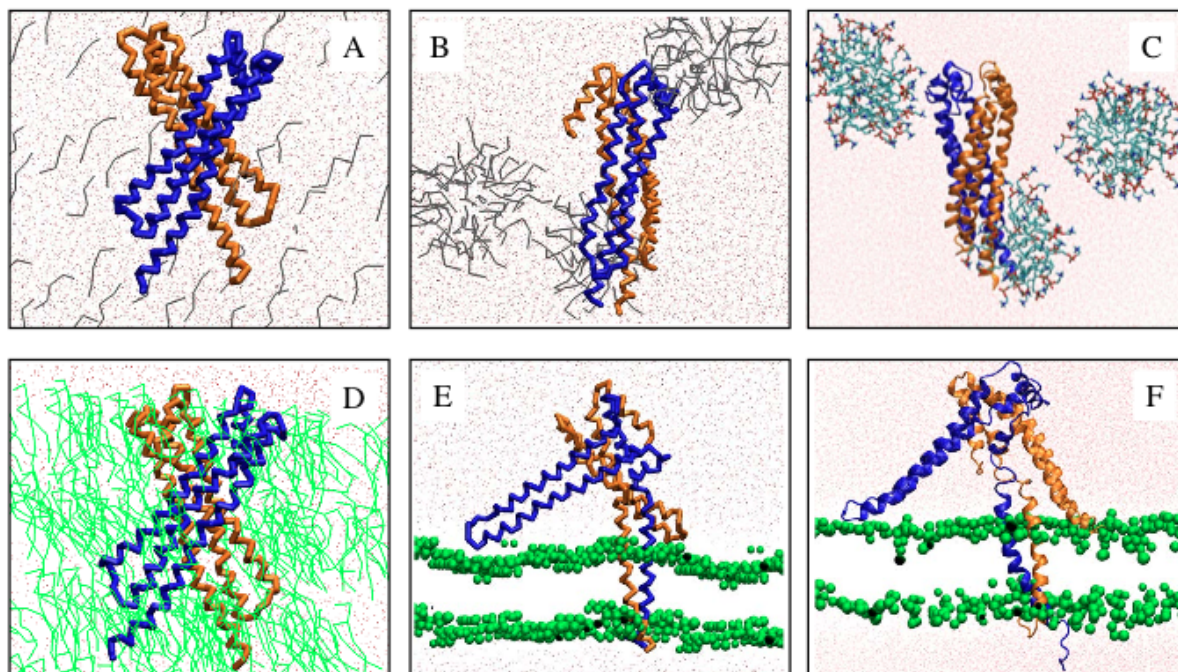
**Figure 17: Crystal contacts involving the TM domains of two symmetry-related dimers of NccX.** *Panel A*, the TM domain of dimeric NccX is made of the N-terminal half of both H1 helices (blue and cyan or red and yellow). The two H1 helices of a TM domain join

together at Tyr29, and segments upstream (toward the N-terminus) form a fork with an angle of about 55°. This angle is the result of crystal contacts involving two TM domains that are interlocked, and run antiparallel to each other (cyan and yellow or blue and red, respectively) (lower panel). **Panel B**, ribbon representation of residues 4 to 36 of two interlocked dimers corresponding to a close-up view of the area shaded in blue in the lower panel A. The N-terminal and C-terminal halves of each TM domain engage in different types of helix:helix interactions. The N-terminal halves of TM helices down to the GxxxG motif form antiparallel dimers with their counterparts from symmetry-related NccX proteins (cyan and yellow or blue and red). The C-terminal halves of TM helices (downstream the GxxxG motif) form parallel dimers within NccX dimers. Color code: purple, GxxxG motifs; pink, atom contacts between helices with parallel docking within the NccX dimer; brick red, atom contacts between helices with anti-parallel docking between interlocked NccX dimers. Some residues making the interface boundaries are indicated.

Actually, the crystal structure of NccX that was determined reflects many artifactual constraints and represents a model that is incompatible with the insertion and the dimerization of TM domains inside the membrane. In the following, we used *in silico* simulations to investigate the link between the NccX crystal structure and the significant body of data obtained with CnrX periplasmic domain (9,10,12).

**Structure-based models of NccX** - In a first series of *in silico* experiments, NccX<sub>DPC</sub>, the CG model derived from the X-ray structure of NccX, was mixed with single molecules of either DPC or POPC in random orientations and positions. Several experiments with similar outcomes were performed and **Figure 18** illustrates a representative MD simulation with DPC (**panels A to C**) or with POPC (**panels D to F**). With DPC, the end of the simulation shows that the DPC molecules organized as micelles and that the structure of NccX remained in a collapsed state (**Figure 18B**). This suggests that the crystalline structure of NccX reflects a conformation of the DPC-solubilized protein. An all-atom simulation was performed, starting from the output of the CG simulation. This improved the representation of the system and the results yielded by the CG approach were comforted. For instance, **Figure 18C** shows the association of a micelle of DPC with the hydrophobic tips of the helices H1 of each NccX protomer. With POPC, the lipid molecules quickly formed a lipid bilayer between two layers of water, thus mimicking the inner membrane (**Figure 18E**). After few microseconds, NccX was found inserted in the lipid bilayer by the two long N-terminal helices predicted to anchor the protein in the biological membrane. The parts of NccX that correspond to the TM and periplasmic domains clearly tend to partition in the membrane bilayer and the solvent, respectively. However, the hairpin formed by the loop connecting H2

and H3, which participates in the hydrophobic dimer interface (see above and **Figure 22**), still contacted the lipid bilayer after 12  $\mu$ s. Here as well, the all-atom simulation (**Figure 18F**) confirmed the pertinence of the CG approach.



**Figure 18:** Simulations performed with  $\text{NccX}_{\text{DPC}}$ . Upper panels,  $\text{NccX}_{\text{DPC}}$  was mixed with DPC molecules in random position (**A**) and the result of a simulation of 5  $\mu$ s is shown (**B**). Lower panels,  $\text{NccX}_{\text{DPC}}$  was mixed with POPC molecules in random position (**D**) and the result of a simulation of 12  $\mu$ s is shown (**E**). **Panels C and F** correspond to 10 and 50 ns prolongations in a united-atom representation of the CG simulations shown in **panels B and E**, respectively. The united-atom simulation showed that the systems did not bring to light a quirk in the CG models. The two protomers of NccX are depicted in blue and brown. The water and DPC molecules are colored red and grey, respectively. POPC molecules are colored green and represented as sticks (**D**) or as beads (**E and F**). The  $\text{Na}^+$  and  $\text{Cl}^-$  ions have been purposely omitted for clarity.

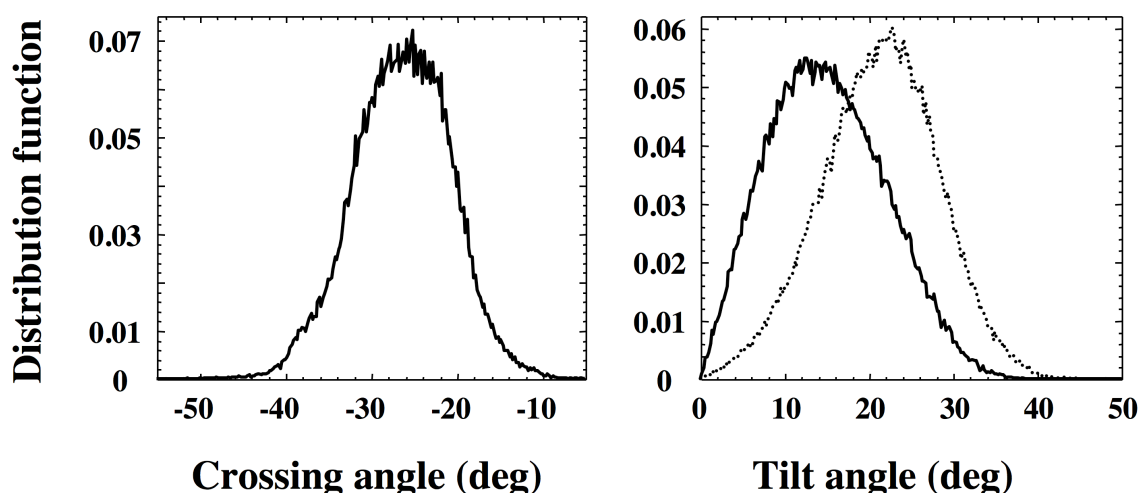
The conditions of simulation are as follows.

**Upper panels**, number of DPC: 116; number of  $\text{Na}^+$ : 31; number of  $\text{Cl}^-$ : 18; number of atoms (coarse-grained): 27,900, number of atoms (all-atom): 326,184; initial box dimension (nm): 13.0, 13.0, 17.9; final box dimension (nm): 13.7, 13.7, 17.9

**Lower panels**, number of POPC: 167; number of  $\text{Na}^+$ : 13; number of  $\text{Cl}^-$ : 0; number of atoms (coarse-grained): 7,108, number of atoms (all-atom): 89,102; initial box dimension (nm): 9.5, 9.5, 10.5; final box dimension (nm): 7.7, 7.7, 13.6

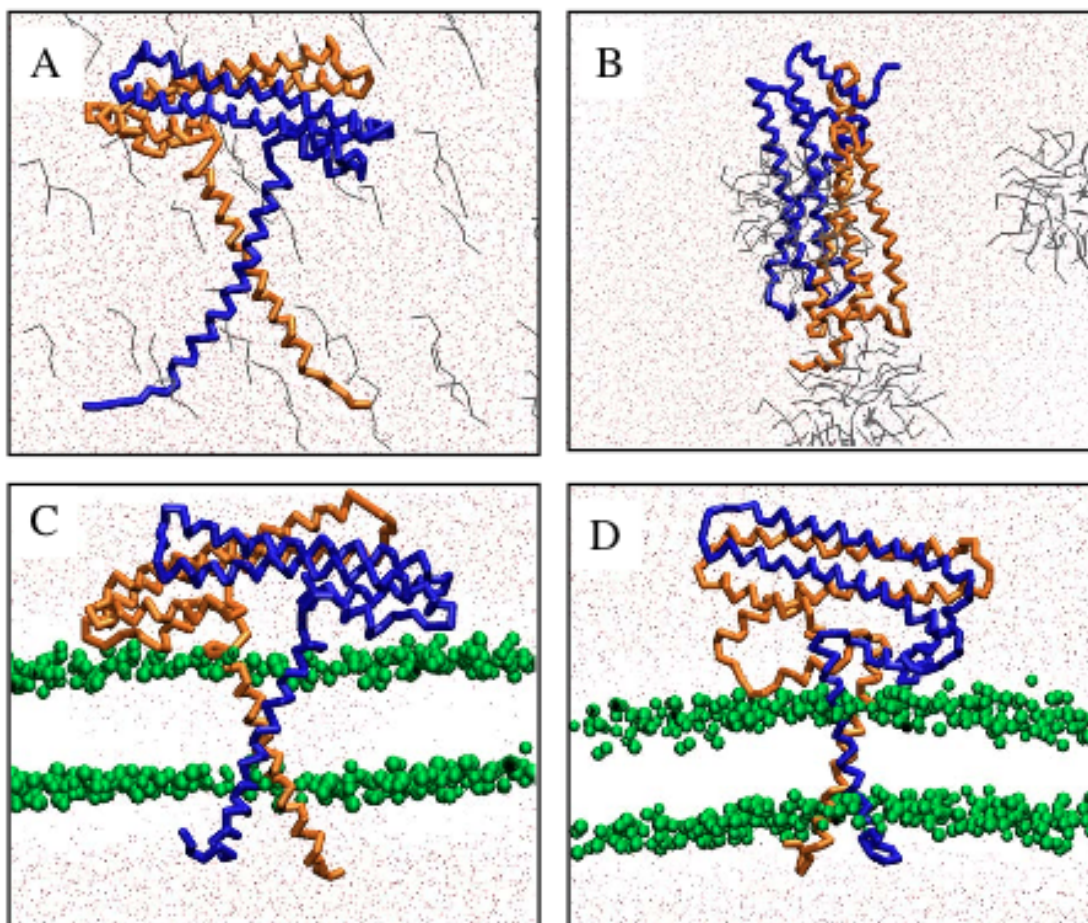


Membrane insertion of NccX<sub>DPC</sub> was examined throughout the simulation experiment. The crossing angles of H1 helices,  $\Omega$ , showed a right-handed population with average negative values of about  $-26^\circ$  (**Figure 19**). The distribution of the tilt angles of the two TM helices with average values with respect to the bilayer normal of  $22^\circ \pm 6$  and  $12^\circ \pm 7$ , respectively (**Figure 19**) indicates that the anchoring helices were in dynamic equilibrium while they remained inserted in the membrane over the full time of simulation.



**Figure 19:** Distribution of crossing angles (*left panel*) and tilt angles (*right panel*) of the two protomers averaged over the time course (12  $\mu$ s) of the simulation of NccX<sub>DPC</sub> in POPC shown in Fig. 4DE.

As a control, a second series of simulations was performed with NccX<sub>pred</sub>, a plausible CG model for the native conformation of membrane-embedded NccX, *i.e.* with the TM helices running perpendicular with respect to the globular soluble domain (10) and packed with a right-handed negative crossing angle of about  $-35^\circ$ , as defined by Chothia and *al.* (36). NccX<sub>pred</sub> was either mixed with DPC in random orientation or inserted in preformed POPC bilayer. **Figure 20** shows that the NccX<sub>pred</sub> model was not stable in DPC and that, similar to NccX<sub>DPC</sub>, it displayed a collapsed structure. In contrast NccX<sub>pred</sub> remained inserted in the POPC membrane, keeping a similar global fold. These two series of simulations confirmed both the destabilization of the inter-protomer hydrophobic interactions by DPC and the relevance of the model of TM insertion of the H1 helices.



**Figure 20:** Simulations performed with NccX<sub>pred</sub>. Upper panels, NccX<sub>pred</sub> was mixed with DPC molecules in random position (A) and the result of a simulation of 5  $\mu$ s is shown (B). Lower panels, NccX<sub>pred</sub> was mixed with POPC equilibrated as a membrane patch (C) and the result of a simulation of 14  $\mu$ s is shown (D). Color code as in **Figure 18**. The Na<sup>+</sup> and Cl<sup>-</sup> ions have been purposely omitted for clarity.

The conditions of simulation are as follows:

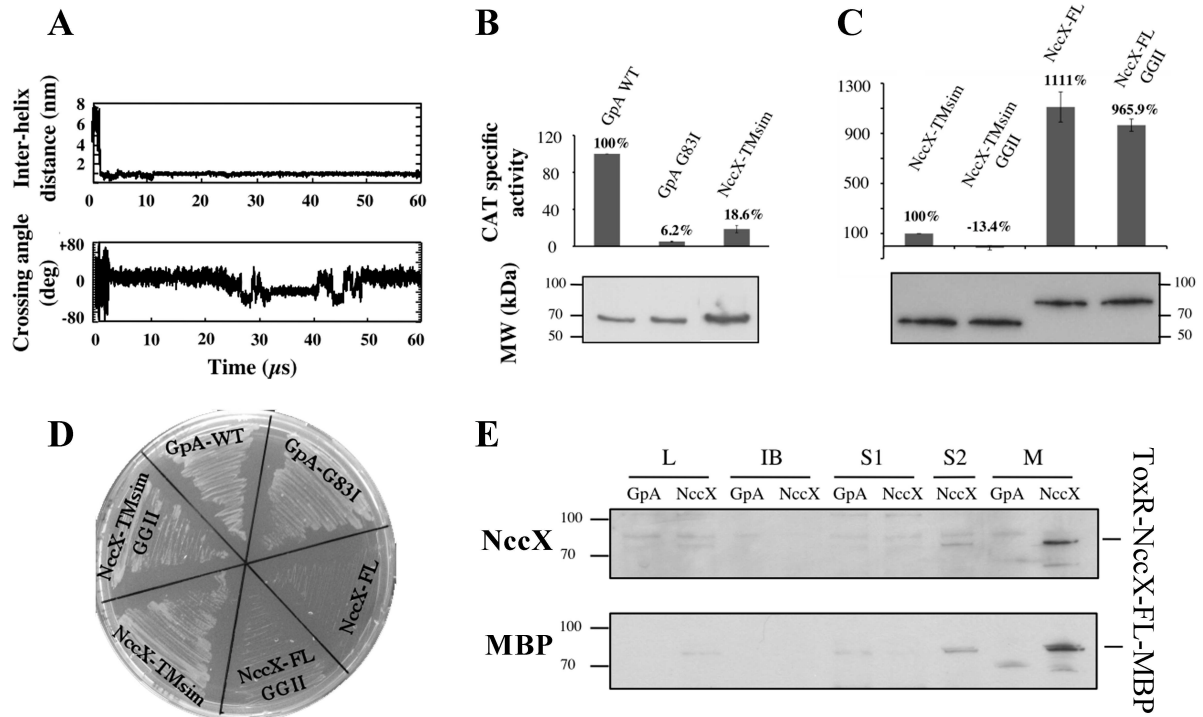
**Upper panels**, number of DPC: 117; number of Na<sup>+</sup>: 28; number of Cl<sup>-</sup>: 22; number of atoms: 27,899; initial box dimension (nm): 13.0, 13.0, 17.0; final box dimension (nm): 13.7, 13.7, 17.9

**Lower panels**, number of POPC: 267; number of Na<sup>+</sup>: 13; number of Cl<sup>-</sup>: 22; number of atoms: 10,125; initial box dimension (nm): 10.5, 12.5, 9.0; final box dimension (nm): 8.9, 10.6, 12.3

*In silico and in vivo transmembrane segments interactions* - To gain further insight in the possible interaction between TM segments, self-assembling simulations of the NccX TM segments embedded in POPC bilayer were performed using CG-MD. For these calculations, a 35-residue long peptide called NccX-TMsim, encompassing the 21-

residue long TM segment and flanking residues, was used (**Figure 15**). Two NccX-TMsim were inserted into a preformed POPC bilayer 6 nm away from each other. **Figure 21A** shows that they self-assembled in less than 2  $\mu$ s and that a long-lasting helix dimer was spontaneously formed. Examination of the structure of this dimer during a typical 60  $\mu$ s simulation revealed that the helices explored different conformational spaces (**Figure 21A**). While the main helix packing was left-handed (average value of the helix-crossing angle  $\Omega$  of about 10°), a right-handed populations also existed ( $\Omega$  of about -20°), which is in agreement with the right-handed crossing angle measured with NccX<sub>DPC</sub> in POPC. There is no clear period of alternation between both states. As NccX-TMsim contains the putative G<sub>16</sub>xxxG<sub>20</sub> interaction motif, the double-mutant G16I-G20I (referred to as GGII) was chosen to further analyze its role in dimerization by the TOXCAT assay in both NccX-TMsim and full-length NccX (NccX-FL). The results of the TOXCAT assays are displayed in **Figure 21BC** and those of the corresponding controls in Fig. 7DE. The expression level of each constructs assessed by Western blotting against MBP antibodies in whole-cell lysates (**Figure 21BC**) was used to normalize the specific CAT activities.

In a first experiment, NccX-TMsim was compared to the widely used controls GpA-wt and GpA-G83I (14). Consistent with CG simulations, the dimerization of NccX-TMsim was detected *in vivo* by recording a CAT activity that represented only 18.6 % of the CAT activity induced by the dimerization of GpA-wt (**Figure 21B**). Then, NccX-TMsim was compared to NccX-FL, both constructs bearing the GGII double mutation or not (**Figure 21C**). Interestingly, the CAT activity yielded by NccX-FL was 10-fold higher than NccX-TMsim taken as a reference. The GGII double mutation cancelled the protein interaction in NccX-TMsim while it had no significant effect in the full-length construct. These results demonstrate that the two N-terminal ends of dimeric full-length NccX are close to each other, in agreement with the results of the simulations reported above (**Figure 18**). They also indicate that, while the GxxxG motif is important for NccX-TMsim interaction, it has a minor impact for the dimerization of full-length NccX protomers, which is mainly driven by the periplasmic domains.

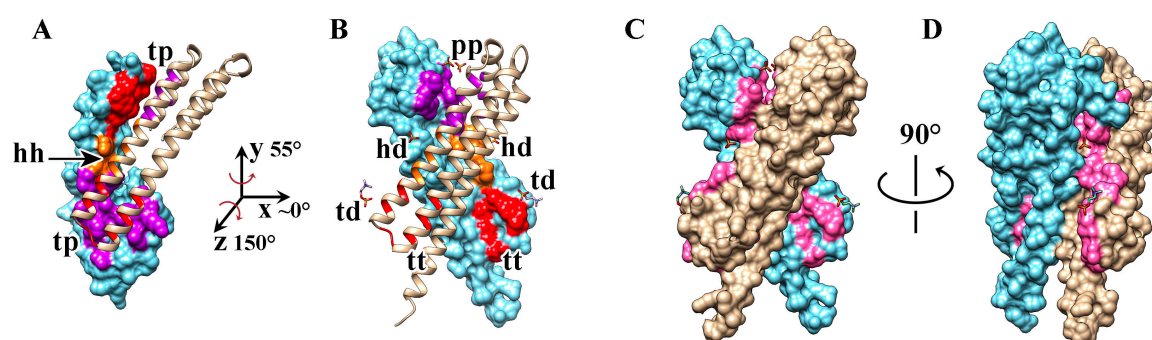


**Figure 21:** NccX transmembrane segment homodimerization revealed by CG-MD simulations and the TOXCAT assay. **Panel A, upper panel**, time-course of the inter-helix distance over a typical 60  $\mu$ s simulation. **Lower panel**, time-course of the helix crossing angle  $\Omega$  over the same simulation. Positive and negative crossing angles correspond to left-handed and right-handed packings, respectively. **Panel B**, the ability of NccX TM helices (NccX-TMsim) to homodimerize was assessed with the TOXCAT reporter assay: GpA-WT and GpA-G83I were respectively used as positive and negative controls. For each construct, CAT specific activity was normalized against GpA-WT, averaged over three independent experiments and corrected for the production levels of the chimeric proteins ToxR-GpA-MBP or ToxR-NccX-TMsim-MBP that were assessed by Western blot with anti-MBP antibodies (**lower panel**). **Panel C**, NccX-TMsim, taken as a reference, was compared to NccX-FL and to the double GGII mutants of either construct. Normalization of the CAT activities as in panel B. **Panels D and E** are controls of topology of the constructs used for the TOXCAT assays. **Panel D**, MaIE complementation assays confirmed the correct topology of all constructs but the full-length proteins. For the latter, the bulky metal-sensor part of NccX likely hampered the ability of MBP to interact with the maltose transporter. **Panel E**, the topology of both NccX-FL constructs was found to be correct as they were specifically detected in the membrane fraction. High CAT activity (**Figure 21C**) demonstrated the correct orientation of the chimeric protein, with ToxR in the cytoplasm. Total proteins were separated by SDS-PAGE and blotted onto a PVDF membrane. ToxR-NccX-FL-MBP was revealed by Western blot analysis with antibodies to CnrXs (**upper panel**) that cross-react with NccX (10) or with antibodies to MBP (**lower panel**). The ToxR-GpA-MBP construct was used as a negative control. A very faint band corresponding to ToxR-NccX-FL-MBP was visible in the lanes L and S2 while this protein was massively detected in the lane M corresponding to the membrane fraction. L, lysate obtained by sonic oscillation; IB, inclusion bodies; S1, supernatant of the first centrifugation used to pellet the membrane. The pellet was suspended in sonication buffer and centrifuge again. The new supernatant (S2) and the membrane fraction (M) were used for further analysis.

***Interpretation of the detergent-induced structural rearrangements*** - The crystal structure of DPC-solubilized NccX differs markedly from the prediction shown in **Figure 16A**. Despite the clear correspondence between the secondary structure elements of NccX and CnrXs, the cradle-shaped CnrXs dimer (**Figure 16A**) is not reproduced, even partially by the NccX dimer. Structural rearrangements probably occurred after solubilization by DPC. However, the 3.12 Å resolution was insufficient to accurately locate DPC molecules in the electron density. Still, the polar head of a detergent molecule was located near the side chains of Trp94 and Glu99 of each protomer, *i.e.* at the tip of the H2-H3 hairpin (**Figure 22**). Furthermore, two phosphate moieties from DPC were observed close to the loop connecting H3 and H4, and two more moieties were located near Arg114 residues, *i.e.* near the middle of the H3 helices. Both sites belong to CnrXs dimer interface (10), thus suggesting that DPC destabilized the interface of the native periplasmic domain. This was examined in more detail.

Compared to CnrXs, the dimer interface of NccX is noticeably different (**Figure 22**). In CnrXs, H4' is part of the helix bundle and contributes to the dimer interface by binding the partner protomer *via* the tip of its H2'-H3' hairpin (10). This contact is no longer possible in NccX in which H4 sticks out of the helix bundle. Hence, the residues of CnrXs dimer interface are pulled apart in NccX where the dimer interface buries 1943 Å<sup>2</sup> which is much less than 3054 Å<sup>2</sup> in CnrXs. A homology model of the periplasmic sensor domain of NccX, called NccXs (**Figure 22A**), was built using the crystal structure of Zn-bound CnrXs as a template (the residues present in PDB entry 2y3d of CnrXs share 80% identity with the corresponding residues in NccX). In this model, the dimer interface consists in pairs of residues carried by both H3' helices (hh contacts in Fig. 8A) or carried by H2' and H4' helices from either protomer (tp contacts in **Figure 22A**), like in the CnrXs template (10). These interactions no longer existed in DPC-solubilized NccX as H2 and H1 of each protomer form an antiparallel coiled-coil where H2 residues that were engaged in NccXs dimer interface now interact with TM residues in H1 of the same protomer (tt contacts in **Figure 22B**). A wider permutation of the residues in contact occurred as follows: the H2':H4' dimer interface in NccXs (tp contacts) was redistributed between the H1-H2 coiled-coil dimerization motif in NccX and a new contact area where H4 residues face H3 residues that were previously part of the soluble NccXs dimer interface (pp contacts in **Figure 22B**). **Figure 22** shows that DPC moieties co-localized with regions that were at least partially buried in NccXs and that are exposed to the solvent in NccX, thus emphasizing the role of

DPC in the stabilization of the topology seen in NccX crystal structure. Altogether these data show that the residues contributing to the dimer interface of the native NccX periplasmic domain are redistributed upon DPC-solubilization. One can conclude that the detergent, in addition to solubilizing the membrane bilayer, destabilized the hydrophobic interactions between subunits, thus allowing the two periplasmic domains to rearrange and collapse against the hydrophobic helices H1.



**Figure 22:** Redistribuition of the residues contributing to the dimer interface of the NccX periplasmic domain upon DPC-solubilization. A homology model of NccXs, the periplasmic domain of NccX, was built with Modeller (35), using the crystal structure of Zn-bound CnrXs as a template (PDB entry 2y3d). Chains A (*tan*) of NccXs model and NccX crystal structure have been superimposed and are presented side-by-side as ribbons in **A** and in **B**, respectively. Chains B are shown as molecular surfaces (*blue*). The dimer interface of each protomer of NccXs was considered as a collection of three pieces: the tip of the H2-H3 hairpin (*t, red*), the body of helix H3 (*h, orange*) and the platform built by H3-H4 (*p, purple*). These pieces (*t, h, p*) were colored on the molecular surface representation of protomer B of NccXs (**A**) or NccX (**B**). The phosphate moiety of some DPC molecules bound to the protein is depicted as orange sticks. Full polar heads of DPC are shown as orange (phosphate moiety) and blue (choline moiety) sticks. **A**, in NccXs, two protomers stick together via tip-to-platform contacts (*tp*) and helix-to-helix contacts (*hh*). **B**, in NccX, *tp* contacts have been replaced by tip-to-DPC (*td*), tip-to-TM (*tt*) and platform-to-platform (*pp*) contacts. The *hh* contacts have also been disrupted and replaced by helix-to-DPC (*hd*) contacts. As a result, the dimer interface in crystallized NccX has been completely remodeled as compared to that of NccXs, with DPC and the TM N-terminal half sticking in (see text). Rotations about the *y* and *z* axes make NccXs chain B superimposable on NccX chain B. **C**, molecular surface representation of NccX with DPC moieties as sticks. The orientation of NccX is the same as in **B**. The residues colored in pink are those of the NccXs dimer interface. As a result, regions that were at least partially buried in NccXs and that are exposed to the solvent in NccX do appear in pink on a blue or tan background. **D**, the same representation as in **C** was rotated 90° as indicated. DPC moieties co-localized with regions in pink, thus emphasizing the role of DPC in the stabilization of the topology seen in NccX crystal structure, which is at odds with the topology of the soluble periplasmic domain of NccX, as inferred by homology with CnrXs (10).

## DISCUSSION

Although the single-pass membrane proteins represent almost half of all membrane proteins, few examples of structures containing an isolated TM helix already exist. This is the case for the rat and the human monoamine oxidase A (37,38), whose membrane attachment not only involves a C-terminal TM helix but also hydrophobic patches at the surface of the protein. Such is also the case of the membrane-bound cytochrome *c* quinol dehydrogenase NrfH from *Desulfovibrio vulgaris*: NrfH is dimeric and both TM helices pack together in a right-handed fashion (39). Most recently, the structure of a cytochrome P450 from yeast showed a single TM domain standing alone and playing an unforeseen role in soluble domain orientation outside the membrane, indicating that there are lessons to be learned by characterizing more single-pass membrane proteins (40). To our knowledge, such is the scarce repertoire of structures available for bitopic membrane proteins. Specific difficulties in both purification and ordered crystallization can explain this fact. The structural characterization of full-length NccX, the membrane-anchored Ni, Co and Cd sensor protein of the NccYXH transmembrane signaling complex also contributes to fill this void in the PDB. NccX was efficiently extracted regardless of the detergent used (*i.e.* DPC, DDM, OG, C8E4), but only DPC allowed achieving both purification and crystallization. Although widely used in membrane protein investigation (41, 42), DPC has been seldom used for crystallization whereas DPC micelles can be considered a membrane mimetic of choice for NMR studies, owing to the number of structures observed (41). Very recently, the effect of DPC on protein conformation was illustrated in two instances, where solution structures obtained in DPC micelles and crystal structures obtained either in LAPAO micelles or lipid *meso*-phase were significantly different. In one example, the solution structure of *E. coli* diacylglycerol kinase showed a domain-swapped trimer whereas the same enzyme and two variants crystallized in a lipid bilayer showed no sign thereof: the implication is such that the architecture and the catalytic properties of the active sites are different in both structures (43,44). In the other case, when the solution structure of the uncoupling protein UCP2 solubilized and purified in DPC was described (45), comparison with the high-resolution X-ray structure of the bovine

ATP/ADP mitochondrial carrier crystallized in LAPAO micelles (46) suggested that it carried biases affecting biological interpretation (47). This raises the question of the behavior of a membrane protein once extracted from its native membrane environment and of the techniques to be used to explore this behavior. Regarding UCP2, the physiological relevance of its solution structure was assessed with the help of theoretical calculations. Indeed, one major conclusion of this study was that combining functional and computational investigations is a promising general strategy that would be beneficial to analyze the structures of numerous membrane proteins (47). As a general guideline, alternative methods are required for structure validation, quality assessment and enhancement (48). This is particularly true for the membrane proteins with an activity difficult to assess *in vitro*.

Such a strategy was adapted to the 3.12 Å resolution X-ray structure of NccX. Extensive work on the soluble domain alone made it obvious that NccX crystal structure was at odds with the functional native form (9,10,12). We have performed MD simulations and TOXCAT assays to rationalize the X-ray structure of DPC-solubilized NccX. Solubilization of the membrane bilayer and detergent destabilization of the main hydrophobic contacts of the dimer interface led to the crystallization of a protein in which the periplasmic domains of both protomers parted from each other and collapsed against the hydrophobic TM segments within H1 helices. That the detergent is responsible for the structural collapse of NccX can be inferred from the redistribution of hydrophobic interactions among the TM helix and a pair of hydrophobic patches within the periplasmic domain. While the exposure of the TM helix followed membrane solubilization, it seems that DPC played an active role in dissociating the dimer of periplasmic domains because head groups of DPC were found in close vicinity of the hydrophobic patches that kept them together. The sensitivity of NccX structure to DPC solubilization might be a consequence of the intrinsic plasticity of this sensor, the conformation of which was proposed to be tuned from the metal-binding site *via* hydrophobic packing in the core of the periplasmic domain (10).

Additional constraints affected the TM segments. In the crystal, the TM segments of one dimer form a hydrophobic fork that interacts with the symmetry-related hydrophobic fork of another dimer to compose an interlocked tail-to-tail arrangement of dimers. This is a case of crystal packing with extreme consequences, where the antiparallel docking of TM helices displaced the parallel, possibly native docking, *i.e.* an aberrant topology was produced, that is incompatible with membrane insertion. Crystal packing was also supported by several zinc ions that were present in excess in the crystallization conditions. Interestingly, NccX



protomers best superimposed with chain A of Zn-bound CnrXs, *i.e.* the chain that contained two zinc ions (10). This shows that each NccX protomer considered individually, retained the characteristic fold of the protomers of CnrXs extended by most of the residues of the long N-terminal membrane-anchoring helix. From its C-terminus to the N-terminus, helix H1 contains several distinctive features. His46 and His42, at the C-terminal end, are strictly conserved residues in the CnrX protein family and are essential residues of the metal-binding site (9,10,12). The residue His38 was found to chelate a Zn(II) ion in both NccX and protomer A of CnrXs crystallized in the presence of an excess of Zn(II). This residue is not strictly conserved (10) and the phenotype of CnrX is poorly affected by the His38-Arg substitution (8). His38 belongs to a stretch of about 10 residues (Residues 31 to 40) that was not resolved in the crystal structure of CnrXs and was proposed to make a flexuous junction between CnrXs and the TM segment, thus tethering the sensor domain to the membrane. The flexibility of this stretch is illustrated by its poor definition in the electron density maps of CnrXs, and by secondary structure predictions suggesting a random coil in the full-length protein. That NccX or CnrX contains a proline or a glycine, respectively, at position 37 further argues for this hypothesis as these residues are well-known helix-breakers. We can thus speculate that the kink observed at Pro37 is the result of the constraints that arose with membrane dislocation and helix repacking in the presence of detergent. When the membrane was solubilized by DPC, constraints imposed by hydrophobic surface exclusion, crystal packing, and zinc-mediated interactions yielded a topology that does not fit seamlessly the sequence of NccX.

Going up to the N-terminus, the membrane-inserted sequence of 21 residues, extending from residue 29 to 9, contains a GxxxG motif that often mediates interactions between TM helices (14,49-54). In addition, this TM sequence includes 11  $\beta$ -branched residues (L, V, I) and 7 small residues (A, G, S). The strength of the interaction mediated by the GxxxG motif is enhanced by nearby  $\beta$ -branched residues (51) and a short interhelical distance due to the presence of small residues favors the TM interactions (52). However, the repartition of such residues is crucial (52). For instance, based on the seminal work of Engelman and coworkers on the dimerization of GpA, the complete motif of interactions LIxxGVxxGVxxT was identified, with a minimal consensus motif for stability defined as LxxxGxxxGxxxT (55). This shows that the GxxxG motif requires to be stabilized on both sides (52,55). Both NccX and CnrX display a very incomplete motif in the TM region with a corresponding sequence as STxxG<sub>16</sub>xxxG<sub>20</sub>VxxA. This is in agreement with the minor role of

the GxxxG motif in TM segment interactions observed as far as the full-length protein was considered. Finally, the rest of the N-terminal sequence extending from residue 8 to 1 is predicted to be located in the cytoplasm (Scheme 1), in agreement with its content in several neutral or polar residues.

MD simulations are of great interest to study lateral association and oligomerization of TM helices (56,57). CG simulations performed in the presence of DPC resulted in NccX conformations matching the collapsed structure of DPC-solubilized NccX regardless of the initial model, be it NccX<sub>DPC</sub> or NccX<sub>pred</sub>, suggesting that the crystalline structure stems from solubilization of the membrane bilayer as well as from destabilization of the inter-protomer hydrophobic interactions. Reciprocally, CG simulations provided a nice picture of the transition from NccX<sub>DPC</sub> to NccX<sub>pred</sub>, when the former was placed in a POPC environment. The partition between the hydrophobic helices in the POPC membrane and the sensor domain in the aqueous phase is fast and easily reproducible, even if the coordinate RMSD of the soluble part with respect to NccX<sub>pred</sub> is still high at the end of the longest simulations. Relaxing the constraints of GROMACS on the assignment of the secondary structures during CG dynamics was not sufficient to allow complete NccX refolding. Altogether our results strongly suggest that the periplasmic and TM domains of NccX form independent folding units, and that the soluble domain already characterized is representative of the periplasmic domain in full-length NccX/CnrX (10). The main determinant that brings the dimer together is the periplasmic domain: this keeps NccX TM segments close to each other and could potentialize their dimerization *via* the GxxxG motif, as exemplified by the *in silico* and *in vivo* characterization of NccX-TMsim interactions. Moreover, the interaction of the TM segments has been proved to be highly dynamic in the simulations reported here. CnrX/NccX and CnrY/NccY can engage into both periplasmic and transmembrane interactions (8,10). How transmembrane signaling is performed by CnrYXH/NccYXH complexes once CnrX/NccX has bound nickel or cobalt is not known. Whether the dynamic interaction of TM segments is related to the function of CnrX/NccX in signaling and how these TM segments may help propagating the signal to CnrH/NccH with the relay of CnrY/NccY are two questions that require further work.

Even though it is imperfect, the structure of full-length NccX extends the very limited catalog of structures available for fully characterized single-pass transmembrane proteins. Besides, the results presented here emphasize the benefits of implementing complementary

approaches in order to reach physiologically relevant information from detergent-solubilized proteins.

*Acknowledgments* - We thank Eve de Rosny, Juliette Trepreau and Adeline Veyret for their help during the first steps of purification of NccX. The beamline staff at Proxima-1 (synchrotron SOLEIL, Gif Sur Yvette) is thanked for his help with data collection and structure determination. We also thank Dr. P. Fuchs (Paris Diderot University) for providing useful Python scripts for trajectory analyses with Gromacs. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

## REFERENCES

1. Mergeay, M., Nies, D., Schlegel, H.G., Gerits, J., Charles, P., and Van Gijsegem, F. (1985) *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J. Bacteriol.* **162**, 328-334.
2. Monchy, S., Benotmane, M. A., Janssen, P., Vallaeys, T., Taghavi, S., van der Lelie, D., and Mergeay, M. (2007) Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* are specialized in the maximal viable response to heavy metals. *J. Bacteriol.* **189**, 7417-7425.
3. von Rozycki, T., and Nies, D.H. (2009) *Cupriavidus metallidurans* evolution of a metal:resistant bacterium. *Antonie Van Leeuwenhoek* **96**, 115-139.
4. Janssen, P. J.; Van Houdt, R.; Moors, H.; Monsieurs, P.; Morin, N.; Michaux, A., Benotmane, M.A., Leys, N., Vallaeys, T., Lapidus, A., Monchy, S., Médigue, C., Taghavi, S., McCorkle, S., Dunn, J., van der Lelie, D., and Mergeay, M. (2010) The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. *PLoS One* **5**, e10433.
5. Nies, D.H. (2003) Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.* **27**, 313-339.
6. Grass, G., Grosse, C., and Nies, D. H. (2000) Regulation of the *cnr* cobalt and nickel resistance determinant from *Ralstonia* sp. strain CH34. *J. Bacteriol.* **182**, 1390-1398.
7. Tibazarwa, C., Wuertz, S., Mergeay, M., Wyns, L., and van der Lelie, D. (2000) Regulation of the *cnr* cobalt and nickel resistance determinant of *Ralstonia eutropha* (*Alcaligenes eutrophus*) CH34. *J. Bacteriol.* **182**, 1399-1409.
8. Grass, G., Fricke, B., and Nies, D. H. (2005) Control of expression of a periplasmic nickel efflux pump by periplasmic nickel concentrations. *Biometals* **18**, 437-448.
9. Pompidor, G., Maillard, A. P., Girard, E., Gambarelli, S., Kahn, R., and Covès, J. (2008) X-ray structure of the metal-sensor CnrX in both the apo- and copper-bound forms. *FEBS Lett.* **582**, 3954-3958.
10. Trepreau, J., Girard, E., Maillard, A.P., de Rosny, E., Petit-Haertlein, I., Kahn, R., and Covès, J. (2011) Structural basis for metal sensing by CnrX. *J. Mol. Biol.* **408**, 766-779.
11. Trepreau, J., de Rosny, E., Duboc, C., Sarret, G., Petit-Haertlein, I., Maillard, A.P., Imbert, A., Proux, O., and Covès, J. (2011) Spectroscopic characterization of the metal-binding sites in the periplasmic metal-sensor domain of CnrX from *Cupriavidus metallidurans* CH34. *Biochemistry* **50**, 9036-9045.
12. Trepreau, J., Grosse, C., Mouesca, J.-M., Sarret, G.; Girard, E., Petit-Haertlein, I., Kuennemann, S., Desbourdes, C., de Rosny, E., Maillard, A.P., Nies, D.H., and Covès, J. (2014) Metal sensing and signal transduction by CnrX from *Cupriavidus metallidurans* CH34: role of the only methionine assessed by a functional, spectroscopic, and theoretical study. *Metallomics* **6**, 263-273.
13. Maillard, A.P., Girard, E., Ziani, W., Petit-Haertlein, I., Kahn, R., and Covès, J. (2014) The crystal structure of the anti- $\sigma$  factor CnrY in complex with the  $\sigma$  factor CnrH shows a new structural class of anti- $\sigma$  factors targeting extracytoplasmic function  $\sigma$  factors. *J. Mol. Biol.* **426**, 2313-2327.

14. Russ, W.P., and Engelman, D.M. (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci. USA* **96**, 863-868.
15. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
16. Kabsch, W. (2010) XDS. *Acta Cryst.* **D66**, 125-132.
17. Sheldrick, G. M. (2008) A short history of SHELX. *Acta Cryst.* **A64**, 112-122.
18. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658-674.
19. Kowan, K. (2010) Recent developments in classical density modification. *Acta Cryst.* **D66**, 470-478.
20. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025.
21. Emsley, P., Lohkamp, B., Scott, W., and Cowtan, K. (2010) Features and development of Coot. *Acta Cryst.* **D66**, 486-501.
22. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vornrhein, C., and Womack, T.O. (2011) BUSTER version 2.11.15. Cambridge, United Kingdom: Global Phasing Ltd.
23. Karplus, P.A., and Diederichs, K. (2012) Linking crystallographic model and data quality *Science* **336**, 1030-1033.
24. Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010) Graphical tools for macromolecular crystallography in PHENIX *Acta Cryst.* **D66**, 12-21.
25. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics.* **14**, 33-38.
26. Brooks, B.R., Brooks III, C.L., Mackerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., and *al.* (2009) CHARMM: The Biomolecular Simulation Program. *J. Comp. Chem.* **30**, 1545-1614.
27. Neria, E., Fisher, S., and Karplus, M. (1996) Simulation of activation free energies in molecular systems. *J. Chem. Phys.* 1996, **105**, 1902-1921.
28. de Jong, D.H., Singh, G., Bennett, W.F.D., Arnarez, C., Wassenaar, T.A., Schäfer, L.V., Periole, X., Tieleman, D.P., and Marrink, S.J. (2013) Improved parameters for the MARTINI coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687-697.
29. Van der Spoel, D., Lindahl, B., Hess, G., Groenhof, G., Mark, A.E., and Berendsen, H.J.C. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718.
30. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem Theory Comp.* **4**, 435-447.
31. Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H. (2007) The MARTINI force field: coarse-grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812-7824.

32. Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.J. (2008) The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comp.* **4**, 819-834.
33. Wassenaar, T.A., Pluhackova, K., Böckmann, R.A., Marrink, S.J., and Tieleman, D.P. (2014) Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. *J. Chem. Theory Comput.* **10**, 676-690.
34. Wei, P., Liu, X., Hu, M.-H., Zuo, L.-M., Kai, M., Wang, R., and Luo, S.-Z. (2011) The dimerization interface of the glycoprotein Ib $\beta$  transmembrane domain corresponds to polar residues within a leucine zipper motif. *Protein Sci.* **20**, 1814-1823.
35. Sali, A.; Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.
36. Chothia, C., Levitt, M. and Richardson, D. (1981) Helix to helix packing in proteins. *J. Mol. Biol.* **145**, 215-250.
37. Ma, J., Yoshimura, M., Yamashita, E., Nakagawa, A., Ito, A., and Tsukihara, T. (2004) Structure of rat monoamine oxidase A and its specific recognitions for substrates and inhibitors. *J. Mol. Biol.* **338**, 103-114.
38. Son, S.Y., Ma, J., Kondou, Y., Yoshimura, M., Yamashita, E., and Tsukihara, T. (2008) Structure of human monoamine oxidase A at 2.2 Å resolution: the control of opening the entry for substrates/inhibitors. *Proc. Natl. Acad. Sci. USA* **105**, 5739-5744.
39. Rodrigues, M.L., Oliveira, T.F., Pereira, I., and Archer, M. (2006) X-ray structure of the membrane-bound cytochrome *c* quinol dehydrogenase NrfH reveals novel haem coordination. *EMBO J.* **25**, 5951-5960.
40. Monk, B.C., Tomasiak, T.M., Keniya, M.V., Hushmann, F.U., Tyndal, J.D.A., O'Connell, J.D., Cannon, R.D., McDonald J.G., Rodriguez, A., Finer-Moore, J.S, and Stroud, R.M. (2014) Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. *Proc. Natl. Acad. Sci. USA* **111**, 3865-3870.
41. Warschawski, D.E., Arnold, A.A., Beaugrand, M., Gravel, A., Chartrand, E., and Marcotte, I. (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochem. Biophys. Acta* **1808**, 1957-1974.
42. Tulumello, D.V., and Deber, C.M. (2012) Efficiency of detergents at maintaining membrane protein structures in their biologically relevant forms. *Biochem. Biophys. Acta* **1818**, 1351-1358.
43. Van Horn, W.D., Kim, H.J., Ellis, C.D., Hadziselimovic, A., Sulistijo, E.S., Karra, M.D., Tian, C., Sönnichsen, F.D., and Sanders, C.R. (2009) Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science* **324**, 1726-1729.
44. Li, D., Lyons, J.A., Pye, V.E., Vogeley, L., Aragão, D., Kenyon, C.P., Shah, S.T., Doherty, C., Aherne, M., and Caffrey, M. (2013) Crystal structure of the integral membrane diacylglycerol kinase. *Nature* **497**, 521-524.
45. Berardi, M.J., Shih, W.M., Harrison, S.C., and Chou, J.J. (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* **476**, 109-113.

46. Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J.-M., and Brandolin, G. (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426**, 39-44.
47. Zoonens, M., Comer, J.R., Masscheleyn, S., Pebay-Peyroula, E., Chipot, C.J., Miroux, B., and Dehez, F. (2013) Dangerous liaisons between detergents and membrane proteins. The case of mitochondrial uncoupling protein 2. *J. Am. Chem. Soc.* **135**, 15174-15182.
48. Zhou, H-X., and Cross, T.A. (2013) Influences of membrane mimetic environments on membrane protein structures. *Ann. Rev. Biophys.* **42**, 361-392.
49. Lemmon, M.A., Flanagan, J.M., Treulein, H.R., Zhang, J., and Engelman, D.M. (1992) Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry* **31**, 12719-12725.
50. Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002) GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* **41**, 5990-5997.
51. Senes, A., Gerstein, M., and Engelman, D.M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921-936.
52. Senes, A.; Engel, D.E.; DeGrado, W.F. (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* **14**, 465-479.
53. Moore, D.T., Berger, B.W., and DeGrado, W.F. (2008) Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Structure* **16**, 991-1001.
54. Cymer, F., Veerappan, A., and Schneider, D. (2012) Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochem. Biophys. Acta* **1818**, 963-973.
55. Melnik, R.A., Kim, S., Curran, A.R., Engelman, D.M., Bowie, J.U., and Deber, C.M. (2004) The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *J. Biol. Chem.* **279**, 16591-16597.
56. Psachoulia, E., Marshall, D.P., and Sansom, M.S. (2010) Molecular dynamics simulations of the dimerization of transmembrane alpha-helices. *Acc. Chem. Res.* **43**, 388-396.
57. Aci-Sèche, S., Sawma, P., Hubert, P., Sturgis, J.N., Bagnard, D., Jacob, J., Genest, M., and Garnier, N. (2014) Transmembrane recognition of the semaphorin co-receptors neuropilin 1 and plexin A1: coarse-grained simulations. *PLoS One* **9**, e97779.

L'analyse de la structure de NccX ainsi que la réalisation d'expériences *in silico* et *in vivo* ont permis de mettre en évidence l'effet délétère du détergent DPC sur la structure native de la protéine. En effet, en solubilisant les membranes plasmiques, le détergent a détruit l'environnement lipidique dans lequel étaient enfouis les segments TM. Il en a résulté une forme où les éléments de structure secondaire sont conservés mais réarrangés de façon spectaculaire. Des simulations par dynamique moléculaire ont été réalisées en vue de tester l'insertion dans une bicouche lipidique du modèle de NccX tel qu'il apparaît dans la structure cristalline en présence de DPC. Ces expériences réalisées par Serge Crouzy ont permis de montrer que les portions de NccX correspondant aux domaines TM et périplasmique ont respectivement tendance à partitionner dans les phases lipidique et aqueuse, conformément au modèle attendu. Sachant que la protéine NccX est un dimère, et qu'un motif d'homodimérisation est présent dans le domaine TM de NccX et CnrX, la capacité du domaine TM de NccX à s'homodimériser a fait l'objet d'une autre campagne de simulations par dynamique moléculaire (collaboration avec Nobert Garnier). Ces expériences *in silico* ont confirmé que l'homodimérisation peut avoir lieu, tout en mettant en évidence la polydispersité de l'état dimérique au cours du temps, donc le caractère dynamique de l'interaction entre deux segments TM.

Les domaines TM de NccX et de CnrX possèdent en effet un motif GxxxG (séquence indiquée dans la **figure 15**). Dans la structure de NccX obtenue, les motifs GxxxG des deux hélices TM sont distants de 18 Å. Ces motifs n'interagissent pas entre eux mais avec le domaine périplasmique de NccX et avec le domaine TM d'un protomère de NccX adjacent dans le cristal (**Figure 17**). La capacité de ce motif à participer à des interactions protéine:protéine se manifeste de façon aberrante dans les conditions de cristallisation : il s'est substitué à une interaction stabilisant le dimère des domaines périplasmique de NccX natif et il est impliqué dans l'édifice cristallin.

Les expériences TOXCAT effectuées sur les protéines chimériques NccX et NccX-TMsim (segment TM de NccX flanqué de quelques résidus) (**Figure 21**) indiquent que NccX dimérise fortement alors que NccX-TMsim ne s'homodimérise que faiblement. La mutation des glycines du motif GxxxG n'a pas d'effet significatif sur la dimérisation de NccX alors qu'elle rompt l'interaction entre deux NccX-TMsim. Ces résultats nous ont ainsi permis d'identifier le domaine périplasmique de NccX comme déterminant majeur de

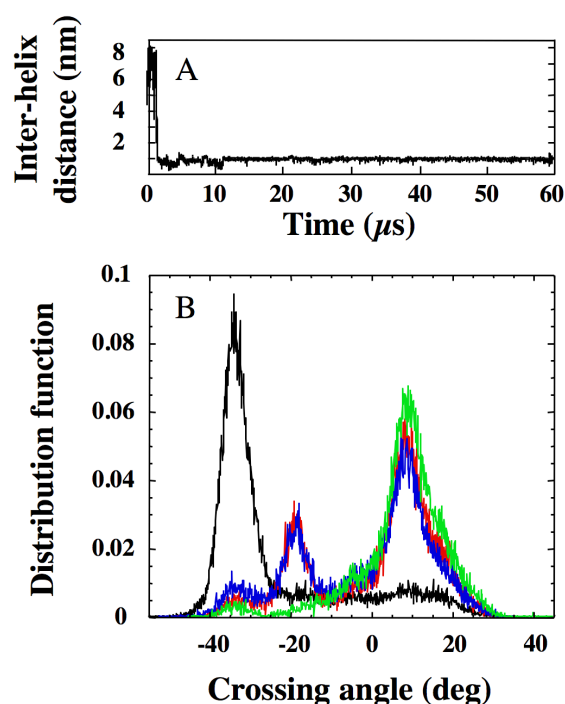


l'homodimérisation de NccX. Il apparaît, cependant, que le motif GxxxG permet le rapprochement des segments TM de NccX dans la membrane dans une moindre mesure.

## C. Résultats complémentaires

Une analyse plus fine des interactions des segments TM a été réalisée en complément des résultats présentés dans l'article. L'interaction des hélices TM est souvent médiée par les motifs GxxxG ou de type GxxxG (une glycine ou les deux sont remplacées par des petits résidus tels que les résidus alanine et sérine) (Cymer *et al.*, 2012; Kleiger *et al.*, 2002; Lemmon *et al.*, 1992; Moore *et al.*, 2008; Russ et Engelman, 1999; Senes *et al.*, 2000, 2004). Les interactions *via* GxxxG sont renforcées, de part et d'autre de ce motif, par des résidus dont la chaîne latérale est branchée au carbone  $\beta$  (L, V et I) et par de petits résidus (A, G, S). La position de ces résidus est très importante pour la stabilisation de l'interaction. Pour la GpA, le motif complet de dimérisation LIXXGVXXGVXXT a été identifié (Melnyk *et al.*, 2004).

Le segment TM de NccX, composé de 21 résidus, contient 11 résidus dont la chaîne latérale est branchée au carbone  $\beta$  et 7 petits résidus. Afin d'identifier des résidus potentiellement importants pour l'interaction des segments TM, des simulations ont été effectuées à partir de NccX-TMsim, par Norbert Garnier. Quatre simulations indépendantes ont été considérées comme représentatives de l'ensemble. Elles ont données des résultats différents, illustrant la liberté d'interaction de ce système (NccX-TMsim). Par ailleurs, chaque simulation révèle que les hélices TM de NccX-TMsim interagissent de manière dynamique. En effet, lorsque deux modèles d'hélices NccX-TMsim sont disposés à 6 nm l'un de l'autre dans une bicouche d'un phospholipide, le POPC (1-Palmitoyl-2-OleoylPhosphatidylCholine), les hélices interagissent quasi instantanément (0,2  $\mu$ s) (**Figure 23A**) et de façon dynamique : la distribution des angles de croisement des hélices indique que les hélices se croisent *via* trois angles majoritaires (**Figure 23B**). Les hélices TM de NccX-TMsim se croisent principalement *via* deux angles négatifs ( $\Omega$  :  $-35^\circ$  et  $-20^\circ$ ) et un angle positif ( $\Omega$  :  $+10^\circ$ ). La valeur de ces angles est comparable avec les mesures effectuées dans d'autres protéines TM obtenues par simulations suggérant une topologie similaire des hélices TM (**Tableau 3**).



**Figure 23** : Simulation gros grains de l'interaction des hélices NccX-TMsim. Deux monomères ont été insérés dans une bicouche de POPC, le long de l'axe Z qui est normal au plan de la bicouche lipidique, à une distance de 6 nm l'un de l'autre. **A** : Représentation de la distance entre les deux hélices en fonction du temps pour l'une des 4 simulations effectuées. Les hélices viennent au contact l'une de l'autre quasi instantanément (moins de 2  $\mu$ s). **B** : Représentation de la distribution des angles de croisement des hélices de NccX mesurée pour les 4 simulations effectuées indépendamment.

**Tableau 3** : Exemples d'angles de croisement d'hélices TM

Protéine	Etat oligomérique	Angles de croisement majoritaires ( $\Omega$ )	Références
<b>Glycophorine A</b>	homodimère	-25°	(Psachoulia <i>et al.</i> , 2010)
<b>Syndecan-2</b>	homodimère	-30°	(Psachoulia <i>et al.</i> , 2010)
<b><math>\alpha</math>IIb/<math>\beta</math>3</b>	hétérodimère	-25° et +28°	(Psachoulia <i>et al.</i> , 2010)
<b>DAP12</b>	homodimère	-35° et +20°	(Wei <i>et al.</i> , 2014a)
<b>Plexin-A1</b>	homodimère	-35°, -33° et +20°	(Aci-Sèche <i>et al.</i> , 2014)
<b>NccX</b>	homodimère	-35°, -20° et +10°	(Ziani <i>et al.</i> , 2014)

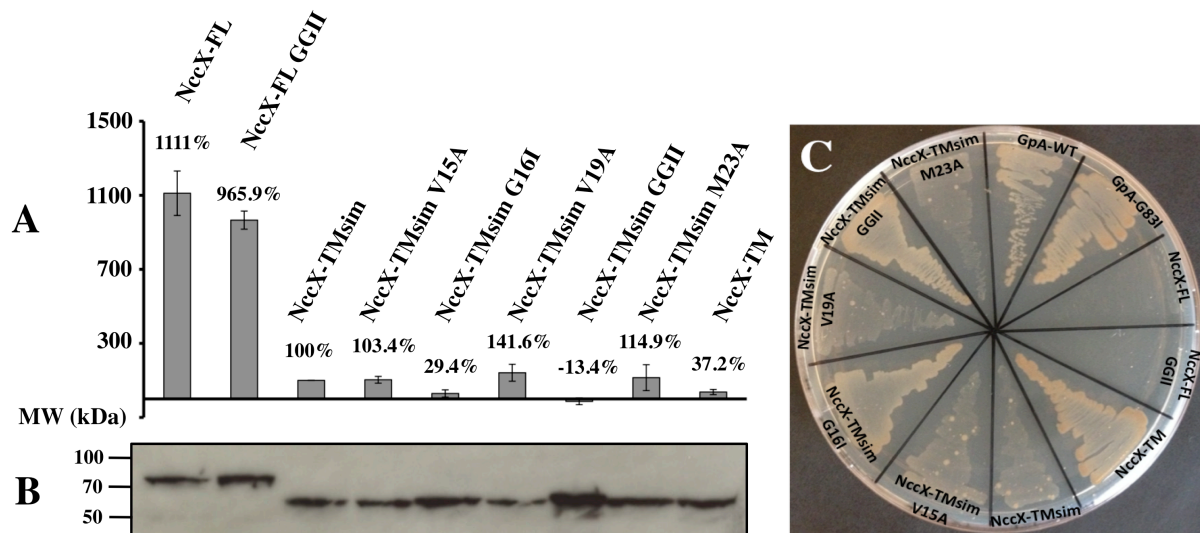
L'analyse de ces 4 simulations a permis de détecter les paires de résidus (chaque résidu sur un segment TM) dont la distance était la plus courte et donc les résidus qui forment l'interface

lors de l'homodimérisation de NccX-TMsim. Le **tableau 4** représente le bilan des 4 simulations.

**Tableau 4** : Liste des paires des résidus dont la distance (entre parenthèse) est la plus courte pour les principaux angles de croisement formés par les hélices.

Angle de croisement ( $\Omega$ )		
$\Omega = -35^\circ$ (0,5 - 0,6 nm)	$\Omega = -20^\circ$ (0,5 - 0,6 nm)	$\Omega = +10^\circ$ (0,6 - 0,7 nm)
G16 G16	G16 L14	G16 S12
G20 G20	G20 L14	G16 V15
A24 A24	G20 L18	V19 V15
	M23 L18	M23 V19
	G27 L22	M23 L22
		V30 V26
		V30 Y19
		T34 S33

Les résidus V15, G16, V19, G20 et M23 détectés comme importants lors de la première simulation ont été remplacés par une alanine. Leur rôle dans l'homodimérisation de NccX-TMsim a ensuite été testé par TOXCAT. Les activités CAT mesurées pour les mutants V15A, V19A et M23A de NccX-TMsim sont similaires à celle de NccX-TMsim indiquant qu'aucun de ces résidus pris isolément n'est déterminant pour le rapprochement des segments TM de NccX (**Figure 24**).



**Figure 24 :** L'homodimérisation de NccX s'effectue principalement par ses domaines périplasmiques. **A :** Mesure de l'activité CAT induite par les constructions ToxR-NccX-MBP. Les résidus révélés lors de la première simulation comme importants pour l'homodimérisation ont été mutés en alanines. Seuls les glycines du motif GxxxG semblent être importants pour l'interaction des domaines TM. Les activités CAT sont normalisées par rapport à celle mesurée pour NccX-TMsim. **B :** Test de complémentation *malE* pour vérifier la topologie des différentes protéines chimériques. Sur un milieu minimum M9, où le maltose est la seule source carbonée, seules les cellules exprimant la MBP dans le périplasma et donc la chimère correctement orientée survivent. Les protéines chimériques sont correctement orientées et insérées dans la membrane. La topologie des deux chimères contenant NccX a été testée indépendamment par fractionnement cellulaire (**Figure 21E**). Les résultats indiquent que NccX est correctement orientée dans la membrane.

Par ailleurs, la longueur du segment TM et les résidus qui font la jonction entre ce segment TM, la MBP et ToxR influent sur le niveau d'expression de l'enzyme CAT dans le test TOXCAT (Russ et Engelman, 1999; Wei *et al.*, 2011). Il semble que plus la longueur de l'hélice TM est importante, plus l'activité CAT est importante. Toutefois, cette activité peut être la conséquence d'une forte expression de la protéine chimérique. Il s'agit donc de faire un compromis entre le niveau d'expression de CAT et la production de la protéine chimérique qui sont des paramètres dépendant de la protéine d'intérêt. Nous avons donc mesuré l'activité CAT de la protéine chimérique contenant un segment TM de NccX (NccX-TM, L9-Y29), plus petit que NccX-TMsim (T6-S40). L'activité mesurée pour NccX-TM (normalisée par rapport à la quantité de chimères produites) est plus de deux fois inférieure à celle de NccX-TMsim (**Figure 24A**). Ce résultat confirme que l'activité CAT est sensible à la longueur du segment TM inséré dans la protéine chimérique TOXCAT et indique que les résidus de part et d'autre du segment TM seraient importants pour le rapprochement des domaines TM.

Enfin, la longueur de NccX-TMsim est plus adéquate que celle de NccX-TM pour observer des différences d'activités mesurées avec les chimères de mutants de NccX-TMsim et c'est pour cela que nous avons choisi NccX-TMsim comme référence pour l'étude de l'interaction des domaines TM de NccX.

A l'issue des 4 simulations (**Tableau 4**), d'autres résidus sont apparus comme importants pour l'interaction de NccX-TMsim. Cependant, leur rôle n'a pas été testé *in vivo* puisque nos résultats indiquent que l'homodimérisation de NccX est majoritairement guidée par le domaine périplasmique.

La description du senseur NccX, de son organisation dans la membrane était l'objectif principal de ce travail. Une fois constatée l'incohérence de la structure cristallographique obtenue en présence de DPC avec celle de CnrXs, déjà connue et déjà validée par des approches physiologiques et biophysiques, une analyse approfondie des structures obtenues et la mise en œuvre de quelques approches expérimentales complémentaires (*in vivo* et *in silico*) a permis de trouver un sens à la structure déterminée en présence de DPC et d'appuyer un message de prudence à l'égard des résultats générés en présence de détergent.

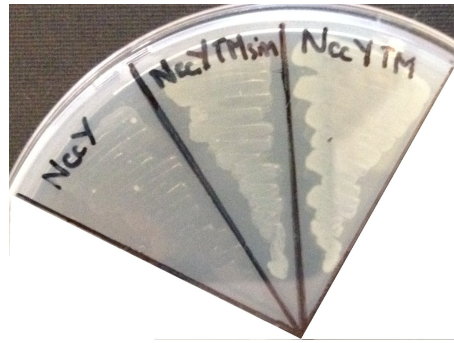
## D. Homodimérisation de NccY

Le signal, perçu par NccX dans le périplasme, est transmis jusqu'à NccH, dans le cytoplasme. NccY, qui inhibe NccH en la séquestrant à la membrane, est donc essentielle à la propagation du signal du périplasme jusqu'au cytoplasme. Cela suggère que NccY doit interagir avec NccX au niveau périplasmique et/ou membranaire. Dans un premier temps, nous avons testé l'homodimérisation de NccY en absence de NccX *in vivo*, par le test TOXCAT. Nous avons utilisé deux fragments de NccY en plus de la protéine complète : un fragment correspondant au segment TM (NccY-TM), et un autre pour lequel le segment TM est flanqué de quelques résidus (NccY-TMsim) (**Figure 25**). Le protocole expérimental est identique à celui qui a été utilisé pour étudier l'homodimérisation de NccX.

```
      10      20      30      40      50      60      70
MENIDEWL VQ AKKVTYEASR EPGLGRIQQR LSREPSQMV L TARHDILRAV CCAALASLVA FTAIDRIAVG LYQKQ
      80      90      95
QPTWV AAPSAASPFG LLIGK
```

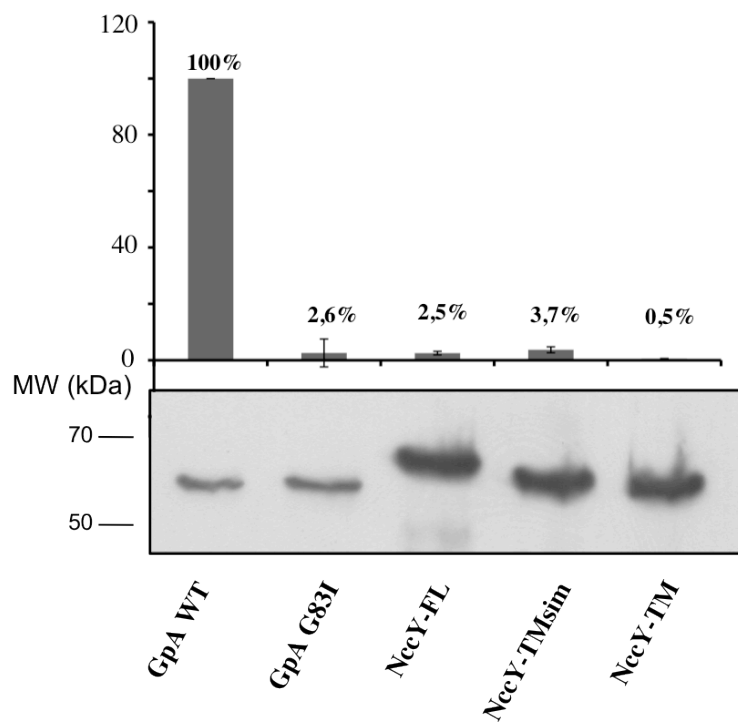
**Figure 25** : Séquence protéique de NccY. Les fragments de NccY (NccY-TMsim et NccY-TM) testés par TOXCAT sont colorés en orange. La séquence correspondant à NccY-TM est en gras et soulignée.

Les résultats du test de complémentation *malE* indiquent que la MBP est localisée dans le périplasme et donc que toutes les protéines chimériques sont correctement orientées (**Figure 26**).



**Figure 26 :** Test de complémentation *maleE* pour vérifier la topologie des protéines chimériques ToxR-NccY-MBP. Sur un milieu minimum M9, où le maltose est la seule source carbonée, seules les cellules exprimant la MBP dans le périplasma et donc la chimère correctement orientée survivent.

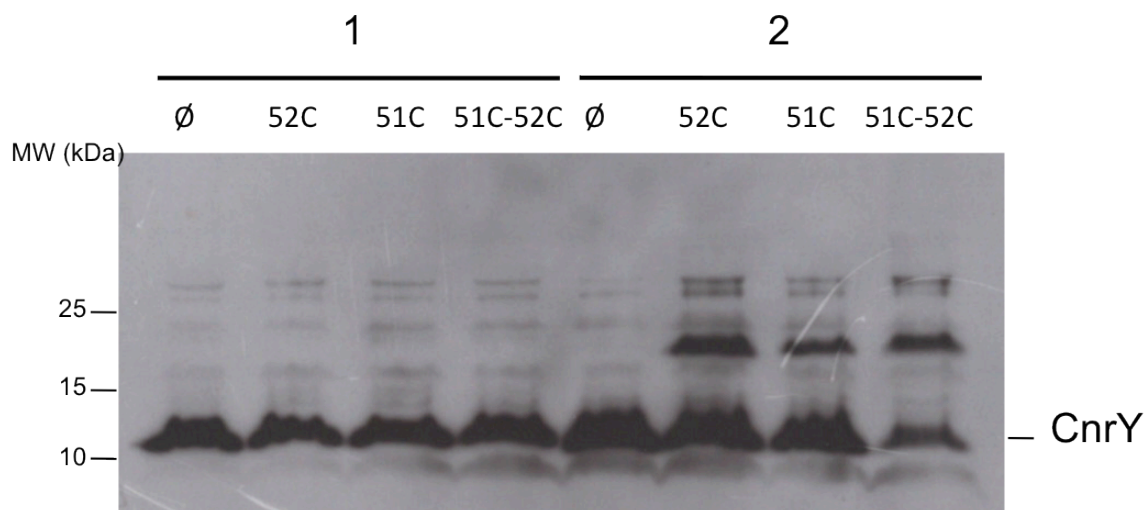
Les activités spécifiques de la CAT ont été normalisées par rapport à celle du contrôle positif, GpA WT. Elles sont similaires à l'activité spécifique CAT du contrôle négatif, GpA G83I, indiquant que l'homodimère de NccY, s'il existe, ne se forme pas spontanément *in vivo* (Figure 27).



**Figure 27 :** Activités spécifiques CAT mesurées pour les chimères ToxR-NccY-MBP. Ces activités sont normalisées par rapport à celle de la chimère GpA WT (contrôle positif) et par rapport à la quantité de protéines chimériques produites détectée par Western Blot anti-MBP. L'activité CAT mesurée pour la protéine chimérique GpA G83I sert de contrôle négatif.

## E. Interaction entre NccX et NccY

Alors que les résultats obtenus avec le test TOXCAT indiquent sans ambiguïté que NccY est produite sous forme monomérique dans la membrane, les cystéines 51 ou 52 de deux CnrY sont suffisamment proches pour former un pont disulfure en présence de CnrX (**Figure 28**). Cela suggère que la présence du partenaire membranaire dans le complexe favorise l'interaction des protéines CnrY. Les similarités de séquence, de fonction et de structure permettent de penser que cela est possible aussi dans le cas du système Ncc. Pour le vérifier, nous avons donc coproduit NccX et la chimère ToxR-NccY-MBP afin d'étudier l'influence de NccX sur l'organisation de NccY (**Figure 29a**).



**Figure 28 :** Mise en évidence du rapprochement de deux protéines CnrY par "cysteine-scanning". Les expériences de pontage chimique ont été effectuées à partir d'une souche d'*E. coli* surexprimant *his-cnrYX*. CnrY possède 3 cystéines : une dans son domaine cytoplasmique (C30) et deux au niveau du segment TM (C51 et C52). Quatre mutants de CnrY ont été testés. Pour tous les mutants, la cystéine 30 a été mutée en sérine. En plus de cette mutation (mutant 51C-52C), la cystéine 51 (52C) ou la cystéine 52 (51C) ou les deux (Ø) ont été mutées en sérine. Après induction à l'IPTG, les bactéries ont été traitées avec 2 mM de  $\text{CuSO}_4$  pendant 15 minutes (2) ou pas (1) avant que les groupes thiols ne soient alkylés avec du N-ethylmaleimide à la concentration de 0,26 g/L. Les protéines du lysat ont été analysées par Western Blot anti-his. CnrY est détectée quel que soit le mutant et les conditions expérimentales. Une bande à environ 20 kDa correspondant au dimère de CnrY est détectée lorsque les cellules sont traitées avec du  $\text{CuSO}_4$  (2) et lorsque la cystéine 51 ou 52 ou les deux sont présentes dans la séquence de CnrY (52C, 51C, 51C-52C, respectivement).

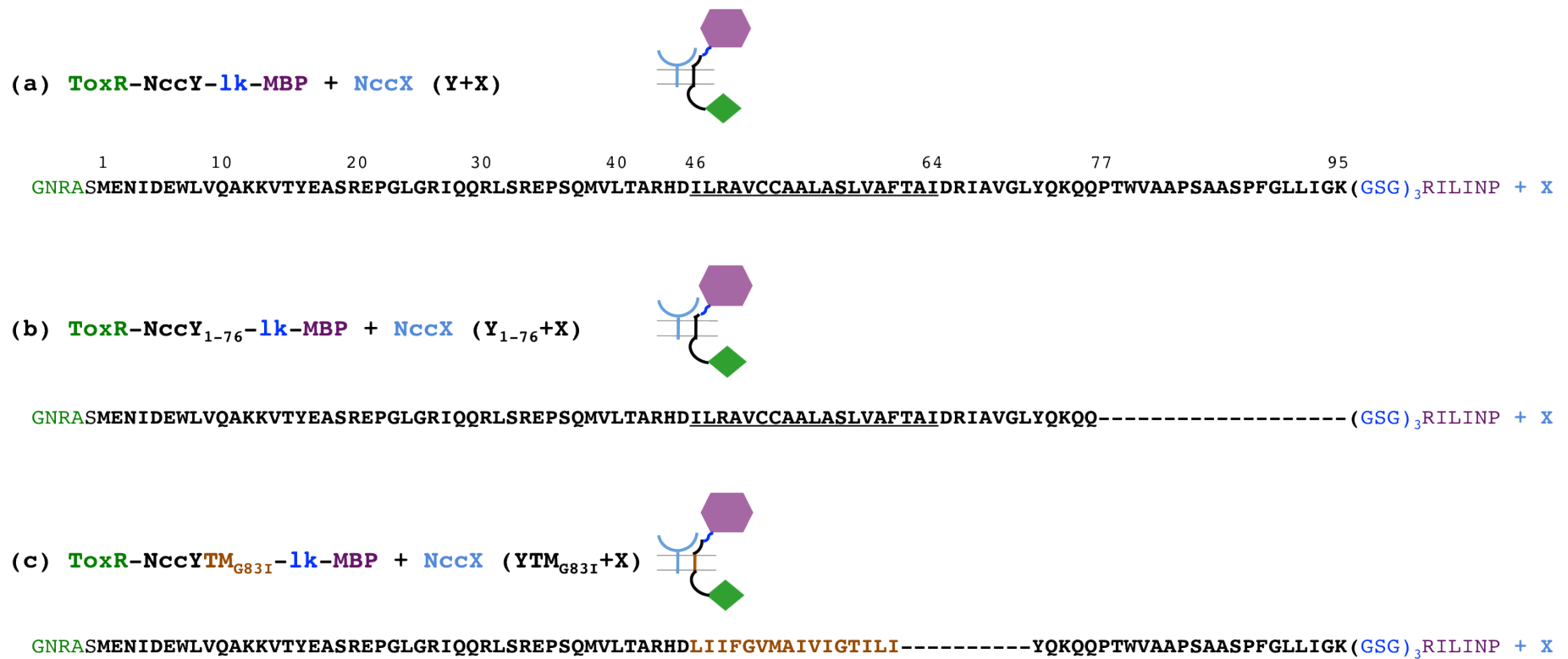
L'activité CAT mesurée pour ToxR-NccY-MBP en présence de NccX (Y+X) est de 0,083  $\Delta A_{412}$ /min/mg alors que celle de la chimère contenant NccY, qui rappelons-le ne



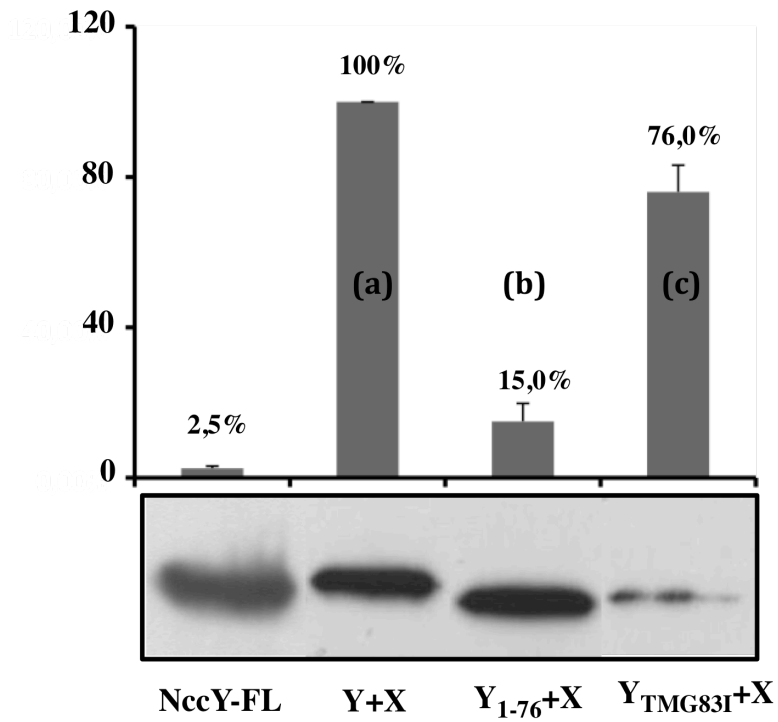
dimérise pas, est de  $0,004 \Delta A_{412}/\text{min}/\text{mg}$ . L'activité CAT est plus 20 fois plus importante lorsque la chimère de NccY est coproduite avec NccX. Cela indique donc que deux protéines NccY sont rapprochées en présence de NccX, et donc qu'au moins deux molécules de NccY interagissent avec un dimère de NccX. Par la suite, nous nous sommes servis de Y+X comme contrôle positif d'interaction (**Figure 30**).

Comme l'activité CAT mesurée avec ToxR-NccY-MBP dépend de sa capacité à interagir avec NccX, nous disposons d'un moyen d'identifier les déterminants de l'interaction entre NccX et NccY. L'interaction entre NccX et NccY peut avoir lieu dans le périplasma et/ou dans la membrane. Afin de savoir si l'interaction s'effectuait au niveau périplasmique, nous avons supprimé dans la protéine chimérique ToxR-NccY-MBP, une grande partie du domaine périplasmique de NccY et testé sa capacité à se rapprocher en présence de NccX. Lorsque le domaine périplasmique de NccY a été tronqué de 19 résidus et qu'il ne reste donc que 12 résidus périplasmiques (Y<sub>1-76</sub>+X) (**Figure 29b**), l'activité CAT mesurée a été divisée par un facteur supérieur à 6 (**Figure 30b**), ce qui indique que ToxR-NccY<sub>1-76</sub>-MBP ne se rapproche plus en présence de NccX et donc que sans une grande partie de son domaine périplasmique, NccY n'interagit plus avec NccX. Le domaine périplasmique joue donc un rôle majeur dans son interaction avec NccX.

La question suivante était de savoir si l'interaction entre NccY et NccX s'effectuait seulement dans le périplasma ou si les segments TM interagissaient aussi. Pour vérifier si le domaine périplasmique de NccY, en plus d'être nécessaire, est suffisant pour l'interaction avec NccX, nous avons remplacé le segment TM de NccY par un segment TM connu pour être monomérique. Nous avons ainsi remplacé le domaine TM de NccY et 5 résidus juxtamembranaires de NccY par le domaine TM de GpA G83I (**Figure 29c**). L'activité mesurée pour cette protéine chimérique (ToxR-Y<sub>TMG83I</sub>+X-MBP) (**Figure 30c**) est de 76%. L'activité est maintenue à un niveau comparable à celle du type sauvage (ToxR-Y+X-MBP) (**Figure 30a**). Il semble donc, qu'en présence de NccX, deux NccY se rapprochent dans la membrane bien que le segment TM de G83I, substitué au TM de NccY, ne soit pas prédisposé à dimériser ni à interagir avec NccX. On peut donc penser que le domaine TM intervient peu dans l'interaction entre NccX et NccY et donc que le domaine périplasmique de NccY est suffisant pour l'interaction avec NccX (**Figure 30**). Les domaines périplasmiques de NccX et de NccY semblent donc contenir les déterminants majeurs de l'interaction entre NccX et NccY.



**Figure 29** : Séquences protéiques et représentation schématique des chimères de NccY testées par TOXCAT.



**Figure 30 :** Activités spécifiques CAT mesurées pour les chimères ToxR-NccY-MBP en absence (NccY-FL) ou en présence de NccX (**a**, **b** et **c**). Ces activités sont normalisées par rapport à celle de ToxR-NccY-MBP + NccX (Y+X) et par rapport à la quantité de protéines chimériques produite, détectée par Western Blot anti-MBP.

# III. Etude de l'interaction entre CnrX et CnrY dans le périplasma

## A. Mise en évidence de l'interaction entre les parties périplasmique de CnrX et CnrY

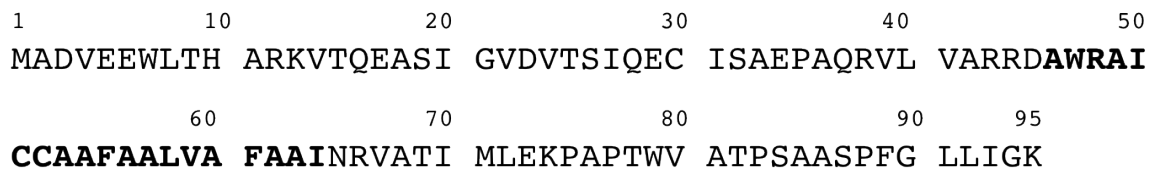
Les résultats précédents (partie II.E.) démontrent que l'interaction entre NccX et NccY s'effectue essentiellement au niveau périplasmique. Or, CnrX/NccX doit transmettre le signal à CnrY/NccY qui est le seul lien entre le senseur de stress (CnrX/NccX) et le régulateur de réponse (CnrH/NccH). L'étude de l'interaction entre CnrX/NccX et CnrY/NccY dans le périplasma apparaît donc être une étape importante pour une meilleure compréhension de tels systèmes de signalisation. Afin d'étudier l'interaction entre les protéines CnrX/NccX et CnrY/NccY dans le périplasma, nous avons à nouveau basculé sur le système Cnr pour continuer la caractérisation, déjà bien avancée, du domaine périplasmique de CnrX (CnrXs). Nous avons aussi utilisé le domaine périplasmique de CnrY (CnrYp).

Avant mon arrivée dans le laboratoire, la protéine CnrXs (résidus 31-148, 13,3 kDa) représentant le domaine senseur périplasmique soluble de CnrX avait été produite dans *E. coli*, purifiée et cristallisée en complexe avec plusieurs métaux (Pompidor *et al.*, 2008; Trepreau *et al.*, 2011b) (**Introduction** partie III.D.4.c.i.). La protéine a été cristallisée sous forme de dimère. Chaque monomère, d'environ 13 kDa, est formé par quatre hélices  $\alpha$  (H1 à H4), le tout formant une épingle à cheveux et possède un site de fixation au métal formé par 5 résidus appartenant aux hélices H1, H2 et H3 (**Figure 31**).



**Figure 31** : Séquence et structure de CnrXs. **(a)** Séquence de CnrXs. La protéine a été tronquée et le premier résidu est la sérine 31. Les hélices H1, H2, H3 et H4 sont colorées en cyan, bleu, vert et orange, respectivement. **(b)** Structure de CnrXs (code PDB : 2y3d). Bien que présents, les résidus 31 à 38 ne sont pas déterminés dans la structure. Le code couleur est identique à celui de (a).

CnrY est une protéine bitopique composée d'un domaine cytosolique (résidus 1 à 45), d'un segment TM prédit (résidus 46 à 64) et d'une partie périplasmique (CnrYp, résidus 65 à 95) (**Figure 32**).



**Figure 32** : Séquence protéique de CnrY. La partie périplasmique (CnrYp) est soulignée, le segment TM, en gras. Le reste de la séquence forme le domaine cytosolique.

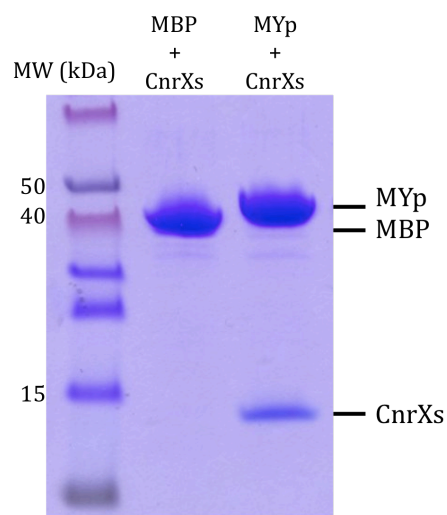
CnrYp compte 31 résidus dont 21 sont hydrophobes ce qui suggère que cette protéine est très peu soluble dans une solution aqueuse. Afin de pouvoir la produire sous forme soluble et la purifier, nous avons fusionné l'extrémité N-terminale de CnrYp à l'extrémité C-terminale de la Maltose Binding Protein (MBP) qui est un des meilleurs acteurs de solubilité décrits (Fox *et al.*, 2003). Entre les séquences codant ces deux protéines, nous avons introduit la séquence codant un site de clivage reconnu par la protéase du "Tobacco Etch Virus" (TEV) dans le cas où on voudrait isoler CnrYp. La séquence de la protéine de fusion MBP-CnrYp (MYp), dont la masse moléculaire est d'environ 45 kDa, est illustrée dans la **figure 33**.

MKTEEGKLV I W I N G D K G Y N G L A E V G K K F E K D T G I K V T V E H P D K L E E K F P Q V A A T G D G P D I I F  
W A H D R F G G Y A Q S G L L A E I T P D K A F Q D K L Y P F T W D A V R Y N G K L I A Y P I A V E A L S L I Y N K D L L P  
N P P K T W E E I P A L D K E L K A K G K S A L M F N L Q E P Y F T W P L I A A D G G Y A F K Y E N G K Y D I K D V G V D N  
A G A K A G L T F L V D L I K N K H M N A D T D Y S I A E A A F N K G E T A M T I N G P W A W S N I D T S K V N Y G V T V L  
P T F K G Q P S K P F V G V L S A G I N A A S P N K E L A K E F L E N Y L L T D E G L E A V N K D K P L G A V A L K S Y E E  
E L A K D P R I A A T M E N A Q K G E I M P N I P Q M S A F W Y A V R T A V I N A A S G R Q T V D E A L K D A Q T D Y D I P  
T T E N L Y F Q G H **M N R V A T I M L E K P A P T W W A T P S A A S P F G L L I G K**

**Figure 33** : Séquence protéique de MYp, protéine de fusion entre la MBP et CnrYp. La séquence correspondant à la MBP est colorée en violet et celle de CnrYp en orange. La séquence de liaison issue de la polyprotéine du TEV, contenant le site de clivage protéolytique est en noir.

Pour mettre en évidence *in vitro* l'interaction entre CnrXs et CnrYp, nous avons procédé par coprécipitation sur des billes d'amylose. L'amylose interagit spécifiquement avec la MBP. On teste donc l'interaction directe entre CnrYp et CnrXs en suivant la capacité de la MYp à retenir CnrXs sur la résine d'amylose. Pour ces expériences, des lysats clarifiés de bactéries ayant surproduit un dérivé de la MBP (MBP, MYp, mutant) ou CnrXs ont été mélangés à une suspension de billes d'amylose. Ces résultats ont été reproduits avec les protéines CnrXs et MYp pures mais l'étude de l'interaction CnrXs:CnrYp a été poursuivie en utilisant les lysats de bactéries, à la fois par souci de simplicité et pour constater que la spécificité de l'interaction permet de purifier CnrXs dans un mélange complexe. Toutes les expériences de coprécipitation présentées ci-dessous ont été effectuées au moins trois fois.

La **figure 34** représente un gel SDS-PAGE sur lequel les protéines retenues sur les billes d'amylose ont été déposées. La MBP a migré sous forme d'une bande à environ 40 kDa et MYp sous forme d'une bande à environ 45 kDa. L'intensité des bandes correspondant à la MBP et à MYp permet de contrôler le caractère quantitatif de l'expérience : ces protéines étant présentes en quantités équivalentes, les quantités de CnrXs coprécipitée (bande à environ 13 kDa) peuvent être comparées directement. Ces expériences montrent que CnrXs coprécipite avec MYp et pas avec la MBP. Cela signifie donc qu'il n'y a pas d'interaction non spécifique entre la MBP et CnrXs et que CnrYp interagit avec CnrXs. Ces résultats confirment ceux obtenus avec NccY et NccX *in vivo* (partie **II.E**).



**Figure 34** : Coprécipitation de CnrXs avec la MBP ou la MYp. La piste de gauche contient les marqueurs de masse moléculaire.

## B. Caractérisation de l'interaction entre CnrYp et CnrXs

### 1. Identification du déterminant de l'interaction de CnrYp

Pour identifier le ou les déterminants d'interaction de CnrYp avec CnrXs, plusieurs approches étaient possibles : le pontage chimique, la coprécipitation, la résonance plasmonique de surface, l'échange hydrogène/deutérium couplé à la spectrométrie de masse, la RMN, la cristallographie... A la suite de la mise en évidence de l'interaction entre les domaines périplasmiques de CnrX et CnrY, nous avons commencé par utiliser la méthode de coprécipitation après avoir modifié la séquence de CnrYp par mutagenèse. Cette approche est favorisée par la taille relativement courte de CnrYp et son manque présumé de structure secondaire (l'extrémité C-terminale pourrait cependant former une hélice). L'alignement de la séquence de CnrY avec des séquences homologues (**Introduction Figure 19**) met en évidence la conservation des résidus en C-terminal de CnrYp et notamment celle des deux leucines L<sub>91</sub>L<sub>92</sub> de CnrY. Nous avons alors testé l'interaction de CnrXs avec deux mutants de

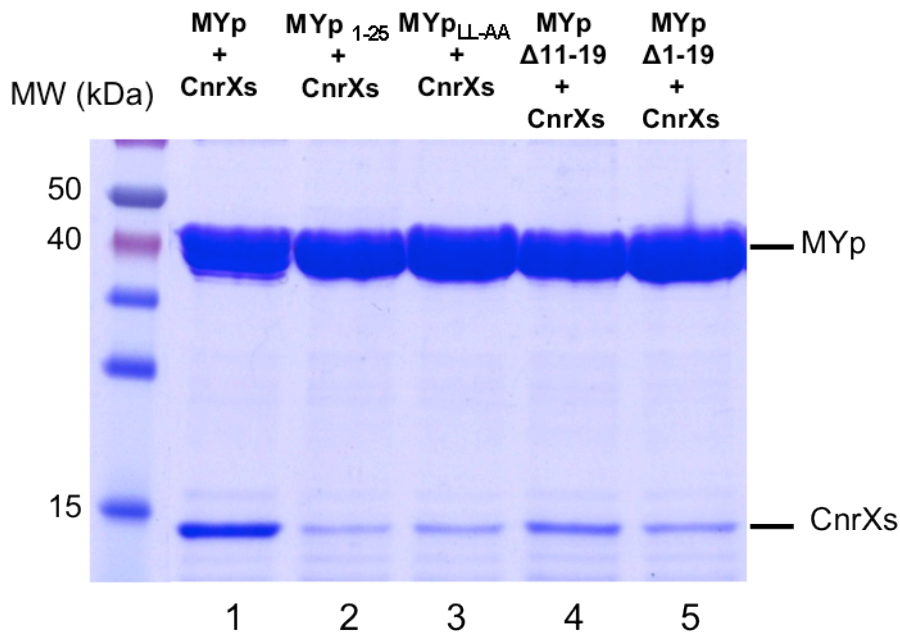
MYp : une forme tronquée, MYp<sub>1-25</sub>, où la portion 3' du gène *cnrYp*, codant les 6 résidus (GLLIGK) de l'extrémité C-terminale a été supprimée et MYp<sub>LL-AA</sub> où les 2 leucines L<sub>91</sub>L<sub>92</sub> ont été remplacées par deux alanines (**Figure 35**). Le résultat des expériences de coprécipitation réalisées avec ces mutants est présenté dans la **figure 36**.



**Figure 35** : Séquence de MYp et de ses dérivés testés par coprécipitation.

Les bandes correspondant à CnrXs sur le gel SDS-PAGE sont 8 et 7 fois moins intenses pour les mutants MYp<sub>1-25</sub> et MYp<sub>LL-AA</sub>, respectivement (**Figure 36, pistes 2 et 3**). L'interaction entre les mutants de MYp et CnrXs est fortement perturbée. Les 6 résidus de l'extrémité carboxyterminale (GLLIGK) et plus précisément les 2 leucines sont donc nécessaires pour l'interaction avec CnrXs.





**Figure 36** : Coprécipitation de CnrXs avec MYp ou ses dérivés. La piste de gauche contient les marqueurs de masse moléculaire.

Afin de savoir si les résidus GLLIGK sont suffisants pour l'interaction avec CnrXs, nous avons testé, toujours par coprécipitation, deux dérivés de la MYp dans lesquels les résidus en amont de ces acides aminés ont été remplacés par une séquence contenant le motif glycine-serine-glycine (GSG) répété plusieurs fois. Pour le mutant MYp  $\Delta$ 11-19, les résidus 11 à 19 de CnrYp ont été remplacés par la séquence de liaison (GSG)<sub>3</sub>, de même longueur. Le mutant MYp  $\Delta$ 1-19 a été construit en remplaçant les résidus 1 à 19 par le linker (GSG)<sub>6</sub>G. Nous avons choisi de ne pas substituer les 6 résidus en amont de GLLIGK (SAASPF) pour éviter de déstabiliser l'éventuelle structure secondaire de l'extrémité C-terminale (**Figure 35**).

L'intensité de la bande de CnrXs coprécipitée avec MYp  $\Delta$ 11-19 et MYp  $\Delta$ 1-19 (**Figure 36, pistes 4 et 5**) est respectivement 3 et 5 fois moindre que celle de CnrXs coprécipitée avec MYp de type sauvage (**Figure 36, piste 1**). Ce résultat indique que la substitution des résidus 11 à 19 et 1 à 19 de CnrYp déstabilise l'interaction avec CnrXs et que l'interaction est d'autant plus déstabilisée que le nombre de résidus substitués est important. La présence des 12 derniers résidus de l'extrémité carboxyterminale de CnrYp, comprenant notamment le motif GLLIGK n'est donc pas suffisante pour l'interaction avec CnrXs.

## 2. Modulation de l'interaction entre CnrYp et CnrXs par les métaux

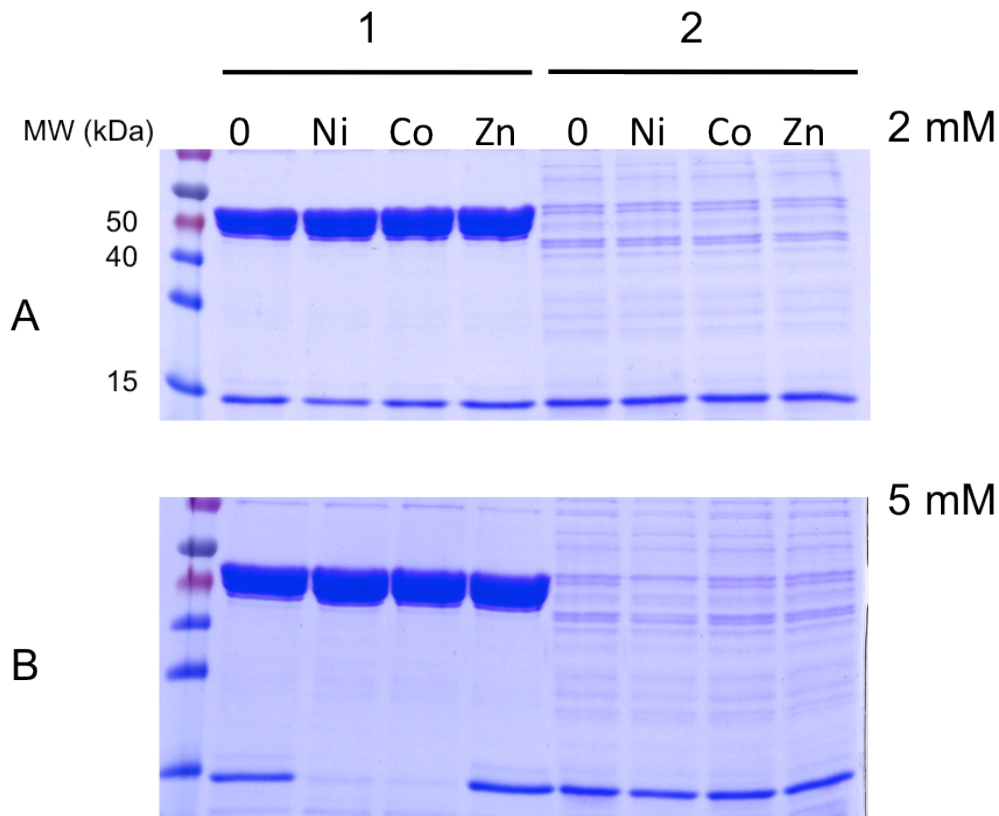
Rappelons que *Cupriavidus metallidurans* CH34 est capable de résister au nickel et au cobalt grâce à une pompe d'efflux dont la production est spécifiquement induite en présence d'un excès de nickel, de cobalt ou, dans une moindre mesure, de cuivre. Par ailleurs, du zinc est retrouvé dans la forme inactive de la protéine CnrX. La partie périplasmique du senseur métallique, CnrXs, perçoit l'excès de métaux. La façon dont CnrX alerte CnrY de la présence de métal est encore inconnue. Après avoir montré que les régions périplasmiques de ces deux protéines interagissent, nous avons vérifié l'effet des métaux sur l'interaction entre CnrXs et CnrYp.

Le protocole de coprécipitation a été adapté pour tester l'effet des métaux sur l'interaction entre CnrXs et CnrYp. Les différents sels de métaux lourds ont été ajoutés au lysat clarifié des bactéries surproduisant CnrXs. Le lysat de CnrXs a aussi été utilisé sans métaux en guise de contrôle positif. Après 15 min d'incubation, le lysat en présence ou en absence de métaux est centrifugé afin de sédimenter les éventuels précipités de protéines. Le surnageant de ce lysat est ensuite ajouté aux protéines MYp retenues par les billes d'amylose (amylose:MYp). La quantité de CnrXs présente dans ce surnageant a été vérifiée sur gel SDS-PAGE (**Figure 37\_2**). Par la suite, le protocole est similaire à celui utilisé précédemment.

Tout d'abord, l'expérience a été effectuée avec 2 mM de nickel, de cobalt et de zinc (**Figure 37A**). Certains métaux tels que le zinc sont connus pour faire précipiter les protéines. Dans notre expérience, la quantité de CnrXs ajoutée aux billes d'amylose est la même quelque soit le métal utilisé (**Figure 37A\_2**). Ainsi, nous pouvons comparer les quantités de CnrXs retenues sur la résine d'amylose.

On détecte une bande, correspondant à CnrXs, d'intensité similaire quel que soit le métal utilisé mais aussi sans métal (**Figure 37\_1**). Cela suggère que ces métaux n'ont aucun effet sur l'interaction entre CnrXs et CnrYp ou que le seuil de sensibilité de CnrXs n'a pas été atteint. Pour vérifier cette hypothèse, nous avons augmenté la concentration d'ions métalliques à 5 mM (**Figure 37B\_1**). Cette concentration ne précipite pas CnrXs, qui est en quantité comparable quel que soit le métal utilisé (**Figure 37B\_2**). A cette concentration, le signal correspondant à CnrXs disparaît lorsque CnrXs est incubé avec du nickel et du cobalt alors

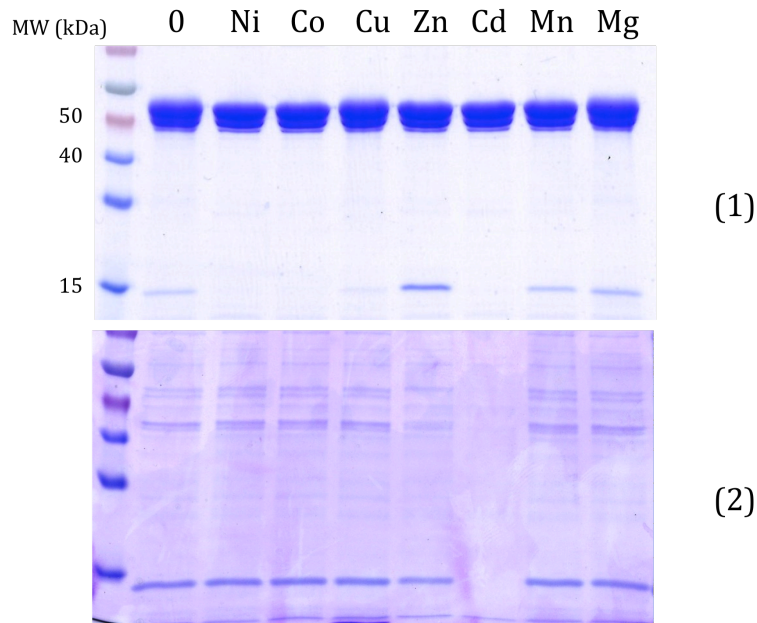
qu'il est similaire au contrôle positif lorsque CnrXs est incubé avec du zinc. Les résultats indiquent d'une part que la concentration des métaux réelle disponible pour CnrXs lorsqu'ils sont ajoutés au lysat bactérien riche en protéines est surestimée et que le nickel et le cobalt empêchent l'interaction entre CnrXs et CnrYp contrairement au zinc qui ne déstabilise pas l'interaction.



**Figure 37** : Effet des métaux sur la coprécipitation de CnrXs avec MYp. Les pistes de gauche (**1**) correspondent aux protéines retenues sur les billes d'amylose après incubation du lysat de CnrXs avec soit 2 mM (**A**) soit 5 mM (**B**) de nickel sous forme  $\text{NiSO}_4$  (Ni), de cobalt,  $\text{CoCl}_2$  (Co) et de zinc,  $\text{ZnSO}_4$  (Zn). Sur les pistes 0, CnrXs n'est pas incubé avec des métaux (contrôle positif). Les pistes de droite (**2**) correspondent aux protéines du surnageant de lysat de CnrXs incubé avec les différents métaux.

L'expérience précédente a été renouvelée afin de tester d'autres métaux lourds divalents : le cuivre, le cadmium, le manganèse et le magnésium. Sur la **figure 38**, nous pouvons observer qu'il n'y a pas de signal correspondant à CnrXs avec le nickel, le cobalt alors qu'une même quantité de protéine CnrXs a été ajoutée aux billes d'amylose:MYp

(**Figure 38\_2**). Aucune bande correspondant à CnrXs n'est détectée avec le cadmium. Cependant, il apparaît que le cadmium précipite les protéines (**Figure 38\_2**). Avec le cuivre, le signal est plus faible (environ trois fois) que le contrôle. Des bandes d'intensités similaires au contrôle sont détectées avec le manganèse et le magnésium alors qu'avec le zinc, l'intensité mesurée est environ trois fois plus importante que celle du contrôle (**Figure 38**).



**Figure 38** : Effet des métaux sur la coprécipitation de CnrXs avec MYp. Le panneau (1) correspond au gel SDS-PAGE sur lequel a été déposé les protéines retenues sur les billes d'amylose, après incubation du lysat de CnrXs avec 5 mM de nickel ( $\text{NiSO}_4$ , Ni), de cobalt ( $\text{CoCl}_2$ , Co), de cuivre ( $\text{CuSO}_4$ , Cu), de zinc ( $\text{ZnSO}_4$ , Zn), de cadmium ( $\text{CdSO}_4$ , Cd), de manganèse ( $\text{MnSO}_4$ , Mn) et de magnésium ( $\text{MgSO}_4$ , Mg). Sur les pistes 0, CnrXs n'est pas incubé avec des métaux (contrôle positif). Le panneau (2) correspond au gel SDS-PAGE sur lequel a été déposé le surnageant de lysat de CnrXs incubé avec les différents métaux, en l'absence de MYp.

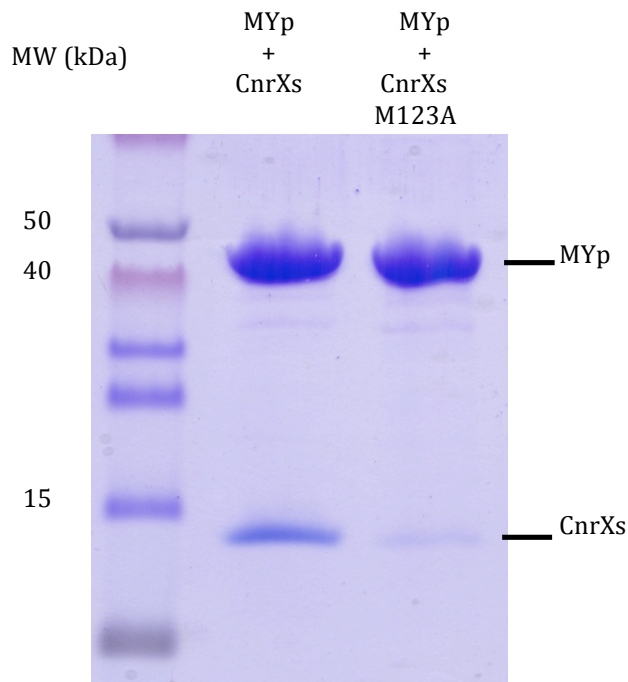
Il ressort de ces résultats que l'interaction entre CnrYp et CnrXs est bien modulée de façon spécifique par différents cations métalliques. Les métaux les plus inducteurs de la réponse au stress métallique induite par CnrYXH que sont le nickel et le cobalt ont l'effet le plus marqué sur l'interaction entre CnrXs et CnrYp. Le cuivre, dont le pouvoir d'induction est plus faible que le cobalt et le nickel, déstabilise l'interaction entre CnrXs et CnrYp contrairement au zinc (métal non inducteur) qui la stabilise. Le manganèse et le magnésium ne modifient pas l'interaction entre les deux protéines. Nous ne pouvons pas conclure sur l'effet du cadmium sur l'interaction puisqu'il précipite CnrXs à la concentration utilisée.

La conclusion de cette série d'expériences est que l'interaction entre CnrXs et CnrYp est favorisée en présence de zinc et déstabilisée en présence de nickel, de cobalt ou de cuivre. Nous savons que la substitution du zinc par le nickel ou le cobalt au site métallique de CnrX induit des modifications conformationnelles qui déclenchent probablement le processus de signalisation. Que ces changements de conformation affectent l'interaction entre les domaines périplasmiques de CnrX et CnrY est une hypothèse séduisante, dans la perspective de la transmission d'un signal de CnrX à CnrY.

### 3. Identification des régions de CnrXs interagissant avec CnrYp

#### a. Identification par coprécipitation

Le rôle de la méthionine M123 (ligand du cobalt, du cuivre et du nickel) a été défini précédemment (Trepreau *et al.*, 2011b, 2014). Il serait le déclencheur de la transduction du signal. Le mutant de CnrXs dans lequel la méthionine 123 est remplacée par une alanine (M123A) a été précédemment préparé au laboratoire. Nous avons testé sa capacité à interagir avec CnrYp par coprécipitation. MYp interagit beaucoup moins avec le mutant M123A qu'avec CnrXs non mutée (**Figure 39**). Comme la production de CnrXs-M123A est semblable à celle du type sauvage, cela peut signifier que M123 joue un rôle direct dans l'interaction avec CnrYp ou un rôle indirect, par exemple en altérant la structure de CnrXs. La structure cristallographique de CnrXs-M123A a été obtenue en présence de nickel et est superposable à la structure de CnrXs de type sauvage métallée par le nickel (RMSD calculé sur 106 atomes inférieur à 0,5 Å) (Trepreau *et al.*, 2014). Cela montre que la mutation n'affecte pas le repliement de CnrXs. Sur la base des données disponibles, et sachant que le résidu M123 est enfoui, on ne peut cependant pas exclure que la dynamique conformationnelle de CnrXs-M123A soit altérée : la conformation cristallisée dans les deux cas résulterait d'un phénomène de sélection pendant la cristallogenèse. Les deux hypothèses envisagées pour interpréter les données récoltées sur CnrXs-M13A sont donc considérées comme valides.



**Figure 39** : Coprécipitation de CnrXs ou CnrXs M123A avec MYp.

Dans l'étude de l'interaction dans le périplasma, nous venons de démontrer que CnrXs interagit avec CnrYp. L'extrémité C-terminale de CnrYp est nécessaire mais pas suffisante pour l'interaction. De plus, cette interaction est dépendante des métaux. En effet, la forme inactive de CnrXs en complexe avec du zinc interagit avec CnrYp. Au contraire, le cobalt, le nickel et dans une moindre mesure, le cuivre, déstabilisent l'interaction. Nous avons aussi montré le rôle important de la M123 de CnrXs dans l'interaction avec CnrYp mais la région de CnrXs avec laquelle CnrYp interagit est encore inconnue.

Afin d'identifier la ou les régions de CnrXs impliquées dans l'interaction avec CnrYp, nous avons d'abord essayé de cocristalliser un mélange équimolaire de CnrXs et de MYp. L'analyse des données de diffraction collectées avec ces cristaux (collaboration avec Anne Volbeda, IBS) a malheureusement révélé que seule la protéine CnrXs avait cristallisé. Cette approche aurait pu être poursuivie de différentes manières : changement de la stœchiométrie du mélange, mutation de la MYp pour réduire l'entropie de surface (Moon *et al.*, 2010) ou clivage de la protéine de fusion. Toutefois, compte tenu du temps disponible pour la

réalisation de mon projet de thèse, nous avons préféré une approche plus rapide et qui elle-aussi est très informative : la méthode d'échange Hydrogène-Deutérium associé à la spectrométrie de masse (HDX-MS).

## **b. Identification par échange Hydrogène/Deutérium (HDX)**

Le but de la méthode HDX-MS est de localiser la surface d'interaction entre deux molécules. Elle repose sur l'altération des conditions de l'échange des protons des liaisons peptidiques avec le solvant. Les protons présents dans les protéines sont échangeables plus ou moins rapidement selon la valeur du pH, le groupe chimique qui les porte et leur disponibilité. Lorsqu'une protéine protonée est diluée dans un tampon deutéré, les hydrogènes des chaînes latérales sont échangés rapidement tandis que ceux des liaisons amides sont échangés plus lentement. Le nombre de protons échangés peut-être mesuré par MS (pour un deutérium incorporé, la masse de la protéine augmente de 1 Da). Les hydrogènes liés aux atomes de carbone des acides aminés ne sont pas échangeables. La disponibilité d'un proton échangeable dépend de son accessibilité au solvant et de sa participation à des structures secondaires : les ions deutérium vont s'incorporer très rapidement dans des régions accessibles et peu structurées alors que l'échange est plus lent dans les régions structurées et/ou enfouies dans la protéine. Dès lors, la comparaison des vitesses d'échange des protons d'une protéine dans différentes conditions (conditions physicochimiques, mutations, en complexe) apporte des informations sur des changements de conformation ou d'accessibilité éventuels. Combinée à une protéolyse contrôlée, la technique HDX-MS permet d'obtenir une information localisée. C'est ainsi que la technique HDX-MS permet de caractériser les régions de protéines impliquées dans des interactions : après incubation dans un tampon deutéré, celles-ci présentent une incorporation relativement réduite de deutérium dans la forme complexée par rapport à la forme libre. Ce projet a bénéficié d'une collaboration avec Eric Forest (IBS).

Les expériences de HDX-MS ont été réalisées sur un complexe reconstitué avec des protéines purifiées. Dans ces expériences, une nouvelle construction de CnrXs a été utilisée, dans laquelle les résidus N-terminaux de CnrXs, non résolus dans la densité électronique, ont été supprimés. Les structures cristallographiques de CnrXs indiquent en effet que les résidus

S31-G37 (**Figure 31**) forment une extrémité flexible, dont l'agitation entropique pourrait perturber l'interaction avec MYp. Afin de favoriser l'interaction avec CnrYp, nous avons construit un mutant de CnrXs dans lequel les résidus S31-G37 de CnrXs ont été supprimés (CnrXs $\Delta$ Nter) (**Figure 40**).

```

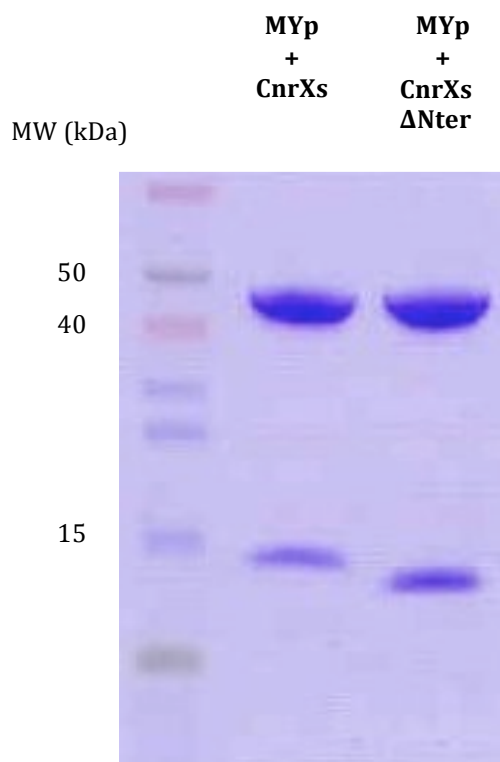
1           10           20           30           40           50           60
MHGDLHEILH EAVPLDANER EILELKEDAF AORRREIETR LRAANGKLAD AIAKNPAWSP EVEAAT
      H1                               H2
70           80           90           100           110 112
OEVE RAAGDLORAT LVHVFEMRAG LKPEHRPAYD RVLIDALRRG SQ
      H3                               H4

```

**Figure 40** : Séquence protéique de CnrXs $\Delta$ Nter. Les peptides correspondant aux 4 hélices (H1 à H4) sont soulignés. La numérotation correspond à celle qui sera utilisée par la suite pour les expériences de cartographie.

La capacité de CnrXs $\Delta$ Nter à interagir avec CnrYp a été testée par coprécipitation. La **figure 41** montre que CnrXs $\Delta$ Nter interagit environ 1,5 fois plus avec CnrYp que CnrXs. Nous avons donc choisi d'utiliser cette protéine pour les expériences d'échanges H/D. Le protocole de purification de CnrXs $\Delta$ Nter est similaire à celui de CnrXs (Pompidor *et al.*, 2008). Le test au 4-(2-pyridylazo)resorcinol (PAR) (McCall et Fierke, 2000) a permis de montrer qu'ainsi préparée, CnrXs $\Delta$ Nter est sous forme apo. En effet, aucune trace de nickel, cobalt, cuivre, zinc, cadmium ou manganèse n'a été détectée dans l'échantillon de protéine purifiée.





**Figure 41** : Coprécipitation de CnrXs ou CnrXs $\Delta$ Nter avec MYp.

Par ailleurs, la formation du complexe CnrXs $\Delta$ Nter:MYp, a été favorisée en introduisant le ligand (MYp) en excès par rapport à la cible (CnrXs). Des expériences ont été faites avec une stœchiométrie de 2 ou de 6 moles de MYp pour 1 mole de CnrXs $\Delta$ Nter.

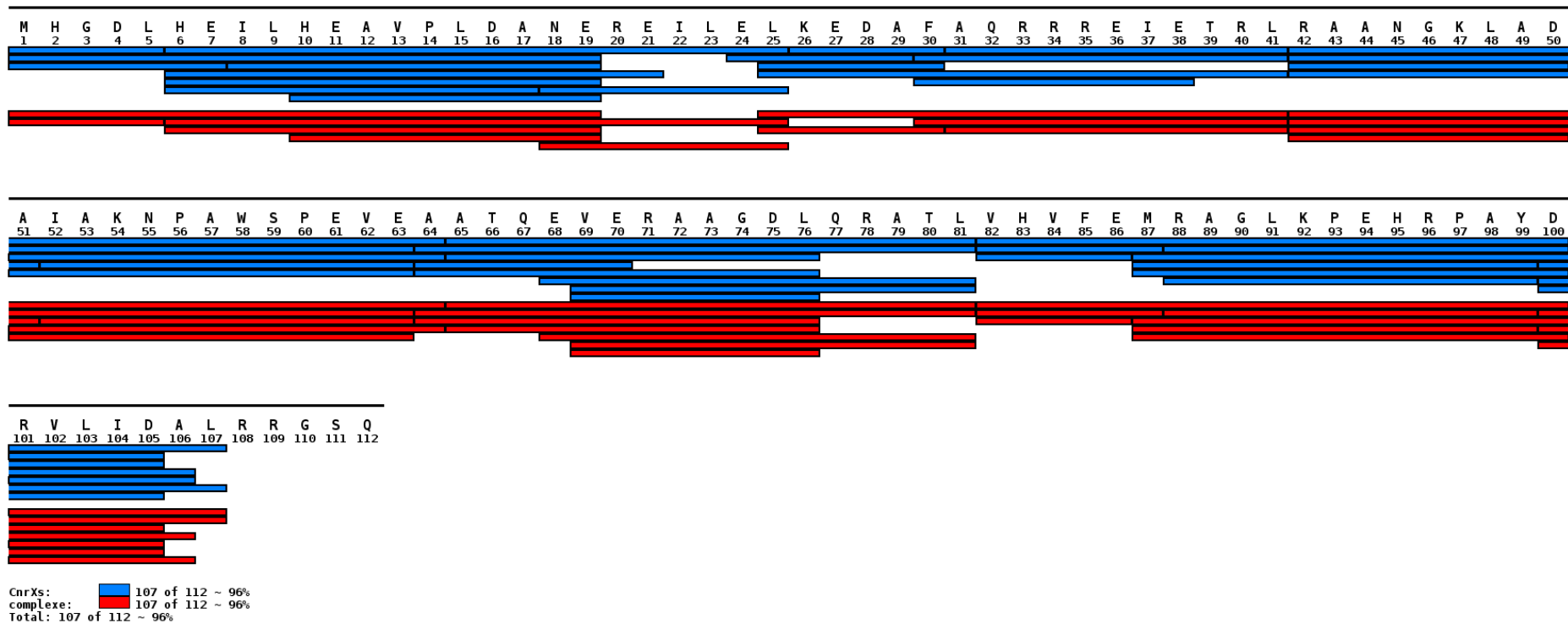
## **ii. Confirmation de l'interaction entre CnrXs et MYp par la méthode HDX-MS**

Nous avons d'abord vérifié que la méthode HDX permet effectivement d'observer l'interaction entre CnrXs $\Delta$ Nter et MYp. Pour ce faire, CnrXs $\Delta$ Nter libre ou en complexe avec MYp a été incubée dans un tampon deutéré. Les résultats de cette expérience indiquent qu'il y a moins de deutériums dans CnrXs $\Delta$ Nter en complexe avec MYp que dans sa forme libre : après 15 secondes de deutération, on détecte 7,2  $\pm$  0,8 D de différence entre la protéine CnrXs $\Delta$ Nter seule et en complexe avec son partenaire. Après 120 secondes de deutération, on détecte 7,7  $\pm$  0,6 D de différence entre la protéine CnrXs $\Delta$ Nter seule et en complexe avec MYp. En plus de démontrer la faisabilité de l'expérience de HDX-MS avec l'étude du complexe CnrXs $\Delta$ Nter:MYp, ces résultats nous ont permis de confirmer que

CnrXs $\Delta$ Nter interagit avec CnrYp. L'identification des résidus impliqués dans cette interaction a pu être entreprise.

### iii. Cartographie de CnrXs

La première étape dans une étude HDX-MS consiste à trouver les conditions de protéolyse de la protéine donnant la meilleure couverture de séquence (bon taux de couverture de séquence et taille des peptides relativement courte) avec les peptides obtenus, dans des conditions limitant le rééchange (bas pH et basse température). La meilleure cartographie, a été obtenue avec la pepsine immobilisée (collaboration avec Petr Man, Institut de Microbiologie à Prague, République tchèque). Nous avons donc utilisé la pepsine immobilisée pour faire une cartographie de CnrXs $\Delta$ Nter avant et après deutération mais aussi en présence de MYp. La protéolyse de CnrXs $\Delta$ Nter libre conduit à la formation de 46 peptides partiellement chevauchants détectés et identifiés par spectrométrie de masse (MS). Ces peptides couvrent 96% de la séquence de CnrXs $\Delta$ Nter (seuls les 5 derniers résidus n'ont pas été recouverts) (**Figure 42**, traits bleus). En présence de son partenaire MYp, de nombreux peptides de CnrXs $\Delta$ Nter ne peuvent plus être identifiés par MS. Cette différence s'explique par le fait que la quantité et la masse de MYp sont supérieures à celles de CnrXs $\Delta$ Nter. Après l'étape de protéolyse, les peptides de MYp détectés par MS sont beaucoup plus nombreux que ceux de CnrXs $\Delta$ Nter et sont susceptibles de masquer le signal correspondant aux peptides de CnrXs $\Delta$ Nter. L'expérience a été réitérée en utilisant 1 mole de CnrXs $\Delta$ Nter pour 2 moles de MYp. Dans ce cas, 37 peptides de CnrXs $\Delta$ Nter ont été détectés par MS (**Figure 42**, traits rouges). Bien que le signal correspondant à CnrXs $\Delta$ Nter a été amélioré dans ces conditions, le nombre de peptides issus de la protéine en complexe est resté inférieur à celui de la protéine libre. Les peptides issus de la protéolyse de CnrXs $\Delta$ Nter en complexe avec MYp détectés par MS assuraient toujours une couverture à 96% de la séquence de CnrXs $\Delta$ Nter. Cette nouvelle expérience a donc essentiellement reproduit la précédente.



**Figure 42** : Couverture de la séquence de CnrXs $\Delta$ Nter par les peptides résultant de la protéolyse par la pepsine immobilisée. Les peptides de CnrXs $\Delta$ Nter libre (bleu) et en complexe avec MYp (rouge) identifiés en MS, sont disposés au dessous de la séquence peptidique de CnrXs $\Delta$ Nter. La cartographie a été établie avec le logiciel <http://ms.biomed.cas.cz/MSTools/>.

#### iv. Echanges H/D couplés à la MS

Pour interpréter les résultats d'une expérience d'échanges H/D couplés à la MS, on compare le nombre de deutériums (#D) et le pourcentage de deutération (%D) pour chacun des peptides de CnrXs $\Delta$ Nter, libre ou en complexe avec MYp, ainsi que leur évolution avec la durée du marquage au deutérium (temps de deutération). Ces données sont présentées dans le **tableau 5**.

Des différences significatives ( $\Delta$ %D supérieur à 20%) ont été mesurées pour les peptides 25-41, 30-41 et 31-41 (**Tableau 5**, lignes surlignées en rouge) alors qu'aucune différence n'a été détectée pour le peptide 25-30. Par conséquent, la vitesse d'échange H/D n'a été ralentie qu'au niveau de la région 31-41 (région 67-77 de la protéine complète CnrX). Ces résultats indiquent que la région 31-41 (AQRREIETRL, **figure 43**) de CnrXs $\Delta$ Nter est impliquée dans l'interaction avec MYp. De plus, nous avons démontré précédemment par coprécipitation que CnrXs n'interagit pas avec la MBP (**Figure 34**). C'est donc CnrYp qui interagit avec la région 31-41 de CnrXs $\Delta$ Nter.

Les expériences de coprécipitation et de mutagenèse ont montré que l'extrémité C-terminale de CnrYp, plus particulièrement les leucines 91 et 92, est fortement impliquée dans l'interaction avec CnrXs (**Figure 36**). L'interaction entre MYp et CnrXs serait donc de nature hydrophobe. Or, la région 31-41 a pour principale caractéristique d'être fortement chargée (4 arginines et 2 glutamates) ce qui n'est pas compatible avec une interaction de nature hydrophobe. Les déterminants d'interaction identifiés sur CnrXs et sur CnrYp par des techniques différentes pourraient donc ne pas correspondre entre eux.

Rappelons que les échanges H/D détectés par la méthode HDX-MS concernent les hydrogènes des liaisons amides. Ces hydrogènes peuvent être enfouis dans la protéine alors que les chaînes latérales des résidus correspondants sont accessibles au solvant. La protection de la région 31-41 peut donc s'expliquer de différentes manières. La région 31-41 pourrait interagir avec CnrYp : dans ce cas, l'interaction serait probablement de nature électrostatique. Elle pourrait aussi interagir de manière non spécifique avec la MBP ou être masquée du fait de la présence de la MBP. Contrairement aux expériences de coprécipitation où la MBP est liée aux billes d'amylose, en HDX-MS, la MBP a un plus fort degré de liberté (elle n'est fusionnée qu'à CnrYp). On peut ainsi très bien imaginer que la fixation de CnrYp dans la cavité relativement hydrophobe de CnrXs, dont les résidus ne sont que partiellement exposés, donc partiellement accessibles au solvant en l'absence comme en la présence de CnrYp,

positionnerait la MBP à proximité de cette zone d'interaction et que la MBP perturberait les échanges H/D non pas dans la zone de contact mais à proximité de celle-ci. Des expériences de mutagenèse permettront de clarifier l'implication de cette région dans l'interaction avec CnrYp.

D'autres différences moins marquées ( $\Delta\%D$  entre 10 et 15%) sont observées pour les peptides couvrant la région C-terminale de CnrXs $\Delta$ Nter (69 à 107) (peptides surlignés en rose dans le **tableau 5**). Comme aucune différence n'est détectée pour le peptide 69-76, la vitesse d'échange H/D n'est ralentie qu'au niveau de la région 77-107 de CnrXs $\Delta$ Nter (région 113-143 de la protéine complète CnrX) (**Figure 43**). Plusieurs hypothèses peuvent expliquer cette petite différence d'incorporation du deutérium : (i) une interaction avec CnrYp a ralenti l'échange, (ii) la proximité de la MBP a ralenti l'échange, (iii) ce domaine C-terminal se structure, devenant moins disponible pour un échange H/D. L'hypothèse de l'interaction avec CnrYp est peu probable : CnrYp est plus susceptible d'être impliquée dans l'interaction avec la région 31-41 de CnrXs $\Delta$ Nter où la différence de deutérium entre la forme libre et la forme complexée de CnrXs $\Delta$ Nter est plus marquée. L'accessibilité à cette région est donc perturbée soit par la MBP (protéine volumineuse de 45 kDa) soit par une structuration du domaine C-terminal de CnrXs $\Delta$ Nter liée à son interaction avec CnrYp.

En conclusion, La méthode HDX-MS a permis de mettre en évidence un ralentissement significatif de l'échange H/D dans la région 31-41 de CnrXs $\Delta$ Nter lorsque cette dernière est incubée en présence de MYp. Cette région pourrait donc soit interagir directement avec CnrYp soit se structurer suite à la fixation de CnrYp soit être masquée par la MBP. La nature hydrophobe de CnrYp semble incompatible avec une interaction entre CnrYp et cette région fortement chargée de CnrXs $\Delta$ Nter. Les deux autres hypothèses envisagées sont donc considérées comme plus prometteuses.

**Tableau 5 :** Peptides (la séquence et la position du premier et du dernier résidu sont indiqués) analysés par SM après 15 ou 120 secondes de deutération. Le nombre de deutérium et le pourcentage de deutération pour la forme libre de CnrXsΔNter (#D free, %D free) et en complexe avec MYp (#D complex, %D complex) sont indiqués ainsi que la différence du nombre de deutérium du pourcentage de deutération ( $\Delta$ #D et  $\Delta$ %D).

Start	End	Sequence	Deut Time (sec)	#D free	%D free	#D complex	%D complex	D#D*	D%D*
1	5	MHGDL	15	0,93	29,1	0,66	20,6	0,27	8,4
1	5	MHGDL	120	0,9	28	0,71	22	0,19	6
1	7	MHGDLHE	15	1,54	32,1				
1	7	MHGDLHE	120	1,5	31,3				
1	19	MHGDLHEILHEAVPLDANE	15	5,09	37,4	4,18	30,7	0,91	6,7
1	19	MHGDLHEILHEAVPLDANE	120	6,48	47,6	5,74	42,2	0,74	5,4
6	17	HEILHEAVPLDA	15	2,68	33,5				
6	17	HEILHEAVPLDA	120	3,83	47,9				
6	19	HEILHEAVPLDANE	15	3,79	39,5	3,2	33,4	0,59	6,1
6	19	HEILHEAVPLDANE	120	5,02	52,3	4,26	44,4	0,76	8
6	21	HEILHEAVPLDANERE	15	4,92	44				
6	21	HEILHEAVPLDANERE	120	6,21	55,4				
6	25	HEILHEAVPLDANEREILEL	15	6,37	44,2	6	41,7	0,37	2,6
6	25	HEILHEAVPLDANEREILEL	120	8,64	60	8,32	57,8	0,32	2,2

8	19	ILHEAVPLDANE	15	3,61	45,1				
8	19	ILHEAVPLDANE	120	4,03	50,4				
10	19	HEAVPLDANE	15	3,55	55,5	3,24	50,6	0,31	4,9
10	19	HEAVPLDANE	120	3,72	58	3,51	54,8	0,21	3,2
18	25	NEREILEL	15	2,6	46,5	1,95	34,8	0,65	11,7
18	25	NEREILEL	120	3,5	62,5	3,21	57,3	0,3	5,3
24	29	ELKEDA	15	2,26	56,5				
24	29	ELKEDA	120	2,31	57,6				
25	30	LKEDAF	15	1,84	46,1	1,61	40,2	0,24	5,9
25	30	LKEDAF	120	2,04	51	2,05	51,3	-0,01	-0,3
<b>25</b>	<b>41</b>	<b>LKEDAFQAQRREIETRL</b>	<b>15</b>	<b>8,26</b>	<b>64,5</b>	<b>1,92</b>	<b>15</b>	<b>6,33</b>	<b>49,5</b>
<b>25</b>	<b>41</b>	<b>LKEDAFQAQRREIETRL</b>	<b>120</b>	<b>9,84</b>	<b>76,9</b>	<b>3,51</b>	<b>27,4</b>	<b>6,33</b>	<b>49,5</b>
26	30	KEDAF	15	1,42	44,3				
26	30	KEDAF	120	1,51	47,1				
30	38	FAQRRREIE	15	2,02	31,6				
30	38	FAQRRREIE	120	3,08	48				
30	41	FAQRRREIETRL	15	1,77	20,1	1,02	11,6	0,74	8,4
<b>30</b>	<b>41</b>	<b>FAQRRREIETRL</b>	<b>120</b>	<b>3,58</b>	<b>40,6</b>	<b>1,63</b>	<b>18,6</b>	<b>1,94</b>	<b>22,1</b>

31	41	AQRRREIETRL	15	1,45	18,1	0,99	12,3	0,46	5,7
<b>31</b>	<b>41</b>	<b>AQRRREIETRL</b>	<b>120</b>	<b>3,12</b>	<b>39</b>	<b>1,48</b>	<b>18,5</b>	<b>1,64</b>	<b>20,5</b>
42	50	RAANGKLAD	15	0,39	6	0,37	5,8	0,01	0,2
42	50	RAANGKLAD	120	0,72	11,3	0,46	7,1	0,27	4,2
42	51	RAANGKLADA	15	0,5	6,9	0,36	5	0,14	1,9
42	51	RAANGKLADA	120	0,82	11,4	0,34	4,7	0,49	6,7
42	63	RAANGKLADAIAKNPAWSPEVE	15	2,69	17,7	2,75	18,1	-0,06	-0,4
42	63	RAANGKLADAIAKNPAWSPEVE	120	4,14	27,2	3,09	20,4	1,05	6,9
42	64	RAANGKLADAIAKNPAWSPEVEA	15	2,87	17,9	2,69	16,8	0,18	1,1
42	64	RAANGKLADAIAKNPAWSPEVEA	120	4,42	27,6	3,2	20	1,22	7,7
51	63	IAKNPAWSPEVE	15	2,46	30,8	2,58	32,3	-0,12	-1,5
51	63	IAKNPAWSPEVE	120	3,46	43,2	3,01	37,7	0,44	5,6
51	64	IAKNPAWSPEVEA	15	2,44	27,8	2,96	33,6	-0,51	-5,8
51	64	IAKNPAWSPEVEA	120	3,67	41,7	3,88	44,1	-0,21	-2,4
52	63	IAKNPAWSPEVE	15	2,39	33,1	2,28	31,6	0,11	1,5
52	63	IAKNPAWSPEVE	120	3,3	45,9	2,69	37,3	0,61	8,5
64	70	AATQEVE	15	0,43	9				
64	70	AATQEVE	120	0,9	18,8				

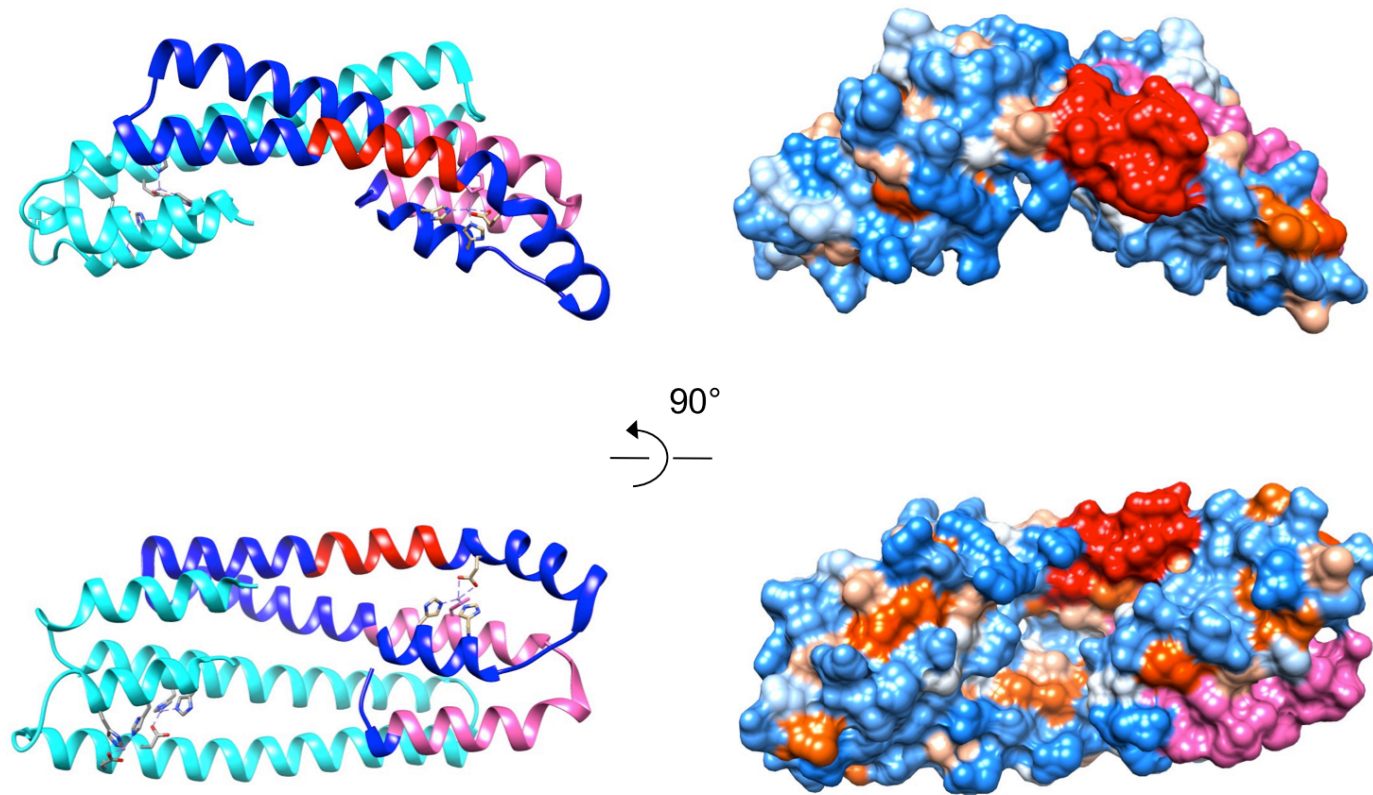


64	76	AATQEVERAAGDL	15	0,69	7,2	0,46	4,8	0,23	2,4
64	76	AATQEVERAAGDL	120	1,31	13,7	0,71	7,4	0,6	6,3
64	81	AATQEVERAAGDLQRATL	15	1,11	8,1	0,63	4,7	0,47	3,5
64	81	AATQEVERAAGDLQRATL	120	2,07	15,2	0,83	6,1	1,24	9,1
65	76	ATQEVERAAGDL	15	0,65	7,4	0,61	6,9	0,04	0,5
65	76	ATQEVERAAGDL	120	1,33	15,1	0,89	10,1	0,44	5
65	81	ATQEVERAAGDLQRATL	15	0,98	7,6	0,57	4,5	0,4	3,1
65	81	ATQEVERAAGDLQRATL	120	1,78	13,9	0,7	5,4	1,09	8,5
68	81	EVERAAGDLQRATL	15	0,71	6,8	0,44	4,2	0,27	2,6
68	81	EVERAAGDLQRATL	120	1,38	13,3	0,47	4,5	0,91	8,8
69	76	VERAAGDL	15	0,4	7,1	0,41	7,3	-0,01	-0,2
69	76	VERAAGDL	120	0,64	11,4	0,4	7,1	0,25	4,4
69	81	VERAAGDLQRATL	15	0,78	8,1	0,4	4,1	0,38	4
69	81	VERAAGDLQRATL	120	1,45	15,1	0,39	4,1	1,06	11
82	86	VHVFE	15	0,19	6	0,11	3,3	0,09	2,6
82	86	VHVFE	120	0,43	13,3	0,1	3,1	0,33	10,3
82	87	VHVFEM	15	0,27	6,7	0,17	4,2	0,1	2,5
82	87	VHVFEM	120	0,58	14,4	0,19	4,7	0,39	9,7

82	100	VHVFEMRAGLKPEHRPAYD	15	2,14	16,7	1,75	13,7	0,4	3,1
82	100	VHVFEMRAGLKPEHRPAYD	120	3,76	29,3	2,27	17,8	1,48	11,6
87	99	MRAGLKPEHRPAY	15	2,38	29,7	2,02	25,2	0,36	4,5
87	99	MRAGLKPEHRPAY	120	3,55	44,3	2,67	33,4	0,88	10,9
87	100	MRAGLKPEHRPAYD	15	2,47	28,1	1,98	22,5	0,49	5,6
87	100	MRAGLKPEHRPAYD	120	3,71	42,2	2,47	28	1,24	14,1
87	105	MRAGLKPEHRPAYDRVLID	15	1,96	15,3	1,84	14,3	0,13	1
87	105	MRAGLKPEHRPAYDRVLID	120	3,85	30,1	2,33	18,2	1,52	11,9
88	99	RAGLKPEHRPAY	15	2,27	31,5	1,92	26,7	0,35	4,8
88	99	RAGLKPEHRPAY	120	3,17	44,1				
88	100	RAGLKPEHRPAYD	15	2,3	28,7				
88	100	RAGLKPEHRPAYD	120	3,3	41,2				
100	105	DRVLID	15	0,26	6,4	0,19	4,7	0,07	1,8
100	105	DRVLID	120	0,61	15,2	0,24	6	0,37	9,2
100	106	DRVLIDA	15	0,3	6,2	0,2	4,2	0,1	2,1
100	106	DRVLIDA	120	0,8	16,6	0,21	4,3	0,59	12,4
100	107	DRVLIDAL	15	0,31	5,5	0,14	2,5	0,17	3
100	107	DRVLIDAL	120	0,85	15,2	0,15	2,7	0,7	12,5

101	105	RVLID	15	0,16	5,1	0,09	2,8	0,07	2,2
101	105	RVLID	120	0,38	11,8	0,1	3	0,28	8,8
101	106	RVLIDA	15	0,22	5,6	0,12	2,9	0,11	2,7
101	106	RVLIDA	120	0,6	14,9	0,14	3,4	0,46	11,5
101	107	RVLIDAL	15	0,28	5,7	0,28	5,8	-0,01	-0,1
101	107	RVLIDAL	120	0,79	16,5	0,21	4,3	0,59	12,2

\* Ces colonnes présentent les différences du nombre de deutérium ( $\Delta\#D$ ) et du pourcentage de deutération ( $\Delta\%D$ ) entre la forme libre et la forme complexée.



**Figure 43** : Structure du dimère de CnrXs (code PDB : 2Y3D). **A gauche** : Les monomères sont représentés en bleu et cyan. Sur le monomère bleu, la région 31-41 de CnrXs $\Delta$ Nter, qui correspond aux résidus 67 à 77 de la protéine complète CnrX, est colorée en rouge, la région 77-107 (113-143 de CnrX) est en rose. **A droite** : Représentation de la surface moléculaire du dimère de CnrXs. Plus la surface est hydrophobe, plus elle est orange et plus elle est hydrophile, plus elle est bleue. La surface des peptides 31-41 et 77-107 (visible que sur un monomère) est en rouge et rose, respectivement.

# **Conclusions et perspectives**

Le complexe CnrYXH contrôle à la fois l'expression des gènes de régulation et de résistance au cobalt et au nickel chez *C. metallidurans* CH34. Les bases moléculaires de la perception du stress métallique par le domaine senseur de CnrX (CnrXs) ont été posées à la suite d'études structurales, spectroscopiques et physiologiques. Les résultats de ces expériences ont permis de proposer que l'activation de la voie de signalisation Cnr résulte de la substitution du zinc par le nickel ou du cobalt dans le site métallique de CnrXs. Alors que le zinc est pentacoordiné dans une géométrie bipyramidale à l'étal basal, le nickel et le cobalt sont capables de recruter le soufre de la seule méthionine (M123) de CnrXs comme sixième ligand. Le nickel et le cobalt induisent donc un changement considérable dans la géométrie du site métallique (bipyramidale à octaédrique) qui fait évoluer le domaine senseur de CnrX vers une conformation active. La réorganisation du site métallique autour des ions métalliques agonistes entraîne des modifications allostériques aboutissant à un dimère plus compact dont la cavité hydrophobe, probablement orientée face à la membrane, est plus étroite. Il s'agit de la première étape de la cascade de transduction du signal TM. Cette étape est similaire à celle observée pour les métallorégulateurs transcriptionnels (**Introduction** partie **I.C.2.**) à la différence près que ces derniers assurent leurs fonctions de senseur et de régulateur transcriptionnel dans le même compartiment, et que la détection du métal et la régulation de la transcription sont assurées par deux domaines d'une même protéine. Dans le cas du système CnrYXH, la topologie et la distribution des trois protéines du complexe ajoutent une complexité supplémentaire : le signal de détection du métal perçu dans le périplasme doit franchir la membrane interne pour aboutir à la libération du facteur  $\sigma$  CnrH dans le cytoplasme. Obtenir la cartographie des interactions entre les protéines du complexe CnrYXH est donc nécessaire pour mieux comprendre les ressorts de la transduction du signal par ce complexe.

Le projet de thèse était de décrire la chaîne de signalisation à travers la membrane en caractérisant les interactions entre les trois protéines du complexe dans les différents compartiments cellulaires : périplasme, membrane interne et cytoplasme.

### **- Interactions dans le périplasme :**

La topologie de CnrX, senseur du système, est telle qu'il est quasiment absent du cytoplasme. Le signal peut donc être transmis de CnrX à CnrY au niveau des domaines périplasmiques et/ou transmembranaires de ces deux protéines. L'interaction des domaines

périplasmiques de CnrX (CnrXs) et de CnrY (CnrYp) a été démontrée par des expériences de coprécipitation et d'échanges hydrogène/deutérium (H/D).

Nous avons démontré que la partie carboxyterminale (GLLIGK) de CnrYp est nécessaire pour que CnrXs et CnrYp coprécipitent mais que cette interaction est faible en dehors du contexte de la protéine entière. La méthode d'échange hydrogène/deutérium couplée à la spectrométrie de masse (HDX-MS) a permis de mettre en évidence un ralentissement significatif de l'échange H/D dans la région (A<sub>67</sub>QRRREIETRL<sub>77</sub>) de CnrXs lorsque CnrXs est incubée en présence de MBP-CnrYp. Cette région pourrait donc interagir directement avec CnrYp. Des mutations au niveau de cette région suivi d'expériences de coprécipitation permettront de savoir si ces résidus sont impliqués dans l'interaction avec CnrYp. Le ralentissement des échanges H/D dans cette région de CnrXs pourrait aussi être la conséquence de structuration de la protéine suite à la fixation de CnrYp. Toutefois, la nature hydrophobe de CnrYp semble incompatible avec une interaction entre CnrYp et cette région fortement chargée de CnrXs. Comme la technique HDX-MS ne peut mettre en évidence une interaction que si elle concerne une zone où les hydrogènes des liaisons amides sont accessibles au solvant, il est possible de supposer que l'effet mesuré sur la région A<sub>67</sub>QRRREIETRL<sub>77</sub> résulterait de la présence de la MBP (42 kDa) à proximité de la zone d'interaction (effet de masquage) et non d'une interaction directe avec CnrYp. L'interaction pourrait avoir lieu dans la cavité hydrophobe faisant face à la membrane (Trepreau *et al.*, 2011b) : il suffirait que les protons des amides des résidus concernés ne soient pas échangeables (pour cause d'enfouissement) pour que cette interaction passe inaperçu en HDX-MS. Cette hypothèse peut être vérifiée par HDX-MS en utilisant CnrYp fusionnée à une protéine moins volumineuse que la MBP comme l'étiquette Strep (Strep-CnrYp).

Nous avons aussi confirmé, par des expériences de RMN (collaboration avec Dominique Marion, IBS), l'interaction entre CnrXs et MBP-CnrYp. Il ne s'agit que de résultats préliminaires qui n'ont pas été présentés dans ce manuscrit. L'identification des régions de CnrXs impliquées dans l'interaction a échoué du fait de la présence de la MBP qui masque le signal correspondant à CnrXs. Un moyen de s'affranchir du problème lié à la MBP est d'utiliser la Strep-CnrYp, comme proposé pour les expériences HDX-MS ou encore de libérer CnrYp de l'influence de la MBP par clivage protéolytique. La dissociation des deux protéines est possible car nous avons introduit entre MBP et CnrYp, une séquence de clivage à la TEV.

Par ailleurs, nous avons démontré que l'interaction entre CnrXs et CnrYp est déstabilisée par le nickel, le cobalt et dans une moindre mesure par le cuivre alors que le zinc la stabilise. A l'état basal, CnrXs qui est probablement liée au zinc interagit avec CnrYp. Les modifications conformationnelles de CnrXs liées à la fixation du cobalt et du nickel permettent la libération de CnrYp. La déstabilisation de l'interaction entre CnrX et CnrY dans le périplasme serait donc la seconde étape de la transduction de signal TM.

### **- Interactions dans la membrane :**

Les domaines TM constituent un lien physique entre les compartiments extra- et intracellulaire. Alors que certains TM n'ont pas de rôle dans la transduction du signal, d'autres sont emblématiques de la fonction de signalisation à travers la membrane. Parmi les cas les mieux connus, on trouve la protéolyse régulée des domaines TM qui permet la libération d'un domaine de signalisation comme cela a été décrit pour RseA ou pour SREBP. C'est aussi le cas pour le récepteur ErbB4 dont la protéolyse au niveau membranaire permet la libération d'un fragment qui agit comme répresseur de la transcription des gènes impliqués dans la différenciation des astrocytes (Lal et Caplan, 2011). Le changement de conformation des domaines TM a aussi une grande importance pour la signalisation cellulaire. Différents types de mouvements ont été décrits parmi lesquels les translations latérales, les mouvements de piston ou encore de rotations parallèlement ou perpendiculairement à la membrane (Matthews *et al.*, 2006). La rupture de symétrie des hélices TM est aussi un changement de conformation décrit notamment pour les récepteurs couplés aux protéines G (GPCR). Malgré la diversité structurale et fonctionnelle des GPCR du génome humain, elles présentent un mécanisme commun de transduction du signal qui passe par les changements de conformation des segments TM et plus précisément par des mouvements en sens opposés de deux hélices TM suite à la fixation du Na<sup>+</sup> au centre de ses 7 segments TM (Katritch *et al.*, 2014). L'interaction des segments TM *via* des motifs consensus tels que le motif GxxxG ou le motif "leucin zipper" est aussi un processus essentiel pour la transduction du signal à travers la membrane chez certaines protéines comme pour le récepteur ErbB2 qui possède deux motifs GxxxG et dont la dimérisation est nécessaire à son activité (Cymer et Schneider, 2010; Hubert *et al.*, 2010).



La caractérisation des interactions dans la membrane a été engagée afin de découvrir l'organisation des hélices TM de CnrX et CnrY et ainsi de mieux comprendre leur rôle dans la transduction du signal.

Si la purification de la forme entière de CnrX a été un échec, celle de la protéine homologue NccX fut couronnée de succès. C'est donc le modèle Ncc que nous avons utilisé pour caractériser les possibles interactions dans la membrane entre NccX et NccY. La structure cristallographique de NccX a été déterminée avec une résolution de 3,12 Å. Cette dernière vient enrichir la banque de données des protéines membranaires encore très pauvre notamment en structure de protéines possédant un seul segment TM. Cependant, l'utilisation du détergent DPC nécessaire à la solubilisation des membranes a altérée la structure de NccX qui n'est pas compatible avec l'insertion des hélices TM dans la membrane. Cette structure a été réconciliée avec le modèle de la protéine basé sur la structure de CnrXs *via* l'utilisation de simulations couplées aux expériences *in vivo* TOXCAT.

Des simulations par dynamique moléculaire ont aussi permis de mettre en évidence le caractère dynamique des interactions entre les segments TM (plusieurs angles de croisement des segments TM) commun à d'autres protéines TM comme la glycophorine A (Psachoulia *et al.*, 2010) ou la plexin-A1 (Aci-Sèche *et al.*, 2014).

Les expériences menées *in vivo* et utilisant la méthodologie TOXCAT ont permis de démontrer que l'homodimérisation de NccX est guidée principalement par la dimérisation du domaine périplasmique de NccX. Elles ont aussi montré que, malgré un contexte non favorable, le motif de dimérisation GxxxG contribue à l'homodimérisation de l'hélice TM de NccX. Cette interaction pourrait être potentialisée par l'homodimérisation du domaine périplasmique de NccX.

Des expériences similaires avec NccY ont montré que le rapprochement des protéines NccY dépend du domaine périplasmique de NccY et de la présence de NccX. Le domaine périplasmique de NccY/CnrY est donc le déterminant principal de l'interaction entre les protéines NccX/CnrX et NccY/CnrY. Le seul rôle des segments TM dans la transduction du signal serait ainsi d'assurer l'ancrage à la membrane pour CnrX/NccX et d'établir un lien physique entre le périplasme et le cytoplasme pour la protéine CnrY/NccY.

Afin de vérifier si les observations sur les domaines périplasmiques de CnrX et CnrY sont applicables aux protéines complètes, il sera intéressant de muter les résidus identifiés comme importants pour l'interaction CnrXs:CnrYp sur les protéines complètes et de vérifier leur

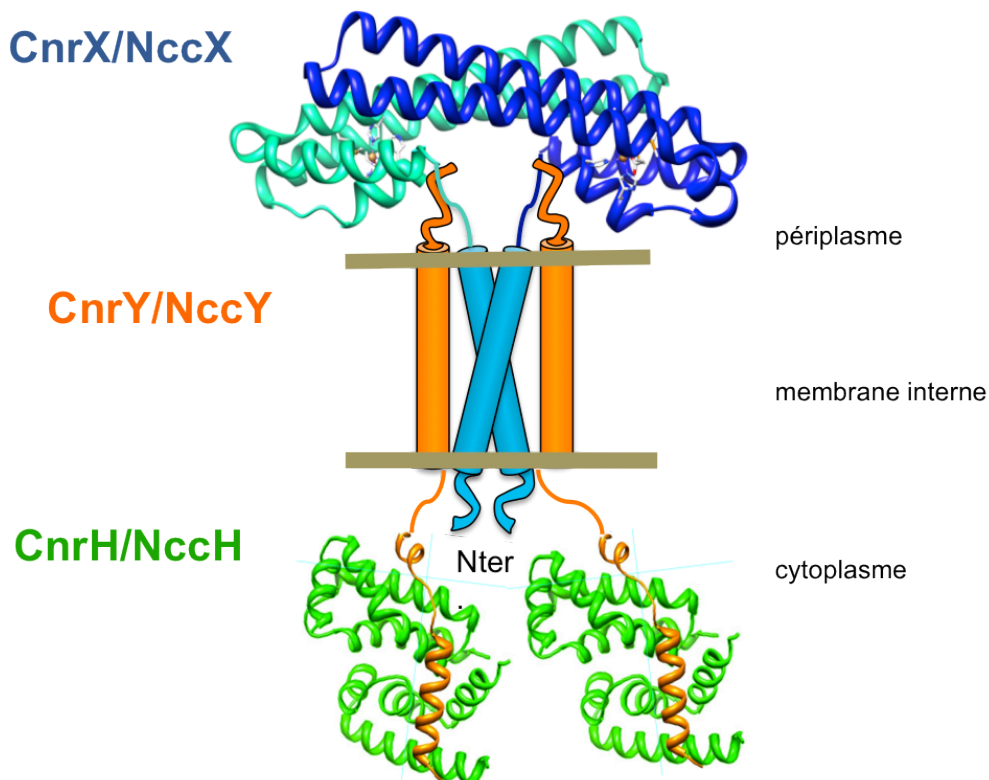
implication *in vivo* pour l'interaction avec NccX. En effet, il serait intéressant de supprimer les résidus de l'extrémité carboxyterminale de NccY et vérifier si ils sont importants pour l'interaction avec NccX. Ce test pourra aussi servir à estimer l'effet des métaux inducteurs (nickel et cobalt) ou non (zinc). En cultivant les bactéries qui coproduisent NccY et NccX en présence de concentrations croissantes en différents métaux, on pourra évaluer leurs effets sur l'interaction NccX:NccY.

### **- Interactions dans le cytoplasme :**

Les bases moléculaires de l'interaction entre un facteur  $\sigma$  et son facteur anti- $\sigma$  ont déjà été décrites dans quelques cas, comme celui des complexes RseA: $\sigma^E$  d'*E. coli* (Campbell *et al.*, 2003) et ChrR: $\sigma^E$  de *R. sphaeroides* (Campbell *et al.*, 2007). Ces structures ont permis de mettre en évidence un mode d'inhibition commun à de nombreux facteurs anti- $\sigma$ , caractérisés par la présence d'un premier type d'ASD. Le domaine cytosolique de CnrY (CnrYc) ne présente aucune homologie avec les facteurs anti- $\sigma$  déjà caractérisés structuralement et est l'un des plus petits facteurs anti- $\sigma$  connus. Caractériser l'interaction entre le facteur  $\sigma$  CnrH et le domaine cytosolique du facteur anti- $\sigma$  CnrY (CnrYc) a permis de montrer que les trente premiers résidus de CnrYc se lient aux mêmes régions fonctionnelles clés que les autres facteurs  $\sigma$  bien qu'il soit au moins deux fois plus court. Ceci est rendu possible par le fait qu'il interagit avec une conformation compacte de CnrH, dont les deux domaines interagissent directement, contrairement à  $\sigma^E$  d'*E. coli* et *R. sphaeroides*. De manière surprenante, la structure du complexe entre CnrH et CnrYc est comparable avec celle du complexe PhyR:NepR (**Introduction** partie **I.C.3.b** et **Résultats** partie **I**). Cette convergence imprévisible entre ces structures révèle un nouveau mode d'inhibition des facteurs  $\sigma$  par les ASD des facteurs anti- $\sigma$ . Une nouvelle classe d'ASD est ainsi définie dans laquelle les ASD dont la taille est plus petite (30 résidus contre environ 70), forment deux hélices qui maintiennent le facteur  $\sigma$  dans une conformation fermée.

De plus, nous avons identifié le dipeptide W<sub>7</sub>L<sub>8</sub> comme déterminant de l'inhibition de CnrH par CnrY. Ces résidus interagissent dans une poche hydrophobe du domaine  $\sigma_4$  de CnrH. La mutation de ces résidus dans le système CnrYXH de CH34 permettra d'en savoir plus sur le rôle de ces résidus dans la transduction du signal. Ces expériences pourront être effectuées par l'équipe de Dietrich H. Nies (Halle, Allemagne) qui maîtrise la génétique microbienne de *C. metallidurans* CH34.

La cartographie des interactions nous a permis d'établir un modèle structural quasi complet du complexe CnrYXH/NccYXH (**Figure 1**) et a permis de découvrir une nouvelle famille de facteurs anti- $\sigma$ . D'autres part, nous avons démontré que l'interaction entre CnrY et CnrX dans le périplasma est modulée par les métaux inducteurs du système Cnr mais nous ne savons toujours pas comment s'effectue la déstabilisation de l'interaction CnrY:CnrH qui permet d'activer CnrH.



**Figure 1** : Modèle de l'organisation supposée des complexes CnrYXH/NccYXH.

Les différents mécanismes permettant d'activer les facteurs de transcription, présentés en introduction, regroupent la protéolyse régulée du facteur anti- $\sigma$ , la cascade de phosphorylation, l'échange de partenaire ou encore les changements de conformation. L'activation des facteurs  $\sigma$  peut même découler du couplage de plusieurs de ces mécanismes comme cela est décrit pour le système Fec (**Introduction** partie **I.C.4.b.iii.**).

Un exemple de voie de signalisation reposant sur l'échange de partenaire est celui de PhyK:PhyR:NepR, où le facteur  $\sigma$  est libérée par suite de l'interaction du facteur anti- $\sigma$  avec une protéine mimant un facteur  $\sigma$ . Ce mécanisme-ci est peu probable dans le cas de la voie

Cnr du fait de l'absence d'homologues de PhyR dans *C. metallidurans* CH34. Quant à l'hypothèse d'un changement de conformation du facteur anti-sigma, elle peut être écartée car CnrYc n'a pas de cœur hydrophobe capable de médier un changement de conformation. De fait, le domaine cytosolique de CnrY est entièrement accessible au solvant, ce qui fait que CnrYc est accessible à une modification post-traductionnelle (phosphorylation) ou à une dégradation protéolytique.

Par ailleurs, l'introduction des gènes *cnrYXHCBAT* dans *E. coli* ne permet pas de transférer la résistance à cette bactérie (Nies, 1989). Dans le laboratoire, nous avons aussi observé que CnrYXH n'est pas fonctionnel dans *E. coli* bien que CnrH de CH34 soit compatible avec l'ARNP d'*E. coli*. Cela indique qu'un ou plusieurs facteurs qui permettent à CnrYXH d'opérer dans *C. metallidurans* CH34 sont absents d'*E. coli*. Ces partenaires pourraient être des kinases et/ou des protéases. La protéolyse régulée et/ou la phosphorylation de CnrY apparaissent comme deux mécanismes possibles de transduction du signal, pouvant même intervenir ensemble (Barik *et al.*, 2010).

D'après le logiciel de prédiction NetPhosBac (Miller *et al.*, 2009), CnrYc possède plusieurs sites potentiels de phosphorylation notamment deux sérines et trois thréonines. La phosphorylation d'un site dépend de plusieurs facteurs (spécificité des kinases, contexte du motif putatif de phosphorylation, (Newman *et al.*, 2014)) et les outils de prédiction sont considérés comme peu fiables. Des expériences préliminaires effectuées en collaboration avec Virginie Molle (DIMNP, Montpellier II) indiquaient que CnrYc est phosphorylée par la kinase PknD de *M. tuberculosis*. Deux sérine/thréonine protéine kinases ont été identifiées par similarité de séquence avec PknD de *M. tuberculosis*, dans le génome de *C. metallidurans* CH34 : Rmet\_0854 et Rmet\_5309. La recherche d'une preuve expérimentale de la phosphorylation de CnrY par Rmet\_0854 a été entreprise dans le laboratoire de Virginie Molle et ont montré que CnrYc n'est pas un substrat de cette enzyme. Il serait tout de même intéressant de tester la fonctionnalité du système Cnr dans *E. coli* en présence de la kinase de CH34. Un autre moyen de vérifier si la phosphorylation est le mécanisme permettant de libérer CnrH serait de muter les résidus potentiellement phosphorylables de CnrYc (Ser ou Thr) en un résidu mimant la phosphorylation comme le glutamate.

Tout comme la détection de la production de CnrX dans des conditions où *C. metallidurans* CH34 est cultivée en présence des métaux inducteurs (**Introduction** partie

**III.D.4.b.)**, il serait intéressant de détecter la production de CnrY dans les différentes conditions de cultures de CH34. Nous avons mis au point le protocole de détection de CnrY par Western Blot à l'aide d'anticorps dirigés contre CnrYc-his<sub>6</sub> purifié en présence d'urée. Il s'agit d'un outil nouveau qui permettra peut-être de vérifier si CnrY est surproduite dans des conditions où *C. metallidurans* CH34 est stressée avec du nickel ou du cobalt et ce qu'il devient dans ces conditions. On peut en effet imaginer que CnrY soit dégradée et que des produits de dégradation soient éventuellement détectés, ou bien que la bande correspondant à CnrY comigre avec un produit phosphorylé, lui-même détecté à l'aide d'anticorps appropriés. Dans un cas comme dans l'autre, cela fournirait une première piste pour comprendre le principe du fonctionnement de la voie de signalisation Cnr.

En conclusion, les acteurs du complexe CnrYXH sont maintenant en place et la majorité de leurs interactions est identifiée. Si de nombreuses questions restent posées, les résultats présentés dans ce manuscrit permettent d'envisager les démarches expérimentales menant aux réponses correspondantes et balisent le chemin des recherches futures.

# **Matériels et méthodes**

# I. Production et purification des protéines d'intérêt

## A. Milieux de culture

*E. coli* a été cultivée dans un milieu liquide complet Luria-Bertani (LB) ou minimum (M9) dont la composition est détaillée dans le **tableau 1**. Le milieu LB peut être aussi utilisé sous forme solide après ajout de 15% d'agar. Tous les milieux ont été stérilisés à l'autoclave. Les antibiotiques ont été ajoutés, si nécessaire, aux concentrations finales suivantes : ampicilline (Amp) ( $100 \mu\text{g.mL}^{-1}$ ), kanamycine (Kan) ( $50 \mu\text{g.mL}^{-1}$ ), chloramphénicol (Cm) ( $34 \mu\text{g.mL}^{-1}$ ) et streptomycine (Str) ( $10 \mu\text{g.mL}^{-1}$ ).

**Tableau 1** : Composition des milieux de culture

Milieu LB	Milieu M9	Milieu M9 "enrichi"
Extraits de levure (5 g/L)	$\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ (10 g/L)	$\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ (10 g/L)
NaCl (10 g/L)	$\text{KH}_2\text{PO}_4$ (3 g/L)	$\text{KH}_2\text{PO}_4$ (3 g/L)
Bacto-tryptone (10 g/L),	NaCl (0,5 g/L)	NaCl (0,5 g/L)
	$\text{NH}_4\text{Cl}$ (1 g/L)	$\text{NH}_4\text{Cl}$ (1 g/L)
	Glucose (2 g/L)	Glucose (4 g/L)
	$\text{MgSO}_4$ (1 mM)	$\text{MgSO}_4$ (1 mM)
	$\text{CaCl}_2$ (0,1 mM)	$\text{CaCl}_2$ (0,1 mM)
	Thiamine (2 mM)	$\text{MnCl}_2$ (0,1 mM)
		$\text{ZnSO}_4$ (0,05 mM)
		$\text{FeCl}_3$ (0,05 mM)
		Cocktail de vitamines* (2 mL)

\* pyridoxine (0,5 g/L), biotine (0,5 g/L), panthothenate hemi calcium (0,5 g/L), acide folique (0,5 g/L), chlorure de choline (0,5 g/L), niacinamide (0,5 g/L), riboflavine (0,5 g/L) et thiamine (2,5 g/L).

## **B. Souches bactériennes et plasmides**

Tout ou une partie de l'opéron *cnrYXH* a été amplifiée par PCR à partir d'ADN génomique de *C. metallidurans* CH34. Les gènes *nccX* et *nccY* ont été amplifiés par PCR à partir d'ADN génomique de *C. metallidurans* 31A.

Plusieurs dérivés des produits de ces gènes, protéines étiquetées ou mutées, ont été construits pour les besoins du travail présenté ici. Ils seront présentés au fur et à mesure de leur apparition. Le détail des vecteurs et des souches d'expression se trouve dans le **tableau 2**.



**Tableau 2 :** Récapitulatif des protéines, des plasmides dans lesquels les gènes codant les protéines d'intérêt ont été clonés, des souches de clonage et d'expression et du marqueur de résistance de la souche d'expression transformée par le plasmide recombinant.

<b>Protéines</b>	<b>plasmide de destination</b>	<b>souche de clonage</b>	<b>souche d'expression</b>	<b>résistance</b>
CnrXs (Pompidor <i>et al.</i> , 2009)	pET30b (Novagen)	<i>E. coli</i> XL10 (Agilent)	BL21 DE3 (Stratagène)	Kan
MBP-CnrYp	pETM-40 (Novagen)	<i>E. coli</i> XL10	Rosetta (Stratagène)	Kan et Cm
NccX	pccKAN (Engelmann,1999)	<i>E. coli</i> XL10	MM39 (Russ et Engelman, 1999)	Amp
NccY	pccKAN	<i>E. coli</i> XL10	MM39	Amp
Strep-CnrH + CnrYc	pASK7+ (IBA)  pET-Duet-1 (Novagen)	<i>E. coli</i> XL10	Rosetta	Amp et Cm
CnrYXH + GFP	pTrc99a  pACYC-Duet1 (Novagen)	<i>E. coli</i> XL10  <i>E. coli</i> XL10	TOP10 (Invitrogen)	Amp et Cm
his-CnrY- CnrX	pET30b	<i>E. coli</i> XL10	Rosetta	Kan et Cm

## C. Clonage

Les gènes d'intérêts se trouvant sur l'ADN plasmidique des bactéries *C. metallidurans* CH34 ou 31A ont été amplifiés par la réaction de polymérisation en chaîne (*Polymerase Chain Reaction*, PCR) (détaillée dans la partie suivante **I.D.**) à partir d'amorces nucléotidiques appropriées.

Une fois purifiés à l'aide de micro-colonnes à centrifuger (kit "QIAquick PCR Purification" de Qiagen ou équivalent), les produits de PCR ont été digérés avec les enzymes de restriction appropriées, qui seront présentées cas par cas. Les plasmides de destinations ont été digérés et déphosphorylés. Les gènes d'intérêt ont été placés sous le contrôle du promoteur fort et inductible du plasmide choisi, à l'aide de la T4 DNA ligase, conditionnée dans le "Rapid ligation kit" de Fermentas ou un kit équivalent.

## D. Mutagenèse dirigée

Les mutations ont été réalisées en utilisant le kit "QuikChange Lightning Site-Directed Mutagenesis" (Agilent) ou selon un protocole similaire en utilisant la "Phusion High-Fidelity Polymerase" (Thermo Scientific).

### 1. Mutagenèse par le protocole "QuikChange"

L'amplification d'ADN par PCR est réalisée avec le mélange suivant :

- 2,5 µL du mélange des oligonucléotides 3' et 5' à 12,5 µM chacun
- 3 µL de plasmide matrice à environ 5 ng/µL
- 1 µL du mélange des dNTP à 10 mM chacun
- 5 µL de QuikChange Lightning Buffer
- 1,5 µL de QuikSolution
- 1 µL d'enzyme ADN polymérase Pfu
- qsp 50 µL d'eau distillée stérile

Protocole d'amplification de l'ADN :

- Dénaturation : 95°C, 2min
  - Dénaturation courte : 95°C, 50 sec
  - Hybridation : 50 à 65°C, 50 sec 18 cycles
  - Elongation : 68°C, temps dépendant de la taille du plasmide ou du gène à amplifier (vitesse de polymérisation de la Pfu : 1000 pb/minute)
- Elongation finale : 68°C, 5 min
- Conservation : 8°C

2 µL d'enzyme de restriction DpnI ont ensuite été ajoutés au mélange PCR afin de digérer l'ADN matrice (ADN méthylé) pendant 1 heure à 37°C.

## **2. Mutagenèse avec la polymérase Phusion**

Mélange des composants pour la PCR :

- 2,5 µL du mélange des oligonucléotides 3' et 5' à 10 µM chacun
- 2 µL de plasmide matrice à environ 5 ng/µL
- 1 µL de dNTP à 10 µM
- 10 µL de tampon Phusion
- 1 µL d'enzyme Phusion
- qsp 50 µL d'eau distillée stérile

Protocole d'amplification de l'ADN :

- Dénaturation : 95°C, 2min
  - Dénaturation courte : 95°C, 50 sec
  - Hybridation : 50 à 65°C, 50 sec 30 à 35 cycles
  - Elongation : 72°C, temps dépendant de la taille du plasmide ou du gène à amplifier (vitesse de polymérisation de la Phusion : 1000 pb/30 sec)
- Elongation finale : 72°C, 5 min
- Conservation : 8°C

2 µL d'enzyme de restriction DpnI ont ensuite été ajoutés au mélange PCR afin de digérer l'ADN matrice (ADN méthylé) pendant 1 heure à 37°C.

## **E. Transformation bactérienne**

Le produit de PCR est transformé dans les cellules compétentes de clonage (*E. coli* XL10). Pour cela, un dixième (v/v) de plasmide obtenu par mutagenèse est ajouté aux cellules. Ce mélange est incubé 30 minutes dans la glace puis 1 minute à 42°C. Ce choc thermique est arrêté par retour du mélange dans la glace. On y ajoute ensuite 200 µL de milieu LB qu'on incube 1 heure à 37°C à 180 rpm. La suspension cellulaire est enfin étalée sur un milieu LB gélosé supplémenté en antibiotique(s) qui est placé dans une étuve à 37°C toute une nuit.

L'ADN plasmidique, contenu dans les bactéries ayant poussé, est purifié à l'aide du kit "EZ-10 Spin Column" (Bio Basic Inc.) afin d'être envoyé à séquencer (Eurofins).

Ce processus de transformation est similaire pour les souches d'expression.

## **F. Culture et expression**

La souche d'expression contenant le gène codant la protéine d'intérêt est mise en préculture dans 5 mL de milieu LB ou M9 supplémenté en antibiotique(s) et 0,6% de glucose durant toute une nuit à 37°C. Cette préculture permet d'inoculer la culture dont la croissance à 37°C est suivie par l'absorbance à 600 nm (A600). Lorsque A600 atteint 0,4-0,6, l'expression de la protéine est induite avec 0,5 mM d'isopropyl β-D-1-thiogalactopyranoside (IPTG) durant 3 à 4 heures à 37°C (30°C pour CnrY et ses dérivés).

## **G. Purification de protéines**

### **1. Purification de CnrXs**

Les dérivés de CnrXs (CnrXsM123A, CnrXsΔNter) sont purifiés selon le même protocole. Après induction, la culture cellulaire est centrifugée à 6000 rpm (rotor SLC-6000, Sorvall) pendant 20 minutes à 4°C. Le culot bactérien est repris dans 20 mL de tampon de lyse A

(Hepes-NaOH 50 mM pH 8, NaCl 100 mM, EDTA 2 mM) avant d'être centrifugé à 6000 rpm pendant 20 minutes à 4°C. Les culots peuvent être stockés à - 80°C.

### **a. Lyse cellulaire**

Le culot est resuspendu dans 20 mL de tampon de lyse A supplémenté d'un cocktail d'antiprotéases (Complete, Roche, 1 pastille/50 mL). Les bactéries sont lysées par sonication (3 min efficaces : 5 secondes de sonication active, 15 secondes de refroidissement, 40% d'amplitude) puis centrifugées à 45000 rpm (rotor TI70, Beckman) pendant 50 minutes à 4°C. Le surnageant est incubé avec 2% de sulfate de streptomycine pendant une heure sous agitation à 4°C. Ceci permet de faire précipiter les acides nucléiques qui sont ensuite éliminés après 20 minutes de centrifugation à 10000 rpm à 4°C (rotor F0850, Beckman).

### **b. Précipitation des protéines**

Le surnageant est incubé en présence de 80% de sulfate d'ammonium pendant une heure sous agitation à 4°C. Après 20 minutes de centrifugation à 10000 rpm à 4°C (rotor F0850, Beckman), le culot est resuspendu avec 5 mL de tampon B (Hepes-NaOH 50 mM pH 8, EDTA 2 mM) et dialysé contre 500 mL de tampon B sur 18 heures (limite de taille: 3500 Da, membrane de dialyse Spectrum) afin d'éliminer le sulfate d'ammonium. Le bain de dialyse est changé trois fois au bout de 3h minimum.

### **c. Chromatographie sur colonne échangeuse d'anions**

L'échantillon est injecté (0,5 mL/min) sur une colonne Q sepharose (HiTrap-HP, 5 mL, GE Healthcare) préalablement équilibrée en tampon B. La colonne est lavée par 5 volumes de colonne (VC) de tampon B à 1 mL/min. La protéine est éluée par un gradient linéaire de 0 à 10% de tampon C (Hepes-NaOH 50 mM pH 8, EDTA 2 mM, NaCl 1 M). Les fractions à retenir sont choisies d'après les chromatogrammes (absorbance à 280 nm) et les analyses électrophorétiques en conditions dénaturantes (SDS-PAGE). Les fractions retenues sont concentrées jusqu'à un volume final de 2 mL par ultrafiltration sur une membrane en cellulose régénérée (limite de taille 3 kDa, Amicon, Millipore).

## **d. Filtration sur gel**

La dernière étape de purification consiste à injecter l'extrait de protéines (2 mL) sur une colonne d'exclusion par la taille Superdex 75 (XK26/60, 120 mL, GE Healthcare) préalablement équilibrée en tampon D (Hepes-NaOH 50 mM pH 8, NaCl 100 mM). Les protéines sont éluées dans ce même tampon à un débit de 0,7 mL/min. Les fractions éluées sont, comme précédemment, vérifiées par absorbance à 280 nm et SDS-PAGE. Les fractions contenant la protéine d'intérêt sont ensuite concentrées jusqu'à 1-2 mL par ultrafiltration sur une membrane de cellulose régénérée (limite de taille: 3 kDa). La protéine est conservée à -80°C.

Une moyenne de 30 à 40 mg de CnrXs pure ou de ses dérivés est obtenue à partir de 1 litre de milieu de culture LB.

## **2. Purification de NccX**

Le culot des bactéries ayant produit NccX est resuspendu dans du tampon A (Hepes-NaOH 50 mM pH 7,5, NaCl 500 mM, MgCl<sub>2</sub> 10 mM) supplémenté d'un cocktail d'antiprotéases (Complete, Roche, 1 pastille/50 mL) et d'une petite quantité de DNase avant d'être lysé par sonication. Les cellules intactes et les débris cellulaires sont éliminés par centrifugation pendant 20 minutes à 4000 rpm (rotor SLC-6000, Sorvall) et les membranes sont isolées en centrifugeant les surnageants pendant 90 minutes à 45000 rpm (rotor TI70, Beckman). La fraction membranaire est lavée deux fois dans le tampon A. Le dernier culot est resuspendu dans un volume de tampon B correspondant à 7 mL par litre de culture (Hepes-NaOH 50 mM pH 7,5, NaCl 500 mM) et stocké à -80°C.

### **a. Préparation de la fraction membranaire**

Les préparations de membrane sont diluées 10 fois dans le tampon C (Hepes-NaOH 50 mM pH 7,5, NaCl 500 mM, EDTA 10 mM, 0,1% m/v de dodécylphosphocholine (DPC) (2 CMC) (Anagrade, Anatrace) avant d'être homogénéisées par le dispositif Dounce. L'homogénéisation est achevée par le passage du mélange au travers du microfluidizer M110-

P (Microfluidics) à 1000 psi. L'extrait membranaire est ensuite incubé 3 heures dans la glace. Les éléments insolubles sont enlevés par centrifugation pendant 30 minutes à 20000 rpm et le surnageant est concentré jusqu'à 4 mL en utilisant une cellule Diaflo équipée d'une membrane YM-10 (limite de taille 10 kDa, Amicon).

### **b. Filtration sur gel**

Les protéines solubilisées sont injectées sur une colonne Superdex200 (GE-Healthcare) qui a été au préalable équilibrée avec du tampon C. L'élution est conduite à 0,8 mL/min. Les fractions retenues sont vérifiées par spectrophotométrie (absorbance à 280 nm) et analysées sur gel SDS-PAGE. Les fractions contenant NccX sont rassemblées et concentrées jusqu'à 2 mL par centrifugation dans un micro-concentrateur Vivaspin-6 (limite de taille: 30 kDa, Vivascience).

### **c. Chromatographie sur colonne échangeuse d'anions**

La solution de protéine est diluée 10 fois dans le tampon D (Hepes-NaOH 50 mM pH 7,5, Foscholine-12 0,1% m/v) et chargée sur une colonne Q HiTrap (GE-Healthcare). Les protéines sont éluées par un gradient linéaire de NaCl, elles sont rassemblées et concentrées jusqu'à 15 mg/mL par centrifugation dans un micro-concentrateur Vivaspin-6. Les protéines sont aliquotées et stockées à -80°C. A cette étape, NccX est pure sur SDS-PAGE.

Le rendement de purification est d'environ 0,5 à 1 mg de NccX pure à partir de 1 litre de milieu de culture LB.

## **3. Purification de MBP-CnrYp**

La protéine de fusion entre la "Maltose Binding Protein" et CnrYp (MBP-CnrYp ou MYp) ainsi que ses dérivés se purifient selon le même protocole. Toutes les étapes de purification s'effectuent à 4°C.

## **a. Lyse cellulaire**

Le culot des cellules récoltées après induction de la production de MBP-CnrYp est resuspendu dans du tampon de lyse (Hepes-NaOH 100 mM pH 7,5, NaCl 200 mM, EDTA 4 mM) supplémenté d'un mélange d'antiprotéases (Complete, Roche).

Les cellules sont lysées par sonication (3 minutes efficaces : 5 secondes de sonication active, 15 secondes de refroidissement, à une amplitude de 70%). Le lysat est ensuite centrifugé à 45000 rpm pendant 45 minutes à 4°C (rotor Ti70, Beckman).

## **b. Chromatographie d'affinité : "Maltose Binding Protein" et résine d'amylose**

Le lysat clarifié est dilué 5 fois dans du tampon A (Hepes-NaOH 50 mM pH 7,5, NaCl 20 mM, EDTA 2 mM, DTT 2 mM) avant d'être déposé sur 4 colonnes de 5 mL (résine d'amylose, Biolabs) préalablement équilibrées en tampon A. La résine est lavée par 3 volumes de colonne (VC) de tampon A. L'élution s'effectue par 5 VC de tampon B (Hepes-NaOH 100 mM pH 7,5, NaCl 200 mM, EDTA 4 mM, Maltose 10 mM). Les fractions retenues sont choisies après analyse par SDS-PAGE.

## **c. Chromatographie sur colonne échangeuse d'anions**

L'échantillon est injecté (0,5 mL/min) sur une colonne Q sepharose (HiTrap-HP, 5 mL, GE Healthcare) préalablement équilibrée en tampon A. La colonne est lavée par 5 VC de tampon A à 1 mL/min. La protéine est éluée par un gradient linéaire de 0 à 100% de tampon C (Hepes-NaOH 50 mM pH 7,5, NaCl 1 M, EDTA 2 mM, DTT 2 mM). Les fractions à retenir sont choisies d'après les chromatogrammes (absorbance à 280 nm) et les analyses par SDS-PAGE. Les fractions retenues sont concentrées jusqu'à un volume final de 2 mL par ultrafiltration sur une membrane de cellulose régénérée (limite de taille 3 kDa, Millipore).

## **d. Filtration sur gel**

La dernière étape de purification consiste à injecter l'extrait de protéines (2 mL) sur une colonne d'exclusion par la taille Superdex 75 (XK26/60, 120 mL, GE Healthcare) préalablement équilibrée en tampon D (Hepes-NaOH 20 mM pH 7,5, NaCl 100 mM). Les protéines sont éluées dans ce même tampon à un débit de 0,7 mL/min. Les fractions éluées sont, comme précédemment, vérifiées par absorbance à 280 nm et sur gel SDS-PAGE. Les



fractions contenant la protéine d'intérêt sont ensuite concentrées jusqu'à 1-2 mL par ultrafiltration sur une membrane de cellulose régénérée (limite de taille : 3 kDa). La protéine est conservée à -80°C.

Une moyenne de 10 mg de protéine (MBP-CnrYp ou dérivés) est obtenue après la dernière de purification à partir de 1 litre de milieu de culture LB.

## **4. Purification de Strep-CnrH:CnrYc**

Le culot de cellules ayant surproduit CnrH fusionnée à une étiquette Strep (WSHPQFEK) (Strep-CnrH) et CnrYc est repris dans du tampon A (Tris-HCl 20 mM pH 8, NaCl 40 mM, KCl 20 mM, Glycérol 5%) supplémenté d'un cocktail d'antiprotéases CLAPA (Chymostatine 1 µg/mL, Leupeptine 1 µg/mL, Antipaine 1 µg/mL, Pepstatine 1 µg/mL, Aprotinine 5 µg/mL, Sigma-Aldrich). Les cellules sont ensuite lysées.

Le complexe CnrH:CnrYc a été obtenu pur à partir du lysat clarifié (30000 rpm, 20 min, rotor Ti70, Beckman) après 3 étapes de purification.

### **a. Chromatographie d'affinité (par gravité) : étiquette Strep:StrepTactine**

Le lysat clarifié est déposé sur une colonne StrepTactine (10 mL, IBA) préalablement équilibrée en tampon A. La colonne est lavée par 3 CV de tampon A. L'éluion s'effectue par 10 CV de tampon B (Tris-HCl 20 mM pH 8, NaCl 40 mM, KCl 20 mM, Glycérol 5%, Desthiobiotine 2,5 mM). Les fractions retenues sont choisies après analyse par SDS-PAGE.

### **b. Chromatographie d'affinité (interactions hydrophobes)**

L'échantillon est injecté sur une colonne HiTrap Phenyl sepharose (1 mL, GE Healthcare) préalablement équilibrée en tampon C (Hepes-NaOH 50 mM pH 7, KCl 1 M, DTT 2 mM). La colonne est lavée avec du tampon C. Le complexe Strep-CnrH:CnrYc est présent dans la fraction non retenue. Les fractions choisies sont concentrées par ultrafiltration sur une membrane de cellulose régénérée (limite de taille 5 kDa, Millipore).

### **c. Filtration sur gel**

L'échantillon est enfin injecté sur une colonne Superdex75 (GE Healthcare) préalablement équilibrée en tampon D (Hepes-NaOH 50 mM pH 8, NaCl 200 mM, KCl 20 mM, DTT 2 mM). Les protéines sont éluées dans ce même tampon. Les fractions éluées sont, comme précédemment, vérifiées par absorbance à 280 nm et par SDS-PAGE. Les fractions contenant les protéines d'intérêt sont ensuite concentrées et conservées un mois à 4°C.

Une moyenne de 1 mg de complexe est obtenue après la dernière de purification à partir de 1 litre de milieu de culture LB.

## **II. Techniques biochimiques**

### **A. Séparation des protéines par électrophorèse en conditions dénaturantes (SDS-PAGE)**

Nous avons utilisé la technique décrite par Schägger *et al.*, (Schägger et von Jagow, 1987) (SDS-PAGE Tris-Tricine) qui permet une très bonne séparation des protéines de masses moléculaires comprises entre 5 et 20 kDa ainsi que celle décrite par Laemmli *et al.*, (Laemmli, 1970) (SDS-PAGE Tris-Glycine) qui permet une bonne séparation des protéines de 20 kDa et plus.

### **B. Révélation**

#### **1. Coloration au bleu de Coomassie**

Les gels de polyacrylamide sont incubés dans une solution composée de 10% (v/v) d'acide acétique, de 40% (v/v) d'éthanol et de 0,1% (m/v) de Serva blue R250 (bleu de Coomassie). On lave ensuite le colorant présent dans le gel mais non fixé aux protéines avec de l'éthanol 30%.

## 2. Western Blot

Une fois séparées sur gel SDS-PAGE, les protéines sont transférées sur une membrane de PolyVinylidene Fluoride (PVDF, 0,45  $\mu\text{m}$ , Millipore) par un transfert liquide ("Trans-Blot Cell", Biorad). La membrane PVDF est trempée dans du méthanol 100% avant d'être placée du côté anode. Le gel de polyacrylamide est placé du côté cathode. Une feuille de papier filtre et une éponge sont placées de chaque côté. Le transfert des protéines chargées négativement se fait de la cathode vers l'anode.

La migration s'effectue à 100 V et 350 mA pendant 35 minutes dans le tampon de transfert (Tris-HCl 0,03 M, Glycine 0,2 M, Méthanol 20%). La membrane est ensuite réhydratée par une solution de PBS-Tween (Phosphate de sodium 10 mM, NaCl 140 mM, Tween-20 0,03% (v/v), pH 7,2), saturée 30 minutes dans du PBS-Tween-Lait 5% (m/v). La membrane est ensuite incubée pendant 1 heure en présence d'anticorps primaires anti-CnrXs (dilués au 1/10000 dans du PBS-Tween) ou anti-CnrYc (dilués au 1/5000 dans du PBS-Tween) ou d'anti-énolase (dilués au 1/10000 dans du PBS-Tween) ou d'anti-MBP (dilués au 1/30000 dans du PBS-Tween-Lait 5%, New England Biolabs). Après lavage dans du PBS-Tween, la membrane est incubée pendant 1 heure en présence d'anticorps secondaires couplés à une peroxydase (dilués au 1/10000 dans du PBS-Tween). Après plusieurs lavages au PBS-Tween, les anticorps s'étant couplés aux protéines d'intérêt sont détectés grâce au réactif ECL (Luminata Forte, Millipore). Le signal de chimioluminescence est détecté par exposition d'un film radiographique (Amersham Hyperfilm ECF, GE Healthcare) sur la membrane de PVDF.

Les intensités des bandes révélées par Western Blot peuvent être quantifiées via le logiciel "ImageJ" si le signal n'est pas saturé (signal gris plutôt que noir, absence de halo).

## C. Dosage des protéines

Le dosage protéique est effectué selon la méthode de Bradford (Biorad) avec la sérum albumine bovine comme standard (Bradford, 1976).

## **D. Dosage colorimétrique des métaux avec du 4-(2-pyridylazo)resorcinol (PAR)**

Le contenu métallique de CnrXs $\Delta$ Nter après purification a été déterminé par un dosage colorimétrique avec du 4-(2-pyridylazo)resorcinol (PAR) (McCall et Fierke, 2000).

Dans cette expérience, la protéine d'intérêt est dénaturée afin de libérer les métaux en complexe avec la protéine dans la solution. Le PAR chélate les métaux sous la forme d'un complexe (2 PAR pour 1 métal). Le PAR en complexe avec du manganèse, du nickel, du zinc et du cadmium voit son absorbance augmenter à 500 nm alors que la longueur d'onde maximale du PAR en complexe avec du cobalt ou du cuivre est de 514 nm (McCall et Fierke, 2000). L'absorbance augmente de 0,032 par  $\mu$ M de métal fixé au PAR.

CnrXsdNter (20  $\mu$ M) a été incubée avec 4 M de guanidine-HCl. La solution de PAR utilisée contient 0,4 mM de PAR, 50 mM d'Hepes pH 7,5 traité au chelex (résine permettant d'éliminer les traces de métaux) et 0,1 mM de NaOH.

Dans un premier temps le spectre (solution de PAR) de référence est mesuré. L'absorbance entre 220 nm et 700 nm est ensuite mesurée après ajout de 0,5  $\mu$ M de CnrXsdNter et 3 minutes d'attente de l'échantillon dans le noir. Des spectres d'échantillon avec des concentrations croissantes de CnrXsdNter (1  $\mu$ M par 1  $\mu$ M) sont ensuite mesurés. Pour CnrXsdNter, les spectres dans les différentes conditions sont similaires ce qui indique que la protéine n'est pas complexée avec du métal.

### III. Tests d'interaction *in vivo*

#### A. Test TOXCAT

##### 1. Constructions utilisées

L'homo-oligomérisation des segments transmembranaires (TM) de NccX, NccY et leurs mutants est testée en utilisant le test TOXCAT (Russ et Engelman, 1999). Les vecteurs d'expression pccKAN, pccGpA-WT, pccGpA-G83I et la souche *E. coli* MM39 déficiente en *malE* codant la Maltose Binding Protein (MBP) nous ont été aimablement fournis par Donald Engelman (Yale University, New Haven, CT). NccX et NccY et leurs dérivés (longueurs différentes ou mutants) sont exprimés en fusion entre la MBP du côté périplasmique (extrémité C-terminale) et le domaine N-terminal de la protéine ToxR du côté cytoplasmique (N-terminale).

Le protocole utilisé est une variante de celui utilisé par Wei *et al.*, (Wei *et al.*, 2011). *nccX* et *nccY* sont amplifiés par PCR à partir des plasmides pET30b *nccX* et pET30b *nccY* et clonés dans pccKAN entre les sites de restriction NheI et BamHI.

Les plasmides obtenus (pccNccX-FL, pccNccY-FL et leurs dérivés) sont transformés dans les souches de clonage XL-10 (Agilent). Les séquences des différents inserts sont vérifiées par séquençage de l'ADN. La séquence des différentes chimères est présentée dans la deuxième partie du chapitre Résultats. Le plasmide résultant est ensuite transformé dans les cellules MM39.

##### 2. Vérification de la topologie des protéines chimériques

###### a. Test de complémentation *malE*

La topologie des protéines chimériques est vérifiée avec le test de complémentation *malE* qui consiste à étaler des cellules MM39 transformées par les plasmides contenant le gène codant la protéine chimérique d'intérêt sur une boîte gélosée contenant du milieu M9 supplémenté en ampicilline et avec 0,4% de maltose comme unique source de carbone. Ces

boîtes sont ensuite incubées à 37°C pendant 2 jours. Les cellules exprimant la MBP dans le périplasme poussent ce qui indique que les protéines chimériques ont une bonne orientation et insertion dans la membrane. Dans le cas contraire (pas de croissance des bactéries) comme pour la chimère ToxR-NccX-FL-MBP, nous avons effectué un fractionnement cellulaire pour vérifier si la protéine chimérique s'insère dans la membrane (voir ci-dessous).

## **b. Fractionnement cellulaire**

Le lysat des cellules d'*E. coli* MM39 contenant les constructions ToxR-NccX-FL-MBP ou ToxR-GpA-WT-MBP utilisé pour le test TOXCAT est centrifugé à 4500 x g pendant 20 minutes à 4°C (Thermo Scientific). Le culot (cellules intactes, débris cellulaires et corps d'inclusion) est repris dans du tampon de sonication (Tris-HCl 25 mM pH 8, EDTA 2 mM) et lysé une nouvelle fois pour qu'il y ait le minimum de cellules intactes. Le culot obtenu par centrifugation à 4500 x g pendant 20 minutes à 4°C est lavé une première fois avec du tampon de sonication contenant 0,5% de Triton X-100 et une deuxième fois sans le détergent. Les corps d'inclusion ("Inclusion Body", IB) sont obtenus par centrifugation à 4500 x g pendant 20 minutes à 4°C. Le surnageant issu de la première centrifugation est centrifugé à 118000 x g pendant 1 heure à 4°C (Beckmann) résultant en une fraction membranaire (culot) et un surnageant (S1). La fraction membranaire est resoumise à un cycle de basse et haute vitesse de centrifugation. Le nouveau surnageant (S2) et la fraction membranaire (M) ainsi que S1 et IB sont analysés par Western Blot anti-MBP.

## **3. Culture et expression**

Une fois la topologie des chimères vérifiée, une colonie de cellules MM39 transformées par le plasmide d'intérêt est inoculée dans 3 mL de LB supplémentés d'ampicilline. La culture et l'induction sont détaillées dans la partie Culture et Expression (partie I.F). Un volume correspondant à une A600 de 90 est centrifugé à 4000 rpm (Heraeus megafuge 16, ThermoScientific) pendant 30 minutes à 4°C. Le culot est lavé avec 5 mL de tampon de sonication (Tris-HCl 25 mM pH 8, EDTA 2 mM) avant d'être centrifugé comme précédemment. Le culot peut être conservé à - 80°C. Le culot est resuspendu avec 30 mL de tampon de sonication et d'un cocktail d'anti-protéases (CLAPA) puis lysé (3 min efficaces : 1

min de sonication active, 5 min de refroidissement, 45% d'amplitude). Le lysat est centrifugé à 13000 x g (Sorvall) pendant 10 minutes à 4°C. Le surnageant est utilisé pour mesurer l'activité Chloramphénicol Acétyl-Transférase (CAT) qui rend compte de la dimérisation des ToxR médiée par les segments TM.

## 4. Mesure de l'activité CAT

L'enzyme CAT catalyse l'O-acétylation du chloramphénicol (Cm) en utilisant l'acétylCoA comme donneur d'acétyl. Les produits de cette réaction sont le Cm-3-acétate et du Coenzyme A. A pH 7,8, le groupe thiol du coenzyme A va former un pont disulfure avec une partie du dinitrobenzoate. Les produits de cette réaction sont un dérivé du CoA et le thionitrobenzoate qui absorbe à 412 nm. L'activité CAT est alors facilement mesurée en suivant l'absorbance à 412 nm au cours du temps.

L'activité CAT est mesurée, à température ambiante, de la manière suivante : 12 µL de surnageant sont ajoutés à 300 µL de tampon de réaction (Tris-HCl 0,1 M pH 7,8, acétyl-CoA 0,1 mM, acide 5,5'-dithiobis-(2-nitrobenzoïque) (DTNB) 0,4 mg/mL), et l'absorbance à 412 nm est mesurée pendant 4 minutes pour établir la vitesse basale d'hydrolyse de l'acétyl-CoA. 12 µL de chloramphénicol à 2,5 mM sont ensuite ajoutés au mélange et l'absorbance est suivie pendant 6 minutes. La pente correspondant à la ligne de base est soustraite à la pente après ajout du dernier substrat. L'activité CAT ( $\Delta A_{412}/\text{min}$ ) est mesurée sur le spectrophotomètre UV-visible (Cary 50 BIO). Les mesures sont effectuées sur trois expériences indépendantes et trois fois pour chaque expérience. Les activités sont rapportées à la quantité totale de protéines (activité spécifique,  $\Delta A_{412}/\text{min}/\text{mg}$ ).

En parallèle, les cellules en suspension ont été utilisées pour mesurer la quantité de protéines chimériques par Western Blot anti-MBP.

## **B. Test *in vivo* de l'inhibition de CnrH par la partie cytoplasmique de CnrY (CnrYc)**

Des cellules d'*E. coli* TOP 10 sont cotransformées avec deux plasmides : le plasmide pTrc contenant les gènes *cnrYXH* sous le contrôle du promoteur *trc*, inductible par l'IPTG et un second plasmide pACYC contenant le gène codant la GFP (Green Fluorescent Protein) en aval du promoteur de *cnrY*, inductible par CnrH.

Les bactéries sont cultivées toute une nuit à 30°C. Après une induction de *cnrYXH* avec 0,5 mM d'IPTG pendant 3 heures à 30°C, 1 mL de suspension cellulaire est centrifugée à 13000 g pendant une minute. Le culot est ensuite lavé trois fois avec 1 mL de PBS. Le culot est enfin resuspendu de sorte à obtenir une absorbance à 600 nm de 0,2. Des aliquots de 100 µL à cette absorbance sont déposés en double sur une plaque 96 puits à fond noir pour la mesure de la fluorescence. La fluorescence de la GFP est mesurée à 520 nm dans chaque puits par un FluoStar Optima (BMGLabtech) après excitation à 490 nm.

Les différentes souches sont comparées dans une même expérience et chaque expérience est effectuée au moins trois fois. La souche surproduisant CnrH seule, a été utilisée pour évaluer le taux maximal de fluorescence.

En parallèle, les cellules en suspension issues des lavages ont été utilisées pour contrôler quantitativement la production de CnrX par Western Blot avec des anticorps dressés contre CnrX.

## **IV. Test d'interaction *in vitro***

### **A. Expérience de coprécipitation**

La coprécipitation permet d'étudier les interactions entre deux macromolécules : protéine:protéine par exemple. Pour étudier l'interaction entre CnrX et CnrY dans le



périplasme, nous avons utilisé les domaines périplasmiques de CnrX (CnrXs) et de CnrY (CnrYp). CnrYp étant hydrophobe, nous l'avons fusionné à la MBP (MBP-CnrYp) qui est un des meilleurs acteurs de solubilité décrits (Fox *et al.*, 2003).

L'amylose interagit spécifiquement avec la MBP. On teste donc l'interaction directe entre CnrYp et CnrXs en suivant la capacité de la MBP-CnrYp à retenir CnrXs sur la résine d'amylose.

Différentes constructions basées sur la MBP ont été testées : la MBP seule, la protéine de fusion MBP-CnrYp-wt, et différents mutants de CnrYp. Les billes d'amylose préalablement équilibrées en tampon A (Hepes-NaOH 20 mM pH 7,5 et NaCl 100 mM) sont incubées avec du lysat de MBP-CnrYp-wt ou leurs dérivés pendant 20 minutes dans la glace en homogénéisant de temps en temps. Les protéines du surnageant (protéines non fixées) sont éliminées après centrifugation à 3100 x g (Sigma) pendant 5 minutes à 4°C. Les billes d'amylose sont ensuite lavées avec 500 µL de PBS. Les protéines non fixées sont éliminées comme précédemment.

Du lysat de CnrXs ou ses dérivés est ensuite ajouté aux billes:MBP-CnrYp. Le mélange est agité sur une roue pendant 1 heure en chambre froide, puis centrifugé pendant 5 minutes à 3100 x g. Les billes restantes sont lavées deux fois avec 1 mL de PBS. Les protéines retenues sont solubilisées par ajout de tampon de dénaturation sur les billes avant d'être déposées sur gel SDS-PAGE.

## **B. Expérience d'échange hydrogène:deutérium couplé à la spectrométrie de masse (HDX-MS)**

### **1. Mise en évidence de l'interaction CnrXs $\Delta$ Nter:MBP-CnrYp**

Pour vérifier si les protéines CnrXs $\Delta$ Nter et MBP-CnrYp interagissent, les expériences HDX-MS ont été effectuées avec la protéine CnrXs $\Delta$ Nter seule et en complexe avec MBP-CnrYp.

10  $\mu\text{L}$  de CnrXs $\Delta\text{Nter}$  (5,4 mM) ou 10  $\mu\text{L}$  du complexe CnrXs $\Delta\text{Nter}$ :MBP-CnrYp (0,18 mM:1,14 mM, stœchiométrie 1:6) est dilué dans 40  $\mu\text{L}$  de tampon deutéré (5 mM Hepes pD 7,4, 1 mM EDTA, 2 mM DTT, 200 mM NaCl, préparation des solutions stocks dans du  $\text{D}_2\text{O}$ ). Le potentiel deutérium (pD) est relié au pH par l'équation :  $\text{pD} = \text{pH} + 0,4$ .

Les échanges isotopiques sont réalisés pendant 15 secondes ou 120 secondes à température ambiante. L'échange est arrêté par dilution de l'échantillon dans 150  $\mu\text{L}$  de tampon de quenching (1M de glycine-HCl pH 2,5). De ce mélange, 10  $\mu\text{L}$  (injection de 270 pmol de CnrXs $\Delta\text{Nter}$  ou de 90 pmol de CnrXs $\Delta\text{Nter}$  et 507 pmol de MBP-CnrYp) sont ensuite dessalés avec 0,03% TFA (tampon A) à 100  $\mu\text{L}/\text{min}$  sur une colonne Trap (Michrom Bioresources). Les protéines sont ensuite éluées avec 50% de tampon B (90% d'acétonitrile et 0,0% de TFA) à 300  $\mu\text{L}/\text{min}$  sur une colonne HPLC C4. Pour éviter le rééchange, les tampons, les vannes et les colonnes sont conservées à 4°C. La totalité du débit est ensuite amenée jusqu'au spectromètre de masse (électrospray-TOF, 6210 Agilent).

## **2. Cartographie de CnrXs $\Delta\text{Nter}$**

Les expériences de cartographie ont été effectuées, en essayant différentes conditions de protéolyse, à l'institut de microbiologie à Prague (République tchèque), par spectrométrie de masse tandem (MS-MS) à l'aide d'un spectromètre de masse Bruker APEX-Q FTMS (9.4 T). Outre les logiciels de Bruker et Agilent, le logiciel HDExaminer (Sierra Analytics) a été utilisé pour le traitement des données.

## **3. Identification des peptides de CnrXs $\Delta\text{Nter}$ impliqués dans l'interaction**

L'expérience HDX-MS est similaire à celle effectuée avec les protéines entières à quelques différences près :

Après avoir arrêté l'échange H/D avec la solution de quenching, les protéines sont digérées sur une colonne de pepsine immobilisée. Le dessalage s'effectue à l'aide d'une colonne Microtrap (Michrom Bioresources). Les peptides sont ensuite séparés sur une colonne HPLC C18 avec un gradient de 15 à 50% de tampon B à 50  $\mu\text{L}/\text{min}$ .

Cette expérience a aussi été réalisée à partir d'une autre concentration du complexe CnrXs $\Delta$ Nter:MBP-CnrYp : 0,38mM pour Xs et 0,77mM pour MBP-Yp (injection de 190 pmol de CnrXs et 385 pmol de MBP-CnrYp, stœchiométrie 1:2)

Toutes les expériences HDX-MS ont été effectuées au moins trois fois et ont été automatisées à l'aide d'un robot PAL System (CTC Analytics) qui permet de piloter toutes les étapes de l'expérience : échange H/D, digestion protéolytique, séparation HPLC et spectrométrie de masse.

# Références bibliographiques

Ades, S.E., Connolly, L.E., Alba, B.M., and Gross, C.A. (1999). The *Escherichia coli* sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. *Genes Dev.* *13*, 2449–2461.

Aldon, D., Brito, B., Boucher, C., and Genin, S. (2000). A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. *EMBO J.* *19*, 2304–2314.

Alpuche Aranda, C.M., Swanson, J.A., Loomis, W.P., and Miller, S.I. (1992). *Salmonella typhimurium* activates virulence gene transcription within acidified macrophage phagosomes. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 10079–10083.

Angerer, A., and Braun, V. (1998). Iron regulates transcription of the *Escherichia coli* ferric citrate transport genes directly and through the transcription initiation proteins. *Arch. Microbiol.* *169*, 483–490.

Antelmann, H., and Helmann, J.D. (2011). Thiol-Based Redox Switches and Gene Regulation. *Antioxid. Redox Signal.* *14*, 1049–1063.

Argüello, J.M., Eren, E., and González-Guerrero, M. (2007). The structure and function of heavy metal transport P1B-ATPases. *BioMetals* *20*, 233–248.

Bader, M.W., Sanowar, S., Daley, M.E., Schneider, A.R., Cho, U., Xu, W., Klevit, R.E., Le Moual, H., and Miller, S.I. (2005). Recognition of Antimicrobial Peptides by a Bacterial Sensor Kinase. *Cell* *122*, 461–472.

Bae, J.-B., Park, J.-H., Hahn, M.-Y., Kim, M.-S., and Roe, J.-H. (2004). Redox-dependent

changes in RsrA, an anti-sigma factor in *Streptomyces coelicolor*: zinc release and disulfide bond formation. *J. Mol. Biol.* *335*, 425–435.

Barik, S., Sureka, K., Mukherjee, P., Basu, J., and Kundu, M. (2010). RseA, the SigE specific anti-sigma factor of *Mycobacterium tuberculosis*, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol. Microbiol.* *75*, 592–606.

Booth, I.R. (2002). Stress and the single cell: Intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. *Int. J. Food Microbiol.* *78*, 19–30.

Borremans, B., Hobman, J.L., Provoost, A., Brown, N.L., and van Der Lelie, D. (2001). Cloning and functional analysis of the pbr lead resistance determinant of *Ralstonia metallidurans* CH34. *J. Bacteriol.* *183*, 5651–5658.

Bourret, R.B., and Silversmith, R.E. (2010). Two-component signal transduction. *Curr. Opin. Microbiol.* *13*, 113–115.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* *72*, 248–254.

Braun, V. (1997). Surface signaling: novel transcription initiation mechanism starting from the cell surface. *Arch. Microbiol.* *167*, 325–331.

Braun, V., Mahren, S., and Sauter, A. (2006). Gene Regulation by Transmembrane Signaling. *Biometals* *19*, 103–113.

Brown, K.L., and Hughes, K.T. (1995). The role of anti-sigma factors in gene regulation. *Mol. Microbiol.* *16*, 397–404.

Brown, M.S., Ye, J., Rawson, R.B., and Goldstein, J.L. (2000). Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell* *100*, 391–398.

Bury-Moné, S., Nomane, Y., Reymond, N., Barbet, R., Jacquet, E., Imbeaud, S., Jacq, A., and Bouloc, P. (2009). Global Analysis of Extracytoplasmic Stress Signaling in *Escherichia coli*. *PLoS Genet.* *5*, e1000651.

Campagne, S., Damberger, F.F., Kaczmarczyk, A., Francez-Charlot, A., Allain, F.H.-T., and Vorholt, J.A. (2012). Structural basis for sigma factor mimicry in the general stress response

of Alphaproteobacteria. *Proc. Natl. Acad. Sci. U. S. A.* *109*, E1405–E1414.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* *9*, 527–539.

Campbell, E.A., Tupy, J.L., Gruber, T.M., Wang, S., Sharp, M.M., Gross, C.A., and Darst, S.A. (2003). Crystal Structure of *Escherichia coli*  $\sigma$ E with the Cytoplasmic Domain of Its Anti- $\sigma$  RseA. *Mol. Cell* *11*, 1067–1078.

Campbell, E.A., Greenwell, R., Anthony, J.R., Wang, S., Lim, L., Das, K., Sofia, H.J., Donohue, T.J., and Darst, S.A. (2007). A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol. Cell* *27*, 793–805.

Campbell, E.A., Westblade, L.F., and Darst, S.A. (2008). Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr. Opin. Microbiol.* *11*, 121–127.

Chen, G., and Zhang, X. (2010). New insights into S2P signaling cascades: regulation, variation, and conservation. *Protein Sci. Publ. Protein Soc.* *19*, 2015–2030.

Chillappagari, S., Seubert, A., Trip, H., Kuipers, O.P., Marahiel, M.A., and Miethke, M. (2010). Copper Stress Affects Iron Homeostasis by Destabilizing Iron-Sulfur Cluster Formation in *Bacillus subtilis*. *J. Bacteriol.* *192*, 2512–2524.

Cho, U.S., Bader, M.W., Amaya, M.F., Daley, M.E., Klevit, R.E., Miller, S.I., and Xu, W. (2006). Metal Bridges between the PhoQ Sensor Domain and the Membrane Regulate Transmembrane Signaling. *J. Mol. Biol.* *356*, 1193–1206.

Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011). HTRA proteases: regulated proteolysis in protein quality control. *Nat. Rev. Mol. Cell Biol.* *12*, 152–162.

Coudray, N., Valvo, S., Hu, M., Lasala, R., Kim, C., Vink, M., Zhou, M., Provasi, D., Filizola, M., Tao, J., et al. (2013). Inward-facing conformation of the zinc transporter YiiP revealed by cryoelectron microscopy. *Proc. Natl. Acad. Sci.* *110*, 2140–2145.

Cowan, S.W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R.A., Jansonius, J.N., and Rosenbusch, J.P. (1992). Crystal structures explain functional properties of two *E. coli* porins. *Nature* *358*, 727–733.

- Cymer, F., and Schneider, D. (2010). Transmembrane helix-helix interactions involved in ErbB receptor signaling. *Cell Adhes. Migr.* 4, 299–312.
- Cymer, F., Veerappan, A., and Schneider, D. (2012). Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim. Biophys. Acta* 1818, 963–973.
- Dalbey, R.E., Wang, P., and Kuhn, A. (2011). Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* 80, 161–187.
- Darst, S.A., Kubalek, E.W., and Kornberg, R.D. (1989). Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature* 340, 730–732.
- Deng, W., Li, C., and Xie, J. (2013). The underlying mechanism of bacterial TetR/AcrR family transcriptional repressors. *Cell. Signal.* 25, 1608–1613.
- Dian, C., Vitale, S., Leonard, G.A., Bahlawane, C., Fauquant, C., Leduc, D., Muller, C., de Reuse, H., Michaud-Soret, I., and Terradot, L. (2011). The structure of the *Helicobacter pylori* ferric uptake regulator Fur reveals three functional metal binding sites. *Mol. Microbiol.* 79, 1260–1275.
- Diels, L., Dong, Q., van der Lelie, D., Baeyens, W., and Mergeay, M. (1995). The *czc* operon of *Alcaligenes eutrophus* CH34: from resistance mechanism to the removal of heavy metals. *J. Ind. Microbiol.* 14, 142–153.
- Diels, L., Van Roy, S., Taghavi, S., and Van Houdt, R. (2009). From industrial sites to environmental applications with *Cupriavidus metallidurans*. *Antonie Van Leeuwenhoek* 96, 247–258.
- DiRita, V.J. (1992). Co-ordinate expression of virulence genes by ToxR in *Vibrio cholerae*. *Mol. Microbiol.* 6, 451–458.
- DiRita, V.J., and Mekalanos, J.J. (1991). Periplasmic interaction between two membrane regulatory proteins, ToxR and ToxS, results in signal transduction and transcriptional activation. *Cell* 64, 29–37.
- Duffus, J. (2002). “Heavy metals” - a meaningless term? *Pure Appl Chem* 74, 793–807.

- Egler, M., Grosse, C., Grass, G., and Nies, D.H. (2005). Role of the extracytoplasmic function protein family sigma factor RpoE in metal resistance of *Escherichia coli*. *J. Bacteriol.* *187*, 2297–2307.
- Eitinger, T., Suhr, J., Moore, L., and Smith, J.A.C. (2005). Secondary Transporters for Nickel and Cobalt Ions: Theme and Variations. *Biometals* *18*, 399–405.
- Enz, S., Mahren, S., Stroeder, U.H., and Braun, V. (2000). Surface signaling in ferric citrate transport gene induction: interaction of the FecA, FecR, and FecI regulatory proteins. *J. Bacteriol.* *182*, 637–646.
- Escolar, L., Pérez-Martín, J., and de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. *J. Bacteriol.* *181*, 6223–6229.
- Espenshade, P.J., and Hughes, A.L. (2007). Regulation of Sterol Synthesis in Eukaryotes. *Annu. Rev. Genet.* *41*, 401–427.
- Fairman, J.W., Noinaj, N., and Buchanan, S.K. (2011). The structural biology of  $\beta$ -barrel membrane proteins: a summary of recent reports. *Curr. Opin. Struct. Biol.* *21*, 523–531.
- Faller, P., Hureau, C., and La Penna, G. (2014). Metal Ions and Intrinsically Disordered Proteins and Peptides: From Cu/Zn Amyloid- $\beta$  to General Principles. *Acc. Chem. Res.* *47*, 2252–2259.
- Farrugia, M.A., Macomber, L., and Hausinger, R.P. (2013). Biosynthesis of the Urease Metallocenter. *J. Biol. Chem.* *288*, 13178–13185.
- Faruque, S.M., Asadulghani, Rahman, M.M., Waldor, M.K., and Sack, D.A. (2000). Sunlight-Induced Propagation of the Lysogenic Phage Encoding Cholera Toxin. *Infect. Immun.* *68*, 4795–4801.
- Ferguson, A.D., Chakraborty, R., Smith, B.S., Esser, L., van der Helm, D., and Deisenhofer, J. (2002). Structural basis of gating by the outer membrane transporter FecA. *Science* *295*, 1715–1719.
- Ferguson, A.D., Amezcua, C.A., Halabi, N.M., Chelliah, Y., Rosen, M.K., Ranganathan, R., and Deisenhofer, J. (2007). Signal transduction pathway of TonB-dependent transporters. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 513–518.



- Fillat, M.F. (2014). The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Arch. Biochem. Biophys.* *546*, 41–52.
- Flynn, J.M. (2004). Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev.* *18*, 2292–2301.
- Fox, J.D., Rutzahn, K.M., Bucher, M.H., and Waugh, D.S. (2003). Maltodextrin-binding proteins from diverse bacteria and archaea are potent solubility enhancers. *FEBS Lett.* *537*, 53–57.
- Francez-Charlot, A., Frunzke, J., Reichen, C., Ebnetter, J.Z., Gourion, B., and Vorholt, J.A. (2009). Sigma factor mimicry involved in regulation of general stress response. *Proc. Natl. Acad. Sci.* *106*, 3467–3472.
- Gao, R., and Stock, A.M. (2010). Molecular strategies for phosphorylation-mediated regulation of response regulator activity. *Curr. Opin. Microbiol.* *13*, 160–167.
- García Vescovi, E., Soncini, F.C., and Groisman, E.A. (1996). Mg<sup>2+</sup> as an extracellular signal: environmental regulation of *Salmonella* virulence. *Cell* *84*, 165–174.
- Garvis, S.G., Beuzón, C.R., and Holden, D.W. (2001). A role for the PhoP/Q regulon in inhibition of fusion between lysosomes and *Salmonella*-containing vacuoles in macrophages. *Cell. Microbiol.* *3*, 731–744.
- Goris, J., De Vos, P., Coenye, T., Hoste, B., Janssens, D., Brim, H., Diels, L., Mergeay, M., Kersters, K., and Vandamme, P. (2001). Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis* Steinle et al. 1998 emend. *Int. J. Syst. Evol. Microbiol.* *51*, 1773–1782.
- Grass, G., Große, C., and Nies, D.H. (2000). Regulation of the *cnr* Cobalt and Nickel Resistance Determinant from *Ralstonia* sp. Strain CH34. *J. Bacteriol.* *182*, 1390.
- Grass, G., Fricke, B., and Nies, D.H. (2005). Control of Expression of a Periplasmic Nickel Efflux Pump by Periplasmic Nickel Concentrations. *Biometals* *18*, 437–448.
- Groisman, E.A. (2001). The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol.* *183*, 1835–1842.

- Grosse, C., Grass, G., Anton, A., Franke, S., Santos, A.N., Lawley, B., Brown, N.L., and Nies, D.H. (1999). Transcriptional organization of the *czc* heavy-metal homeostasis determinant from *Alcaligenes eutrophus*. *J. Bacteriol.* *181*, 2385–2393.
- Grosse, C., Friedrich, S., and Nies, D.H. (2007). Contribution of extracytoplasmic function sigma factors to transition metal homeostasis in *Cupriavidus metallidurans* strain CH34. *J. Mol. Microbiol. Biotechnol.* *12*, 227–240.
- Gruber, T.M., and Gross, C.A. (2003). Assay of *Escherichia coli* RNA Polymerase: Sigma–Core Interactions. In *Methods in Enzymology*, Sankar Adhya and Susan Garges, ed. (Academic Press), pp. 206–212.
- Gunn, J.S., and Miller, S.I. (1996). PhoP-PhoQ activates transcription of *pmrAB*, encoding a two-component regulatory system involved in *Salmonella typhimurium* antimicrobial peptide resistance. *J. Bacteriol.* *178*, 6857–6864.
- Guo, L., Lim, K.B., Poduje, C.M., Daniel, M., Gunn, J.S., Hackett, M., and Miller, S.I. (1998). Lipid A Acylation and Bacterial Resistance against Vertebrate Antimicrobial Peptides. *Cell* *95*, 189–198.
- Gupta, S., Chai, J., Cheng, J., D’Mello, R., Chance, M.R., and Fu, D. (2014). Visualizing the kinetic power stroke that drives proton-coupled zinc(ii) transport. *Nature advance online publication*.
- Hantke, K. (2001). Iron and metal regulation in bacteria. *Curr. Opin. Microbiol.* *4*, 172–177.
- Hayden, J.D., and Ades, S.E. (2008). The extracytoplasmic stress factor, sigmaE, is required to maintain cell envelope integrity in *Escherichia coli*. *PloS One* *3*, e1573.
- Heddle, J., Scott, D.J., Unzai, S., Park, S.-Y., and Tame, J.R.H. (2003). Crystal structures of the liganded and unliganded nickel-binding protein NikA from *Escherichia coli*. *J. Biol. Chem.* *278*, 50322–50329.
- Heinrich, J., and Wiegert, T. (2009). Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Res. Microbiol.* *160*, 696–703.
- Helmann, J.D. (2002). The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* *46*, 47–110.

Helmann, J.D., and Chamberlin, M.J. (1988). Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* 57, 839–872.

Herrou, J., Foreman, R., Fiebig, A., and Crosson, S. (2010). A structural model of anti-anti- $\sigma$  inhibition by a two-component receiver domain: the PhyR stress response regulator. *Mol. Microbiol.* 78, 290–304.

Herrou, J., Rotskoff, G., Luo, Y., Roux, B., and Crosson, S. (2012). Structural basis of a protein partner switch that regulates the general stress response of  $\alpha$ -proteobacteria. *Proc. Natl. Acad. Sci.* 109, E1415–E1423.

Higgins, C.F., and Linton, K.J. (2001). Structural biology. The xyz of ABC transporters. *Science* 293, 1782–1784.

Hubert, P., Sawma, P., Duneau, J.-P., Khao, J., Hénin, J., Bagnard, D., and Sturgis, J. (2010). Single-spanning transmembrane domains in cell growth and cell-cell interactions: More than meets the eye? *Cell Adhes. Migr.* 4, 313–324.

Imlay, K.R., and Imlay, J.A. (1996). Cloning and analysis of *sodC*, encoding the copper-zinc superoxide dismutase of *Escherichia coli*. *J. Bacteriol.* 178, 2564.

Iwig, J.S., and Chivers, P.T. (2010). Coordinating intracellular nickel–metal-site structure–function relationships and the NikR and RcnR repressors. *Nat. Prod. Rep.* 27, 658.

Jacquamet, L., Traoré, D. a. K., Ferrer, J.-L., Proux, O., Testemale, D., Hazemann, J.-L., Nazarenko, E., El Ghazouani, A., Caux-Thang, C., Duarte, V., et al. (2009). Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Mol. Microbiol.* 73, 20–31.

Janssen, P.J., Van Houdt, R., Moors, H., Monsieurs, P., Morin, N., Michaux, A., Benotmane, M.A., Leys, N., Vallaey, T., Lapidus, A., et al. (2010). The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. *PloS One* 5, e10433.

Jongeneel, C.V., Bouvier, J., and Bairoch, A. (1989). A unique signature identifies a family of zinc-dependent metallopeptidases. *FEBS Lett.* 242, 211–214.

Juhnke, S., Peitzsch, N., Hübener, N., Grosse, C., and Nies, D.H. (2002). New genes involved in chromate resistance in *Ralstonia metallidurans* strain CH34. *Arch. Microbiol.* 179, 15–25.

- Kang, J.G., Paget, M.S., Seok, Y.J., Hahn, M.Y., Bae, J.B., Hahn, J.S., Kleanthous, C., Buttner, M.J., and Roe, J.H. (1999). RsrA, an anti-sigma factor regulated by redox change. *EMBO J.* *18*, 4292–4298.
- Katritch, V., Fenalti, G., Abola, E.E., Roth, B.L., Cherezov, V., and Stevens, R.C. (2014). Allosteric sodium in class A GPCR signaling. *Trends Biochem. Sci.* *39*, 233–244.
- Kim, I., Stiefel, A., Plantör, S., Angerer, A., and Braun, V. (1997). Transcription induction of the ferric citrate transport genes via the N-terminus of the FecA outer membrane protein, the Ton system and the electrochemical potential of the cytoplasmic membrane. *Mol. Microbiol.* *23*, 333–344.
- Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002). GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry (Mosc.)* *41*, 5990–5997.
- Koebnik, R., Locher, K.P., and Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* *37*, 239–253.
- Kolmar, H., Hennecke, F., Götze, K., Janzer, B., Vogt, B., Mayer, F., and Fritz, H.J. (1995). Membrane insertion of the bacterial signal transduction protein ToxR and requirements of transcription activation studied by modular replacement of different protein substructures. *EMBO J.* *14*, 3895.
- Kräutler, B. (2005). Vitamin B12: chemistry and biochemistry. *Biochem. Soc. Trans.* *33*, 806.
- Kulathila, R., Kulathila, R., Indic, M., and van den Berg, B. (2011). Crystal structure of *Escherichia coli* CusC, the outer membrane component of a heavy metal efflux pump. *PloS One* *6*, e15610.
- Kutsukake, K., and Iino, T. (1994). Role of the FliA-FlgM regulatory system on the transcriptional control of the flagellar regulon and flagellar formation in *Salmonella typhimurium*. *J. Bacteriol.* *176*, 3598–3605.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* *227*, 680–685.
- Lal, M., and Caplan, M. (2011). Regulated intramembrane proteolysis: signaling pathways and biological functions. *Physiol. Bethesda Md* *26*, 34–44.

- De Las Peñas, A., Connolly, L., and Gross, C.A. (1997). SigmaE is an essential sigma factor in *Escherichia coli*. *J. Bacteriol.* *179*, 6862–6864.
- Lebrette, H., Iannello, M., Fontecilla-Camps, J.C., and Cavazza, C. (2013). The binding mode of Ni-(L-His)<sub>2</sub> in NikA revealed by X-ray crystallography. *J. Inorg. Biochem.* *121*, 16–18.
- Van der Lelie, D., Corbisier, P., Baeyens, W., Wuertz, S., Diels, L., and Mergeay, M. (1994). The use of biosensors for environmental monitoring. *Res. Microbiol.* *145*, 67–74.
- Van der Lelie, D., Schwuchow, T., Schwidetzky, U., Wuertz, S., Baeyens, W., Mergeay, M., and Nies, D.H. (1997). Two-component regulatory system involved in transcriptional control of heavy-metal homeostasis in *Alcaligenes eutrophus*. *Mol. Microbiol.* *23*, 493–503.
- Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J., and Engelman, D.M. (1992). Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry (Mosc.)* *31*, 12719–12725.
- Li, C., Peck, H.D., LeGALL, J., and Przybyla, A.E. (1987). Cloning, Characterization, and Sequencing of the Genes Encoding the Large and Small Subunits of the Periplasmic [NiFe]hydrogenase of *Desulfovibrio gigas*. *DNA* *6*, 539–551.
- Li, W., Bottrill, A.R., Bibb, M.J., Buttner, M.J., Paget, M.S.B., and Kleanthous, C. (2003). The Role of zinc in the disulphide stress-regulated anti-sigma factor RsrA from *Streptomyces coelicolor*. *J. Mol. Biol.* *333*, 461–472.
- Li, X., Wang, B., Feng, L., Kang, H., Qi, Y., Wang, J., and Shi, Y. (2009). Cleavage of RseA by RseP requires a carboxyl-terminal hydrophobic amino acid following DegS cleavage. *Proc. Natl. Acad. Sci.* *106*, 14837–14842.
- Liesegang, H., Lemke, K., Siddiqui, R.A., and Schlegel, H.G. (1993). Characterization of the inducible nickel and cobalt resistance determinant *cnr* from pMOL28 of *Alcaligenes eutrophus* CH34. *J. Bacteriol.* *175*, 767–778.
- Llamas, M.A., Imperi, F., Visca, P., and Lamont, I.L. (2014). Cell-surface signaling in *Pseudomonas*: stress responses, iron transport, and pathogenicity. *FEMS Microbiol. Rev.* *38*, 569–597.
- Lonetto, M., Gribskov, M., and Gross, C.A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J. Bacteriol.* *174*, 3843–3849.

- Lonetto, M.A., Brown, K.L., Rudd, K.E., and Buttner, M.J. (1994). Analysis of the *Streptomyces coelicolor* sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. U. S. A. *91*, 7573–7577.
- Long, F., Su, C.-C., Lei, H.-T., Bolla, J.R., Do, S.V., and Yu, E.W. (2012). Structure and mechanism of the tripartite CusCBA heavy-metal efflux complex. Philos. Trans. R. Soc. B Biol. Sci. *367*, 1047–1058.
- De Lorenzo, V., Wee, S., Herrero, M., and Neilands, J.B. (1987). Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor. J. Bacteriol. *169*, 2624–2630.
- Lourenço, R.F., Kohler, C., and Gomes, S.L. (2011). A two-component system, an anti-sigma factor and two paralogous ECF sigma factors are involved in the control of general stress response in *Caulobacter crescentus*. Mol. Microbiol. *80*, 1598–1612.
- Lu, M., Chai, J., and Fu, D. (2009). Structural basis for autoregulation of the zinc transporter YiiP. Nat. Struct. Mol. Biol. *16*, 1063–1067.
- Maddocks, S.E., and Oyston, P.C.F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiol. Read. Engl. *154*, 3609–3623.
- Mahren, S., Schnell, H., and Braun, V. (2005). Occurrence and regulation of the ferric citrate transport system in *Escherichia coli* B, *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Photobacterium luminescens*. Arch. Microbiol. *184*, 175–186.
- Maillard, A.P., Girard, E., Ziani, W., Petit-Härtlein, I., Kahn, R., and Covès, J. (2014). The Crystal Structure of the Anti- $\sigma$  Factor CnrY in Complex with the  $\sigma$  Factor CnrH Shows a New Structural Class of Anti- $\sigma$  Factors Targeting Extracytoplasmic Function  $\sigma$  Factors. J. Mol. Biol. *426*, 2313–2327.
- Mascher, T. (2013). Signaling diversity and evolution of extracytoplasmic function (ECF)  $\sigma$  factors. Curr. Opin. Microbiol. *16*, 148–155.
- Mascher, T., Helmann, J.D., and Udden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol. Mol. Biol. Rev. MMBR *70*, 910–938.
- Massé, E., and Arguin, M. (2005). Ironing out the problem: new mechanisms of iron

homeostasis. *Trends Biochem. Sci.* 30, 462–468.

Matthews, E.E., Zoonens, M., and Engelman, D.M. (2006). Dynamic Helix Interactions in Transmembrane Signaling. *Cell* 127, 447–450.

McCall, K.A., and Fierke, C.A. (2000). Colorimetric and Fluorimetric Assays to Quantitate Micromolar Concentrations of Transition Metals. *Anal. Biochem.* 284, 307–315.

Mecenas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J., and Gross, C.A. (1993). The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev.* 7, 2618–2628.

Melnyk, R.A., Kim, S., Curran, A.R., Engelman, D.M., Bowie, J.U., and Deber, C.M. (2004). The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *J. Biol. Chem.* 279, 16591–16597.

Mergeay, M., Houba, C., and Gerits, J. (1978). Extrachromosomal inheritance controlling resistance to cadmium, cobalt, copper and zinc ions: evidence from curing in a *Pseudomonas* [proceedings]. *Arch. Int. Physiol. Biochim.* 86, 440–442.

Mergeay, M., Nies, D., Schlegel, H.G., Gerits, J., Charles, P., and Gijsegem, F.V. (1985). *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J. Bacteriol.* 162, 328.

Mergeay, M., Monchy, S., Vallaeyts, T., Auquier, V., Benotmane, A., Bertin, P., Taghavi, S., Dunn, J., van der Lelie, D., and Wattiez, R. (2003). *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. *FEMS Microbiol. Rev.* 27, 385–410.

Mettrick, K.A., and Lamont, I.L. (2009). Different roles for anti-sigma factors in siderophore signalling pathways of *Pseudomonas aeruginosa*. *Mol. Microbiol.* 74, 1257–1271.

Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B., and Mijakovic, I. (2009). NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* 9, 116–125.

Miller, S.I., Kukral, A.M., and Mekalanos, J.J. (1989). A two-component regulatory system (phoP phoQ) controls *Salmonella typhimurium* virulence. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5054–5058.

Miller, V.L., Taylor, R.K., and Mekalanos, J.J. (1987). Cholera toxin transcriptional activator toxR is a transmembrane DNA binding protein. *Cell* 48, 271–279.

Misra, T.K. (1992). Bacterial resistances to inorganic mercury salts and organomercurials. *Plasmid* 27, 4–16.

Missiakas, D., Betton, J.M., and Raina, S. (1996). New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* 21, 871–884.

Monchy, S., Benotmane, M.A., Wattiez, R., van Aelst, S., Auquier, V., Borremans, B., Mergeay, M., Taghavi, S., van der Lelie, D., and Vallaes, T. (2006). Transcriptomic and proteomic analyses of the pMOL30-encoded copper resistance in *Cupriavidus metallidurans* strain CH34. *Microbiol. Read. Engl.* 152, 1765–1776.

Monchy, S., Benotmane, M.A., Janssen, P., Vallaes, T., Taghavi, S., van der Lelie, D., and Mergeay, M. (2007). Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* Are Specialized in the Maximal Viable Response to Heavy Metals. *J. Bacteriol.* 189, 7417–7425.

Moon, A.F., Mueller, G.A., Zhong, X., and Pedersen, L.C. (2010). A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci. Publ. Protein Soc.* 19, 901–913.

Moore, D.T., Berger, B.W., and DeGrado, W.F. (2008). Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Struct. Lond. Engl.* 1993 16, 991–1001.

Mulrooney, S.B., and Hausinger, R.P. (2003). Nickel uptake and utilization by microorganisms. *FEMS Microbiol. Rev.* 27, 239–261.

Murakami, K.S., Masuda, S., and Darst, S.A. (2003). Crystallographic analysis of *Thermus aquaticus* RNA polymerase holoenzyme and a holoenzyme/promoter DNA complex. *Methods Enzymol.* 370, 42–53.

Navarro, C., Wu, L.-F., and Mandrand-Berthelot, M.-A. (1993). The nik operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol. Microbiol.* 9, 1181–1191.

Newman, R.H., Zhang, J., and Zhu, H. (2014). Toward a systems-level view of dynamic



phosphorylation networks. *Front. Genet.* *5*, 263.

Nies, D.H. (1995). The cobalt, zinc, and cadmium efflux system CzcABC from *Alcaligenes eutrophus* functions as a cation-proton antiporter in *Escherichia coli*. *J. Bacteriol.* *177*, 2707–2712.

Nies, D.H. (1999). Microbial heavy-metal resistance. *Appl. Microbiol. Biotechnol.* *51*, 730–750.

Nies, D.H. (2000). Heavy metal-resistant bacteria as extremophiles: molecular physiology and biotechnological use of *Ralstonia* sp. CH34. *Extremophiles* *4*, 77–82.

Nies, D.H. (2003). Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.* *27*, 313–339.

Nies, D.H. (2004). Incidence and function of sigma factors in *Ralstonia metallidurans* and other bacteria. *Arch. Microbiol.* *181*, 255–268.

Nies, D.H. (2007). Bacterial Transition Metal Homeostasis. In *Molecular Microbiology of Heavy Metals*, D.H. Nies, and S. Silver, eds. (Springer Berlin Heidelberg), pp. 117–142.

Nies, D.H., Rehbein, G., Hoffmann, T., Baumann, C., and Grosse, C. (2006). Paralogs of genes encoding metal resistance proteins in *Cupriavidus metallidurans* strain CH34. *J. Mol. Microbiol. Biotechnol.* *11*, 82–93.

Nikaido, H., and Vaara, M. (1985). Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* *49*, 1–32.

Ochs, M., Veitinger, S., Kim, I., Welz, D., Angerer, A., and Braun, V. (1995). Regulation of citrate-dependent iron transport of *Escherichia coli*: fecR is required for transcription activation by FecI. *Mol. Microbiol.* *15*, 119–132.

Paget, M.S.B., and Helmann, J.D. (2003). The sigma70 family of sigma factors. *Genome Biol.* *4*, 203.

Palmgren, M.G., and Nissen, P. (2011). P-Type ATPases. *Annu. Rev. Biophys.* *40*, 243–266.

Parkin, E.T., Trew, A., Christie, G., Faller, A., Mayer, R., Turner, A.J., and Hooper, N.M. (2002). Structure-activity relationship of hydroxamate-based inhibitors on the secretases that cleave the amyloid precursor protein, angiotensin converting enzyme, CD23, and pro-tumor

- necrosis factor-alpha. *Biochemistry (Mosc.)* 41, 4972–4981.
- Parkinson, J.S. (1993). Signal transduction schemes of bacteria. *Cell* 73, 857–871.
- Pohl, E., Haller, J.C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E., and Vasil, M.L. (2003). Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol. Microbiol.* 47, 903–915.
- Pompidor, G., Maillard, A.P., Girard, E., Gambarelli, S., Kahn, R., and Covès, J. (2008). X-ray structure of the metal-sensor CnrX in both the apo- and copper-bound forms. *FEBS Lett.* 582, 3954–3958.
- Pompidor, G., Girard, E., Maillard, A., Ramella-Pairin, S., Bersch, B., Kahn, R., and Covès, J. (2009). Biostructural analysis of the metal-sensor domain of CnrX from *Cupriavidus metallidurans* CH34. *Antonie Van Leeuwenhoek* 96, 141–148.
- Psachoulia, E., Marshall, D.P., and Sansom, M.S.P. (2010). Molecular Dynamics Simulations of the Dimerization of Transmembrane  $\alpha$ -Helices. *Acc. Chem. Res.* 43, 388–396.
- Ratledge, C., and Dover, L.G. (2000). Iron metabolism in pathogenic bacteria. *Annu. Rev. Microbiol.* 54, 881–941.
- Reith, F., Etschmann, B., Grosse, C., Moors, H., Benotmane, M.A., Monsieurs, P., Grass, G., Doonan, C., Vogt, S., Lai, B., et al. (2009). Mechanisms of gold biomineralization in the bacterium *Cupriavidus metallidurans*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 17757–17762.
- Reyes-Caballero, H., Campanello, G.C., and Giedroc, D.P. (2011). Metalloregulatory proteins: metal selectivity and allosteric switching. *Biophys. Chem.* 156, 103–114.
- Rhodijs, V.A., Suh, W.C., Nonaka, G., West, J., and Gross, C.A. (2006). Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol.* 4, e2.
- Rowley, G., Spector, M., Kormanec, J., and Roberts, M. (2006). Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nat. Rev. Microbiol.* 4, 383–394.
- Rozycki, T. von, and Nies, D.H. (2009). *Cupriavidus metallidurans*: evolution of a metal-resistant bacterium. *Antonie Van Leeuwenhoek* 96, 115–139.
- Russ, W.P., and Engelman, D.M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci. U. S. A.* 96, 863–868.

- Saier, M.H., Tam, R., Reizer, A., and Reizer, J. (1994). Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. *Mol. Microbiol.* *11*, 841–847.
- Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* *166*, 368–379.
- Schmidt, T., and Schlegel, H.G. (1994). Combined nickel-cobalt-cadmium resistance encoded by the ncc locus of *Alcaligenes xylosoxidans* 31A. *J. Bacteriol.* *176*, 7045–7054.
- Aci-Sèche, S., Sawma, P., Hubert, P., Sturgis, J.N., Bagnard, D., Jacob, L., Genest, M., and Garnier, N. (2014). Transmembrane Recognition of the Semaphorin Co-Receptors Neuropilin 1 and Plexin A1: Coarse-Grained Simulations. *PLoS ONE* *9*, e97779.
- Senes, A., Gerstein, M., and Engelman, D.M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* *296*, 921–936.
- Senes, A., Engel, D.E., and DeGrado, W.F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* *14*, 465–479.
- Sheikh, M.A., and Taylor, G.L. (2009). Crystal structure of the *Vibrio cholerae* ferric uptake regulator (Fur) reveals insights into metal co-ordination. *Mol. Microbiol.* *72*, 1208–1220.
- Silhavy, T.J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. *Cold Spring Harb. Perspect. Biol.* *2*, a000414.
- Silver, S., and Phung, L.T. (1996). Bacterial heavy metal resistance: new surprises. *Annu. Rev. Microbiol.* *50*, 753–789.
- Sohn, J., Grant, R.A., and Sauer, R.T. (2007). Allosteric activation of DegS, a stress sensor PDZ protease. *Cell* *131*, 572–583.
- Soncini, F.C., García Véscovi, E., Solomon, F., and Groisman, E.A. (1996). Molecular basis of the magnesium deprivation response in *Salmonella typhimurium*: identification of PhoP-regulated genes. *J. Bacteriol.* *178*, 5092–5099.
- Sorenson, M.K., and Darst, S.A. (2006). Disulfide cross-linking indicates that FlgM-bound

and free sigma28 adopt similar conformations. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 16722–16727.

Sorenson, M.K., Ray, S.S., and Darst, S.A. (2004). Crystal Structure of the Flagellar  $\sigma$ /Anti- $\sigma$  Complex  $\sigma$ 28/FlgM Reveals an Intact  $\sigma$  Factor in an Inactive Conformation. *Mol. Cell* *14*, 127–138.

Springgate, C.F., and Loeb, L.A. (1973). Mutagenic DNA Polymerase in Human Leukemic Cells. *Proc. Natl. Acad. Sci. U. S. A.* *70*, 245.

Staroń, A., and Mascher, T. (2010). General stress response in  $\alpha$ -proteobacteria: PhyR and beyond. *Mol. Microbiol.* *78*, 271–277.

Staroń, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H., and Mascher, T. (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol. Microbiol.* *74*, 557–581.

Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000). Two-Component Signal Transduction. *Annu. Rev. Biochem.* *69*, 183–215.

Su, C.-C., Long, F., Zimmermann, M.T., Rajashankar, K.R., Jernigan, R.L., and Yu, E.W. (2011). Crystal structure of the CusBA heavy-metal efflux complex of *Escherichia coli*. *Nature* *470*, 558–562.

Taghavi, S., Mergeay, M., Nies, D., and van der Lelie, D. (1997). *Alcaligenes eutrophus* as a model system for bacterial interactions with heavy metals in the environment. *Res. Microbiol.* *148*, 536–551.

Tibazarwa, C., Corbisier, P., Mench, M., Bossus, A., Solda, P., Mergeay, M., Wyns, L., and van der Lelie, D. (2001). A microbial biosensor to predict bioavailable nickel in soil and its transfer to plants. *Environ. Pollut. Barking Essex 1987* *113*, 19–26.

Trepreau, J., de Rosny, E., Duboc, C., Sarret, G., Petit-Hartlein, I., Maillard, A.P., Imberty, A., Proux, O., and Covès, J. (2011a). Spectroscopic Characterization of the Metal-Binding Sites in the Periplasmic Metal-Sensor Domain of CnrX from *Cupriavidus metallidurans* CH34. *Biochemistry (Mosc.)* *50*, 9036–9045.

Trepreau, J., Girard, E., Maillard, A.P., de Rosny, E., Petit-Haertlein, I., Kahn, R., and Covès, J. (2011b). Structural basis for metal sensing by CnrX. *J. Mol. Biol.* *408*, 766–779.

- Trepreau, J., Grosse, C., Mouesca, J.-M., Sarret, G., Girard, E., Petit-Haertlein, I., Kuennemann, S., Desbourdes, C., de Rosny, E., Maillard, A.P., et al. (2014). Metal sensing and signal transduction by CnrX from *Cupriavidus metallidurans* CH34: role of the only methionine assessed by a functional, spectroscopic, and theoretical study. *Met. Integr. Biometal Sci.* *6*, 263–273.
- Tsai, K.-J., Lin, Y.-F., Wong, M.D., Yang, H.H.-C., Fu, H.-L., and Rosen, B.P. (2002). Membrane topology of the p1258 CadA Cd(II)/Pb(II)/Zn(II)-translocating P-type ATPase. *J. Bioenerg. Biomembr.* *34*, 147–156.
- Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2012). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat. Rev. Microbiol.* *10*, 123–136.
- Valko, M., Morris, H., and Cronin, M.T.D. (2005). Metals, Toxicity and Oxidative Stress. *Curr. Med. Chem.* *12*, 1161–1208.
- Valls, M., and de Lorenzo, V. (2002). Exploiting the genetic and biochemical capacities of bacteria for the remediation of heavy metal pollution. *FEMS Microbiol. Rev.* *26*, 327–338.
- Vandamme, P., and Coenye, T. (2004). Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int. J. Syst. Evol. Microbiol.* *54*, 2285–2289.
- Vanechoutte, M., Kämpfer, P., De Baere, T., Falsen, E., and Verschraegen, G. (2004). *Wautersia* gen. nov., a novel genus accommodating the phylogenetic lineage including *Ralstonia eutropha* and related species, and proposal of *Ralstonia* [*Pseudomonas*] *syzygii* (Roberts et al. 1990) comb. nov. *Int. J. Syst. Evol. Microbiol.* *54*, 317–327.
- Wei, P., Liu, X., Hu, M.-H., Zuo, L.-M., Kai, M., Wang, R., and Luo, S.-Z. (2011). The dimerization interface of the glycoprotein Ib $\beta$  transmembrane domain corresponds to polar residues within a leucine zipper motif. *Protein Sci. Publ. Protein Soc.* *20*, 1814–1823.
- Wei, P., Xu, L., Li, C.-D., Sun, F.-D., Chen, L., Tan, T., and Luo, S.-Z. (2014a). Molecular Dynamic Simulation of the Self-Assembly of DAP12-NKG2C Activating Immunoreceptor Complex. *PloS One* *9*, e105560.
- Wei, W., Liu, X., Sun, P., Wang, X., Zhu, H., Hong, M., Mao, Z.-W., and Zhao, J. (2014b). Simple whole-cell biodetection and bioremediation of heavy metals based on an engineered

lead-specific operon. *Environ. Sci. Technol.* *48*, 3363–3371.

Welz, D., and Braun, V. (1998). Ferric citrate transport of *Escherichia coli*: functional regions of the FecR transmembrane regulatory protein. *J. Bacteriol.* *180*, 2387–2394.

Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. (2004). Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell* *117*, 483–494.

Williams, R.J.P., and Silva, J.J.R.F. da (2005). *The Chemistry of Evolution: The Development of our Ecosystem* (Elsevier).

Wriedt, K., Angerer, A., and Braun, V. (1995). Transcriptional regulation from the cell surface: conformational changes in the transmembrane protein FecR lead to altered transcription of the ferric citrate transport genes in *Escherichia coli*. *J. Bacteriol.* *177*, 3320–3322.

Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A. (2005). Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *J. Biol. Chem.* *280*, 1448–1456.

Ziani, W., Maillard, A.P., Petit-Hartlein, I., Garnier, N., Crouzy, S., Girard, E., and Covès, J. (2014). The X-ray structure of NccX from *Cupriavidus metallidurans* 31A illustrates potential dangers of detergent solubilization when generating and interpreting crystal structures of membrane proteins. *J. Biol. Chem* *114*.586537. doi: 10.1074.

# Résumé

Chez *Cupriavidus metallidurans* CH34, le complexe CnrYXH contribue à réguler l'expression des gènes de régulation et de résistance au cobalt et au nickel en fonction de la concentration environnementale de ces cations. La fixation de cobalt ou de nickel sur le domaine périplasmique senseur de CnrX induit des modifications conformationnelles à l'origine de la transduction du signal transmembranaire. Cela conduit à rendre le facteur sigma CnrH disponible pour sa liaison à l'ARN polymérase (ARNP) dans le cytoplasme. Le complexe CnrH:ARNP se fixe alors spécifiquement sur les promoteurs des gènes *cnrY* et *cnrC* pour initier la transcription des gènes de régulation (*cnrYXH*) et de résistance (*cnrCBAT*). Dans le but de déterminer la nature de ce signal, mon projet de thèse visait à cartographier les interactions protéine:protéine au sein du complexe CnrYXH dans les trois compartiments concernés : le périplasma, la membrane plasmique et le cytoplasme. La documentation des déterminants d'interaction entre CnrX, CnrY et CnrH a permis d'élaborer un modèle structural et fonctionnel pour le complexe CnrYXH et ses homologues soulevant des hypothèses nouvelles sur le fonctionnement du système Cnr.

The CnrYXH complex contributes to regulate the expression of the regulatory genes and resistance genes involved in cobalt and nickel resistance in *Cupriavidus metallidurans* CH34. The binding of nickel or cobalt to CnrX in the periplasm induces conformational modifications that could trigger transmembrane signal transduction. As a result, CnrH is made available in cytoplasm for binding to RNA polymerase. The CnrH:RNA polymerase complex binds specifically the *cnr* promoters and initiates transcription of both the regulatory genes (*cnrYXH*) and resistance genes (*cnrCBAT*). In order to delineate the mechanism of signal transduction through the membrane, I studied the interactions between the three partners in the different cellular compartments: periplasm, plasmic membrane and cytoplasm. The identification of the interaction determinants between CnrX, CnrY and CnrH allowed us to propose a structural and functional model for the CnrYXH complex and its homologues. This model brings up new hypothesis on the function of Cnr system.