



HAL
open science

Développement et applications de protocoles de synthèse in vitro du canal mécanosensible bactérien à large conductance, pour son étude structurale par résonance magnétique nucléaire en phase solide

Alaa Abdine

► To cite this version:

Alaa Abdine. Développement et applications de protocoles de synthèse in vitro du canal mécanosensible bactérien à large conductance, pour son étude structurale par résonance magnétique nucléaire en phase solide. Sciences pharmaceutiques. Université Paris-Diderot - Paris VII, 2011. Français. NNT : . tel-00621192

HAL Id: tel-00621192

<https://theses.hal.science/tel-00621192>

Submitted on 9 Sep 2011

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

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

UNIVERSITÉ PARIS DIDEROT – PARIS 7

U.F.R. DE PHYSIQUE

École doctorale : Interdisciplinaire pour le vivant (I-ViV) - ED 387

THÈSE

Pour l'obtention du diplôme de

DOCTEUR DE L'UNIVERSITÉ PARIS DIDEROT – PARIS 7

Par

Alaa ABDINE

**Développement et applications de protocoles de synthèse *in vitro*
du canal mécanosensible bactérien à large conductance, pour son étude
structurale par résonance magnétique nucléaire en phase solide**

Soutenue le 29 Juin 2011 devant le jury composé de :

Pr. Catherine ETCHEBEST	Présidente du jury
Dr. Jean-Michel BETTON	Rapporteur
Dr. François PENIN	Rapporteur
Pr. Catherine BERRIER	Examinatrice
Dr. Nadège JAMIN	Examinatrice
Dr. Dror WARSCHAWSKI	Directeur

Avant propos

Ce mémoire est le résultat de quatre années de recherche. Je tiens à remercier ici tous ceux qui, de près ou de loin, y ont pris part.

Je remercie très fortement Dror Warschawski. Merci d'avoir dirigé cette thèse en m'associant à ce projet ambitieux, merci pour ta disponibilité au jour le jour, tes conseils et toutes les discussions que nous avons eu.

Je remercie l'ensemble des membres du jury. Sincères remerciements à Catherine Etchebest, présidente de jury. Je suis extrêmement reconnaissant envers Jean-Michel Betton et François Penin, pour avoir accepté d'être rapporteurs de ce manuscrit. Je suis également reconnaissant à Catherine Berrier et Nadège Jamin d'être examinatrices.

Mes années de thèse se sont déroulées au Laboratoire de Biologie Physico-Chimique des Protéines Membranaires - UMR 7099 (Université Paris 7/CNRS). Un très grand merci à Bruno Miroux et Jean-Luc Popot, actuel et précédent directeurs de l'UMR 7099, pour m'avoir accueilli au sein de l'équipe, ainsi que pour les discussions et leurs conseils au cours des quatre années écoulées. Je remercie l'ensemble du personnel du laboratoire pour les quatre années passées au laboratoire !

Je remercie mes parents ainsi que toute ma famille aux quatre coins du monde, mes amis et toutes les personnes que j'ai pu côtoyer durant ces quatre années de m'avoir soutenu pour mener à bien ce projet. Un énorme merci à Claire pour le soutien sans faille, et pour l'aide en bioinformatique.

Ce travail a été financé par le Ministère de l'Enseignement Supérieur et de la Recherche, l'Université Paris Diderot, le Centre National de la Recherche Scientifique, et par la Fondation de la Recherche Médicale.

Développement et applications de protocoles de synthèse *in vitro* du canal mécanosensible bactérien à large conductance, pour son étude structurale par résonance magnétique nucléaire en phase solide

Le génome de différents organismes contient près de 30% de séquences codantes pour des protéines membranaires. Celles-ci sont impliquées dans des processus biologiques tels que la signalisation cellulaire, la transduction énergétique ou le transport des métabolites. Cependant, leur étude structurale se heurte souvent aux problèmes de surexpression, ainsi qu'aux différents inconvénients liés à leur extraction de leur environnement natif.

Le canal mécanosensible à large conductance *MscL* d'*Escherichia coli* est une protéine intrinsèque de la membrane interne de la bactérie. L'activité de cette protéine est fortement dépendante de la membrane ; en effet, le canal s'ouvre lorsque la pression au niveau de la membrane augmente, lors d'un choc hypoosmotique, relarguant de l'eau et différents substrats afin que la bactérie retrouve un état osmotique adéquat à sa survie. Il est donc essentiel de recueillir des données structurales de la protéine dans son environnement natif. Pour cela, nous proposons une étude structurale de la protéine par résonance magnétique nucléaire en phase solide.

La surexpression de la protéine et son marquage aux isotopes ^{13}C et ^{15}N sont deux étapes clés d'un tel projet. Une surexpression bactérienne et un marquage uniforme de la protéine ont été effectués. Les données spectrales obtenues montrent une bonne résolution, mais l'imbrication des corrélations ne permet pas d'en extraire des données structurales. Afin de réduire le nombre de résidus marqués, un marquage spécifique a été appliqué par la voie de la synthèse *in vitro*. Le premier test a montré que l'approche permettait de réduire le temps d'acquisition, tout en diminuant la quantité d'informations. Différentes méthodes ont ensuite été appliquées pour trouver les meilleures combinaisons d'acides aminés à marquer spécifiquement en synthétisant la protéine *in vitro*. Ces approches utilisent les programmes de prédiction de déplacements chimiques ou l'analyse de la séquence à la recherche de paires d'acides aminés uniques. Une approche combinatoire a été également testée. Les différents échantillons synthétisés ont montré une bonne résolution, et ont permis l'identification de plusieurs corrélations.

Les méthodes développées au cours de ce travail pourront aider progressivement à déterminer la structure de la protéine *MscL*. Elles pourront également servir à d'autres protéines membranaires pour leur étude par résonance magnétique nucléaire.

New labeling strategies for a structural study of the mechanosensitive channel of large conductance using cell-free synthesis

Membrane proteins account for almost 30% of the proteome and play essential roles in many cellular mechanisms. The mechanosensitive channel of large conductance *MscL* is an intrinsic membrane protein of the inner membrane of *Escherichia coli*. This protein acts as a valve pressure, in a hypo-osmotic down-shock, to prevent the bacterial lysis. In this work, a study of the structural characteristics of *MscL* in its native environment, the lipid bilayer, by nuclear magnetic resonance in the solid-state is presented. A major problem in determining the three-dimensional structure of membrane proteins by this technique is their overexpression and the analysis of carbon-carbon correlations from uniformly labelled samples.

The analysis of *MscL* by solid-state *NMR* with a uniform labelling showed spectra with a good resolution, but the overlap of the correlations prevented any identification or assignment of the residues. We showed that selective labelling of the membrane protein *MscL* by using cell-free synthesis represents a crucial help in improving data sets obtained by *NMR* in the solid-state. Since the amino acids chemical shifts show a big overlap, different strategies to choose specific labelling patterns to get well-resolved spectra were developed. These strategies are based on chemical shifts prediction, or sequence analysis to isolate unique amino acid pairs. A combined approach was tested. These different approaches showed spectra with a good resolution, and facilitated the identification of amino acids in some cases.

The methods developed in this work may gradually help to determine the three-dimensional structure of the mechanosensitive channel *MscL*, as well as other lipid-dependent membrane proteins.

Liste des abréviations :

AA	Acide aminé
ADP/ATP	Adénosine di/tri phosphate
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNt	ARN de transfert
ARNr	ARN ribosomal
ARNm	ARN messager
MS	Protéine mécanosensible
MscL	canal mécanosensible à large conductance
E. coli	Escherichia coli
M. tuberculosis	Mycobacterium tuberculosis
S. aureus	Staphylococcus aureus
Ec-MscL	MscL de la bactérie E. coli (code Uniprot : P0A742)
Mt-MscL	MscL de M. tuberculosis (code Uniprot : A5U127 ; PDB : 2oar)
Sa-MscL	MscL de S. aureus (code Uniprot : P68806 ; PDB : 3hzq)
SIV	Synthèse in vitro
RTS	Rapid Translation System
RMN	Résonance Magnétique Nucléaire
nS	nanoSiemens
Da/ kDa	Daltons/ kiloDaltons
Å	Angstroms
ppm	parties par million
PG	Phosphatidylglycérol
PI	Phosphatidylinositol
PE	Phosphatidyléthanolamine
PC	Phosphatidylcholine
PS	Phosphatidylsérine
SM	Sphingomyéline
DLPE	1,2-Dilauryl-sn-glycéro-3-phosphatidyl-éthanolamine
DOPE	1,2-Dioleoyl-sn-glycéro-3-phosphatidyl-éthanolamine
DOPC	1,2-Dioleoyl-sn-glycéro-3-phosphatidyl-choline
DPPC	1,2-Dioleoyl-sn-glycéro-3-phosphatidylcholine
DMPE	1,2-Dimyristoyl-sn-glycéro-3-phosphatidyléthanolamine
POPE	1-Palmitoyl-2-oleoyl-sn-glycéro-3-phosphatidyl-éthanolamine
LPC	1-myristoyl-2-hydroxy-sn-glycéro-3-phosphocholine
DPC	1-dodecyl-2-hydroxy-sn-glycéro-3-phosphocholine
OG	Octyl-glucoside
DDM	Dodecyl-maltoside
CMC	Concentration micellaire critique
DARR	Dipolar Assisted Rotational Resonance
PDSF	Proton Driven Spin Diffusion
RFDR	Radio-Frequency-driven Dipolar Recoupling
B ² R ³	BroadBand decoupling of Rotary Resonance Recoupling
CP	Cross Polarization
CP-MAS	Cross Polarization with Magic Angle Spinning
Specific-CP	Spectrally Induced Filtering In Combination with Cross Polarization

Les échantillons synthétisés :

U-MscL/U-¹³C,¹⁵N-MscL = MscL uniformément marqué aux ¹³C et ¹⁵N

IT-MscL/IT-¹³C,¹⁵N-MscL = MscL sélectivement marqué aux ¹³C et ¹⁵N au niveau des isoleucines et thréonines

AKVY-MacL/AKVY-¹³C,¹⁵N-MscL = MscL sélectivement marqué aux ¹³C et ¹⁵N au niveau des alanines, lysines, valines et tyrosine

AGKVY-MscL/AGKVY-¹³C,¹⁵N-MscL = MscL sélectivement marqué aux ¹³C et ¹⁵N au niveau des alanines, glycines, lysines, valines et tyrosine

FIMPR-MscL/FIMPR-¹³C,¹⁵N-MscL = MscL sélectivement marqué aux ¹³C et ¹⁵N au niveau des phénylalanines, isoleucines, méthionines, prolines et arginines

DGLS-MscL/DGLS-¹³C,¹⁵N-MscL = MscL sélectivement marqué aux ¹³C et ¹⁵N au niveau des aspartates, glycines, leucines et sérines

Liste des figures

Introduction

<u>Figure 1</u> : Une membrane plasmique schématisée.....	6
<u>Figure 1.1</u> : Structure chimique de quelques acides gras.....	8
<u>Figure 1.2</u> : Structure chimique des glycérolipides.....	9
<u>Figure 1.3</u> : Structure chimique de la sphingomyéline et de la phosphatidylcholine.....	10
<u>Figure 1.4</u> : Structure chimique d'une molécule de cholestérol et sa disposition par rapport à une molécule de phospholipide.....	10
<u>Figure 1.5</u> : Modes de transport d'une molécule à travers la membrane plasmique.....	12
<u>Figure 2.1</u> : Topologie des protéines membranaires	15
<u>Figure 2.2</u> : Modèle de la mosaïque fluide de <i>Singer & Nicolson</i>	16
<u>Figure 2.3</u> : Mouvement des molécules de lipides au sein de la membrane.....	17
<u>Figure 2.4</u> : Mécanismes d'ouverture des canaux <i>MS</i>	19
<u>Figure 2.5</u> : Hypothèses d'ouverture des canaux <i>MS</i>	20
<u>Figure 2.6</u> : Schéma de la technique de patch clamp.....	21
<u>Figure 3.1</u> : Séquence peptidique d'un monomère de <i>MscL</i>	23
<u>Figure 3.2</u> : Séquence et tracé de l'index d'hydrophilie de la protéine <i>Ec-MscL</i>	24
<u>Figure 3.3</u> : Structures de la protéine mécanosensible <i>Mt-MscL</i> en 1998 et en 2007.....	25
<u>Figure 3.4</u> : Modèle d'ouverture à dix hélices de la protéine <i>MscL</i>	27
<u>Figure 3.5</u> : Les deux modes d'ouverture hypothétiques de <i>MscL</i>	27
<u>Figure 3.6</u> : Structure et composition d'un site de constriction cytoplasmique au niveau de <i>Ec-MscL</i>	29
<u>Figure 3.7</u> : Modèle d'ouverture de la protéine <i>Ec-MscL</i>	31
<u>Figure 3.8</u> : Variation de la conformation du canal <i>MscL</i> en fonction de la tension exercée au niveau de la membrane.....	33
<u>Figure 3.9</u> : Schéma de l'implication de la partie N-terminale dans le mécanisme d'ouverture.....	34
<u>Figure 3.10</u> : Structure de la partie C-terminale de la protéine <i>Mt-MscL</i>	35
<u>Figure 3.11</u> : Structure de la protéine mécanosensible <i>Sa-MscL</i>	37

<u>Figure 4.1</u> : De la séquence à la structure <i>in vivo</i> - cas de la protéine <i>Sa-MscL</i>	40
<u>Figure 4.2</u> : Structures de l'hélice transmembranaire <i>TM2</i> de <i>GlyR</i> dans différents environnements hydrophobes.....	50
<u>Figure 4.3</u> : Structures de peptides ou protéines membranaires déterminées par <i>RMN</i> en phase solide.....	52
<u>Figure 4.4</u> : Spectres obtenus sur différents échantillons de <i>SH3</i>	54
<u>Figure 5.1</u> : Voies de biosynthèse des acides aminés dans <i>E. coli</i>	60
<u>Figure 5.2</u> : Schémas d'enrichissement isotopique des acides aminés dans la souche d' <i>E. coli</i> <i>BL21 (DE3)</i> en présence de glycérol comme source de carbone.....	61
<u>Figure 5.3</u> : Schéma de synthèse d'un polypeptide.....	64
<u>Figure 5.4</u> : Schéma de synthèse d'une protéine membranaire en <i>SIV</i>	67
<u>Tableau 1</u> : Point de fusion de quelques acides.....	11
<u>Tableau 2</u> : Protéines mécanosensibles.....	18
<u>Tableau 3</u> : Caractéristiques des noyaux usuels des molécules biologiques.....	46

Matériels et méthodes expérimentales

<u>Tableau 1</u> : Composition du milieu minimum <i>M9</i>	90
<u>Tableau 2</u> : Composition d'une solution d'oligoéléments <i>TE</i>	90
<u>Tableau 3</u> : Composition d'une solution de tampon <i>FP</i> (pour <i>FRENCH Press</i>).....	91
<u>Tableau 4</u> : Composition du tampon d'équilibration <i>EB</i> (pour <i>Equilibration buffer</i>).....	91
<u>Tableau 5</u> : Tampons de purification utilisés en <i>FPLC</i>	92
<u>Tableau 6</u> : Composition en acides aminés de la protéine <i>MscL</i>	93

Résultats

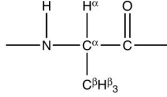
<u>Figure 1</u> : Spectres MAS du 1H des vésicules de DOPC obtenus à différentes températures.....	135
<u>Figure 2</u> : Evolution de la largeur de raie de protons de la DOPC en fonction de la température.....	136
<u>Figure 3</u> : Spectre ^{13}C -MAS unidimensionnel des vésicules de DOPC.....	138
<u>Figure 4</u> : Spectre ^{13}C -MAS bidimensionnel des vésicules de DOPC.....	138
<u>Figure 1.1</u> : Séquence de MscL d' <i>E. coli</i> dans le plasmide pIVEX.....	139
<u>Figure 1.2</u> : Suivi de la purification de la protéine par SDS-PAGE.....	142
<u>Figure 1.3</u> : Spectre ^{13}C - ^{13}C obtenu sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant la séquence DARR et en appliquant un temps de mélange de 1,5 ms.....	143
<u>Figure 1.4</u> : Spectres ^{13}C - ^{13}C obtenus sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant la séquence DARR et en appliquant des temps de mélange de 10, 25, 50 et 100 ms.....	144
<u>Figure 1.5</u> : Evolution de l'intensité des corrélations ^{13}C - ^{13}C en fonction du temps de mélange DARR.....	146
<u>Figure 1.6</u> : Spectres ^{13}C - ^{13}C obtenus sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant les séquences d'impulsion DARR (a), PDSD (b) et B ² R ³ (c) avec un temps de mélange de 50 ms.....	147
<u>Figure 1.7</u> : Spectre ^{13}C - ^{13}C obtenu sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant la séquence d'impulsion DARR en appliquant un temps de mélange de 50 ms et par enregistrement de 256 scans.....	148
<u>Figure 1.8</u> : Spectres ^{13}C - ^{13}C obtenu sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant la séquence d'impulsion DARR en appliquant un temps de mélange de 25 ms à 248 K et un temps de mélange de 20 ms à 268 K.....	149
<u>Figure 1.9</u> : Spectre ^{15}N - ^{13}C obtenu sur l'échantillon μ - ^{13}C - ^{15}N -MscL en utilisant la séquence d'impulsion specific-CP.....	150
<u>Figure 2.1</u> : Suivi de la synthèse de la protéine MscL par SIV en triton-x-100 par SDS-PAGE.....	153
<u>Figure 2.2</u> : Spectre 2D ^{13}C - ^{13}C -DARR obtenu sur l'échantillon (Ile, Thr)- ^{13}C , ^{15}N -MscL avec un temps de mélange de 5 ms à 248 K.....	154
<u>Figure 2.3</u> : Spectres 2D ^{13}C - ^{13}C -DARR obtenus sur l'échantillon (Ile, Thr)- ^{13}C , ^{15}N -MscL avec différents temps de mélange.....	155
<u>Figure 2.4</u> : Évolution du signal de certaines corrélations isolées entre atomes séparés par une seule liaison.....	157

<u>Figure 2.5</u> : Évolution du signal de certaines corrélations isolées entre atomes séparés par deux ou trois liaisons.....	158
<u>Figure 2.6</u> : Spectre ^{15}N - ^{13}C obtenu sur l'échantillon (<i>Ile, Thr</i>)- ^{13}C , ^{15}N - <i>MscL</i> utilisant la séquence d'impulsion <i>specific-CP</i>	159
<u>Figure 2.7</u> : Agrandissement des régions contenant les pics de corrélations entre les différents carbones des <i>Thr</i> de la protéine <i>Ec-MscL</i>	160
<u>Figure 2.8</u> : Agrandissement des régions contenant les pics de corrélations entre les différents carbones de quelques <i>Ile</i> de la protéine <i>Ec-MscL</i>	163
<u>Figure 2.9</u> : Agrandissement des régions contenant les pics de corrélations inter-résiduelles entre une <i>Thr</i> et une <i>Ile</i>	165
<u>Figure 2.10</u> : Superposition des spectres $2D$ ^{13}C - ^{13}C obtenus sur les échantillons <i>IT</i> - ^{13}C , ^{15}N - <i>MscL</i> et <i>u</i> - ^{13}C , ^{15}N - <i>MscL</i>	166
<u>Figure 2.11</u> : Comparaison des spectres uniformément et sélectivement marqués.....	213
<u>Figure 2.12</u> : Comparaison des spectres uniformément et sélectivement marqués.....	214
<u>Figure 3.1</u> : Prédiction des déplacements chimiques sur quinze modèles de <i>Ec-MscL</i>	169
<u>Figure 3.2</u> : Suivi de la synthèse de la protéine <i>MscL</i> par <i>SIV</i> en liposomes de <i>DOPC</i>	170
<u>Figure 3.3</u> : Spectre $2D$ ^{13}C - ^{13}C obtenu sur l'échantillon (<i>Ala, Lys, Val, Tyr</i>)- ^{13}C , ^{15}N - <i>MscL</i> à 248 K avec un temps de mélange de 10 ms.....	171
<u>Figure 3.4</u> : Spectres $2D$ ^{13}C - ^{13}C obtenu sur l'échantillon (<i>Ala, Lys, Val, Tyr</i>)- ^{13}C , ^{15}N - <i>MscL</i> à 248 K avec différents temps de mélange.....	173
<u>Figure 3.5</u> : Évolution du signal de trois corrélations isolées des différents spectres obtenus sur l'échantillon (<i>Ala, Lys, Val, Tyr</i>)- <i>Ec-MscL</i>	174
<u>Figure 3.6</u> : Spectres $2D$ ^{15}N - ^{13}C obtenus sur l'échantillon (<i>Ala, Lys, Val, Tyr</i>)- ^{13}C , ^{15}N - <i>MscL</i> à 248 K.....	175
<u>Figure 4.1</u> : Analyse de la séquence de <i>Ec-MscL</i> pour établir une stratégie de marquage.....	178
<u>Figure 4.2</u> : Spectre $2D$ ^{13}C - ^{13}C obtenu sur l'échantillon (<i>Phe, Ile, Met, Pro, Arg</i>)- ^{13}C , ^{15}N - <i>MscL</i>	180
<u>Figure 4.3</u> : Spectres $2D$ ^{13}C - ^{13}C obtenus sur l'échantillon (<i>Phe, Ile, Met, Pro, Arg</i>)- ^{13}C , ^{15}N - <i>MscL</i> à différents temps de mélange à 238 K (<i>rpl</i> 1/4).....	181
<u>Figure 4.4</u> : Spectres $2D$ ^{13}C - ^{13}C obtenus sur l'échantillon (<i>Phe, Ile, Met, Pro, Arg</i>)- ^{13}C , ^{15}N - <i>MscL</i> à différents temps de mélange à 238 K en 14 h (<i>rpl</i> 1/2).....	182
<u>Figure 4.5</u> : Spectres $2D$ ^{13}C - ^{13}C obtenus sur l'échantillon (<i>Phe, Ile, Met, Pro, Arg</i>)- ^{13}C , ^{15}N - <i>MscL</i> à différents temps de mélange à 238 K en 56 h (<i>rpl</i> 1/2).....	183
<u>Figure 4.6</u> : Évolution du signal de quatre corrélations isolées des différents spectres obtenus sur	

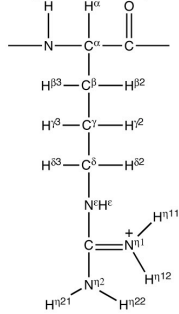
l'échantillon ((<i>Phe, Ile, Met, Pro, Arg</i>)- ¹³ C, ¹⁵ N- <i>MscL</i>	184
<u>Figure 5.1</u> : Alignements de séquence entre <i>Ec-MscL</i> et <i>Mt-MscL</i> en utilisant différents algorithmes....	186
<u>Figure 5.2</u> : Prédiction des déplacements chimiques sur 40 modèles de <i>Ec-MscL</i> . (<i>Ala, Lys, Val</i>).....	187
<u>Figure 5.3</u> : Prédiction des déplacements chimiques sur 40 modèles de <i>Ec-MscL</i> . (<i>Asp, Ser</i>).....	188
<u>Figure 5.4</u> : Spectres 2D ¹³ C- ¹³ C obtenus sur l'échantillon (<i>Asp, Gly, Leu, Ser</i>)- ¹³ C, ¹⁵ N- <i>MscL</i> à 248 K avec différents temps de mélange.....	191
<u>Figure 5.5</u> : Évolution du signal de quatre corrélations issues des différents spectres obtenus sur l'échantillon (<i>Asp, Gly, Leu, Ser</i>)- ¹³ C, ¹⁵ N- <i>Ec-MscL</i>	192
<u>Figure 5.6</u> : Fluctuation des valeurs de déplacements chimiques prédites des C α en fonction de la séquence.....	192
<u>Figure 5.7</u> : Variation de l'écart-type à la moyenne des déplacements chimiques prédits en fonction de la famille d'acides aminés.....	226
<u>Figure 6.1</u> : Spectres 2D ¹³ C- ¹³ C obtenus sur l'échantillon (<i>Ala, Gly, Lys, Val, Tyr</i>)- <i>Ec-MscL</i> dans deux environnements lipidiques différents.....	194
<u>Figure 6.2</u> : Agrandissement de la région des corrélations C α -C β des résidus <i>Ala</i> . (DPPC/DOPC vs. Lipides d' <i>E. coli</i>).....	195
<u>Figure 6.3</u> : Spectres 2D ¹³ C- ¹³ C obtenus sur l'échantillon (<i>Ala, Gly, Lys, Val, Tyr</i>)- <i>Ec-MscL</i> reconstitués en liposomes de DOPC dans de l'eau H ₂ O ou D ₂ O.....	196
<u>Figure 6.4</u> : Agrandissement de la région des corrélations C α -C β des résidus <i>Ala</i> . (DOPC vs. Lipides d' <i>E. coli</i>).....	196
<u>Tableau 2.1</u> : Déplacements chimiques des carbones des 3 <i>Thr</i> de la protéine <i>Ec-MscL</i>	161
<u>Tableau 2.2</u> : Déplacements chimiques des carbones de 8 <i>Ile</i> identifiées de la protéine <i>Ec-MscL</i>	162
<u>Tableau 3.1</u> : Déplacements chimiques des groupes d'acides aminés <i>Ala, Lys et Val</i>	176
<u>Tableau 4.1</u> : Déplacements chimiques des groupes d'acides aminés <i>Phe, Ile, Met, Pro et Arg</i>	184
<u>Tableau 6.1</u> : Déplacements chimiques des carbones de cinq groupes d' <i>Ala</i> de la protéine <i>Ec-MscL</i>	197

Liste des acides aminés, leurs abréviations, ainsi que la nomenclature des noyaux pour la RMN

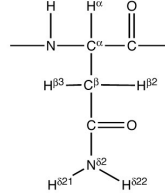
Alanine, Ala, A



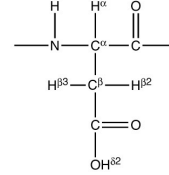
Arginine, Arg, R



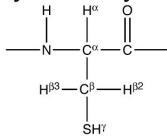
Asparagine, Asn, N



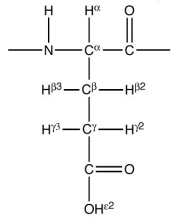
Aspartic acid, Asp, D



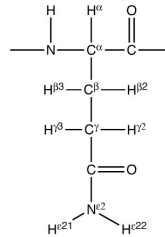
Cysteine, Cys, C



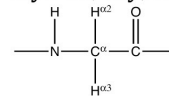
Glutamic acid, Glu, E



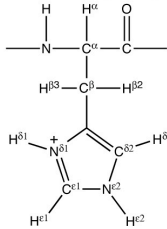
Glutamine, Gln, Q



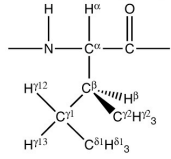
Glycine, Gly, G



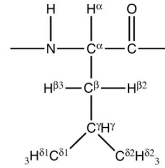
Histidine, His, H



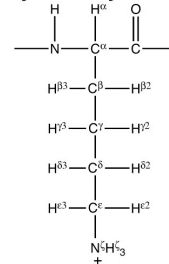
Isoleucine, Ile, I



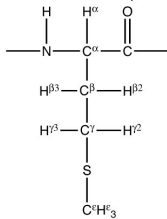
Leucine, Leu, L



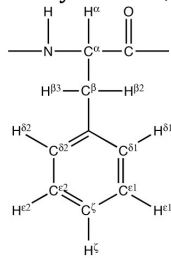
Lysine, Lys, K



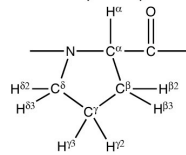
Methionine, Met, M



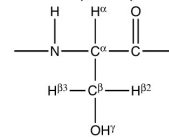
Phenylalanine, Phe, F



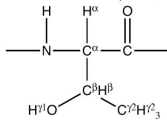
Proline, Pro, P



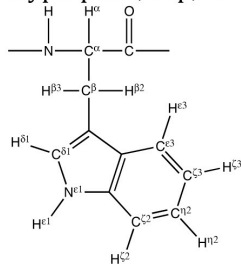
Serine, Ser, S



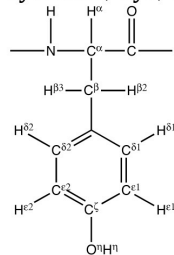
Threonine, Thr, T



Tryptophan, Trp, W



Tyrosine, Tyr, Y



Valine, Val, V

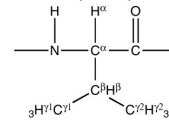


Table des matières

A	Introduction générale.....	6
	I Les lipides.....	7
	II Les protéines membranaires.....	14
	III Le canal mécanosensible bactérien MscL.....	23
	IV L'utilisation de la résonance magnétique nucléaire en phase solide pour étudier MscL.....	39
	V La synthèse in vitro pour l'étude de MscL in situ par RMN en phase solide.....	57
	VI Conclusion - Projet de recherche.....	69
	VII Références.....	71
B	Matériels et Méthodes expérimentales.....	88
	I Matériels.....	89
	II Méthodes expérimentales.....	98
	III Références.....	130
C	Résultats.....	134
	I Analyse de l'échantillon uniformément marqué.....	139
	II Analyse de l'échantillon sélectivement marqué IT-Ec-MscL Apport de la SIV pour l'étude des protéines membranaires par RMN en phase solide.....	151
	III Analyse de l'échantillon sélectivement marqué AKVY-Ec-MscL - Apport des algorithmes de prédiction de déplacement chimiques.....	167
	IV Analyse de l'échantillon sélectivement marqué FIMPR-Ec-MscL - Analyse de la séquence.....	177
	V Analyse de l'échantillon sélectivement marqué DGLS-Ec-MscL - Une combinaison des approches prédictives.....	185
	VI Rôle de l'environnement lipidique - Analyse de l'échantillon AGKVY-Ec-MscL.....	193
	VII Références.....	198
D	Discussion.....	202
	I Analyse de l'échantillon uniformément marqué.....	205
	II Analyse de l'échantillon sélectivement marqué IT-Ec-MscL - Apport de la SIV pour l'étude des protéines membranaires par RMN en phase solide.....	209
	III Analyse de l'échantillon sélectivement marqué AKVY-Ec-MscL - Apport des algorithmes de prédiction de déplacement chimiques.....	215
	IV Analyse de l'échantillon sélectivement marqué FIMPR-Ec-MscL - Analyse de la séquence.....	219
	V Analyse de l'échantillon sélectivement marqué DGLS-Ec-MscL - Une combinaison des approches prédictives.....	223
	VI Rôle de l'environnement lipidique - Analyse de l'échantillon AGKVY-Ec-MscL.....	227
	VII Références.....	230
E	Conclusions.....	234
	I La surexpression bactérienne de la protéine Ec-MscL pour son étude par RMN en phase solide.....	234
	II La synthèse in vitro de la protéine Ec-MscL pour son étude par RMN en phase solide.....	235
	III Conclusion générale.....	240
	IV Références.....	242

F Perspectives.....	245
Références.....	251
G Annexes.....	255
I Structural study of the membrane protein MscL using cell-free expression and solid-state NMR.....	256
II Cell-free expression and labeling strategies for a new decade in solid-state NMR	269
III Cell-free membrane protein expression for solid-state NMR.....	276

Introduction

A Introduction générale

Toute cellule vivante est entourée d'une membrane, et renferme une grande quantité d'autres membranes. Celles-ci remplissent des fonctions diverses. Leur fonction première est la compartimentation. La membrane plasmique isole la cellule de son environnement. Les différentes membranes internes isolent les organites ou les vésicules. Toute membrane biologique est constituée de trois éléments de base : les lipides, les protéines et les glucides. Leur proportion diffère grandement selon les organismes, le type cellulaire et la localisation (Helms, 2002 ; van Meer *et al.* 2008). Les glucides sont généralement associés aux lipides et aux protéines sous forme de glycolipides et de glycoprotéines (Figure 1). Les deux constituants principaux sont donc les lipides et les protéines. Les connaissances concernant les propriétés physico-chimiques de la membrane ainsi que sa composition, découlent de plus d'un siècle d'observations et d'expériences (Edidin ; 2003).

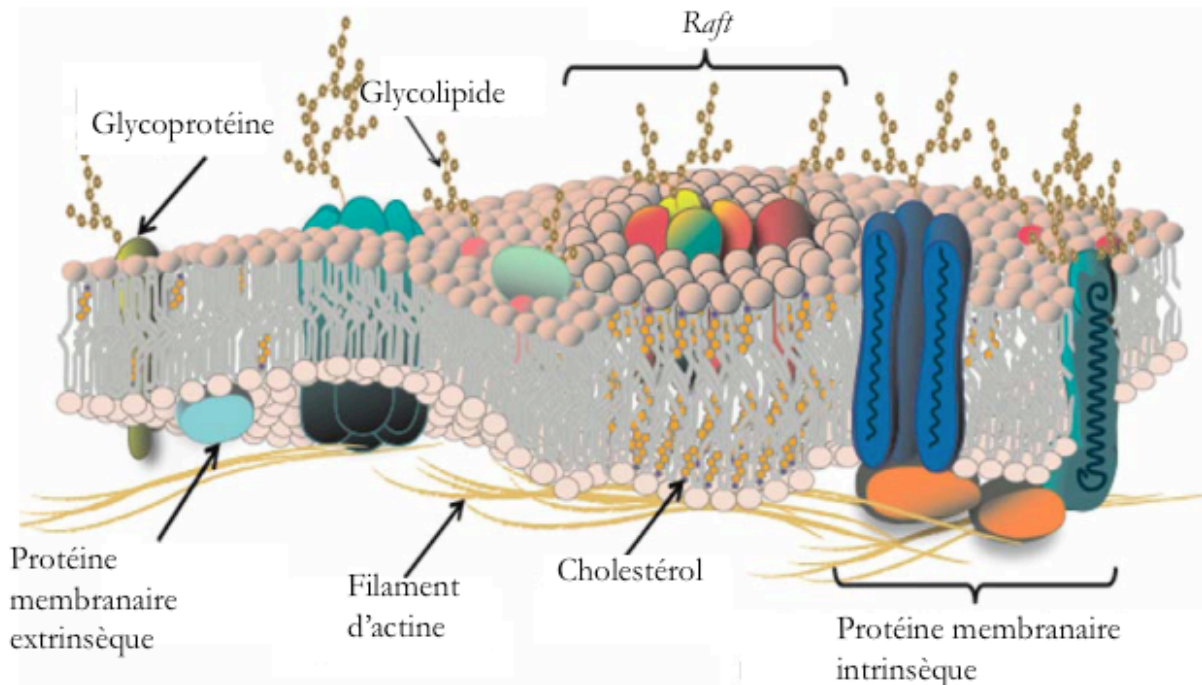


Figure 1 : Une membrane plasmique schématisée
(*Current Topics in Membranes - Collawn & Bebök, 2008*)

I Les lipides

Les lipides constituent le cœur de la membrane. Ce sont des molécules amphiphiles composées d'une partie hydrophile et d'une partie hydrophobe. Ce caractère amphiphile les conduit à s'associer en milieu aqueux, de façon à protéger les parties hydrophobes sous forme de divers assemblages supramoléculaires (micelle, bicelle, vésicule, etc.). Dans les cellules, les lipides sont organisés en bicouche, les parties hydrophobes se retrouvant au centre, entourées des parties hydrophiles, en contact avec l'eau. Les lipides comprennent les phospholipides, les sphingolipides et les stérols. Les phospholipides sont des lipides possédant un groupement phosphate et deux acides gras. Les sphingolipides sont des dérivés de la sphingosine contenant un acide gras. Les stérols sont des lipides dérivant du stérane. Les proportions de ces différents composants varient selon les espèces, les différents tissus dans la même espèce, et selon les différents organites de la même cellule (Christie, 1985 ; Colbeau *et al.*, 1971 ; Wuthier, 1966).

a les acides gras

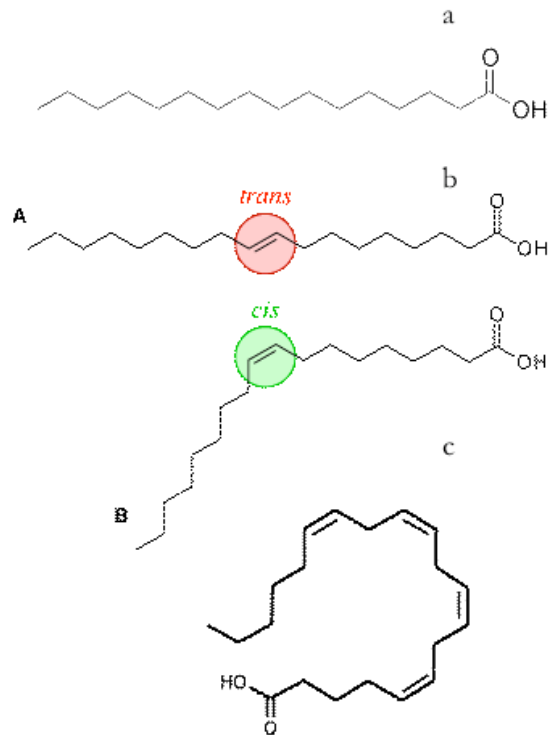
Les acides gras sont des acides carboxyliques à chaînes aliphatiques hydrophobes. Les acides gras naturels possèdent une chaîne carbonée de 4 à 28 atomes de carbone. Les chaînes peuvent comporter une ou plusieurs insaturations. Les acides gras diffèrent donc entre eux non seulement par la longueur des chaînes carbonées, mais aussi par le nombre, la position et la structure spatiale (*cis*, *trans*) des doubles liaisons (MacDonald & Sprecher, 1991 ; [Figure 1.1](#)).

Les acides gras saturés peuvent avoir un grand nombre de conformations possibles. En effet, chaque liaison chimique du squelette étant simple, elle a tout pouvoir de rotation. La chaîne hydrocarbonée est donc flexible. Leur formule générale est $CH_3 - (CH_2)_n - COOH$.

Les acides gras insaturés peuvent contenir une ou plusieurs doubles liaisons et sont dits, selon le cas, mono-insaturés ou poly-insaturés. La présence de ces doubles liaisons confère à ces molécules des propriétés physico-chimiques particulières. À l'état naturel, la majorité des acides gras insaturés ont la configuration *cis*. La présence de double(s) liaison(s) rend toute rotation

entre les atomes impliqués impossible, la structure des acides gras insaturés est donc beaucoup plus rigide que celle des acides gras saturés. La configuration *cis* génère un angle de $\pm 30^\circ$ au niveau d'une double liaison dans la chaîne aliphatique.

Figure 1.1 : Structure chimique de quelques acides gras. Sur la figure sont présentes les structures d'un acide gras saturé (a) l'acide palmitique $C_{16:0}$, d'un acide gras mono-insaturé en configuration *trans* et *cis* (b) l'acide oléique $C_{18:1}$ et finalement la structure d'un acide gras poly-insaturé (c) l'acide arachidonique $C_{20:4}$.



b Les glycérolipides

Des glycérolipides simples (glycérides) jouent essentiellement un rôle énergétique. Des glycérolipides complexes (glycérophospholipides ou phospholipides) participent à la structure des membranes. Les glycérides sont formés par réaction d'estérification du glycérol par des molécules d'acides gras. Les lipides constituant la membrane plasmique sont donc principalement des phospholipides (ou glycérophospholipides). Ce sont des dérivés de l'acide phosphatidique. Les deux premiers atomes de carbone du glycérol sont estérifiés par des chaînes aliphatiques d'acides gras, tandis que le troisième carbone est estérifié à un groupement phosphoryle relié lui-même par une autre liaison ester à un alcool (Figure 1.2). Il s'agit du glycérol, de l'inositol, de l'éthanolamine ou de ses dérivés, la choline ou la sérine, formant respectivement les *PG*, *PI*, *PE*, *PC*, *PS*. À *pH* neutre, le groupement phosphate porte une

charge négative. Le phospholipide est donc chargé négativement lorsque la tête polaire est neutre ou négative (*PG*, *PI* ou *PS*). Il est zwitterionique lorsque la tête polaire est chargée positivement (*PC* et *PE*).

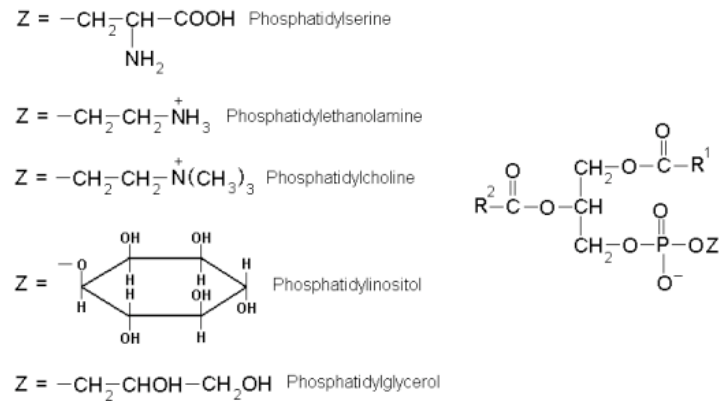


Figure 1.2 : Structure chimique des glycérolipides. À droite, la structure d'un glycérolipide est représentée. Sur la molécule de glycérol sont liées deux molécules d'acide gras R^1 et R^2 et le troisième carbone est lié à un groupement phosphate, relié à la tête polaire Z . Les groupements polaires usuels sont présentés à gauche.

c Les sphingolipides

Les sphingolipides sont constitués d'un élément de base : la sphingosine. La sphingosine est une molécule de 18 à 20 atomes de carbone, contenant des fonctions alcools et une fonction amine. Sur la fonction alcool terminale, un substituant est lié par une liaison phospho-diester. Il peut s'agir de la choline (sphingomyéline) ou d'un groupement glucidique (glycosphingolipides). Sur la liaison amine se lie un acide gras formant ainsi une liaison amide entre le groupement aminé de la sphingosine et le groupement carboxyle de l'acide gras. La sphingomyéline (*SM*) est le sphingophospholipide le plus abondant dans les membranes plasmiques des cellules de mammifères. Bien que la *SM* soit chimiquement différente de la *PC*, leur conformation et distribution de charges sont semblables (**Figure 1.3**). Pour ce qui concerne les glycosphingolipides, lorsque le sucre est simple, il s'agit d'un cérébroside. Lorsque le sucre est un oligosaccharide, il s'agit d'un ganglioside. Ces sphingolipides sont essentiellement présents dans le système nerveux des mammifères (Goñi & Alonso, 2006).

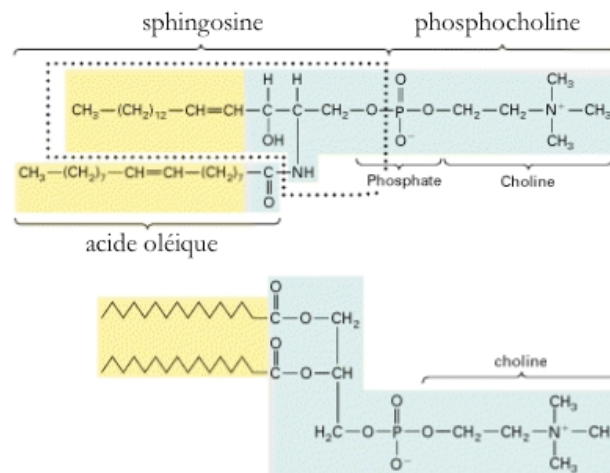


Figure 1.3 : Structure chimique de la sphingomyéline et de la phosphatidylcholine.

d Les stérols

Les stérols se retrouvent dans les membranes des eucaryotes. Le stérol le plus représenté dans les graisses animales est le cholestérol (Róg *et al.*, 2009 ; Figure 1.4). Il s'insère dans la membrane plasmique avec le groupement hydroxyle au niveau de la partie polaire des phospholipides et son corps hydrophobe au contact des chaînes d'acides gras (Demel & de Kruyff, 1976 ; Yeagle, 1985). Les stérols sont également présents dans le monde végétal où ils sont appelés phytostérols (Hartmann, 1999) et dans la levure sous le nom d'ergostérols (Parks & Casey, 1995). Leur structure est très différente de celle des phospholipides, car leurs atomes de carbones forment des chaînes polycycliques rigides. La petite taille du groupement hydrophile leur confère une flexibilité et une mobilité élevée dans la membrane, notamment au niveau de la diffusion transverse.

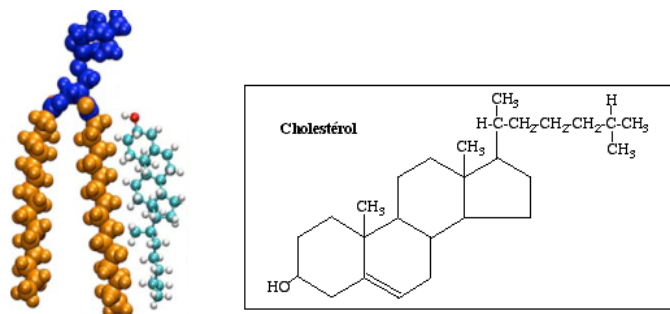


Figure 1.4 : Structure chimique d'une molécule de cholestérol et sa disposition par rapport à une molécule de phospholipide.

e Propriétés des membranes

Dans le vivant, les membranes sont composées de deux feuillets de lipides. Cette organisation est la conséquence du caractère amphiphile de ces molécules. Les parties hydrophobes (chaînes d'acides gras ou chaînes polycycliques) sont au centre de la bicouche et les parties hydrophiles sont en contact avec le milieu aqueux de part et d'autre de la bicouche. Les phospholipides sont les composants majeurs de toutes les membranes biologiques. Le cholestérol est un composant important des membranes des cellules animales uniquement. Les phospholipides sont à la base de la structure principale des membranes, et les propriétés physiques de la membrane découlent de celles de ses composants. Le point de fusion des acides gras augmente avec la longueur de la chaîne carbonée. L'insaturation de la chaîne carbonée diminue la température de fusion (voir le [tableau 1](#) pour les valeurs de température de fusion de quelques acides gras). La température de transition d'une bicouche dépend de sa composition en acide gras. L'état physique de la membrane est également affecté par le cholestérol. À cause de leur orientation dans la bicouche, les molécules de cholestérol brisent l'empaquetage étroit des chaînes d'acides gras et influencent leur mobilité. Mais ce sont les acides gras qui imposent leur état à la majorité des lipides dans la bicouche, ce qui implique des variations dans la nature de ces derniers selon leur fonction dans les organismes vivants. Ainsi, les parois des mycobactéries (Alvarez & Steinbüchel, 2002), les revêtements cireux des végétaux (Samuel *et al.*, 2008) et des insectes sont en général composés d'acides gras saturés à longue chaîne assurant leur rigidité (nombre de carbones supérieur ou égal à 20). Pour des organismes homéothermes, lorsque la température change, des recompositions en acide gras se produisent pour garder la fluidité adéquate des membranes plasmiques.

CATEGORIE	NOM COMMUN	FORMULE	POINT DE FUSION EN °C
Acides gras saturés	Laurique	C12:0	43,5
	Palmitique	C16:0	63
	Stéarique	C18:0	70
	Arachidique	C20:0	76,5
Acides gras mono-insaturés.	Palmitoléique	C12:1 ^{Δ9}	- 0,5
	Oléique	C18:1 ^{Δ9}	13,5
Acides gras poly-insaturés	Linoléique	C18:2 ^{Δ9,12}	- 5
	Linoléinique	C18:3 ^{Δ9,12,15}	- 11
	Arachidonique	C20:4 ^{Δ5,8,11,14}	- 49,5

Tableau 1 : Point de fusion de quelques acides gras.

Les bicouches lipidiques ne sont perméables qu'aux petites molécules hydrophobes (O_2 , N_2 , glycérol, stéroïdes, anesthésiques généraux, etc.), par diffusion simple. Cette diffusion est d'autant plus efficace que la molécule qui traverse la membrane est petite et hydrophobe. Des petites molécules polaires, comme l'eau ou l'urée, diffusent très lentement à travers la membrane. Des molécules neutres de plus grande taille et les ions nécessitent la présence d'un transporteur protéique. Les membranes servent de support à de nombreuses protéines transmembranaires ayant pour rôle de réguler les échanges (par exemple, les canaux ioniques pour les transferts d'ions).

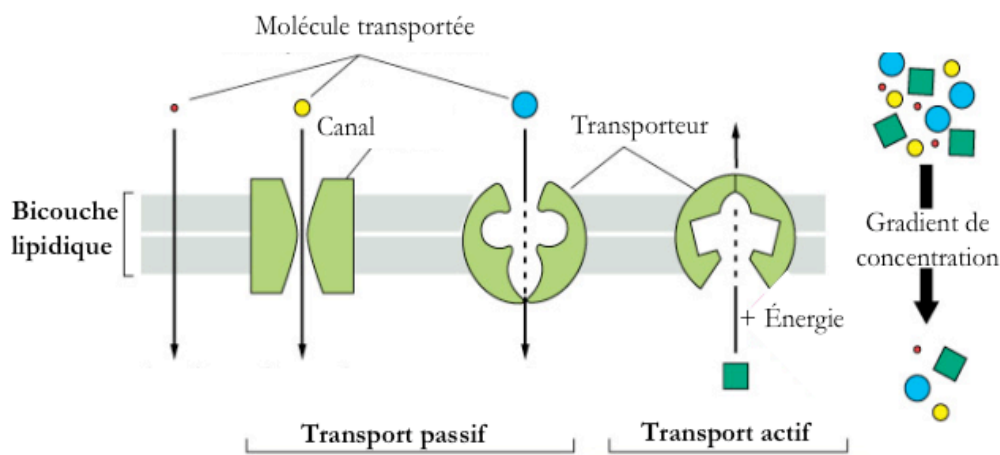


Figure 1.5 : Modes de transport d'une molécule à travers la membrane plasmique.

Ces échanges se déroulent par deux différents modes : le transport passif, et le transport actif (Figure 1.5). Le transport passif, des ions sont véhiculés d'un côté de la membrane à l'autre, le long de leur gradient électrochimique. Le transport des ions et autres molécules neutres contre leur gradient électrochimique nécessite un transport actif, couplé souvent à une hydrolyse d'*ATP*. La membrane sert aussi de support au transfert d'information biologique. Cette information prend la forme d'une hormone, d'un sucre, d'une protéine, captée par des récepteurs spécifiques. Cette interaction entre les facteurs de signalisation et les récepteurs (comme les récepteurs couplés aux protéines *G*) enclenche un mécanisme de signalisation cellulaire aboutissant à une réaction de la cellule face au signal qu'elle a reçu. Dans certains cas, la transmission d'un signal à travers la bicouche lipidique n'est pas médiée par la fixation d'un ligand mais par un phénomène appelé mécanotransduction. Une force mécanique exercée au niveau de la membrane plasmique est convertie en un signal électrique ou biochimique

(Martinac, 2004 ; Orr *et al.*, 2006). La conversion de cette force est souvent accomplie par une protéine présente au niveau de la membrane de manière intrinsèque ou extrinsèque. Une présentation générale des protéines sera effectuée. Les protéines mécanosensibles seront ensuite abordées plus en détail, notamment la protéine mécanosensible bactérienne *MscL*.

II Les protéines membranaires

Chacune des fonctions liées à la membrane met en jeu des protéines associées plus ou moins étroitement à la membrane. Les protéines membranaires jouent un rôle majeur dans le transfert d'information (récepteurs), le transport de matière (transporteurs et canaux) et la transformation d'énergie. Elles peuvent s'insérer partiellement dans un des deux feuillets membranaires en gardant une partie hydrophile : ce sont les protéines extrinsèques (ou périphériques). Cette insertion membranaire est rendue possible par la présence d'une partie hydrophobe, tel qu'un acide gras (myristoylation, palmitoylation ou isoprenylation du côté cytosolique ; ancre glycosylphosphatidylinositol *GPI* du côté extracellulaire) ou par des interactions électrostatiques. Ces dernières sont détachées de la membrane sous des conditions douces qui ne perturbent pas la bicouche lipidique, comme des changements de concentrations ioniques, de *pH* ou la présence d'agents chaotropes ou chélatants (Goñi, 2002).

Les protéines transmembranaires (ou intégrales) possèdent une ou plusieurs régions très fortement hydrophobes intégrées à la membrane. La purification d'une protéine intrinsèque implique donc la rupture de la bicouche lipidique nécessitant l'utilisation de détergent. Ces protéines sont classées en trois groupes : les protéines monotopiques, bitopiques et polytopiques.

a Topologie membranaire

Les protéines monotopiques exposent leur(s) région(s) hydrophile(s) exclusivement sur une des faces de la membrane. Elles ne traversent pas la bicouche lipidique. Elles sont considérées extrinsèques lorsqu'elles interagissent faiblement avec la membrane, à travers une liaison électrostatique, ou en interagissant avec une protéine transmembranaire. Les protéines monotopiques sont considérées intrinsèques lorsque l'utilisation de détergent est nécessaire pour les extraire de la membrane.

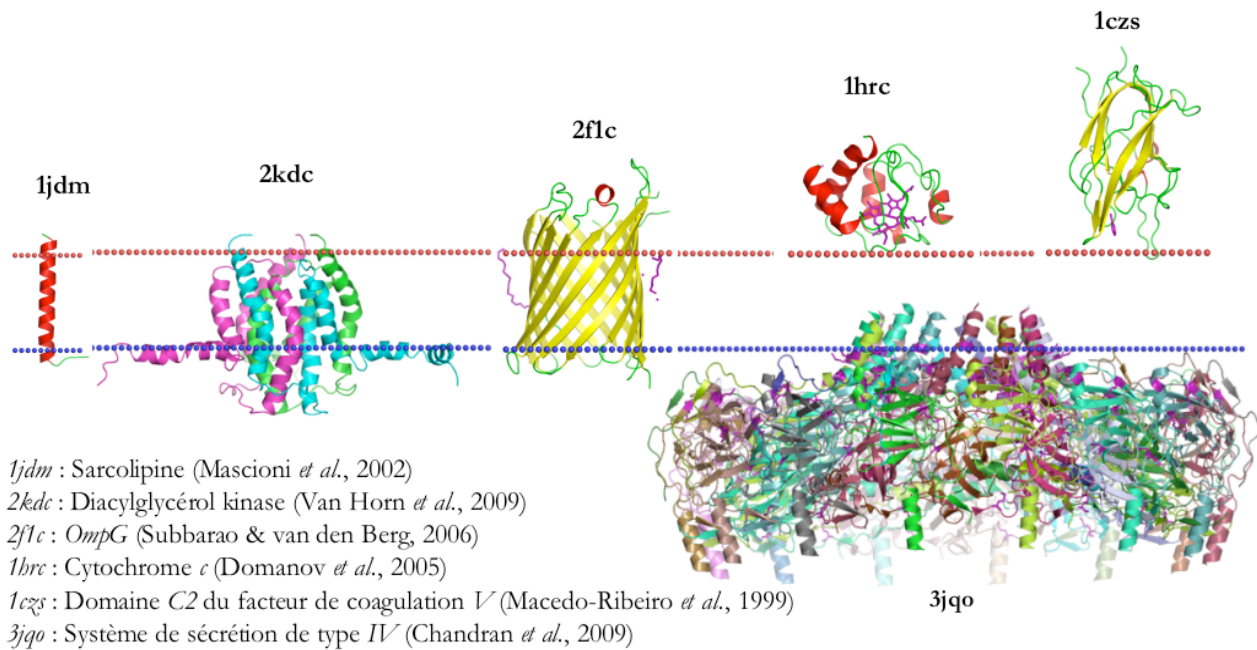


Figure 2.1 : Topologie des protéines membranaires. Six protéines membranaires sont représentées au niveau d'une membrane schématisée par deux traits : le rouge pour le feuillet externe, et le bleu pour le feuillet interne. Parmi ces protéines, trois sont monotopiques, une bitopique et deux polytopiques. De gauche à droite se trouvent les structures *PDB* de : la sarcopiline (bitopique, 1 hélice α), la diacylglycérol kinase (polytopique, 3 X 3 hélices α), la porine *OmpG* (polytopique, 14 feuillets β), le cytochrome *c* (monotopique, 4 hélices α), le domaine C2 du facteur de coagulation *V* (monotopique, 13 brins β) et le système de sécrétion bactérien de type *IV* (monotopique, 33 X 13 brins β et 6 X 13 hélices α).

Les protéines bitopiques et polytopiques exposent leurs régions hydrophiles des deux côtés de la membrane. Les protéines bitopiques ont un seul segment peptidique transmembranaire, les protéines polytopiques en ont plusieurs. Du fait des contraintes structurales imposées par l'hydrophobie au centre de la bicouche, les segments en contact avec la membrane peuvent être sous forme d'une ou de plusieurs hélices α ou sous forme d'une tonneau β (voir la [figure 2.1](#) pour les différentes topologies de protéines membranaires).

b Le modèle de la mosaïque fluide

En l'absence de liaisons covalentes entre l'ensemble des constituants membranaires, leurs interactions sont faibles et la membrane est généralement considérée comme un fluide bidimensionnel (Singer & Nicolson, 1972 ; [Figure 2.2](#)), ce qui confère aux lipides et aux protéines une grande dynamique diffusionnelle.

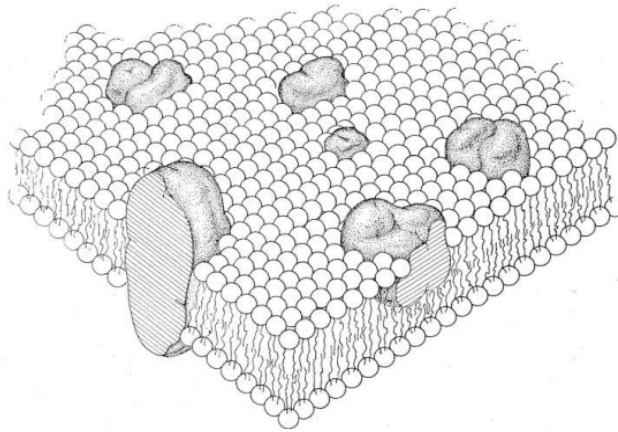


Figure 2.2 : Modèle de la mosaïque fluide de *Singer & Nicolson*.

Les constituants se déplacent plus ou moins librement. Cette fluidité est conditionnée par trois facteurs : la température (son augmentation accélère les mouvements), la quantité de cholestérol (sa présence diminue la fluidité) et la nature des phospholipides. Les acides gras insaturés facilitent la fluidité, les acides gras saturés assurent la rigidité de la membrane.

Quatre types de mouvements rapides et un mouvement lent existent en ce qui concerne les lipides. Les mouvements rapides correspondent aux changements d'orientation des têtes polaires, aux isomérisations trans/gauche des chaînes d'acides gras, à la diffusion latérale (un lipide peut changer de place avec son voisin), et à la rotation (autour de son axe). Le *flip-flop* ou translocation (changement de feuillet avec basculement) est un mécanisme très lent chez la plupart des lipides (Figure 2.3). Les protéines ne peuvent pas effectuer de *flip-flop*, mais elles sont capables de se mouvoir dans cette membrane par rotation et diffusion latérale.

Néanmoins, la grande variété de lipides et de protéines dans la membrane, ainsi que le caractère hors équilibre de ce système en font un système moléculaire complexe fortement hétérogène. Malgré cela, la membrane assure la stabilité de la cellule ou de l'organite qu'elle englobe. L'homéostasie cellulaire est assurée en premier lieu par la membrane plasmique et les différentes protéines membranaires. Ces dernières sont responsables de la transmission des signaux extracellulaires en une réponse électrique ou chimique, par la fixation d'un ligand, mais aussi par la transduction d'un signal mécanique, chimique ou thermique exercé contre la paroi membranaire. Ce phénomène est appelé mécanotransduction. Il regroupe un ensemble de

fonctions permettant à des molécules biologiques d'être mécanosenseurs et/ou mécanosensibles, capables de délivrer une information sur leur environnement et de modifier leurs propriétés sous l'action de contraintes physico-chimiques diverses (Ingber, 2006).

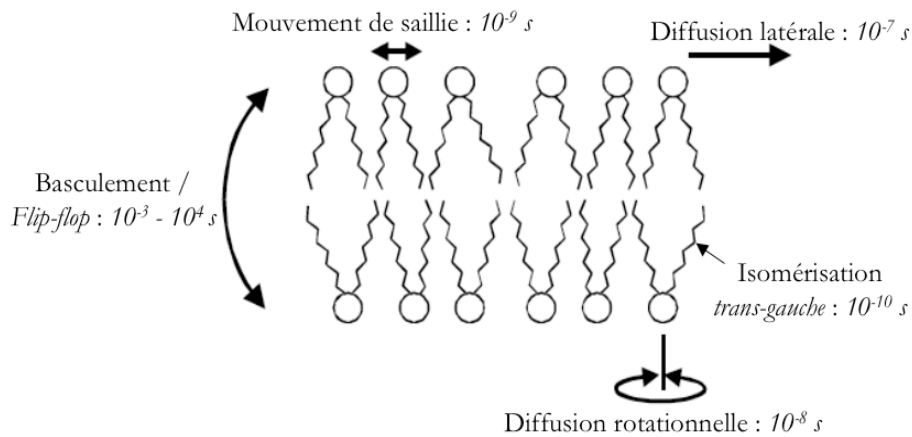


Figure 2.3 : Mouvement des molécules de lipides au sein de la membrane. Les flèches indiquent les mouvements des lipides à l'intérieur de la bicouche. Les temps de corrélations moyens de ces mouvements sont indiqués (Gawrisch, 2005. *The Structure of Biological Membranes* -Yeagle - p. 147).

c les canaux mécanosensibles

La mécanotransduction est un phénomène ubiquitaire qui agit dans plusieurs processus physiologiques sous des formes diverses et variées. Il joue un rôle essentiel dans le développement embryonnaire (Adamo *et al.*, 2009 ; Kücken & Newell, 2005; Wozniak & Chen, 2009), dans les fonctions sensorielles comme l'audition, le toucher et la nociception (Breneman *et al.*, 2009 ; Suzuki *et al.*, 2003). Il agit également dans des fonctions cellulaires comme la prolifération, l'apoptose, l'adhésion, la contractilité et la migration (Grashoff *et al.*, 2010 ; Jaalouk & Lammerding, 2009). Les cellules endothéliales cardiovasculaires agissent également comme des mécanorécepteurs, dont la fonction est d'intégrer les variations de forces hémodynamiques de la circulation et de réguler en conséquence la structure, le remodelage et la tension vasculaires (Davies, 1995 ; Topper & Gimbrone, 1999) (Voir le tableau 2 pour des informations supplémentaires sur les acteurs moléculaires des activités mécanosensibles dans les différents tissus et cellules).

Famille	Protéine	Localisation	Fonction
<i>DEG_s/ENaC_s</i>	<i>MEC-4 - MEC-10 - DGE</i>	Organe sensoriel (<i>C. elegans</i>)	Toucher
	<i>ASIC2a - ASIC3</i>	Nerfs du derme - Ganglion spinal (mammifères)	Toucher - Nociception Senseur d'ischémie
	<i>ENaC</i>	Barorécepteurs de la carotide et de l'aorte Épithélium alvéolaire et rénal (mammifères)	Barorégulation - Transport de Na^+ Osmorégulation
<i>TRP_s</i>	<i>TRPA</i>	Cellules ciliées - Nerfs sensitifs périphériques (mammifères, <i>C. elegans</i> , drosophile)	Jonctions Régulation du Ca^{2+}
	<i>TRPN</i>	Cils - Organes auditifs - neurones mécanosensibles (drosophile, poisson-zèbre, <i>C. elegans</i>)	Audition - Toucher - Proprioception
	<i>TRPV</i>	Épithélium salivaire, rénal et hépatique (mammifères)	Osmorégulation - Thermorégulation Régulation du Ca^{2+}
	<i>TRPP</i>	Cil primaire de l'épithélium rénal - Système vasculaire (mammifères) Membrane interne du vacuole de levure	Jonctions Régulation du Ca^{2+} Osmorégulation
<i>K_{2P}</i>	<i>TREK-1</i>	Système nerveux central - Derme - Système cardiovasculaire (mammifères)	Nociception Vasodilatation Neuroprotection
	<i>TREK-2</i>	Foie, Rate, Pancréas (mammifères)	inconnue
	<i>TRAAK, TRAAK₂</i>	Cerveau - neurones (mammifères)	inconnue

Tableau 2 : Protéines mécanosensibles. Les informations sont issues de (Arnadottir & Chalfie, 2010) et des références citées dans le papier et dans le début du chapitre. *DEG/ENaC* : dégénérine/ canal épithélial à sodium. *TRP* : canal potentiel récepteur transitoire. *K_{2P}* : canal potassium à deux domaines *P*. *MEC-4* et *MEC-10* : canaux mécanosensibles. *DGE* : dégénérine. *ASIC* : canal ionique sensible à l'acide. *TREK* : canal à potassium de la famille *TWIK*. *TRAAK* : canal ionique sensible à l'acide arachidonique de la famille *TWIK*. *TWIK* : canal pour la rectification du flux entrant de potassium, composé de tandem de domaines *P* (motif de structure conservé, participant à la formation du filtre de sélectivité pour le potassium).

La mécanosensibilité est une fonction que possèdent tous les organismes, du plus simple au plus complexe. Cette caractéristique diffère entre les organismes, mais il semble que le mécanisme reste inchangé (Sukharev & Corey, 2004). Elle est véhiculée par des protéines membranaires extrinsèques ou intrinsèques. Des protéines identifiées précédemment comme voltage-dépendantes (Laitko *et al.*, 2006 ; Morris & Juranka, 2007) ou ligand-dépendantes (Casado et Asher, 1998 ; Chemin *et al.*, 2007 ; Patel *et al.*, 1999) sont également mécanosensibles. Le point commun est un changement conformationnel de la protéine lorsqu'un stress mécanique est appliqué au niveau de la membrane (Markin & Martinac, 1991 ; Markin & Sachs, 2004). Deux modes de fonctionnement ont été proposés pour le mécanisme de mécanotransduction : un mode lié à la bicouche (*the bilayer model*) et un autre lié au cytosquelette (*the tethered model*) (Hamill & McBride ; 1997 ; Inoue *et al.*, 2009).

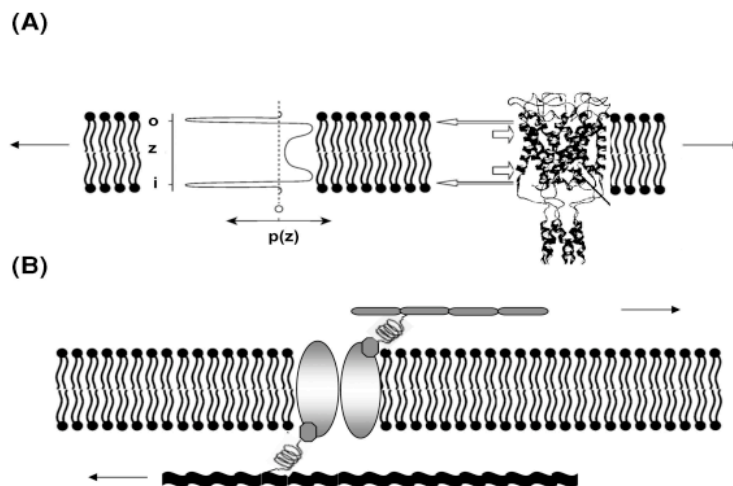
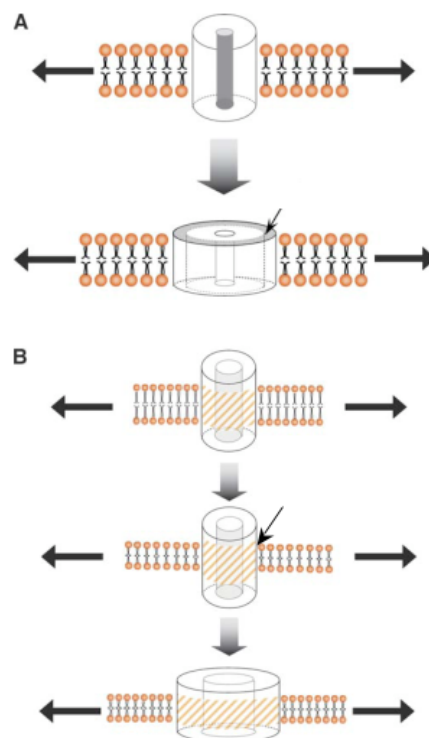


Figure 2.4 : Mécanismes d'ouverture des canaux MS. (A) Le mode lié à la bicouche (*the bilayer model*): l'étirement ou la flexion de la bicouche provoque un changement dans le profil de pression interne, qui entraîne le changement conformationnel du canal MS conduisant à l'ouverture. (B) Le mode lié au cytosquelette (*the tethered model*): la tension exercée au niveau du cytosquelette ou au niveau de la matrice extracellulaire provoque l'ouverture du canal MS sans l'implication de la bicouche. La figure est issue de *Biological Membrane Ion Channels Dynamics, Structure and Application*, par Chung *et al.*, 2006.

Dans le premier modèle, une tension membranaire est nécessaire pour activer les protéines mécanosensibles, alors que dans le second modèle, l'activation des protéines se fait par l'intermédiaire du cytosquelette (Figure 2.4). Dans ce cas, un cisaillement ou le déplacement relatif de la protéine à l'égard du cytosquelette (ou la matrice extracellulaire) entraîne l'ouverture du canal. Pour expliquer le premier modèle, deux hypothèses ont été émises (Figure 2.5): selon la première hypothèse, la protéine ouverte occuperait une plus grande surface au sein de la

bicouche lipidique, ce qui rendrait son énergie libre plus faible (Orr *et al.*, 2006). Selon la deuxième hypothèse, la tension appliquée au niveau de la membrane provoquerait la diminution de l'épaisseur de la bicouche. Pour éviter d'exposer des résidus hydrophobes, les protéines mécanosensibles changeraient de conformation, entraînant leur ouverture (Martinac & Hamill, 2002). La première hypothèse serait en faveur d'une augmentation de la probabilité d'ouverture des protéines mécanosensibles lorsqu'une tension est appliquée au niveau de la membrane, alors que la deuxième privilégie le mésappariement hydrophobe comme responsable de l'ouverture (Lee, 2003 ; Marsh, 2008).

Figure 2.5 : Hypothèses d'ouverture des canaux *MS*. La tension membranaire permet l'ouverture du canal lorsque (A) pour un canal *MS*, l'état ouvert occupe une surface plus grande que l'état fermé, et présente une énergie libre plus faible à l'état ouvert, ou lorsque (B) la tension provoque une diminution de l'épaisseur de la bicouche, provoquant un mésappariement hydrophobe au niveau des hélices transmembranaires du canal *MS*, qui tend à s'ouvrir pour réduire le coût énergétique de l'exposition des résidus hydrophobes au milieu aqueux (Orr *et al.*, 2006).



L'activité mécanosensible (*MS*) a été mise en évidence et caractérisée par des études électrophysiologiques. Mais il est rare que les canaux responsables des activités soient identifiés. Les techniques utilisées pour l'identification sont la purification grâce à un ligand spécifique, ou par l'étude de mutants. En ce qui concerne l'identification par les ligands, les molécules agissant sur les canaux *MS* sont rares. Et en ce qui concerne les mutations, les canaux étant actifs dans des conditions particulières, il est difficile de corrélérer un phénotype à l'absence d'un canal. Néanmoins, cette technique a permis l'identification du canal *MscS* de la bactérie (Levina *et al.*, 1999). Le canal *MS* potassique eucaryote *TREK-1* a été identifié par homologie de séquence (Fink *et al.*, 1996).

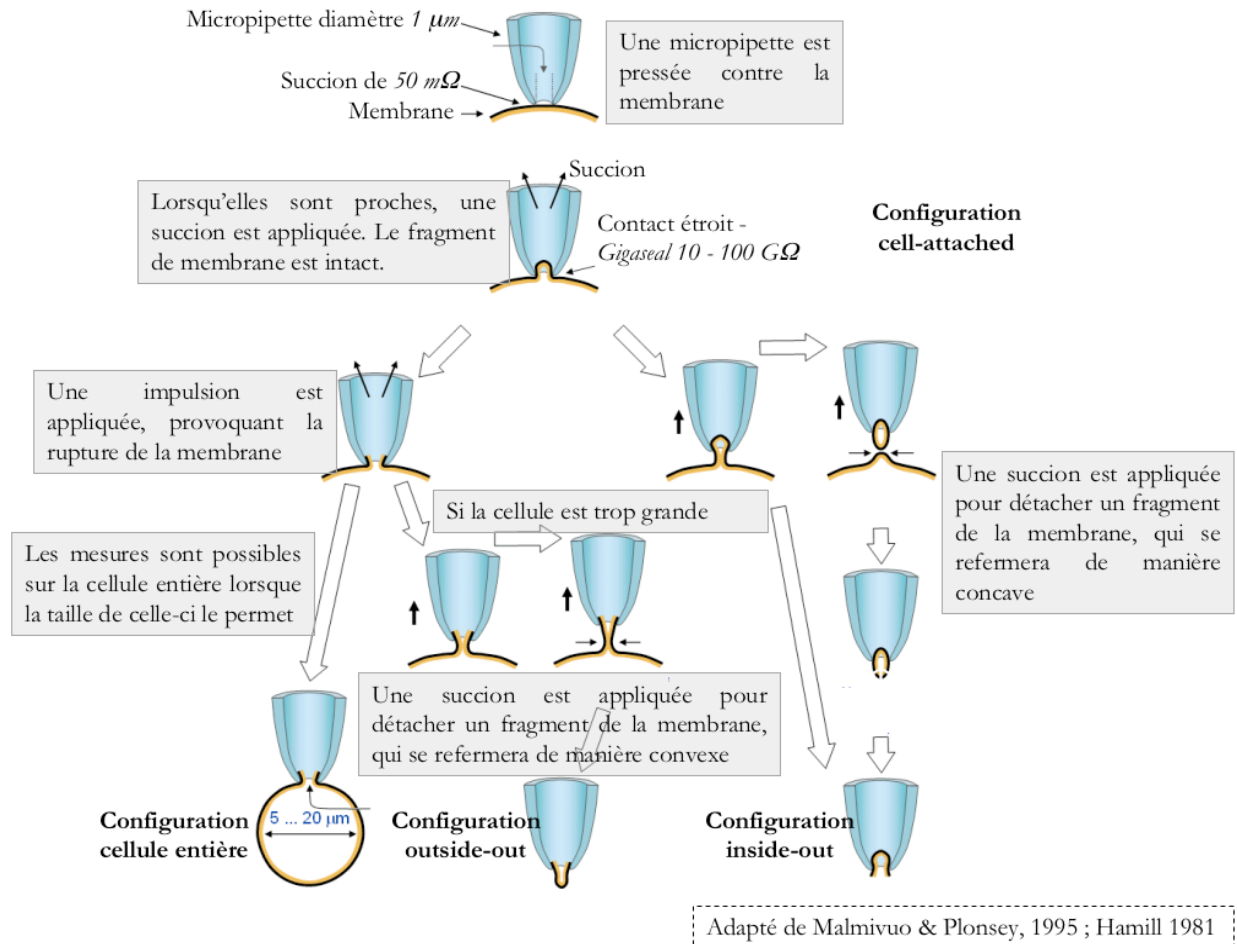


Figure 2.6 : Schéma de la technique de *patch clamp*.

d les canaux mécanosensibles bactériens

Les canaux *MSs* de la bactérie *Escherichia coli* (*E. coli*) ont été les premiers canaux ioniques bactériens identifiés. Ils jouent un rôle dans l'osmorégulation et sont les seuls canaux *MSs* dont la fonction a été clairement démontrée. Le système mécano-senseur de cette bactérie est le représentant de la classe des *MSs* bactériens et constitue un modèle d'étude de la mécanotransduction.

La mise en évidence des canaux *MSs* bactériens a été rendue possible grâce à la méthode de *patch clamp* sur des sphéroplastes (le terme désigne les bactéries dont les parois ont été digérées sous l'action d'un agent externe). Le principe repose sur l'idée suivante : la pointe effilée d'une micropipette de verre est mise en contact étroit avec un très petit morceau (ou

patch) de membrane. Lorsqu'une succion est appliquée, la membrane se scelle hermétiquement à l'extrémité de la pipette. Dès lors, lorsqu'un canal s'ouvre, tout flux ionique passe nécessairement par la pipette. Le courant électrique qui en résulte est quantifié d'après la quantité de courant électrique nécessaire pour maintenir le potentiel de membrane à une valeur imposée (ou *clamped*). (Pour des détails concernant la technique, voir Hamill *et al.*, 1981 ou Neher & Sakmann, 1976). La taille des bactéries rend impossible l'étude en *patch clamp* des canaux *MSs in vivo* puisque la micropipette mesure $1 \mu\text{m}$ de diamètre et que la taille d'une bactérie *E. coli* est $2 \mu\text{m}$ de long sur $1 \mu\text{m}$ de diamètre (Figure 2.6).

L'étude des canaux *MSs* bactériens a été rendue possible en 1987, lorsque B. MARTINAC et ses collaborateurs (*University of Western Australia*) ont appliqué sur des bactéries deux antibiotiques : le mécillinaam, qui inhibe la synthèse des peptidoglycans de la paroi, et la céphalexine qui empêche la division bactérienne. Ces deux molécules combinées favorisent la formation de sphéroplastes. L'ouverture des canaux *MSs* a pu être enregistrée et le phénomène était reproductible. Un canal à large conductance (970 pS) activé par la pression et le voltage y fut décrit. Ce canal reste ouvert pendant quelques secondes lorsqu'il est activé. Le mécanisme observé était *ATP*-indépendant (Martinac *et al.*, 1987). Sur la base de leurs propriétés électrophysiologiques proches de celles des porines, les auteurs avaient supposé que les canaux *MSs* se situeraient dans la membrane externe. En 1989, dans le laboratoire d'A. GHAZI (Institut de Biochimie et Biophysique Moléculaire et Cellulaire – *Université Paris-Sud*), les chercheurs ont isolé les membranes externe et interne d'*E. coli*, puis purifié les fractions membranaires. Chaque fraction a été fusionnée avec des liposomes et testée en *patch clamp*. Des canaux *MSs* de conductance variable ont été détectés essentiellement dans les fractions issues de la membrane interne (Berrier *et al.*, 1989). L'activité de deux systèmes moléculaires a ensuite été identifiée : un système à petite conductance ($0,9 \pm 0,1 \text{ nS}$) de taille moléculaire située entre 200 et 400 *kDa*, et un autre à large conductance ($3,1 \pm 0,3 \text{ nS}$) de taille moléculaire située entre 60 et 80 *kDa*. Ces deux systèmes étaient présents et fonctionnels dans des sphéroplastes, mais aussi après extraction des protéines de la membrane interne et reconstitution en liposomes d'asolectine (phospholipides extraits de graines de soja) dans le laboratoire de C. KUNG (*University of Wisconsin – Madison* ; Sukharev *et al.*, 1993).

III Le canal mécanosensible bactérien *MscL*

a Découverte de *MscL*

C'est au sein du laboratoire de C. KUNG que la protéine responsable de la mécanosensibilité à large conductance dans la bactérie *E. coli* a été isolée et son gène cloné. La séquence nucléotidique prédit une protéine de 136 acides aminés (Figure 3.1), de 14 957 Da de poids moléculaire pour la protéine *MscL* (pour *Mechanosensitive channel of Large conductance*). Elle ne présente pas de modifications post-traductionnelles et son profil d'hydrophobie suggère la présence de deux segments transmembranaires, ce qui ne peut pas correspondre à un arrangement en tonneau β . L'éluion de la protéine à 70 kDa par chromatographie d'exclusion moléculaire en conditions non-dénaturantes suggérait une multimérisation (Sukharev *et al.*, 1994 a). Pour démontrer que le gène *mscL* était seul codant pour la protéine *MscL*, et que celle-ci était seule responsable de l'activité de mécanotransduction, le gène a été exprimé en levures, et dans un système de synthèse *in vitro* (extraits de réticulocytes). Dans les deux cas, l'activité a été retrouvée intacte lors de la reconstitution des protéines dans des vésicules (Sukharev *et al.*, 1994 b ; Blount *et al.*, 1996 a). Des mesures de dichroïsme circulaire (CD) et de spectroscopie infrarouge (FTIR) ont montré que 82% de la protéine était sous forme d'hélices α (Arkin *et al.*, 1998).

MSIIKEFREF AMRGNVVDLA VGVIIGAAFG KIVSSLVADI IMPPLGLLIG GIDFKQFAVT
LRDAQGDIPA VVMHYGVFIQ NVFDLIVAF AIFMAIKLIN KLNRRKEEPA AAPAPTKEEV
LLTEIRDLLK EQNNRS

Figure 3.1 : Séquence peptidique d'un monomère de *MscL*.

Code couleur: acides aminés hydrophiles DENQRHSTK en bleu ; acides aminés soufrés MS en orange ; acides aminés aromatiques FYW en vert et acides aminés hydrophobes AGILPV en rouge.

Les expériences de fusion de la phosphatase alcaline bactérienne (*PhoA*) sur les six acides aspartiques de la protéine *MscL* (D18, D39, D53, D67, D84 et D127) suggéraient fortement que les sous-unités *MscL* contenaient une seule boucle périplasmique, et que leur N et C terminaux étaient orientés du côté cytoplasmique (Blount *et al.*, 1996 a ; Blount *et al.*, 1996 b ; Figure 3.2).

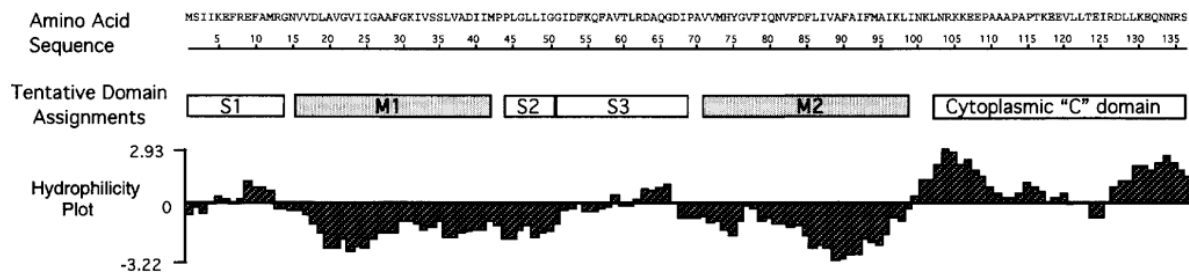


Figure 3.2 : Séquence et tracé de l'index d'hydrophilie de la protéine *Ec-MscL*. Le profil de migration sur gel *SDS-PAGE* en condition dénaturante montrait une bande à 17 *kDa*. Sur gel en conditions natives, la bande se situe entre 60 et 80 *kDa*. Cela suggère une multimérisation de la forme active. Le tracé de l'index d'hydrophilie selon l'échelle de *Kyte* et *Doolittle* indique la présence de deux hélices transmembranaires. Le cœur hydrophobe constitue les 2/3 de la protéine.

En effet, la protéine *PhoA* subit une étape de maturation dans le périplasma pour devenir fonctionnelle. Ces résultats ainsi qu'une analyse du profil d'hydropathie de la séquence de *MscL* ont permis de définir la topologie de la protéine. Le canal mécanosensible bactérien à large conductance *MscL* est un multimère, et chaque monomère est formé de 136 acides aminés. Le côté N-terminal constitué d'une dizaine d'acides aminés est situé dans le cytoplasme (*PhoA* associée à *D18* ne montre aucune activité phosphatase). Une première hélice *TM1* (L19 à A38) traverse la membrane. Une boucle contenant le résidu *D53*, ayant montré le maximum d'activité phosphatase, est périplasmique, et suivie d'une seconde hélice transmembranaire *TM2* (A70 à I96). Finalement la partie C-terminal est cytoplasmique (contenant le résidu *D127*).

La question du nombre de sous-unités subsistait. Des expériences de *cross-linking*, de fusion de deux sous-unités (Blount *et al.*, 1996 b), d'estimation de la taille du pore formé lors de l'ouverture du canal (Cruickshank *et al.*, 1997) et de microscopie électronique sur des cristaux bidimensionnels (Saint *et al.*, 1998) suggéraient une organisation hexamérique. D'autres expériences allaient à l'encontre de cette hypothèse sans pour autant indiquer le nombre de sous-unités formant le complexe (Häse *et al.*, 1997). Des études supplémentaires de *cross-linking* utilisant des agents réticulants de différentes tailles couplées à des expressions de la protéine en double ou triple tandem ont permis de répondre à cette question, et de déterminer que la protéine active était un pentamère (Sukharev *et al.*, 1999).

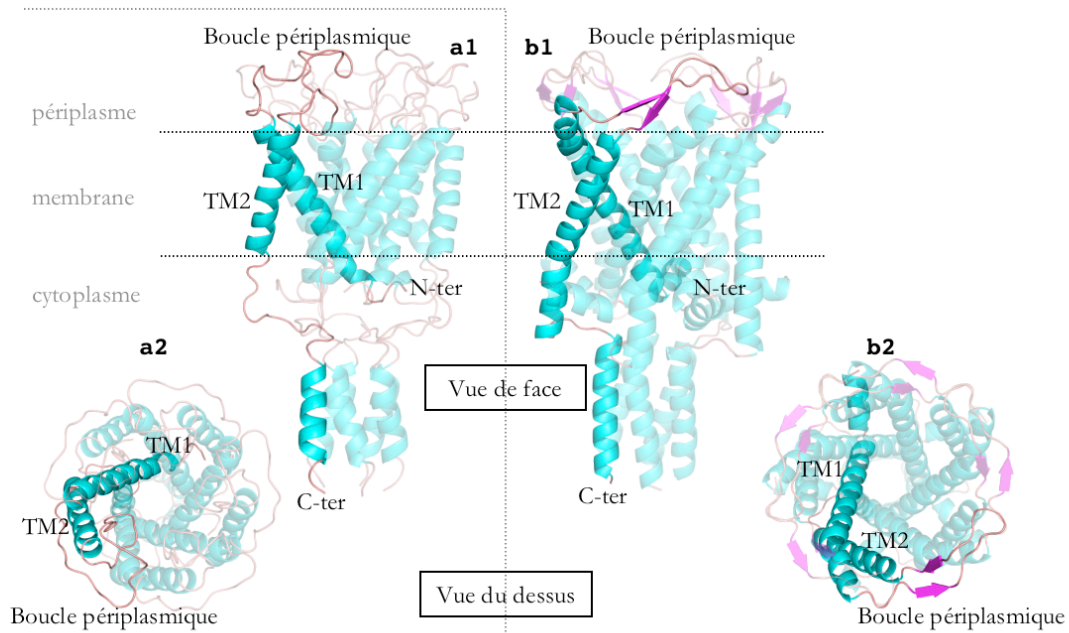
b Structure de *Mt-MscL*

Figure 3.3 : Structures de la protéine mécanosensible *Mt-MscL* en 1998 et en 2007. Dans la première structure obtenue de la protéine *Mt-MscL*, le jeu de données n'a pas permis à *Chang* et ses collaborateurs d'obtenir les coordonnées des 10 premiers résidus ni des 33 derniers (*Chang et al.*, 1998). La structure (code PDB *1msl*) est présentée ici avec une vue de face (a1) et du dessus (a2) montrant un pentamère, chaque monomère étant composé de deux hélices transmembranaires *TM1* et *TM2*. Le pore formé par les cinq hélices *TM1* présente un petit diamètre suggérant un état fermé du canal. En 2007, la structure a été améliorée (code PDB *2oar*). (*Steinbacher et al.*, 2007 dans *Current topics in Membranes*)

La structure tridimensionnelle d'un homologue de la protéine *MscL* du bacille *M. tuberculosis* (*Mt-MscL*) à haute résolution ($3,5 \text{ \AA}$) soutenait les expériences effectuées sur la protéine *MscL* d'*E. coli* (*Ec-MscL*), témoignant d'une organisation pentamérique (*Chang et al.*, 1998 ; [figure 3.3a](#)). Cette structure représente la première obtenue d'un canal *MS*. Elle montre la protéine à un état fermé ou inactif. La protéine *MscL* de *M. tuberculosis* nécessite deux fois plus de tensions que *Ec-MscL* pour son activation ce qui a rendu la cristallogénèse possible (*Moe et al.*, 2000). Elle est composée de 151 acides aminés, présentant 37% d'identité de séquence avec *Ec-MscL*.

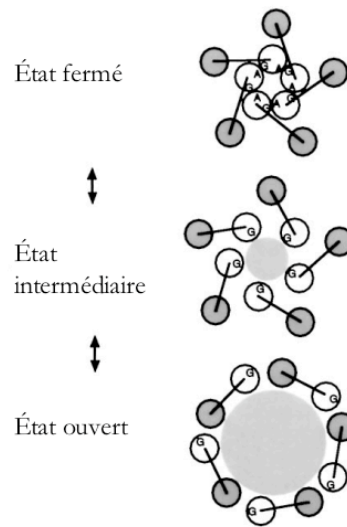
La structure a été déterminée pour les résidus 10 à 118. Les sous-unités sont arrangées autour d'un pore central formé par les hélices transmembranaires *TM1*, et les hélices transmembranaires *TM2* sont au contact avec les lipides. L'hélice *TM1* est inclinée de 37° par rapport à l'axe du pore, l'hélice *TM2* de 28° . Une hélice *TM1* d'un monomère est intercalée entre les hélices *TM1* et *TM2* de la sous-unité voisine. La boucle périplasmique plonge dans le pore dont le diamètre varie entre 2 \AA au niveau du site de constriction et 18 \AA au niveau du pore formé par les hélices cytoplasmiques en C-terminal. La boucle agirait comme un ressort résistant à l'ouverture du canal en s'opposant au mouvement des hélices, et aidant celui-ci à se refermer (Ajouz *et al.*, 2000). Les auteurs soupçonnaient que l'organisation hélicoïdale observée au niveau C-terminal était due à un artéfact expérimental. Les hélices seraient stabilisées par le *pH* dans les conditions de cristallogénèse (0,05% DDM, *pH* 3,6-3,8). Le rôle de ces hélices dans le fonctionnement de la protéine n'était pas clair, surtout que des études antérieures avaient montré que la délétion des 27 derniers résidus de la protéine *Ec-MscL* n'affectait pas la fonction du canal (Blount *et al.*, 1996 b ; Häse *et al.*, 1997).

c Analyse structurale et fonctionnelle de *MscL*

Le diamètre du pore de la structure de *Mt-MscL* montre un canal fermé. La taille du canal *Ec-MscL* ouvert avait été déterminée par des mesures de perméabilité, et il s'avérait que le diamètre du canal ouvert avoisinait les 40 \AA (Ajouz *et al.*, 1998 ; Berrier *et al.*, 2000 ; Cruickshank *et al.*, 1997 ; Sukharev *et al.*, 1999). Le passage de 2 à 40 \AA ne serait permis que par une conformation ouverte impliquant les dix hélices de la protéine (Batiza *et al.*, 1999 ; Yoshimura *et al.*, 1999 ; [figure 3.4](#)).

La structure du domaine N-terminal de *Mt-MscL* n'était pas résolue dans la structure cristallographique, probablement à cause de sa dynamique ou de la présence de plusieurs conformations (Oakley *et al.*, 1999). Cette région est très conservée mais très peu sensible aux mutations, délétions ou insertions. Elle se replierait en hélices α et formerait un faisceau dans le prolongement du pore, agissant comme une seconde barrière lors de l'ouverture du pore (Sukharev *et al.*, 2001 a).

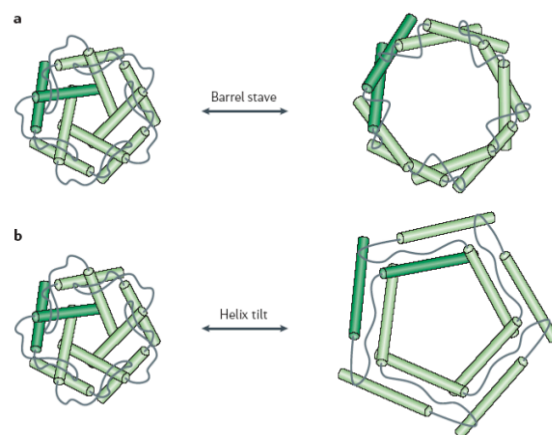
Figure 3.4 : Modèle d'ouverture à dix hélices de la protéine *MscL*. À partir des résultats de mutagenèse et de calcul de la taille du pore du canal ouvert, *Yoshimura* et ses collaborateurs ont proposé un modèle d'ouverture impliquant les dix hélices transmembranaires en insistant sur la rupture d'une interaction *Gly-Ala* intermonomérique provoquant la première étape de l'ouverture (*Yoshimura et al.*, 1999).



i. Rôle des hélices transmembranaires

Par des méthodes de modélisation par homologie basée sur la structure de *Mt-MscL*, ainsi que des critères de prédiction de structure secondaire et de contraintes de liaisons entre résidus conservés, *S. SUKHAREV* et ses collaborateurs (*University of Maryland*) ont émis une hypothèse d'ouverture du canal *Ec-MscL*. En effet, la tension membranaire provoquerait le passage de la protéine par trois conformations intermédiaires et l'ouverture de deux pores, un transmembranaire et un autre cytoplasmique. La première conformation correspondrait à une extension du canal jusqu'aux $2/3$ d'ouverture (*CE*) et serait une étape thermodynamiquement stable. La tension atteindrait ensuite une valeur suffisante pour déplacer un des segments occluant le pore lors de la deuxième étape (*S1*). Ceci engendrerait d'autres sous états au fur et à mesure que la tension augmente et que les segments se déplacent dans le plan pour accroître le diamètre du pore (*S2, S3*) jusqu'à l'ouverture total du canal (*O*) (*Sukharev et al.*, 2001 a, b).

Figure 3.5 : Les deux modes d'ouverture hypothétiques de *MscL*. Le modèle à 10 hélices (a) exposé précédemment, et le modèle d'inclinaison des hélices (b), impliquant uniquement les 5 hélices *TM1* dans l'ouverture du canal.



Le modèle d'ouverture contredisait les travaux antérieurs suggérant l'implication des dix segments transmembranaires dans la formation du pore. Pour les auteurs, ce modèle présentait plusieurs inconvénients : l'exposition lors de l'ouverture de plusieurs résidus hydrophobes des hélices *TM2* à l'eau ; un mode de réarrangement des hélices jamais observé auparavant pour des protéines membranaires ; un pore de diamètre inférieur à ce que suggéraient les expériences de perméabilité. Finalement, ce modèle ne prenait pas en compte le rôle joué par la boucle périplasmique (Blount *et al.*, 1996 b ; Ajouz *et al.*, 2000). Le nouveau modèle proposé présente une prédominance des hélices *TM1* dans la formation du pore, mais n'apporte pas d'informations sur l'implication de la boucle périplasmique (Figure 3.5). Selon ce modèle, l'ouverture du canal serait causée par l'inclinaison des hélices. Ce qui explique pourquoi le mécanisme d'ouverture était *ATP* indépendant, car il n'y aurait pas de changement d'accessibilité à l'eau des résidus hydrophobes, ni l'intégration de résidus chargés dans la membrane. Les résidus constituant les hélices *TM1* qui forment les parois du pore étaient certes hydrophobes, mais leurs chaînes latérales sont de petites tailles, ce qui ferait que le squelette hydrophile de ces acides aminés contribuerait à la formation d'un canal hydrophile.

Néanmoins, lorsque les hélices *TM1* s'inclinent pour ouvrir le canal, suivies des hélices *TM2* au contact de la membrane, des zones hydrophobes de cette dernière s'exposeraient. Pour répondre à ce problème, les auteurs suggéraient deux hypothèses : la première hypothèse propose que la bicouche lipidique est plus fine lorsque la tension est appliquée ce qui expose les zones hydrophobes à un environnement hydrophile ; la deuxième hypothèse suggère qu'au niveau de la protéine des résidus cyto- ou périplasmiques contribuent à recouvrir les zones exposées.

Les modèles de la protéine *Ec-MscL* aux différents états d'ouverture de la protéine ont servi de point de départ à différentes études de dynamique moléculaire au cours des dix dernières années. Des études de dynamique sous contraintes avançaient l'existence d'une fermeture hydrophobe englobant les deux pores suggérés auparavant. Cette fermeture s'étendrait le long des résidus hydrophobes des hélices *TM1* ainsi que ceux de la partie N-terminale (Kong *et al.*, 2002). À partir du modèle fermé, des modèles séquentiels ont été créés en inclinant les hélices *TM1* progressivement, afin d'identifier le rôle de l'hélice N-terminale, et

du degré d'inclinaison de l'hélice *TM1* dans le modèle ouvert (Betanzos *et al.*, 2002).

Les analyses des dynamiques ainsi que des études de mutagenèse ont montré que les hélices *TM2* s'inclinaient simultanément avec les hélices *TM1*. Des résidus hydrophobes de la partie N-terminale étaient très proches de résidus de l'hélice *TM1*. Cet ensemble participe probablement à la formation d'un site de constriction hydrophobe du côté cytoplasmique. Ce qui suggérerait donc que tout remplacement d'un acide aminé de ce site, par un autre de nature hydrophile augmenterait la probabilité d'ouverture du canal. Les connaissances concernant le site de constriction correspondent aux données de mutagenèse dirigée acquises depuis la découverte de la protéine. Différents travaux citaient notamment les résidus : *F7* et *F10* (Oakley *et al.*, 1999 ; Sukharev *et al.* 2001 a) ; *G14* (Gu *et al.*, 1998) ; *L19* (Yoshimura *et al.*, 1999) ; *G22* (Ou *et al.*, 1998 ; Yoshimura *et al.*, 1999 ; Yoshimura *et al.*, 2001) ; *V23*, *G26* et *G30* (Ou *et al.*, 1998). La corrélation d'inclinaison des hélices transmembranaires *TM1* et *TM2* pouvait s'attribuer également à des interactions hydrophobes impliquant les résidus *I3* et *I96* (Sukharev *et al.*, 2001 a ; [Figure 3.6](#)).

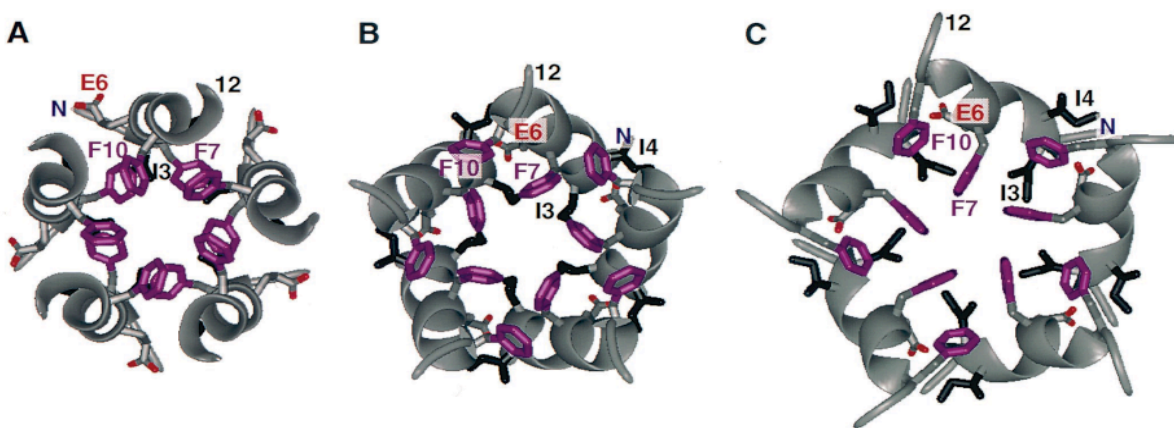


Figure 3.6 : Structure et composition d'un site de constriction cytoplasmique au niveau de *Ec-MscL*. À partir des résultats de mutagenèse et de *cross-linking* réalisés au niveau de la partie N-terminale, Sukharev et ses collaborateurs ont proposé un modèle dans lequel des résidus hydrophobes cytoplasmiques forment un site de constriction (A) dont l'ouverture permet le passage d'un état intermédiaire (B) à un état ouvert (C) (Sukharev *et al.*, 2001 a).

ii. Rôle des parties extra-membranaires

Du fait des conditions expérimentales durant la cristallogénèse, la structure cristallographique révèle un domaine C-terminal formant un faisceau pentamérique. Les hélices y ont une orientation non usuelle dans laquelle les résidus hydrophobes pointent vers l'extérieur

tandis que les résidus chargés s'enfouissent à l'intérieur du faisceau. La séquence de cette partie montre des similitudes au domaine d'oligomérisation d'une protéine présente au niveau du cartilage (*COMP*) (Anishkin *et al.*, 2003). La structure *COMP* montre une super-hélice formée par cinq faisceaux, stabilisée par des interactions hydrophobes internes, et des ponts salins périphériques. Le domaine C-terminal de la protéine *Ec-MscL* a été modélisé par homologie à la protéine *COMP*. Un motif apparaît nécessaire à l'oligomérisation du domaine C-terminal. Ce domaine montre une boucle située à la fin de l'hélice *TM2*, suivie d'une hélice cytoplasmique. La modélisation montre que la superhélice est stable pendant toutes les étapes d'ouverture. Cette partie de la protéine agirait, lors de l'ouverture du canal comme un filtre empêchant la sortie de molécules au-delà d'une certaine taille.

La structure du domaine N-terminal n'a pas été résolue dans la structure cristallographique. D'après les modélisations et les travaux de mutagenèse, elle se replierait en hélice α constituant un faisceau dans le prolongement du pore et agirait comme une seconde barrière pour l'ouverture du pore. L'hypothèse est appuyée par des expériences de mutagenèse sur *Ec-MscL* (Anishkin *et al.*, 2005).

La boucle périplasmique assure la connexion entre les hélices *TM1* et *TM2*. Elle contient quatre glycines et une proline, ce qui lui confère une grande flexibilité. Le clivage de cette boucle par protéolyse facilite l'ouverture du canal (Ajouz *et al.*, 2000). Elle jouerait le rôle d'un ressort contraignant le canal dans sa position fermée et empêchant l'inclinaison des hélices au cours de l'ouverture.

d Mécanisme d'ouverture

L'ouverture du canal forme un pore de 30 à 40 Å de diamètre. Un modèle soutenait l'implication des dix hélices transmembranaires (Blount & Moe, 1999). Ce modèle a été abandonné lorsqu'il s'est avéré que les hélices *TM1* pouvaient former le pore, dans un canal coupé en deux. Les deux entités formaient un canal fonctionnel lorsqu'elles étaient réassociées, et ce canal était plus sensible à la tension (Park *et al.*, 2004). Ce travail confirme également le rôle de ressort de la boucle périplasmique. Le modèle proposé par *SUKHAREV* et ses collaborateurs en 2001 donne la priorité à l'inclinaison des hélices *TM1* pour l'ouverture du

canal. Une faible rotation des hélices *TM1* autour de leur axe est observée. L'inclinaison des hélices serait suivie d'un élargissement global du canal. Les hélices *TM1* et *TM2* s'inclineraient conjointement mais dans deux directions différentes. Le modèle d'ouverture est schématisé par le mécanisme d'ouverture des diaphragmes d'iris (*Iris-like movement* ; Betanzos *et al.*, 2002).

Le faisceau d'hélices N-terminales agirait comme une seconde barrière (Sukharev *et al.*, 2001 b ; [figure 3.7](#)). Le modèle d'ouverture actualisé suggère que le faisceau d'hélices du côté C-terminal agirait comme un filtre pour éviter la fuite de grosses molécules. De nombreuses études confirment le rôle de l'inclinaison des hélices lors de l'ouverture du canal, tant *in vitro* (Chiang *et al.*, 2005) que *in silico* (Gullingsrud & Schulten, 2003 ; Valadié *et al.*, 2003). Tous ces points ont été débattus dans différentes études de dynamique moléculaire et de mutagenèse (Debret *et al.*, 2008 ; Powl *et al.*, 2008 a ; Tang *et al.*, 2008 ; Yefimov *et al.*, 2005).

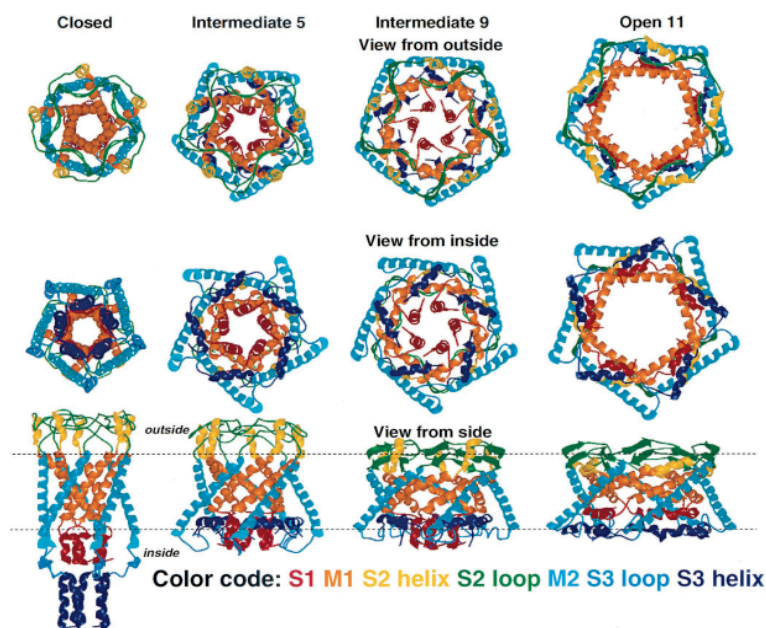


Figure 3.7 : Modèle d'ouverture de la protéine *Ec-MscL*. Deux barrières existent. La première au niveau du site de constriction. La deuxième est formée par les hélices N-terminales dans le cytoplasme. Lors de l'ouverture du canal, la protéine passe par des étapes intermédiaires dans lesquelles la barrière membranaire est ouverte (état intermédiaire 5) puis les hélices N-terminales sont séparées (état intermédiaire 9). L'inclinaison des hélices permet l'ouverture totale de la protéine (Sukharev *et al.*, 2001 b).

i. Les hélices transmembranaires

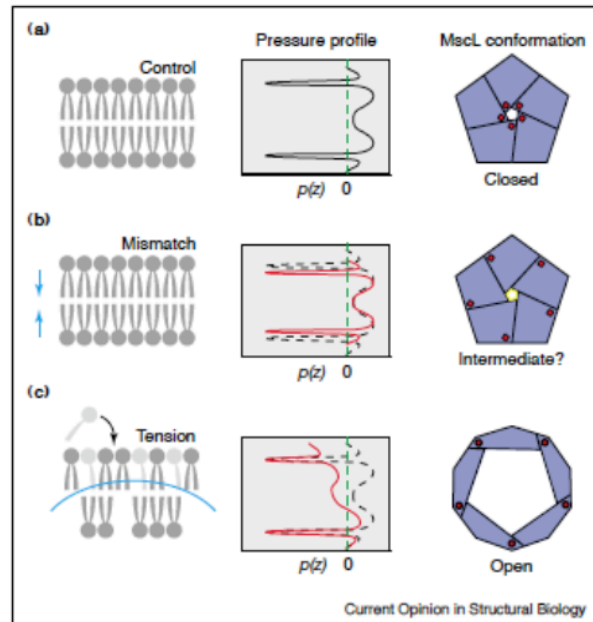
Pour trouver des conditions dans lesquelles la conformation ouverte serait plus stable en vésicules, différentes mutations ont été testées. La glycine *G22* a beaucoup intéressé le groupe de *KUNG* (Ou *et al.*, 1998 ; Yoshimura *et al.*, 1999 ; Yoshimura *et al.*, 2001). Lorsque celle-ci est remplacée par un résidu hydrophile, le canal s'ouvre plus facilement (Bartlett *et al.*, 2006). Ceci a permis d'observer un canal muté ouvert transitoirement, par microscopie électronique (Yoshimura *et al.*, 2008).

Mais les protéines mutées étaient plus difficiles à exprimer, car une activité accrue de *MscL* provoque la mort de la bactérie. Des expériences précédentes ont montré que l'ajout de molécules amphiphiles chargées ou de lysophospholipides dans les vésicules contenant les canaux mécanosensibles diminuait considérablement leur seuil d'activation (Martinac *et al.*, 1990 ; Maingret *et al.*, 2000). En variant la nature des lipides, différentes conformations de *MscL* ont été stabilisées. Ainsi, dans des vésicules contenant des lipides à longue chaîne carbonée (*PC18*) le canal était fermé. La diminution de la longueur des chaînes (*PC14*) engendre des conformations intermédiaires. L'ajout de lysophospholipides (25% *LPC* dans des vésicules de *PC18*) a permis d'obtenir une conformation ouverte plus stable (Figure 3.8). En mesurant les distances entre les résidus des deux hélices *TM1* et *TM2* par RPE, *PEROZO* et ses collaborateurs ont réussi à développer de nouveaux modèles de la protéine, et ont analysé les variations de distance pour élaborer un modèle d'ouverture du canal. Les informations obtenues concernent uniquement les hélices transmembranaires (Perozo *et al.*, 2002 a).

Des analyses comparatives de l'accessibilité des hélices dans les différents états ont été menées, indiquant que les changements dans la structure entre les deux états étaient considérables. Dans l'état ouvert, les segments transmembranaires présentent une plus grande mobilité que dans l'état fermé. En d'autres termes, l'empaquetage des hélices est moins important dans la structure ouverte. Pour les hélices *TM1*, une augmentation de l'exposition à l'eau est notable dans la région qui ferme le pore. Les hélices bougent comme des corps rigides. Lors de la transition entre la conformation fermée et la conformation intermédiaire, les hélices *TM2* ne se déplacent pas. Une inclinaison de 30° est observée pour les hélices *TM1*. Lorsque le canal s'ouvre, les hélices *TM1* poursuivent leur inclinaison tout en se déplaçant dans le plan

pour permettre l'expansion du pore. Les hélices *TM2* s'inclinent dans le sens opposé de 70° . De plus, les analyses des différents mutants ont mis en évidence un site de constriction entre *A20* et *G26* stabilisant la conformation fermée de la protéine (Perozo *et al.*, 2002 b).

Figure 3.8 : Variation de la conformation du canal *MscL* en fonction de la tension exercée au niveau de la membrane (Perozo & Rees, 2003). La protéine est dans sa conformation fermée dans une bicouche stable (a) ; une éventuelle conformation inter-médiaire est stabilisée dans une bicouche contenant des lipides à chaînes d'acides gras plus courtes que dans le précédent cas (b) ; l'ajout de lysophospholipides entraine l'ouverture du canal (c).



Les hélices étaient considérées comme des corps rigides dans le modèle de *SUKHAREV*. Les mouvements de courbure des hélices étaient négligés. Ce qui a amené *C. ETCHEBEST* et ses collaborateurs (Équipe de Bio-informatique Génomique et Moléculaire - *Université Paris-Diderot*) à étudier les déformations subies au niveau des hélices lors de l'ouverture du canal *Ec-MscL* par analyse en modes normaux. Le modèle utilisé était différent de celui de *SUKHAREV*. Il est construit par modélisation par homologie (le modèle de *SUKHAREV* est basé sur des critères de prédiction de structure secondaire et de contraintes de distance entre résidus, obtenues à partir des expériences de mutagenèse). Ce travail suggère l'implication des hélices *TM2* d'abord lors de l'ouverture du canal, en raison du fort taux de résidus hydrophobes. Le site de courbure serait le résidu *F90* d'*Ec-MscL* de l'hélice *TM2* (Valadié *et al.*, 2003). L'ouverture du canal impliquerait des mouvements d'inclinaison des hélices, et une torsion de la structure le long de l'axe du canal. Cette hypothèse a été renforcée par d'autres études *in silico* (Kong *et al.*, 2002 ; Elmore & Dougherty, 2003).

ii. Rôle de la région N-terminale

Le modèle de *SUKHAREV* a proposé que la région N-terminale forme un faisceau d'hélice

dans le prolongement du pore, lorsque le canal était fermé. Cet ensemble constitue une deuxième barrière, qui reste fermée jusqu'à l'expansion totale de la partie transmembranaire de *Ec-MscL* (Sukharev *et al.*, 2001 a).

Toutefois, dans la version révisée de la structure cristallographique de *Mt-MscL*, la région N-terminale a été mieux résolue montrant ainsi une structure en hélice parallèle à la membrane cytoplasmique (Steinbacher *et al.*, 2007). En analysant les résidus de cette partie de *Ec-MscL*, l'équipe de P. BLOUNT (*University of Texas - Southwestern Medical Center*) attribue aux résidus hydrophobes cytoplasmiques un rôle d'ancrage dans la membrane, permettant de définir l'inclinaison des deux hélices transmembranaires (Iscla *et al.*, 2008 ; [figure 3.9](#)). La présence de deux phénylalanines (*F7* et *F10*) conforte cette hypothèse, sachant que ces deux résidus sont conservés au sein de différentes espèces, et que leur mutation ou délétion empêche l'ouverture du canal (Blount *et al.*, 1996 a ; Oakley *et al.*, 1999 ; Sukharev *et al.*, 2001 a).

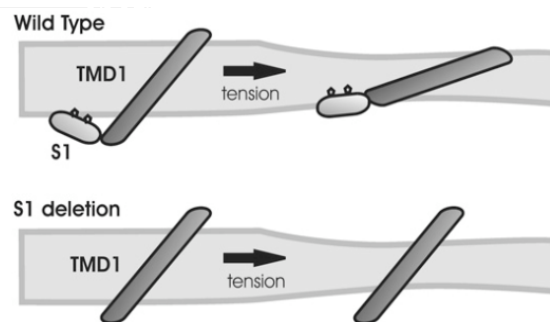


Figure 3.9 : Schéma de l'implication de la partie N-terminale dans le mécanisme d'ouverture. En haut, le segment *S1* contenant des résidus hydrophobes cytoplasmiques (*F7* et *F10* dans *Ec-MscL*), agit comme ancre permettant une inclinaison significative des segments *TM1*, et l'ouverture du canal. La délétion des résidus situés en N-terminal empêche de guider les hélices dans le plan de la membrane, et le canal ne peut s'ouvrir (Schéma et explications issus de Iscla *et al.*, 2008).

iii. Rôle de la région C-terminale

Dans la structure cristallographique de *Mt-MscL*, l'orientation des hélices cytoplasmiques n'était pas correcte à cause d'un *pH* bas pendant la cristallisation. La structure de cette partie au niveau du modèle *Ec-MscL* a été affinée par des études de modélisation et de *cross-linking* (Anishkin *et al.*, 2003). La disposition des hélices lors de l'ouverture du canal leur a attribué le

rôle de filtre bloquant le passage de grosses molécules. La structure révisée de *Mt-MscL* se plaçait en accord avec le modèle d'*Ec-MscL* (Steinbacher *et al.*, 2007 ; [Figure 3.3 b](#)).

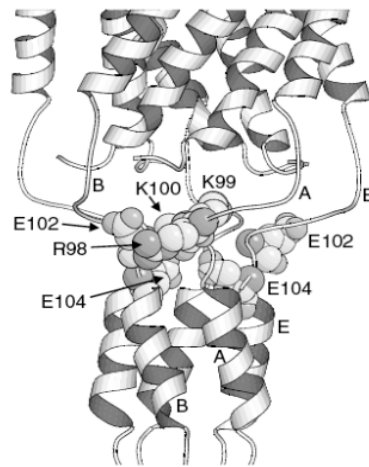


Figure 3.10 : Structure de la partie C-terminale de la protéine *Mt-MscL*. L'emplacement du cluster chargé *RKKxE* est en représentation volumétrique. La figure est issue de Powl *et al.*, 2007.

Les travaux effectués sur la taille du pore et la taille des molécules traversant le canal suggéraient un diamètre de pore du canal ouvert entre 30 et 40 Å (Ajouz *et al.*, 1998 ; Berrier *et al.*, 2000 ; Cruickshank *et al.*, 1997). Ces résultats ont été confirmés par la suite (van den Bogaart *et al.*, 2007). Des travaux de dynamique moléculaire sur le comportement des deux structures de *Mt-MscL* et une analyse par dichroïsme circulaire ont été effectués. Une simulation de 20 ns sur les deux structures montre que l'arrangement des hélices cytoplasmiques est plus stable dans la structure révisée. Le suivi de dénaturation thermique par dichroïsme circulaire confirme ces résultats (Maurer *et al.*, 2008). Les résidus 110-118 de *Mt-MscL* montrent une stabilité particulière et notamment la séquence consensus *LLxEIRD*, mais son rôle dans la stabilité des hélices cytoplasmiques n'est pas élucidé. La délétion d'une partie de la séquence consensus n'altérerait pas le mécanisme d'ouverture, mais diminuait le seuil de sensibilité du canal (Ajouz *et al.*, 2000).

Une autre séquence consensus a été identifiée au sein des *MscL*. Il s'agit du groupe de résidus chargés *RKKxE*, essentiel pour la fonction du canal et qui pourrait agir comme senseur de *pH* (Kloda *et al.*, 2006 ; [Figure 3.10](#)). Toute perturbation de la balance électrique de cette séquence modifierait le mécanisme d'ouverture. Lorsque le *pH* est bas, la protonation de tous

les résidus empêcherait l'ouverture du canal. Les auteurs suggèrent que la présence de ces résidus stabiliserait des conformations intermédiaires lors de l'ouverture du canal. Ces résidus pourraient avoir un effet sur les têtes polaires des phospholipides provoquant des perturbations locales au niveau de la bicouche lipidique pour moduler l'activité du canal. Des expériences d'extinction de fluorescence ont montré des résultats similaires au niveau de *Mt-MscL* (Powl *et al.*, 2005).

iv. Rôle de la boucle périplasmique

La boucle périplasmique des *MscL* joue un rôle important dans l'ouverture du canal. Une coupure de la boucle par des protéases augmente de façon spectaculaire la sensibilité à la pression de *Ec-MscL* (Ajouz *et al.*, 2000). La boucle périplasmique agirait comme un ressort résistant à l'inclinaison des hélices transmembranaires, étape primordiale pour le déclenchement de l'ouverture du canal. Cette proposition a reçu l'appui d'une étude récente montrant que la co-reconstitution des moitiés N et C-terminales de *Ec-MscL* dans des liposomes aboutit à des canaux mécanosensibles présentant une sensibilité accrue à la tension au niveau de la bicouche (Park *et al.*, 2004). Des simulations dans une bicouche courbée mimant le mésappariement hydrophobe ont été effectuées (Meyer *et al.*, 2006). La protéine se trouve dans un système membranaire composé de *DLPE* (Dilauroyl-*PE*) et la courbure est créée en incorporant un lipide à une seule chaîne d'acide gras, le *LPC* (lauroyl-*PC*). La durée des simulations n'a pas dépassé les *10 ns* à cause de la complexité du système. Ce temps était cependant suffisant pour observer des mouvements plus marqués au niveau de la boucle périplasmique qu'au niveau des hélices transmembranaires. Le même phénomène a été observé dans des systèmes composés de *POPE* et *DMPE* (1-palmitoyl-2-oleoyl-*PE* et dimyristoyl-*PE* respectivement) (Debret *et al.*, 2008).

e Structure de *Sa-MscL*

La protéine *MscL* du staphylocoque doré (*Staphylococcus aureus*) a été découverte en 1998 (Moe *et al.*, 1998). La séquence est constituée de 120 acides aminés et présente 40% et 51% d'homologie avec celles de *M. tuberculosis* et *E. coli* respectivement. Le clivage des 26 derniers acides aminés de *Sa-MscL* a engendré une activité accrue. La structure cristallographique de la

protéine clivée a été déterminée (Liu *et al.*, 2009). Un tétramère est observé. Le diamètre global de la structure est de 69 \AA , alors que la structure de *Mt-MscL* montrait un diamètre de 52 \AA (Figure 3.11). Les expériences de *cross-linking* effectuée sur la protéine *Sa-MscL* entière et la protéine clivée montrent le même état oligomérique. Plus récemment, des études de *cross-linking in vivo* ont prouvé que la protéine existe sous forme pentamérique, et présente différents aspects oligomériques *in vitro* (Dorwartt *et al.*, 2010 ; Ghandi *et al.*, 2011).

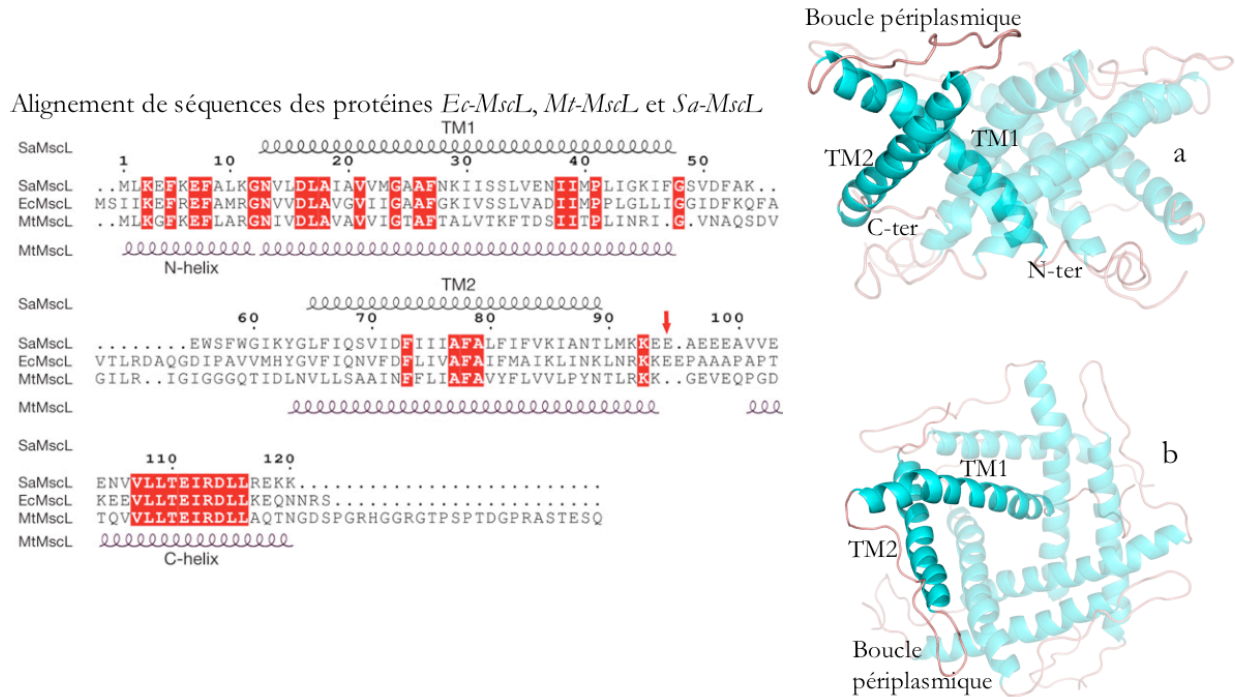


Figure 3.11 : Structure de la protéine mécanosensible *Sa-MscL*. Les séquences des trois protéines *Sa-MscL*, *Ec-MscL* et *Mt-MscL* sont alignées. L'emplacement des hélices des structures expérimentales est indiqué, ainsi que le site de clivage de *Sa-MscL* (flèche rouge). L'alignement montre une conservation de certains résidus clés dans le fonctionnement du canal mécanosensible. La structure obtenue est donc celle de la protéine clivée des 26 résidus en C-terminal. Elle est représentée de face (a) montrant le même agencement que *Mt-MscL*, avec deux hélices transmembranaires séparées par une boucle périplasmique plus courte que celle de *Mt-MscL*. La vue du dessus (b) permet de voir le pore formé par les quatre hélices transmembranaires *TM1* (code PDB : *3hzg*). L'alignement est extrait de Liu *et al.*, 2009.

La boucle périplasmique et l'hélice *TM2* sont plus courtes que celles de *Mt-MscL*. La partie N-terminale n'adopte pas d'arrangement secondaire particulier. Les hélices transmembranaires sont orientées différemment, formant un pore plus élargi que celui de *Mt-MscL*. Le diamètre du site de constriction est de $6 \text{ à } 8 \text{ \AA}$, ce qui est insuffisant pour laisser passer les molécules d'eau. Il s'agit probablement d'une conformation intermédiaire. La structure clivée montre une stabilité accrue des conformations ouvertes lorsque la protéine est reconstituée en liposomes.

f Interactions *MscL*-lipides

Le rôle de la membrane dans le mécanisme d'ouverture des canaux *MS* est primordial. Les différentes études menées sur les interactions entre *MscL* et les lipides environnants ont montré que des changements d'épaisseur de la membrane (Perozo *et al.*, 2002 a), sa fluidité et sa courbure (Kamaraju & Sukharev, 2008, Meyer *et al.*, 2006 ; Nguyen *et al.*, 2005) ou la charge des têtes polaires (Powl *et al.*, 2008 b) pouvaient provoquer l'ouverture du canal ou au moins stabiliser des conformations intermédiaires.

La structure de la protéine à l'état ouvert reste néanmoins inconnue. Diverses méthodes tant biochimiques qu'informatiques ont permis de mieux comprendre le mécanisme d'ouverture, mais plusieurs mystères demeurent quant à l'implication des parties N- et C-terminales, au rôle joué par la boucle périplasmique, ou aux interactions entre les hélices transmembranaires. L'interaction de la protéine avec la membrane se fait à plusieurs niveaux de la protéine, et le rôle joué par l'environnement lipidique fait encore l'objet de différentes hypothèses.

Cependant le maintien de la protéine dans son environnement natif est indispensable pour la garder active et pour étudier ses fonctions. La conformation ouverte est probablement un état métastable du fait qu'une ouverture prolongée du canal est létale pour la bactérie. Garder le canal ouvert pendant la cristallogénèse paraît très difficile. La résonance magnétique nucléaire en phase solide est un bon compromis. Elle permettrait d'obtenir des informations structurales à haute résolution. Elle est adaptée à l'étude de systèmes statiques ou dynamiques. De plus la technique permet l'étude d'un système aussi complexe qu'une protéine membranaire en bicouche lipidique grâce à la technique de rotation à l'angle magique. L'état de l'art actuel de la technique permet d'observer les protons dans des conditions particulières. Mais l'observation des atomes de carbone et d'azote reste plus simple. Et malgré un rapport gyromagnétique plus faible, la technique de polarisation croisée permet d'accroître le signal de ces atomes. Tous ces points seront abordés dans le chapitre suivant concernant l'étude structurale des protéines membranaires.

IV L'utilisation de la résonance magnétique nucléaire en phase solide pour étudier *MscL*

a Structure des protéines

Il existe une large panoplie de méthodes permettant d'étudier les caractéristiques structurales et fonctionnelles d'une protéine. Ainsi, le clonage moléculaire, la surexpression, la mutagenèse dirigée, et diverses approches biochimiques et biophysiques sont indispensables dans la détermination de la structure et la fonction des protéines. Les questions les plus ouvertes dans le domaine des protéines membranaires concernent la compréhension de l'insertion dans la membrane, l'organisation des protéines dans celle-ci mais aussi les mécanismes au travers desquels une protéine membranaire remplit sa fonction au sein de la bicouche lipidique.

La structure primaire de la protéine est la séquence de l'enchaînement des acides aminés. Elle est établie par le gène codant pour la protéine. La détermination de la structure primaire d'une protéine permet naturellement de calculer précisément sa masse moléculaire, et renseigne également sur ses propriétés physico-chimiques, tels le point isoélectrique (Sillero & Ribeiro, 1989), le coefficient d'extinction (Gill & von Hippel, 1989), l'indice aliphatique (Ikai, 1980) ou le profil d'hydrophatie (Kyte & Doolittle, 1982). Pour les protéines membranaires, le profil d'hydrophatie permet également d'avoir une idée sur les segments transmembranaires éventuels (Zhao & London, 2006).

La structure secondaire représente les repliements locaux adoptés par des segments de la protéine, en fonction de leur composition en acides aminés. Le profil d'hydrophatie et les liaisons hydrogènes établies entre les résidus permettent de définir divers arrangements secondaires périodiques: les hélices, les feuilletts et les boucles. Lorsque la conformation locale ne correspond à aucune de ces structures secondaires, il s'agit d'une pelote statistique non-périodique. Des méthodes spectroscopiques permettent de déterminer les arrangements secondaires d'une séquence polypeptidique (Rost & Sander,

1993 ; Surewicz *et al.*, 1993, Wishart *et al.*, 1992). Il existe également des méthodes théoriques pour la détermination de la structuration secondaire d'une protéine à partir de sa composition en acides aminés (Raghava, 2002 ; Rost *et al.*, 2004).

La structure tertiaire ou tridimensionnelle représente l'arrangement spatial des différents segments de la protéine. Elle est stabilisée par des liaisons faibles de type liaisons hydrogènes ou liaisons hydrophobes, par des liaisons covalentes mettant en jeu deux cystéines formant un pont disulfure, ou par des liaisons électrostatiques. Elle dépend fortement de l'environnement de la protéine. Pour une protéine membranaire, la structure tridimensionnelle est conditionnée par la bicouche lipidique.

La structure quaternaire représente l'assemblage de structures tertiaires effectué par les protéines pour devenir fonctionnelle. Elle met en jeu des interactions de faible énergie. La figure 4.1 illustre les différents états d'organisation dans le cas de la protéine *Sa-MscL*.

Séquence de la protéine *Sa-MscL* (Holden *et al.*, 2004)

MLKEFKEFAL KGNVLDLAI A VVMGAAPFKI ISSLVENIIM
 PLIGKIFGSV DPAKEWSPWG IKYGLFIQSV IDFIIFAFAL
 FIFVKIANTL MKKEEAEEEA VVEENVVLLT EIRDLLREKK

Prédiction de la structure secondaire de la protéine par l'algorithme *PSIPRED* (Jones, 1999)

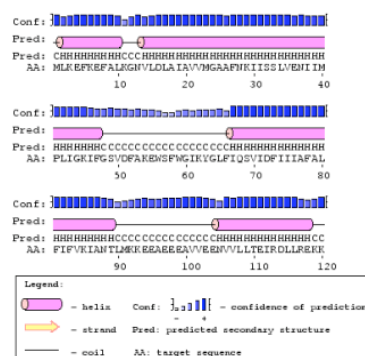
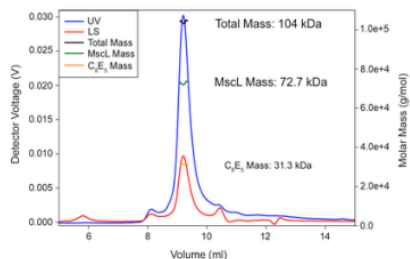
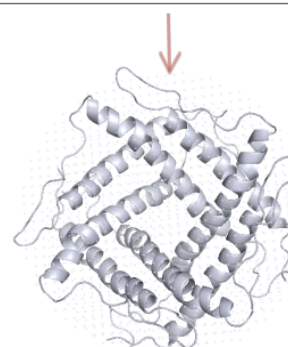


Figure 4.1 : De la séquence à la structure *in vivo* - cas de la protéine *Sa-MscL*. En se basant sur la séquence de la protéine, différents algorithmes permettent actuellement de déterminer la structure secondaire des protéines. La structure tertiaire a été obtenue en 2009 et montrait un tétramère. Une récente étude utilisant différentes techniques révèle une organisation pentamérique *in vivo*, et une organisation tétramérique lorsque la protéine est solubilisée en micelles de détergent.



Structure quaternaire de la protéine déterminée par diverses techniques (centrifugation analytique, SEC, cross-linking, (Dorwart *et al.*, 2010)



Structure tridimensionnelle tétramérique de la protéine déterminée par radiocristallographie (Liu *et al.*, 2009)

b Détermination de la structure tridimensionnelle d'une protéine

Il existe actuellement 72 104 structures de protéines dans la PDB, 87% ont été obtenues par diffraction des rayons X, 12% par résonance magnétique nucléaire. 80% de ces protéines possèdent un poids moléculaire inférieur à 98 kDa, et 12% entre 98 et 196 kDa. La première structure de protéine obtenue fut celle de la myoglobine de *Physeter macrocephalus* (Kendrew *et al.* 1960). La première obtenue par RMN fut celle de l'inhibiteur de la trypsine du pancréas de *Bos taurus* BPTI en 1984 (Williamson *et al.*, 1985). La première structure cristallographique à haute résolution d'une protéine membranaire date de 1985 (Deisenhofer *et al.*, 1984). Il s'agit du centre réactionnel de photosynthèse de *B. viridis*. La structure de la Gramicidine est la première d'un peptide membranaire déterminée par RMN (Ketchum *et al.*, 1993). Depuis 1985, seules 684 structures de protéines membranaires ont été obtenues, dont 53 par RMN. Une base de données contenant toutes les structures de PMs est maintenue par le laboratoire de S. WHITE (*University of California – Irvine* – http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Cette base de données met l'accent sur les structures déterminées par des méthodes de diffraction, bien que certaines structures RMN sont incluses. Une liste des structures de PMs déterminées par RMN est disponible sur une page maintenue par D. WARSCHAWSKI (*Institut de Biologie et de Physico-Chimie* - <http://www.drorlist.com/nmr/MPNMR.html>).

La cristallographie repose sur l'analyse de données de diffraction des rayons X par les cristaux. Son utilisation n'est pas limitée par la taille de la molécule à étudier. En revanche, elle requiert l'obtention de cristaux de bonne qualité présentant un pouvoir diffractant élevé. L'étape de cristallisation constitue souvent une limitation importante à l'utilisation de la technique. La cristallogénèse des protéines reste principalement empirique et se fait par une méthode d'essais-erreurs, où des milliers de conditions expérimentales sont testées. L'automatisation de cette étape a permis de réduire considérablement la quantité d'échantillons nécessaires à chaque essai. La majorité des structures de protéines actuellement résolues par cristallographie aux rayons X correspond à un état statique, peu représentatif de l'ensemble des conformations adoptées en réalité par une protéine en action.

L'utilisation de détergent rajoute une difficulté à la cristallographie des protéines membranaires à cause des caractéristiques physico-chimiques de ces molécules concernant leur concentration micellaire critique et leur agrégation. Néanmoins, la cristallogénèse des protéines membranaires en présence de détergent reste la technique la plus couramment utilisée (Newstead *et al.*, 2008). Une autre difficulté concerne les protéines dont la structure ou la fonction dépendent fortement de la présence de molécules de lipides (Hunte et Richers, 2008 ; Lee, 2004). La cristallisation des protéines membranaires dans un environnement lipidique permet de garder la protéine dans son environnement natif. Les méthodes consistent à placer les protéines dans un environnement lipidique ordonné tel une bicelle, ou une phase cubique (Pebay-Peyroula *et al.*, 1997 ; Ujwal *et al.*, 2008). Mais cette méthode n'a pas encore rencontré beaucoup de succès.

c La résonance magnétique nucléaire

i. Principe de la méthode

La RMN est basée sur l'existence d'un paramagnétisme nucléaire. Seuls les noyaux possédant un nombre quantique de spin I non nul peuvent être étudiés en RMN (voir le [tableau 3](#)). Le phénomène fut découvert en 1946 simultanément à *Harvard* par PURCELL et ses collaborateurs et à *Stanford* par BLOCH et ses collaborateurs. L'approche de BLOCH utilisait les lois de l'induction et donc un formalisme de type vectoriel classique (Bloch *et al.*, 1946). Elle considère que les particules nucléaires se comportent comme des petits aimants lorsqu'ils sont placés dans un champ magnétique intense et contribuent ainsi à une aimantation nucléaire macroscopique. L'approche développée par PURCELL s'appuyait sur l'absorption d'une radiation électromagnétique en utilisant la mécanique quantique et les transferts d'énergie d'un niveau énergétique à l'autre (Purcell *et al.*, 1946). Elle fait appel au spin nucléaire défini comme étant le moment magnétique intrinsèque. Les deux approches furent ensuite unifiées. Ceci explique néanmoins pourquoi deux formalismes sont utilisés en RMN.

Les particules élémentaires (électron, proton, neutron) sont dotées d'un mouvement de spin, qui dans certaines conditions génère un moment magnétique. Les moments magnétiques nucléaires sont responsables du paramagnétisme nucléaire. Le moment magnétique nucléaire $\vec{\mu}$ dépend du moment angulaire de rotation \vec{L} du noyau étudié selon la relation :

$$\vec{\mu} = \gamma \cdot \vec{L}$$

où γ est le rapport gyromagnétique du noyau. Le moment magnétique caractérise l'importance du dipôle nucléaire. Un noyau est observable par RMN s'il présente des propriétés magnétiques caractérisées par l'existence d'un nombre quantique de spin I non nul. Ce nombre de spin est une grandeur quantifiée, positive ou négative et d'incrément $1/2$. Le moment magnétique est relié au nombre quantique de spin par l'équation :

$$\vec{\mu} = \gamma \cdot \hbar \cdot \vec{I}$$

\hbar étant la constante de *PLANCK* réduite $\hbar = h/2\pi$.

Le rapport gyromagnétique γ dépend de la charge et de la masse du noyau. Quand I est égal à $1/2$ (cas de l'hydrogène 1H), le moment magnétique est dipolaire, et les noyaux se comportent comme des sphères uniformément chargées. Concernant les molécules biologiques, la RMN est utilisée dans le cas des noyaux suivants : 1H , 2H , ^{13}C , ^{14}N , ^{15}N , ^{17}O , ^{19}F , et ^{31}P . Les plus couramment étudiés lorsqu'il s'agit de protéines sont 1H , 2H , ^{13}C et ^{15}N . Les isotopes 1H , ^{14}N , ^{19}F , et ^{31}P sont naturellement abondants, cependant l'azote ^{14}N est peu utilisé à cause de son faible rapport gyromagnétique. L'utilisation des isotopes 2H , ^{13}C , ^{15}N et ^{17}O nécessite des méthodes de détection très puissantes ou un enrichissement isotopique de la molécule étudiée. Les isotopes 1H , ^{13}C , ^{15}N , ^{19}F , et ^{31}P possèdent un spin égal à $1/2$, ils ont donc deux niveaux d'énergie lorsqu'ils sont placés dans un champ magnétique statique intense. Les isotopes 2H et ^{14}N possèdent un spin égal à 1 et donc trois niveaux d'énergie. Le noyau ^{17}O a un spin de $5/2$ et possède six niveaux d'énergie.

La description quantique du phénomène de *RMN* permet donc de le comprendre sous l'angle des échanges d'énergie. Ainsi en présence d'un champ magnétique B_0 uniforme, les spins prennent $(2I + 1)$ états quantiques ou niveaux énergétiques. Ce phénomène physique est l'effet *ZEEMAN*. Pour un noyau de spin $1/2$, seuls deux niveaux d'énergie existent. La différence d'énergie entre les deux niveaux est de valeur :

$$\Delta E = E_\beta - E_\alpha = \gamma \cdot \hbar \cdot \vec{B}_0$$

E_β et E_α étant les énergies des deux états quantiques β et α . La distribution de spins est telle que, sur le niveau α , il existe un léger excédent de spins par rapport au niveau β . La statistique de *BOLTZMANN* permet de calculer la population des moments magnétiques. La répartition des spins répond à l'équation :

$$\frac{N_\beta}{N_\alpha} = \exp\left(\frac{\Delta E}{k \cdot T}\right)$$

$\frac{N_\beta}{N_\alpha} < 1$ (par définition). T est la température absolue et k la constante de *BOLTZMANN*. La différence de peuplement des niveaux d'énergies β et α augmente lorsque la différence d'énergie augmente, c'est-à-dire lorsque B_0 augmente et T diminue. L'écart entre les deux énergies est très faible à des températures et champs usuels. C'est pourtant sur cette différence d'énergie très faible que repose la spectroscopie *RMN*. La fréquence de l'onde réalisant la transition entre les deux niveaux est appelée fréquence de *LARMOR*. C'est cette grandeur qui permet l'exploitation de la technique.

Un échantillon placé dans un champ magnétique très puissant est soumis à différentes interactions. La plus importante est celle des moments magnétiques nucléaires avec le champ B_0 . Elle provoque l'effet *ZEEMAN* (*Zeeman et al.*, 1897). L'interaction quadripolaire est la plus importante après l'effet *ZEEMAN*, mais elle est nulle dans le cas d'environnements locaux de symétrie sphérique. Les noyaux sont entourés d'un nuage électronique. Le mouvement des électrons influence le champ statique et y fait écran. Le champ perçu par le spin nucléaire n'est donc pas exactement B_0 , mais un champ modifié par l'environnement électronique du noyau.

Le champ ressenti localement B_{eff} s'écrit :

$$B_{eff} = B_0 - \sigma B_0$$

où σ est la constante d'écran du noyau étudié.

L'observation des fréquences de résonance fournit un outil de caractérisation chimique très puissant, mais l'inconvénient majeur est que la fréquence dépend de la valeur du champ B_0 . L'écart relatif de la fréquence observée par rapport à la fréquence d'un produit de référence permet d'obtenir une échelle indépendante de la valeur de B_0 . Il s'agit de la notion de déplacement chimique définie par :

$$\delta = 10^6 \frac{\nu - \nu_{ref}}{\nu}$$

Il s'agit d'un terme diamagnétique positif, émanant de la circulation des charges autour du noyau. Le champ créé par les électrons s'oppose au champ externe. Le champ effectif est donc inférieur au champ statique, il s'agit du phénomène de blindage. Le champ électronique n'étant pas homogène autour du noyau, cela provoque l'anisotropie du déplacement chimique. La description du déplacement chimique contient dans ce cas deux composantes, l'une isotrope, l'autre anisotrope. En phase liquide, la deuxième est nulle. Cette valeur est fonction de $3\cos^2\theta - 1$. Les variations d'anisotropie du déplacement chimique sont moyennées lorsque le terme $3\cos^2\theta - 1$ est égal à zéro (θ étant l'angle formé entre le spin et le champ magnétique B_0).

Les spins nucléaires créent un champ magnétique dipolaire autour d'eux. Ce champ affecte les spins environnants. C'est le couplage dipolaire. Les interactions dipolaires proviennent donc du couplage dans l'espace des moments magnétiques de deux noyaux. L'énergie d'interaction dipolaire entre deux noyaux i et j est définie par la relation :

$$E_D \propto \left(\frac{\mu}{r_{ij}^3}\right)(3\cos^2\theta - 1)$$

μ étant le moment magnétique, r_{ij}^3 la distance entre les deux noyaux i et j ; et θ l'angle entre le vecteur $|ij|$ et le champ B_0 . Par conséquent, en phase liquide, comme θ varie

constamment et rapidement puisque l'orientation des molécules varie au cours du temps, l'ensemble des couplages dipolaires est moyenné à zéro. Mais en phase solide, θ est fixe et les couplages dipolaires deviennent les interactions principales. Les spectres sont donc obtenus avec des signaux très larges. Pour éliminer le couplage dipolaire, il suffit que $3\cos^2\theta - 1$ soit égal à zéro.

Lorsque les moments magnétiques de deux noyaux sont couplés à travers leurs électrons, il s'agit alors d'un couplage indirect spin-spin appelé couplage J . Cette interaction n'est présente qu'à travers une liaison chimique. Les énergies mises en jeu sont généralement beaucoup plus faibles que pour les autres interactions. Il peut s'agir d'un couplage homonucléaire lorsque les deux noyaux liés sont de même nature, ou d'un couplage hétéronucléaire lorsqu'il s'agit de deux noyaux différents. Dans une solution isotrope, seuls les déplacements chimiques et les couplages indirects (couplages scalaires J) ne sont pas moyennés et sont représentés sur le spectre. En milieu anisotrope, en plus des déplacements chimiques et des couplages indirects, les interactions dipolaires, les différences de susceptibilité magnétique et l'anisotropie de déplacement chimique sont présentes et ne sont pas moyennées (les interactions quadripolaires aussi dans le cas d'un spin supérieur à $1/2$). Chacun de ces phénomènes étant à l'origine d'une augmentation spécifique de la largeur du signal, l'élimination de l'effet d'un phénomène conduit à une

Isotope	Spin I	Abondance naturelle [%]	Rapport gyromagnétique (γ) [$10^7 \cdot \text{rad}/(\text{T} \cdot \text{s})$]
^1H	$1/2$	99,98	26,7519
^2H	1	0,016	4,1066
^{12}C	0	98,9	--
^{13}C	$1/2$	1,108	6,7283
^{14}N	1	99,63	1,9338
^{15}N	$1/2$	0,37	-2,712
^{16}O	0	98,9	--
^{17}O	$5/2$	0,037	-3,6279
^{31}P	$1/2$	100	10,841

Tableau 3 : Caractéristiques des noyaux usuels des molécules biologiques.

amélioration de la résolution. Pour éliminer l'anisotropie du déplacement chimique et le couplage dipolaire il suffit que le terme $3\cos^2\theta - 1$ soit égal à zéro. Pour cela, la valeur de l'angle θ doit être égale à $54,7^\circ$. Cette valeur de l'angle θ est appelée angle magique. La rotation à l'angle magique (*MAGIC ANGLE SPINNING MAS*) permet de réintroduire un mouvement de rotation artificiellement. Ce mouvement permet de rétrécir la largeur de raie en moyennant les interactions moléculaires à zéro (Andrew *et al.*, 1959 ; Lowe, 1959). L'utilisation de séquences de découplage de champs magnétiques intenses permet de moyennner à zéro l'anisotropie des interactions sur la période de rotation *MAS* et d'accéder à des informations à haute résolution sur des échantillons solides non orientés (Bloch, 1958).

Des progrès importants ont été réalisés en *RMN* des solides biologiques. La performance des spectromètres a été accrue. De très hauts champs sont utilisés (>20 *Teslas*). Des sondes sont fabriquées pour permettre une rotation *MAS* de 70 *kHz*. Ceci n'est pas encore suffisant pour obtenir systématiquement une haute résolution au niveau des protons. C'est la raison pour laquelle seuls les isotopes ^{13}C et ^{15}N sont étudiés. Or ces isotopes sont très peu abondants à l'état naturel. Ils sont présents uniquement à $1,1\%$ et $0,37\%$ respectivement. L'observation dans de bonnes conditions de sensibilité des isotopes ^{13}C et ^{15}N est facilitée par deux méthodes : l'enrichissement isotopique qui consiste à remplacer dans la molécule étudiée les isotopes à spin nul par des isotopes observables en *RMN*. Ceci sera discuté dans le chapitre sur le marquage isotopique. L'autre méthode est appelée polarisation croisée ou *CP* (pour *CROSS POLARIZATION*). Elle consiste en un transfert de l'aimantation d'un noyau ayant un rapport gyromagnétique élevé (^1H) vers un noyau de plus faible rapport gyromagnétique (^{13}C ou ^{15}N). Un transfert parfait peut améliorer les signaux du ^{13}C par un facteur de $\gamma_{\text{H}}/\gamma_{\text{C}} = 4$ et ceux du ^{15}N d'un facteur $\gamma_{\text{H}}/\gamma_{\text{N}} = 10$ (Pines *et al.*, 1972 ; Schaefer & Stejskal, 1975). De plus, dans les solides le temps de relaxation des isotopes ^{13}C et ^{15}N est plus grand que celui du proton. La *CP* permet de réduire le temps de répétition au temps de relaxation des protons. La polarisation est transférée du système *I* (le proton) vers le système *S* (le carbone ou l'azote) en irradiant les deux types de spins avec des champs de radiofréquences obéissant à la condition de *HARTMANN* et *HAHN* (Hartmann & Hahn, 1962) :

$$\gamma_S \cdot B_{IS} = \gamma_I \cdot B_{II}$$

Ainsi la technique *CP-MAS* permet de s'affranchir des problèmes liés à l'anisotropie du système solide afin d'étudier des échantillons non orientés dans le champ magnétique. Dans ces conditions, l'attribution des fréquences de résonances, la détermination d'éléments de structures secondaires et tertiaires est réalisable par l'intermédiaire d'expériences multidimensionnelles variées homo ou hétéronucléaires.

Les informations anisotropes moyennées lors de la rotation à l'angle magique contiennent néanmoins des informations importantes. Pour réintroduire sélectivement ces informations dans les spectres, il existe des nombreux moyens de recouplage des interactions. Beaucoup s'appuient sur une synchronisation d'impulsions radiofréquence avec la rotation de l'échantillon. Les techniques concernent notamment la réintroduction des couplages dipolaires, et de ce fait, sont appelées techniques de recouplage dipolaire (Griffin, 1998). Parmi ces techniques, la méthode *DARR* (*¹³C - ¹H DIPOLAR ASSISTED ROTATIONAL RESONANCE*) est une expérience permettant d'observer des corrélations dipolaires entre carbones marqués au ¹³C. Elle utilise l'interaction dipolaire ¹³C - ¹H pour diminuer le temps de l'expérience (le temps de relaxation du ¹H est plus court que celui du ¹³C) et augmenter la sensibilité de résonance et d'interaction entre carbones voisins des protons (la polarisation croisée permet d'augmenter le signal du ¹³C). La méthode *DARR* permet donc d'étudier les interactions entre les carbones proches dans l'espace. La séquence permet des temps de mélange très variables entre les carbones, ce qui est indispensable en biologie structurale des protéines pour observer les corrélations entre les acides aminés à différentes distances (Takegoshi *et al.*, 2001).

ii. La RMN pour l'étude des protéines membranaires en solution

L'apparition de la *RMN* en biologie structurale remonte au milieu des années 1980. Cette technique permet d'accéder à des informations de distances entre paires d'atomes au sein d'une molécule en solution. Ces informations sont utilisées pour construire un modèle tridimensionnel de cette molécule. En solution, au-delà de quelques dizaines de

kDa, les méthodes traditionnelles utilisées s'avèrent inefficaces. Des progrès significatifs tant techniques que méthodologiques ont été réalisés et ont permis de repousser la limite de la taille de la molécule. Au niveau des protéines, la *RMN* permet d'analyser différents aspects structurels et fonctionnels. Outre la détermination des structures tridimensionnelles, elle renseigne sur les interactions protéine-protéine, protéine-ligand, protéine-acide nucléique, et les cinétiques de ces interactions. La dynamique des macromolécules sur une large gamme de temps allant de la picoseconde (mesure de relaxation) à l'heure (cinétique d'échange de protons).

La *RMN* offre la possibilité d'étudier les protéines membranaires dans divers milieux (Figure 4.2). Après la surexpression et le marquage isotopique de la protéine, les membranes peuvent être étudiées en *RMN* en phase solide. Les membranes peuvent être solubilisées, et la protéine membranaire purifiée. À cette étape, la protéine est maintenue en solution par diverses molécules, dans un environnement imitant la membrane, préservant ainsi l'intégrité et la stabilité de la protéine membranaire en dehors de la bicouche lipidique.

La méthode la plus couramment utilisée pour maintenir les protéines membranaires en solution sont des micelles de détergent (Privé, 2007). Celles-ci permettent de mimer l'environnement hydrophobe des chaînes hydrophobes des acides gras, dans les membranes biologiques (Helenius & Simons, 1975 ; Tanford & Reynolds, 1976). Les complexes protéines membranaires/détergent constituent actuellement les organisations les plus étudiées par la *RMN* en phase liquide (Kim *et al.*, 2009 ; Sanders & Sonnischen, 2006). Les détergents sont des molécules amphiphiles composées d'une queue hydrophobe et d'une tête hydrophile. À des concentrations inférieures à leur concentration micellaire critique (*CMC*), les molécules de détergent existent en tant que monomères en solution aqueuse. Au-delà de la *CMC*, des micelles se forment pour enfouir les queues hydrophobes. Les molécules composant les micelles sont en équilibre dynamique avec les monomères de détergent. En présence d'une membrane lipidique et à des concentrations inférieures à la *CMC*, les molécules de détergent s'incorporent dans la bicouche, la perturbent sans la déstructurer. En augmentant la concentration, les

molécules de détergent continuent à s'incorporer jusqu'à la saturation. Au-delà de la *CMC*, la bicouche n'est plus stable. Le détergent solubilise les lipides, et à ce moment-là différentes organisations sont présentes en solution : des micelles de détergent, des micelles binaires détergent/lipides, et des micelles ternaires détergent/lipides/protéines membranaires. L'addition ultérieure de détergent ne fait que diminuer la concentration de lipides dans les micelles de détergent, et peut provoquer la délipidation totale des micelles ternaires, et la formation de micelles binaires détergent/protéines membranaires (Le Maire et al., 2000 ; Popot, 2010). La dynamique élevée des micelles de détergent ainsi que l'instabilité des assemblages provoquent parfois des problèmes lors de la manipulation des protéines membranaires. Des détergents ioniques au pouvoir tensioactif élevé comme le *SDS* sont très efficaces pour solubiliser les membranes, mais entraînent un mauvais repliement, une agrégation irréversible de certaines protéines ou la perte d'un cofacteur (Tate, 2010).

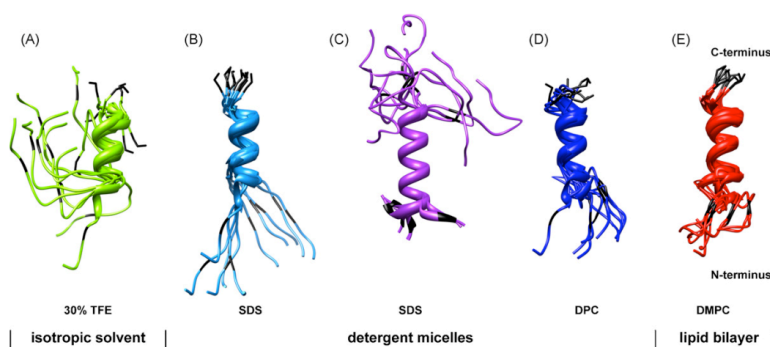


Figure 4.2 : Structures de l'hélice transmembranaire *TM2* de *GlyR* dans différents environnements hydrophobes. La figure est issue de (Bordag & Keller, 2010), et basée sur des travaux de *RMN* (Yushmanov *et al.*, 2003) et de dynamique moléculaire (Johnston *et al.*, 2006).

L'utilisation d'amphipols, de bicelles ou de particules de nanolipoprotéines a été mise au point pour l'étude structurale et fonctionnelle de protéines membranaires en solution aqueuse par *RMN*. Maintenues en amphipols, beaucoup de protéines membranaires conservent leur activité (Gorzelle *et al.*, 2002 ; Pocanschi *et al.*, 2006). La structure tridimensionnelle des protéines de la membrane externe *OmpA* et *OmpX* d'*E. coli* maintenues en amphipols a ainsi été étudiée (Catoire *et al.*, 2010 a; Zoonens *et al.*, 2005). Récemment, le maintien d'un récepteur couplé aux protéines G en amphipols a

permis d'étudier son interaction avec un ligand (Catoire *et al.*, 2010 b). Des systèmes plus ressemblants aux bicouches lipidiques existent. Des particules de lipides peuvent ainsi être utilisées pour maintenir l'intégrité des protéines membranaires. Pour la manipulation des lipides en solution et la formation des vésicules structurées, des molécules de détergent ou de lipides à courte chaîne carbonée sont rajoutées pour former les bicelles (pour des détails concernant la structure des bicelles, voir Triba *et al.*, 2006). L'utilisation de bicelles a permis récemment, de maintenir l'intégrine $\alpha IIb-\beta 3$ sous sa forme dimérique active (Lau *et al.*, 2008) et d'étudier le comportement du peptide antimicrobien alaméthicine (Dittmer *et al.*, 2009). Pour les particules de nanolipoprotéines ou nanodisques, ce sont des protéines amphiphiles qui permettent la structuration des vésicules et leur maintien en solution. Ces particules procurent un environnement hydrophobe de taille limitée, réduisant de manière considérable l'interaction entre les protéines (Bayburt & Sligar, 2010 ; Lyukmanova *et al.*, 2008).

iii. La RMN pour l'étude des protéines membranaires in situ

Cependant, le maintien en solution des protéines membranaires dans des systèmes membranaires mimétiques génère des complexes de masse moléculaire effective élevée. D'une part, l'anisotropie en est accrue et la demi-vie du signal RMN diminue, ce qui provoque une augmentation significative de la largeur de raie. En échange, un temps d'acquisition plus élevé est nécessaire pour augmenter le rapport signal/bruit. D'autre part, l'augmentation de la taille du système provoque la diminution d'efficacité des transferts d'aimantation entre les noyaux observés. Pour cela, des astuces de marquage isotopique, des spectromètres à très haut champ ainsi que des nouvelles expériences hétéronucléaires et multidimensionnelles sont nécessaires.

La RMN en phase solide est la technique idéale pour surmonter l'anisotropie et la complexité des systèmes de protéines membranaires. L'utilisation de cette technique pour l'étude de systèmes désordonnés et solides, tels les micro-cristallins, les fibres amyloïdes ou les protéines membranaires, date du début des années 1990. Concernant les systèmes membranaires, il s'agissait surtout de petits peptides insérés dans des bicouches alignées et

statiques (Figure 4.3). L'étude de la *Gramicidine* (Ketchem *et al.*, 1993) montre quelques différences au niveau des angles de torsion par rapport à la structure obtenue en micelles. La *Magainine* (Bechinger *et al.*, 1993) montre une partie fixe parallèle au plan de la membrane avec quelques résidus mobiles. La protéine du capsid d'un bactériophage filamenteux *Pf1* (Shon *et al.*, 1991) a également été étudiée par RMN en phase solide par des méthodes statiques (Opella *et al.*, 2008). La structure des segments transmembranaires M2 du récepteur nicotinique à l'acétylcholine et du récepteur à glutamate de type *NMDA* a été résolue avec des contraintes obtenues en micelles de détergent et d'autres en vésicules lipidiques (Opella *et al.*, 1999). La conformation de la protéine *v* du virus du sida *Vpu* a été également obtenue par des contraintes en détergent et en membrane (Park *et al.*, 2003). L'organisation pentamérique du *Phospholambane* (30 kDa) a aussi été obtenue par une méthode hybride (Traaseth *et al.*, 2009).

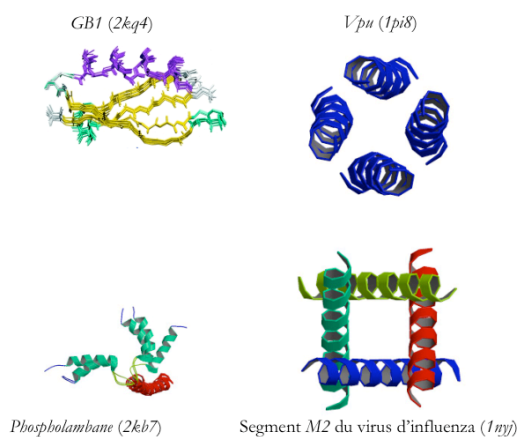


Figure 4.3 : Structures de peptides ou protéines membranaires déterminées par RMN en phase solide.

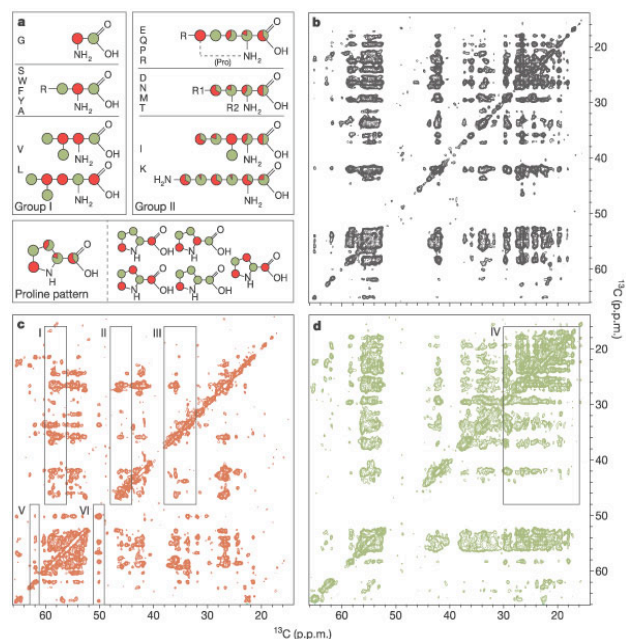
L'orientation à l'angle magique a surtout été appliquée à des protéines à l'état naturellement solide, comme les fibrilles, ou des protéines à un état ordonné micro-cristallin. Le poids moléculaire des systèmes étudiés ne cesse d'augmenter en suivant les améliorations tant en RMN que dans les techniques de production et de marquage de ces protéines. Ainsi, en combinant des images obtenues par microscopie à force atomique (AFM), des clichés de diffraction, ainsi que des corrélations ^{13}C - ^{13}C par RMN en phase solide en MAS, et en utilisant différentes techniques de recouplage, la structure tridimensionnelle du peptide *K3* (22 résidus) sous forme monomérique et tétramérique a

été obtenue. Ce peptide est issu de la digestion de la $\beta 2$ -Microglobuline, composant de l'antigène de type I du complexe majeur d'histocompatibilité, formant des fibres amyloïdes lors de traitement d'hémodialyse prolongé (Iwata *et al.*, 2006). La structure du peptide *K3* montre une organisation en feuillet β avec beaucoup de ressemblance avec la structure tridimensionnelle du protofilament β -amyloïde associé à la maladie d'Alzheimer $A\beta_{1-40}$ (Ferguson *et al.*, 2006 ; Petkova *et al.*, 2006). Le record au niveau des fibrilles est actuellement de 43 kDa. Il s'agit du peptide HET-₅₂₁₈₋₂₈₉ la forme prionique d'une protéine filamenteuse du champignon *Podospora anserina*. L'organisation en solénoïde de feuillet β de ce peptide a été également déterminée en rotation à l'angle magique (Van Melckebeke *et al.*, 2010). Pour les systèmes micro-cristallins, des protéines de taille allant de 1 à 19 kDa ont été étudiées. Des protéines uniformément marquées ont permis de résoudre les structures de l'ubiquitine (9 kDa, Manolikas *et al.*, 2008), l' α -défensine humaine HNP-1 (3 kDa, Zhang *et al.*, 2010) ou encore le dimère de la protéine *Crh* (19 kDa) participant à la répression catabolite dans la bactérie *Bacillus subtilis* (Loquet *et al.*, 2008). Des astuces métaboliques lors du marquage isotopique ont été employées pour les protéines *SH3* (7 kDa ; [figure 4.4](#)) et *GB1* (6 kDa). Au lieu du glucose marqué à l'isotope ^{13}C , la seule source de carbone utilisée pour la surexpression de la protéine était le glycérol $\text{HOCH}_2\text{-CHOH-CH}_2\text{OH}$, avec dans un cas le C^2 en ^{13}C et les carbones C^1 et C^3 en ^{12}C , ou l'inverse.

Le marquage sélectif a permis de résoudre la structure de ces deux protéines (Castellani *et al.*, 2002 ; Franks *et al.*, 2008). Ces travaux ont montré également le rôle que peut jouer le marquage sélectif dans l'augmentation de la résolution au niveau des spectres en RMN en phase solide.

Différents laboratoires ont entrepris l'étude d'échantillons de protéines membranaires dans des bicouches lipidiques en utilisant la rotation à l'angle magique. Ainsi, des spectres montrant la diacyl-glycérol kinase *DGK* d'*E. coli* en micelles de *DM* (n-decyl- β -maltoside) ont été comparés à des spectres de la protéines en liposomes de *POPC* ou *DPPC* à différentes températures (Yamaguchi *et al.*, 2004). Une meilleure résolution a été atteinte lorsque les lipides étaient en phase gel. Sur le même système, les lipides liés à la protéine ont pu être identifiés par l'acquisition de spectres ^{31}P (Lorch *et al.*, 2005).

Figure 4.4 : Spectres obtenus sur différents échantillons de *SH3*. En utilisant du glucose uniformément marqué ou du glycérol marqué au niveau du C^2 , ou au niveau des C^1 et C^3 , H. Oshkinnat et ses collaborateurs ont mis au point une astuce qui leur a permis de résoudre la structure de la protéine *SH3* à l'état micro-cristallin.



L'affinité du transporteur de nucléoside *NupC* et de celui du glucuronide *GusB* d'*E. coli* pour certains de leurs ligands a été suivie en utilisant des récepteurs dans leur environnement natif et des ligands marqués au ^{13}C (Patching *et al.*, 2004). L'interaction du canal à potassium chimérique *KcsA-Kv1.3* (140 kDa) avec la *kaliotoxine* du venin de scorpion *Androctonus mauretanicus*, un peptide de 38 résidus a été étudiée par le laboratoire de M. BALDUS (Bijvoet Center for Biomolecular Research). Dans un premier temps, la structure de la *kaliotoxine* liée au canal a été déterminée (Lange *et al.*, 2005). Dans cette étude, des contraintes ^{13}C - ^{13}C et ^{13}C - ^{15}N ont été obtenues sur la *kaliotoxine* uniformément marquée au ^{13}C et au ^{15}N , alors que la protéine *KcsA-Kv1.3* ne l'était pas. Le changement conformationnel qu'induit la liaison de la toxine au canal a été déterminé en comparant le spectre de la toxine marquée lorsqu'elle est seule, à l'état solide, et lorsqu'elle est liée au canal. La comparaison de spectres de corrélations ^{13}C - ^{13}C de la protéine *KcsA-Kv1.3* uniformément marquée au ^{13}C et au ^{15}N avec ou sans le ligand non marqué a montré des déplacements chimiques différents de certains résidus (Lange *et al.*, 2006) révélant ainsi le site de liaison. L'attribution de la moitié des résidus et le suivi d'inactivation de la protéine par le *pH* ont permis de proposer un modèle d'activation de la protéine par la kaliotoxine, et les réarrangements structuraux des segments transmembranaires que provoque la fixation du ligand (Ader *et al.*, 2008 ; Schneider *et al.*, 2008).

D'autres systèmes membranaires ont été étudiés *in situ* par MAS. Le laboratoire de

H. de GROOT (*Leiden Institute of Chemistry*) s'intéresse aux protéines des systèmes photosynthétiques. La protéine transmembranaire collectrice de lumière LH2 de la bactérie pourpre *Rhodospseudomonas acidophila* a été étudiée. Les premiers résultats montraient une meilleure résolution à basse température sur des spectres ^{13}C 1D et des spectres 2D ^{13}C - ^{15}N de la protéine uniformément marquée au ^{13}C . La basse température provoque la disparition des pics de corrélations des résidus flexibles adoptant une multitude de conformations (Egorova-Zachernyuk *et al.*, 2001). L'attribution de 80% des résidus de la protéine a été effectuée sur des spectres 2D ^{13}C - ^{13}C de plusieurs échantillons par des marquages sélectifs et extensifs (Hong & Jakes, 1999). L'acide succinique $\text{HOOC-CH}_2\text{-CH}_2\text{-COOH}$ est utilisé comme source de carbone. Cette molécule est un intermédiaire dans les cycles d'incorporation du carbone. Trois différents échantillons de LH2 ont été surexprimés en présence d'acide succinique sous forme uniformément marquée au ^{13}C , marquée aux extrémités, ou marquée aux deux carbones C^2 et C^3 à l'isotope ^{13}C (van Gammeren *et al.*, 2004). Un quatrième marquage est effectué. La protéine est surexprimée en présence d'acide succinique non marqué, et des acides aminés marqués au ^{13}C rajoutés en excès dans le milieu (van Gammeren *et al.*, 2005). Des contraintes de distance intra- et inter-résiduelles dans le même monomère ou entre deux monomères, ainsi que des interactions entre la protéine et son cofacteur *Bchl a* (bactériochlorophylle a) ont été obtenues (Ganapathy *et al.*, 2007). L'analyse des déplacements chimiques obtenus et leur écart par rapport aux valeurs théoriques des acides aminés en hélices α ont permis d'identifier différentes régions au sein de la protéine. Lorsque l'écart est très élevé, les résidus sont impliqués dans des arrangements quaternaires, comme des liaisons H intra- ou inter-monomériques, ou représentent un site d'interaction avec *Bchl* (Pandit *et al.*, 2010).

Parmi d'autres exemples de protéines membranaires étudiées *in situ* par RMN en phase solide, la famille des opsines présente l'avantage d'une grande abondance naturelle. Ainsi, la *rhodopsine* bovine (Etzkorn *et al.*, 2007 ; Patel *et al.*, 2005) et la *bactériorhodopsine* bR de *Halobacterium salinarum* (Saitô & Naito, 2007 ; Varga *et al.*, 2008) ont été étudiées dans leur membrane native par différents groupes. Mais la meilleure résolution a été atteinte sur des spectres 2D ^{13}C - ^{13}C , ^{13}C - ^{15}N pour la *protéorhodopsine* pR (27 kDa) dans le laboratoire de

V. LADIZHANSKY (*University of Guelph*) (Shi *et al.*, 2009). Cette protéine est présente dans des bactéries marines vivant dans la zone euphotique de l'océan. La *pR* leur confère la capacité de photosynthèse. La protéine est surexprimée en *E. coli*. Le laboratoire de C. GLAUBITZ (*Goethe University*) avait produit la protéine uniformément marquée, et malgré des pics isolés de 0,5 à 0,7 ppm de largeur, les chercheurs n'ont pas pu isoler les systèmes de spin (Pfleger *et al.*, 2009). Dans le laboratoire de LADIZHANSKY, un marquage inverse a été effectué. Il s'agit d'introduire dans le milieu de culture où la source de carbone est le glucose marqué au ^{13}C , des acides aminés non marqués. Ainsi, par cette approche, les signaux des résidus *Phe*, *Leu* et *Tyr* ont été enlevés. Cela a permis d'identifier à partir de spectres 2D et 3D, et d'attribuer 40% des résidus de la protéine. L'analyse des déplacements chimiques a permis d'établir la structure secondaire de la protéine majoritairement en hélices α , et d'identifier les résidus participant aux boucles extracellulaire et cytoplasmique. L'état de protonation de deux aspartates et deux glutamates a été suivi dans différentes conditions.

D'autres études de protéines membranaires existent, parmi lesquelles l'étude du transporteur de drogue *EmrE* (Agarwal *et al.*, 2007 ; Lehner *et al.*, 2008), de la protéine procaryote *DsbB* impliquée dans la formation de pont disulfure dans le périplasme (Li *et al.*, 2008 ; Sperling *et al.*, 2010), ou aussi une protéine de la membrane externe d'*E. coli* *OmpG* (Hiller *et al.*, 2008 a). Pour le moment, les seules structures obtenues par MAS sont celles de la protéine transmembranaire M2 du virus grippal *influenza A* en interaction avec l'*amantadine*, (Cady *et al.*, 2010) et d'un petit peptide cationique antibactérien, la *protégrine PG-1* des leucocytes de porc (Mani *et al.*, 2006) interagissant avec la membrane bactérienne pour la déstabiliser. Mis à part des peptides n'excédant pas les 4-8 kDa, le marquage uniforme paraît insuffisant pour avoir des spectres permettant la résolution de structure d'une protéine membranaire *in situ* par la RMN en phase solide. Ce domaine en plein essor attend les développements technologiques au niveau de l'appareillage, mais aussi au niveau des séquences utilisées pour la détection des corrélations. En parallèle, les approches biochimiques utilisées ont permis de considérer l'étape de surexpression et de marquage de la protéine comme facteur clef dans le but d'obtenir des données spectroscopiques à haute résolution.

V La synthèse *in vitro* pour l'étude de MscL *in situ* par RMN en phase solide

a Surexpression et marquages isotopiques des PMs

La sensibilité des expériences RMN croît principalement avec l'intensité du champ magnétique B_0 et la quantité de spins identiques présents dans l'échantillon. Pour un champ magnétique donné, le seul moyen d'accroître la sensibilité des expériences est d'augmenter la concentration de la protéine étudiée. La pureté a un impact décisif sur la qualité des spectres, mais aussi sur la longévité de l'échantillon qui peut se trouver dégradé par des protéases présentes. Peu de protéines sont présentes naturellement en grande quantité. De plus, que ce soit pour la RMN en solution ou en solide, les expériences exploitant le 1H sont de moins en moins suffisantes. Un enrichissement isotopique s'avère nécessaire au niveau des atomes ^{13}C et ^{15}N . Le gène codant la protéine d'intérêt est donc cloné dans un vecteur de surexpression. Ce dernier est introduit dans un système procaryote ou eucaryote adapté. Les cellules hôtes vont alors surexprimer la protéine. Pour marquer celle-ci au ^{13}C et au ^{15}N , la surexpression se déroule dans un milieu où les sources de carbone et d'azote contiennent les isotopes ^{13}C et ^{15}N . Parmi les systèmes qui ont été utilisés dans le développement des techniques de surexpression de protéines, il existe les bactéries *E. coli* et *Lactococcus lactis* (Linares *et al.*, 2010), les levures *Pichia pastoris* et *Saccharomyces Cerevisiae* (Egawa *et al.*, 2009 ; Ferndahl *et al.*, 2010), les baculovirus ou cellules d'insectes (Pozza *et al.*, 2009), ou encore dans des cellules mammifères comme les cellules d'ovaires de hamster CHO ou d'autres cellules infectées par le virus de la forêt de Semliki SFV (Casales *et al.*, 2010). Les systèmes de surexpression eucaryotes sont généralement plus difficiles à manipuler (Eifler *et al.*, 2007). 70% des protéines dont la structure se trouve dans la PDB ont été surexprimées dans *E. coli*. Il s'agit d'un organisme modèle. En effet, son génome a été entièrement séquencé en 1997, et témoigne d'une très grande plasticité. La bactérie accepte l'incorporation de gènes par le biais de plasmides, et sa croissance est rapide. Elle permet le marquage uniforme et le marquage sélectif.

Le but principal du marquage d'une protéine pour son étude par *RMN* est d'augmenter la résolution et la sensibilité pendant l'acquisition des spectres. La deutération des protéines ne sera pas abordée dans ce chapitre. Seul sera abordé le remplacement des isotopes ^{12}C et ^{14}N par ceux à spin $1/2$ ^{13}C et ^{15}N dans les protéines en général. Les cas spécifiques des protéines membranaires et de la *RMN* en phase solide seront ensuite exposés.

b Marquage uniforme d'une protéine membranaire dans *E. coli*

L'incorporation des isotopes dans une protéine recombinante surexprimée dans *E. coli* nécessite la croissance de la bactérie sur une source de carbone contenant uniquement l'isotope à incorporer. Le contrôle strict des sources d'azote et de carbone du milieu de croissance des bactéries permet de produire des protéines doublement marquées de façon uniforme avec des rendements de marquage avoisinant les 100%. Le contrôle total du milieu impose l'utilisation d'un milieu minimum, dont le plus fréquent est le *M9* (Sambrook & Russell, 2001). Des milieux modifiés contenant du $^{15}\text{NH}_4\text{Cl}$ ou du $(^{15}\text{NH}_4)_2\text{SO}_4$ comme source d'azote (Muchmore *et al.*, 1989) et du glucose, du glycérol ou de l'acétate marqués au ^{13}C comme source de carbone (Venters *et al.*, 1991) ont été utilisés pour l'obtention de spectres *2D* et *3D*. Par comparaison avec un milieu riche de type *LB*, le *M9* appauvri en métabolites, nécessite une adaptation, et entraîne par conséquent une diminution de la croissance bactérienne, ainsi qu'une baisse du rendement de production. Pour pallier ces inconvénients, l'une des solutions consiste à enrichir le milieu *M9* avec des vitamines et des oligoéléments qui contiennent peu d'atomes de carbone et d'azote (Jansson *et al.*, 1996). Pour réduire le coût du marquage isotopique, la croissance bactérienne dans un milieu minimum contenant les isotopes naturels ^{12}C et ^{14}N , et l'ajout des sources de ^{13}C et ^{15}N au moment de l'induction n'affectent pas le rendement (Cai *et al.*, 1998). Une autre méthode consiste à laisser la masse bactérienne croître dans un milieu riche, puis avant l'induction, centrifuger les bactéries pour les débarrasser des sources de nutriments, et les remettre dans un milieu minimum contenant les sources de ^{13}C et ^{15}N . Cela permet une croissance plus rapide des bactéries avant l'induction, et un meilleur rendement de marquage (Marley *et al.*, 2001).

Le marquage uniforme a été appliqué sur différentes protéines membranaires. Combiné aux méthodes de spectroscopie multidimensionnelle, il permet l'attribution séquentielle de l'ensemble des acides aminés. Dans le cas de petits peptides membranaires, le marquage uniforme a permis d'obtenir des spectres à une résolution suffisante pour obtenir les contraintes structurales nécessaires pour déterminer la structure tridimensionnelle de ces peptides. C'était le cas pour le segment *M2* du récepteur à l'acétylcholine (Opella *et al.*, 1999), du segment *M2* du virus de l'*influenza* (Wang *et al.*, 2001), du peptide *Vpu* du virus de l'*HIV-1* (Park *et al.*, 2003). D'autres peptides ou protéines membranaires ont été uniformément marqués pour leur étude en micelles de détergent par la RMN en phase liquide. Parmi ces protéines, il y a celles de la membrane externe d'*E. coli* *OmpA* et *OmpX* (Fernandez *et al.*, 2001), ou aussi le canal anionique voltage-dépendant humain *VDAC* (Bayrhuber *et al.*, 2008 ; Hiller *et al.*, 2008 b). Mais pour des systèmes étudiés de haut poids moléculaire, pour réduire le temps d'expérience sans affecter la qualité des spectres, et surtout pour simplifier des spectres complexes, le marquage sélectif s'avère être la technique idéale.

c Marquage sélectif d'une protéine membranaire dans *E. coli*

Les bactéries *E. coli* synthétisent les acides aminés à partir des sources de carbone et d'azote présentes. Cependant, si un acide aminé est présent dans le milieu de synthèse, la bactérie utilise la molécule disponible. Sous des conditions d'aérobie, *E. coli* consomme le glucose comme source de carbone, mais peut également consommer le glycérol, le pyruvate, l'acétate et le succinate. Comme source d'azote, les sels d'ammonium sont employés. Pour le marquage uniforme, le choix de la source de carbone est déterminé par deux facteurs : le coût et le rendement. Le glucose est la molécule la plus utilisée. Il se situe en haut des chaînes de biosynthèse des 20 acides aminés. Les voies métaboliques empruntées sont la glycolyse, la voie des pentoses phosphate, le cycle de l'acide citrique, etc. (Figure 5.1).

L'utilisation d'un milieu contenant du glucose marqué au ^{13}C à 100% aboutirait donc à une protéine uniformément marquée. L'utilisation d'un milieu contenant moins de

glucose marqué au ^{13}C et plus de glucose ^{12}C aboutirait à des spectres simplifiés. De plus le marquage n'est pas aléatoire, et sa prédiction est possible (Schubert *et al.*, 2006). La distribution spécifique des atomes de carbone du glucose dans la voie de glycolyse et celle de l'acide citrique permet de tracer plusieurs schémas de marquages spécifiques. Le glucose marqué à des sites spécifiques ($^{13}\text{C}_1$ -glucose et $^{13}\text{C}_2$ -glucose) a été utilisé pour des protéines étudiées en solution (Lundstrom *et al.*, 2007 ; Teilum *et al.*, 2006). Cela permet de sélectivement marquer les chaînes latérales de certains acides aminés.

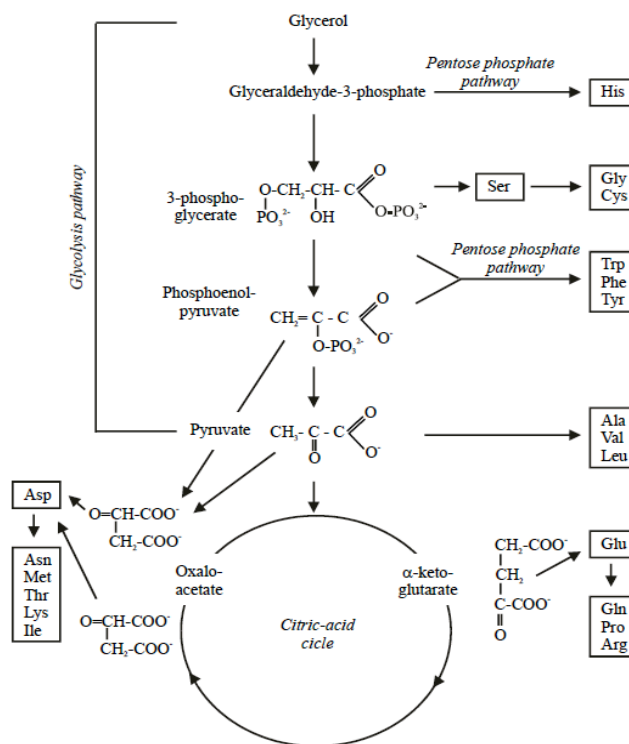


Figure 5.1 : Voies de biosynthèse des acides aminés dans *E. coli*. Les composés précurseurs de chaque acide aminé sont présentés.

L'utilisation du glycérol comme source de carbone permet également le marquage spécifique (LeMaster & Kushlan, 1996). La symétrie de la molécule réduit à deux les formes de glycérol spécifiquement marquées au ^{13}C : le glycérol marqué en C^1 et C^3 : $\text{HO}^{13}\text{CH}_2\text{-}^{12}\text{CHOH-}^{13}\text{CH}_2\text{OH}$ et celui marqué en C^2 $\text{HO}^{12}\text{CH}_2\text{-}^{13}\text{CHOH-}^{12}\text{CH}_2\text{OH}$. La comparaison des spectres obtenus sur deux échantillons synthétisés sur ces deux molécules a permis d'obtenir la structure de *SH3* par RMN en phase solide (Castellani *et al.*, 2002 ; Figure 5.2). De même, la technique a servi à la résolution de la structure de *Crh*

(Loquet *et al.*, 2008) et plus récemment, à l'attribution totale de la protéine *DsbA* en interaction avec la protéine membranaire *DsbB* (Sperling *et al.*, 2010).

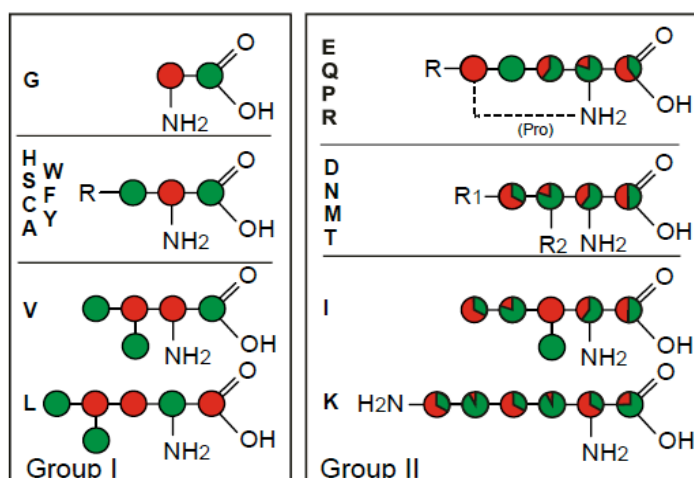


Figure 5.2 : Schémas d'enrichissement isotopique des acides aminés dans la souche d'*E. coli* BL21 (DE3) en présence de glycérol comme source de carbone. Les noyaux représentés en rouge sont ceux qui sont marqués lorsque le glycérol marqué en C² est utilisé; ceux représentés en vert sont ceux qui sont marqués lors de l'utilisation du glycérol marqué en C¹ et C³. Le schéma est extrait du manuscrit de thèse de F. Castellani (Freie Universität Berlin)

Le pyruvate $CH_3-CO-COOH$ est le précurseur métabolique des acides aminés *Ala*, *Val*, *Leu* et *Ile*. L'utilisation de pyruvate ayant le groupement méthyle marqué au ¹³C comme source de carbone aboutit au marquage des groupements méthyles C_β des *Ala*, C_{γ1} et C_{γ2} des *Val*, C_{δ1} et C_{δ2} des *Leu* et C_{γ2} des *Ile* (Lee *et al.*, 1997 ; Ishima *et al.*, 1999).

Les acides acétique et succinique ont été très peu utilisés. Un exemple traitant de l'utilisation de deux formes de l'acide succinique a déjà été exposé dans le chapitre concernant la RMN pour l'étude des protéines membranaires *in situ* (van Gammeren *et al.*, 2004).

Une autre méthode permettant de réduire le recouvrement des pics de corrélation des acides aminés consiste à enlever les signaux de certains résidus (Vuister *et al.*, 1994). En exprimant la protéine dans un milieu minimum contenant du glucose marqué au ¹³C et un ou plusieurs acides aminés où le ¹³C est en abondance naturelle, plusieurs équipes ont pu obtenir des données de meilleure résolution que lorsque les protéines étaient uniformément marquées. Il s'agissait d'enlever les signaux de résidus les plus représentés

dans la protéine (Heise *et al.*, 2005), les résidus hydrophobes (Schneider *et al.*, 2008 ; Shi *et al.*, 2009), ou les résidus possédant de longues chaînes latérales (Seidel *et al.*, 2008). Inversement, l'expression de la protéine sur un milieu contenant du glucose non enrichi en ^{13}C , en présence de certains acides aminés marqués au ^{13}C , permet le marquage sélectif de la protéine au niveau des résidus marqués rajoutés dans le milieu (Fiaux *et al.*, 2004 ; Vosegaard *et al.*, 2008). L'imbrication des réactions métaboliques de biosynthèse des acides aminés et leur réversibilité limitent l'emploi des deux approches. Les réactions de marquages croisés et de transamination sont fréquentes. De plus, un acide aminé présent en excès est utilisé comme source de carbone par la bactérie pour synthétiser d'autres acides aminés. Ainsi, l'incorporation d'un acide aminé en fin d'un cycle de biosynthèse comme les résidus *Leu*, *Val*, *Ile* ou *Lys* est beaucoup plus efficace que l'incorporation de résidus comme les résidus *Ser*, *Asp* ou *Thr* (Lian & Middleton, 2001).

Des techniques hybrides ont été également employées en exploitant les connaissances métaboliques de synthèse des acides aminés. Le domaine transmembranaire de la colicine a été surexprimé en milieu minimum contenant du glycérol marqué en C^2 au ^{13}C , en présence des acides aminés empruntant la même voie métabolique dérivée de l'oxaloacétate $\text{HOOC-CO-CH}_2\text{-CO-COOH}$ (*Asp*, *Asn*, *Met*, *Thr*, *Lys* & *Ile*) ou de l' α -kétoglutarate $\text{HOOC-CO-CH}_2\text{-CH}_2\text{-COOH}$ (*Glu*, *Gln*, *Pro* & *Arg*). Ceci a permis de mesurer des angles de torsion de la chaîne principale du peptide (Hong, 1999), la dynamique des chaînes latérales (Huster *et al.*, 2001) et les changements structuraux entre le peptide en solution et en membrane (Huster *et al.*, 2002).

Les techniques de résolution des structures des *PMs* nécessitent une grande quantité de matériel (1 à 10 mg de protéine pure). Le passage par un milieu minimum pour la surexpression réduit énormément le rendement de la protéine surexprimée et le marquage sélectif n'est pas contrôlé (Paliy & Gunasekera, 2007). Le prix des molécules utilisées pour le marquage isotopique est très élevé. Lorsqu'il s'agit d'une source énergétique, celle-ci est utilisée à un très faible pourcentage pour la protéine cible, et l'enrichissement isotopique se produit à l'échelle des cellules entières. Lorsque des acides aminés enrichis aux isotopes ^{13}C ou ^{15}N sont introduits dans le milieu, ceux-ci sont utilisés en excès. De plus,

l'expérimentateur est confronté à des difficultés spécifiques aux *PMs*, après la surexpression. La solubilisation des membranes, la purification de la protéine cible, la cristallisation ou la reconstitution en membrane sont des étapes au cours desquelles la perte de matériel est inévitable. La probabilité d'inactivation de la protéine augmente lorsqu'elle est gardée en détergent. Les problèmes rencontrés dans les systèmes conventionnels de surexpression *in vivo*, comme la toxicité de la surexpression, les corps d'inclusion, ou la dégradation de la protéine par des protéases bactériennes, sont facilement éliminés par la synthèse *in vitro* (*SIV*).

d La *SIV* pour l'étude des protéines membranaires

Les mécanismes moléculaires derrière la synthèse des protéines sont étudiés depuis la fin des années 1950, lorsque les ribonucléoprotéines ont été identifiées dans les cellules, isolées et leurs propriétés physico-chimiques étudiées (Tissières *et al.*, 1959). Simultanément, la capacité de ces particules à synthétiser des protéines a été démontrée. La découverte des *ARNs* messagers a clarifié le rôle joué par ces particules nommées ribosomes (Jacob & Monod, 1961). Ces connaissances ont permis de développer des systèmes de synthèse *in vitro* des protéines pour pouvoir étudier les ribosomes dans un tube à essai, et de comprendre comment les *ARNs* messagers sont traduits en protéines (Nirenberg & Matthaei, 1961 ; [figure 5.3](#)). Des systèmes issus d'extraits de bactéries, de germes de blé ou de cellules animales ont été utilisés pour tester l'activité d'une protéine ou estimer la toxicité de produits synthétisés (Cohen, 1969 ; Marcu & Dudock, 1974 ; Maxwell, 1962 ; Zubay, 1973 ; Sampson *et al.*, 1972). Le rendement de ces systèmes était faible, et la synthèse du polypeptide était parfois inachevée (Chroboczek, 1985 ; Hunter, 1977 ; Roberts & Paterson, 1973 ; Suryanarayana & Subramanian, 1983). Pour rendre le système plus efficace, et prolonger l'expression, A. SPIRIN (*Lomonosov Moscow State University*) et ses collaborateurs ont pensé à débarrasser le système, au fur et à mesure que la réaction se déroule, des produits réactionnels incluant déchets ou peptides synthétisés, et de restaurer continuellement les molécules d'*ATP*, de *GTP* et les acides aminés (Spirin *et al.*, 1988). À travers une membrane d'ultrafiltration semi-perméable, les produits de synthèse étaient extraits du mélange réactionnel, et les produits d'alimentation introduits à

un débit d'1 mL par heure. Ce système a permis une synthèse continue sur une période allant de 20 à 40 heures, alors que dans les systèmes utilisés auparavant, la surexpression s'arrêtait au bout d'une heure. De plus, le ratio polypeptide synthétisé par molécule d'ARN messenger est passé de 3 à 100.

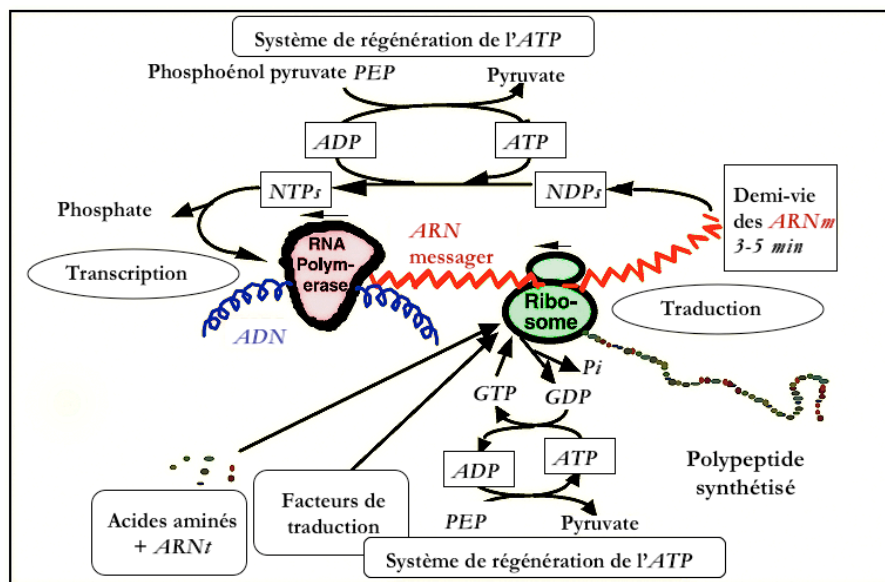


Figure 5.3 : Schéma de synthèse d'un polypeptide. Les extraits utilisés en SIV sont obtenus après croissance des cellules jusqu'à des densités élevées, puis séparation et purification de la machinerie de transcription et de traduction. L'ADN exogène ainsi que le tampon nécessaire pour réactiver l'ARN messenger sont rajoutés, et l'extrait est utilisé pour transcrire l'ADN rajouté et synthétiser la protéine correspondante (Swartz, 2006).

Une autre amélioration fondamentale est parvenue avec l'utilisation d'un système de transcription et de traduction plus efficace. Il s'agit de systèmes exogènes, issus du bactériophage T7 (Kim *et al.*, 1996). L'expansion de l'utilisation de la synthèse *in vitro* a ensuite réellement débuté avec l'automatisation et la commercialisation des systèmes par la société ROCHE DIAGNOSTICS GMBH au début des années 2000. Les premiers systèmes RTS 500 et RTS 9000 étaient composés d'extraits d'*E. coli* et utilisaient la machinerie T7 pour la surexpression. Au cours des années 2000, ROCHE et d'autres sociétés ont développé des systèmes utilisant des lysats de cellules d'insectes, de réticulocytes de lapin, de germes de blé, mais aussi des extraits de cellules de la lignée cellulaire cancéreuse humaine HeLa (Mikami *et al.*, 2008).

La première utilisation des systèmes de *SIV* pour la RMN date de 1995 dans l'équipe de S. YOKOYAMA (RIKEN - Yokohama Institute). La protéine *Ras* est sélectivement marquée à l'azote ^{15}N au niveau des sérines (Kigawa *et al.*, 1995). En *SIV*, les chercheurs ont pu obtenir 1 mg de protéine, en utilisant 0,7 mg de ^{15}N -sérine. En exprimant la protéine recombinante dans *E. coli*, 50 mg de ^{15}N -sérine ont été utilisés, et le rendement était de 17 mg de protéine. Ceci montre l'efficacité de la technique et le rendement d'incorporation vis-à-vis de l'utilisation d'acides aminés marqués. La même équipe a comparé la protéine uniformément marquée au ^{13}C et au ^{15}N , en surexpression bactérienne et en *SIV*. Les spectres ^{15}N -*HSQC* montraient les mêmes déplacements chimiques de tous les résidus, indiquant que la protéine adoptait la même structure tertiaire dans les deux cas (Kigawa *et al.*, 1999). En parallèle, l'application des dernières découvertes leur a permis d'exprimer la protéine *CAT* (*Chloramphénicol acétyl-transférase*) avec un rendement de 6 mg par mL de milieu réactionnel. La régénération d'énergie (créatine phosphate et créatine kinase), la condensation des extraits d'*E. coli* (Kim *et al.*, 1996 ; Nakano *et al.*, 1996), et l'utilisation d'un système de dialyseur à usage unique pour les échanges entre le milieu réactionnel et le milieu d'alimentation (Kim & Choi, 1996) présentent un système idéal pour la surexpression des protéines en vue de leur étude structurale et fonctionnelle. Cette technique a permis l'étude de plusieurs protéines. La peptidyl-prolyl cis-trans *B* isomérase sélectivement marquée ^{15}N -*Arg*/ ^{13}C -*Ala*-*PpiB* (18 kDa) prouvant que le marquage sélectif en *SIV* peut servir à une étude d'un résidu particulier au sein de la protéine (Guignard *et al.*, 2002). La synthèse de 19 échantillons de la *PpiA* ou cyclophiline *A bCypA*, chacun sélectivement marqué en ^{15}N au niveau d'un acide aminé différent, facilite l'attribution des corrélations *HSQC* obtenues (Ozawa *et al.*, 2004). L'application de la technique sur une protéine choisie au hasard dans le génome d'*A. thaliana* (la protéine *At3g010501.1*, 14 kDa) et la résolution de la structure de cette protéine uniformément marquée montre que la technique peut s'inscrire dans le projet de plateforme visant à résoudre les structures de protéines à une grande échelle (Vinarov *et al.*, 2004). Pour réduire le nombre d'échantillons synthétisés, un marquage combinatoire peut être adopté, afin d'éviter le recouvrement des corrélations des acides aminés et d'être plus utile pour les analyses structurales qu'un marquage sélectif résidu par résidu (Ozawa *et al.*, 2006).

En parallèle au développement méthodologique, et à l'application de la *SIV* pour des études structurales, différents laboratoires se sont penchés sur l'application de la technique pour les protéines membranaires, à cause des problèmes évoqués auparavant, concernant l'expression, la solubilisation, la purification, la cristallisation de ces molécules. Les rendements évoqués dans les papiers traitant de la *SIV* dans le cas de protéines solubles, ainsi que la facilité d'incorporation d'acides aminés fluorescents ou marqués spécifiquement rendaient la technique attirante pour l'étude structurale et fonctionnelle des *PMs*. Tous les obstacles concernant la toxicité de la surexpression d'une *PM* en bactérie ou son agrégation sont surmontables en *SIV*. Étant donné que le système est ouvert, la surexpression supporte alors l'introduction de détergent pour avoir la *PM* sous forme soluble, des cofacteurs ou des lipides nécessaires au bon repliement. Ainsi, différents laboratoires ont réussi à synthétiser *in vitro* en 2004, les transporteurs *EmrE*, *SugE* et *TehA* (Klammt *et al.*, 2004), le canal mécanosensible *MscL* (Berrier *et al.*, 2004) à des quantités de l'ordre du *mg*.

Différentes méthodes et différents tests ont été utilisés (Figure 5.4): la *PM* peut être exprimée sous forme insoluble. Le mélange réactionnel est ensuite solubilisé en détergent et la protéine est purifiée et reconstituée en liposomes. Le milieu réactionnel supporte l'introduction de quelques détergents, et la *PM* synthétisée est directement dirigée dans les micelles. Pour cela, les laboratoires de F. BERNHARD (*Goethe Institut*) et A. GHAZI ont testé différents détergents à différentes concentrations plus ou moins proches de la *CMC*. Dans un cas, l'impact chimique du détergent a été testé par un suivi de synthèse d'une protéine soluble, composée de la fusion de *EmrE* et de la *GFP*, et il s'est avéré que des détergents comme la *DPC* ou la β -*OG* empêchaient complètement la synthèse de *EmrE-GFP*. La présence des détergents *Triton X-100* (*TX-100*) et *Triton X-114* (*TX-114*) ne perturbait pas la surexpression de la protéine, mais le rendement était plus faible. La présence de lipides a été testé également. Des lipides à une (*LPC*) ou deux chaînes d'acide gras (*DMPC* et *POPC*) ainsi qu'un mélange de lipides d'*E. coli* ont été testés. La présence de *POPC* perturbait légèrement la synthèse de *EmrE-GFP*. Le rendement n'était pas perturbé en présence de *LPC* et de *DMPC*, et les extraits de lipides d'*E. coli* améliorait la synthèse (Klammt *et al.*, 2004). Dans l'autre cas, l'effet des détergents a été directement testé sur la

PM. Ainsi, la surexpression de *MscL* mais aussi l'incorporation de la *PM* en micelles, et le ratio de *PM* solubilisée sur la fraction précipité ont été suivis (Berrier *et al.*, 2004). Des détergents non-ioniques ayant des faibles *CMC* (*Triton-X-100*, *Tween 20*, *Brij 58p*, *n*-docedyl- β -*D*-maltoside), un détergent non ionique ayant une *CMC* élevée (β -*OG*), un détergent zwitterionique (*CHAPS*) et un détergent ionique (déoxycholate) ont été testés. Ce dernier inhibait complètement la synthèse, probablement à cause de sa charge. L'effet observé du β -*OG* sur la synthèse est le même pour *MscL* que pour *EmrE-GFP*. L'inhibition de synthèse est probablement due à sa *CMC* élevée qui fait qu'il est utilisé à une concentration élevée dans le mélange réactionnel. Les détergents non-ioniques à faible *CMC* semblent adéquats pour la synthèse des *PMs* sous forme soluble. L'oligomérisation et l'activité des *PMs* synthétisées ont été suivies. Dans le cas d'*EmrE*, certaines régions en dehors des 12 hélices transmembranaires seraient déstructurées, mais l'activité de transport de l'éthidium de la protéine est inchangée. Pour *MscL*, l'organisation pentamérique de la protéine, ainsi que son activité *MS* sont identiques à la protéine synthétisée en bactéries.

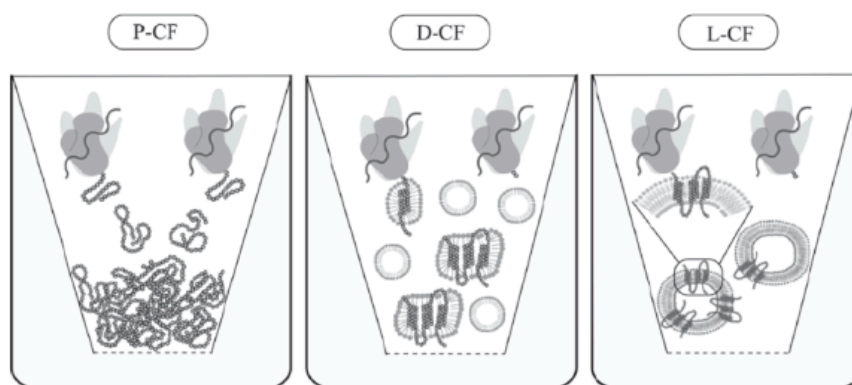


Figure 5.4 : Schéma de synthèse d'une protéine membranaire en *SIV*. Lorsque la synthèse est effectuée en absence de détergents ou liposomes, la *PM* est synthétisée et récupérée sous forme de précipitat *P-CF* ; en présence de micelles de détergent, la *PM* est récupérée dans le surnageant *D-CF* ; en présence de liposomes, diverses *PMs* sont incorporées dans les vésicules, puis récupérées après ultracentrifugation dans le culot (*L-CF*). Figure tirée de Schneider *et al.*, 2010.

L'application de la technique de *SIV* s'est ensuite étendue aux récepteurs couplés aux protéines *G* (*RCPG*) ou des transporteurs eucaryotes (Ishihara *et al.*, 2005 ; Kaiser *et al.*, 2008 ; Keller *et al.*, 2008 ; Klammt *et al.*, 2007 ; Park *et al.*, 2010), et différents

détergents et surfactants ont été testés (Klammt *et al.*, 2005 ; Park *et al.*, 2007). La SIV de la sécrétine de la membrane externe de *Klebsiella oxytoca* PulD et de l'aquaporine AqpZ d'*E. coli* directement en liposomes a été également réalisée (Guilvout *et al.*, 2008 ; Hovijitra *et al.*, 2009). L'insertion membranaire spontanée et l'oligomérisation des PMs synthétisées *in vitro* en présence de liposomes ne seraient efficaces que dans le cas de protéines ne nécessitant pas de cofacteurs pour leur insertion ou leur repliement. Cette technique faciliterait davantage l'utilisation de la SIV pour l'étude de PMs *in situ* sur le plan fonctionnel, mais également pour une étude structurale à haute résolution.

VI Conclusion - Projet de recherche

L'état de l'art actuel de la **résonance magnétique nucléaire en phase solide** nécessite la mise au point de nouvelles techniques de biochimie pour l'étude des **protéines membranaires**. Différentes méthodes d'amélioration du rendement de surexpression et des techniques de marquage ont été testées les vingt dernières années, afin d'obtenir une quantité suffisante de protéines avec un marquage astucieux, et ceci pour l'acquisition de spectres de meilleure résolution. La **synthèse *in vitro*** présente les deux avantages. Cette technique permet la production de protéines recombinantes à des taux de synthèse inégalés dans un système vivant. La synthèse *in vitro* présente également la particularité d'être un **système ouvert**, dans lequel les acides aminés spécifiquement marqués, mais aussi les détergents, les ligands ou les lipides s'incorporent facilement au milieu, sans perturber la synthèse.

Dans ce projet, les deux approches sont utilisées conjointement pour la première fois, pour étudier la protéine membranaire **mécanosensible à large conductance** d'*Escherichia coli*. Cette protéine dépend fortement de son environnement, et son ouverture ne peut se réaliser que dans des vésicules lipidiques. Afin d'apporter des précisions aux connaissances concernant le mode d'ouverture et pour résoudre **la structure du canal à l'état ouvert**, la résonance magnétique nucléaire en phase solide paraît idéale pour suivre les interactions entre la protéine et les lipides, étudier la dynamique de la protéine, et déterminer les caractéristiques structurales de la protéine à l'état ouvert. Les spectres obtenus sur la protéine uniformément marquée présentent un recouvrement des corrélations des acides aminés, malgré le faible poids moléculaire de la protéine. Ceci est notamment dû au pourcentage de structures secondaires en hélices et à la présence d'un grand nombre de résidus hydrophobes. Un **marquage sélectif** contrôlé des résidus de la protéine par synthèse *in vitro* permettrait de réduire le recouvrement des corrélations, afin de faciliter l'attribution des acides aminés et d'obtenir des informations structurales précises, permettant de résoudre la structure de la protéine, et d'apporter des nouvelles connaissances concernant son mode d'ouverture.

La compréhension du mode d'ouverture de la protéine mécanosensible à large conductance pourrait également servir à définir un modèle d'action plus général, concernant d'autres canaux mécanosensibles. L'association de la résonance magnétique nucléaire en phase solide à la synthèse *in vitro* pour l'étude d'une protéine membranaire pourrait ainsi devenir une méthode incontournable pour étudier les protéines membranaires *in situ*, tant sur un aspect fonctionnel que structural. Les méthodes utilisées pour la préparation des échantillons, ainsi que celles employées pour les différents tests quantitatifs et qualitatifs sont détaillées dans le deuxième chapitre (B).

Le troisième chapitre contient les résultats obtenus au cours des quatre dernières années. Il est divisé en cinq parties. Dans un premier temps, les analyses de l'échantillon uniformément marqué sont exposées, de la surexpression bactérienne jusqu'aux spectres $2D\ ^{13}C-^{13}C$ et $^{13}C-^{15}N$ obtenus (Chapitre C.1). Ensuite, la deuxième partie traite de la synthèse *in vitro* d'un échantillon sélectivement marqué au niveau de deux familles d'acides aminés (Chapitre C.2). Des systèmes de spin présents sur les spectres de l'échantillon sélectivement marqué, et invisibles sur les spectres de celui uniformément marqué confirme l'efficacité de la synthèse *in vitro* (Abdine *et al.*, 2010). Après ces résultats, différentes approches de marquage ont été élaborées, et un échantillon a été synthétisé pour chacune des approches. Tout d'abord, l'approche appelée prédictive est détaillée. Elle est basée sur les méthodes de prédiction de déplacements chimiques. Les analyses effectuées sur l'échantillon correspondant sont exposées (Chapitre C.3). Une approche basée sur l'analyse de la séquence de la protéine à la recherche de paires d'acides aminés isolées a permis de synthétiser un échantillon, et les résultats obtenus sont détaillés (Chapitre C.4). À ce stade, ces approches ont fait l'objet d'un deuxième article paru dans la revue *New Biotechnology* (Abdine *et al.*, 2011). Un dernier échantillon a été synthétisé, selon une approche combinant les informations de l'approche prédictive, la séquence mais aussi des données bibliographiques de la protéine *Ec-MscL* (Chapitre C.5).

Le quatrième chapitre est consacré à la discussion des méthodes employées et des résultats obtenus (Chapitre D). Il est suivi par une discussion générale (Chapitre E), et les perspectives envisagées pour ce projet (Chapitre F).

VII Références

- Abdine** A, Verhoeven MV, Park K-H, et al. Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* **2010**;204(1):155-159.
- Abdine** A, Verhoeven MV, Warschawski DE. Cell-free expression and labeling strategies for a new decade in solid-state NMR. *N. Biotechnol.* **2011**;28(3):272-276.
- Adamo** L, Naveiras O, Wenzel PL, et al. Biomechanical forces promote embryonic haematopoiesis. *Nature.* **2009**;459(7250):1131-1135.
- Agarwal** V, Fink U, Schuldiner S, Reif B. MAS solid-state NMR studies on the multidrug transporter EmrE. *Biochim. Biophys. Acta.* **2007**;1768:3036–3043.
- Ajouz** B, Berrier C, Besnard M, Martinac B, Ghazi A. Contributions of the different extramembranous domains of the mechanosensitive ion channel MscL to its response to membrane tension. *J. Biol. Chem.* **2000**;275:1015–1022.
- Ajouz** B, Berrier C, Garrigues A, Besnard M, Ghazi A. Release of thioredoxin via the mechanosensitive channel MscL during osmotic downshock of *Escherichia coli* cells. *J. Biol. Chem.* **1998**;273:26670–26674.
- Andrew** ER, Bradbury A, Eades RG. Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. *Nature.* **1959**;183(4678):1802-1803.
- Anishkin** A, Chiang CS, Sukharev S, Sukharev SI. Gain-of-function mutations reveal expanded intermediate states and a sequential action of two gates in MscL. *J. Gen. Physiol.* **2005**;125:155–170.
- Anishkin** A, Gendel V, Sharifi NA, et al. On the conformation of the COOH-terminal domain of the large mechanosensitive channel MscL. *J. Gen. Physiol.* **2003**;121:227–244.
- Arkin** IT, Blount P, Kung C, Brunger AT, Sukharev SI. Helicity, membrane incorporation, orientation and thermal stability of the large conductance mechanosensitive ion channel from *E. coli*. *Biochim. Biophys. Acta.* **1998**;1369:131–140.
- Arnadóttir** J, Chalfie M. Eukaryotic mechanosensitive channels. *Annu Rev Biophys.* **2010**;39:111-137.
- Bartlett** JL, Li Y, Blount P. Mechanosensitive channel gating transitions resolved by functional changes upon pore modification. *Biophys. J.* **2006**;91:3684–3691.
- Batiza** AF, Rayment I, Kung C. Channel gate! Tension, leak and disclosure. *Structure.* **1999**;7:99–103.
- Bayburt** TH, Sligar SG. Membrane protein assembly into Nanodiscs. *FEBS Lett.* **2010**;584(9):1721-1727.
- Bayrhuber** M, Meins T, Habeck M, et al. Structure of the human voltage-dependent anion channel. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105(40):15370-15375.
- Bechinger** B, Zasloff M, Opella SJ. Structure and orientation of the antibiotic peptide magainin in membranes by solid-state nuclear magnetic resonance spectroscopy. *Protein Sci.* **1993**;2(12):2077-2084.
- Berrier** C, Park K, Abes S, et al. Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry.* **2004**;43(39):12585-12591.
- Berrier** C, Coulombe A, Houssin C, Ghazi A. A patch-clamp study of ion channels of inner and outer membranes and of contact zones of *E. coli*, fused into giant liposomes. Pressure-activated channels are localized in the inner membrane. *FEBS Lett.* **1989**;259(1):27-32.

- Berrier C**, Garrigues A, Richarme G, Ghazi A. Elongation factor Tu and DnaK are transferred from the cytoplasm to the periplasm of *Escherichia coli* during osmotic downshock presumably via the mechanosensitive channel mscL. *J. Bacteriol.* **2000**;182:248–251.
- Berrier C**, Guilvout I, Bayan N, et al. Coupled cell-free synthesis and lipid vesicle insertion of a functional oligomeric channel MscL. MscL does not need the insertase YidC for insertion in vitro. *Biochim. Biophys. Acta.* **2011**;1808(1):41-46.
- Betanzos M**, Chiang CS, Guy HR, Sukharev SI. A large iris-like expansion of a mechanosensitive channel protein induced by membrane tension. *Nat. Struct. Biol.* **2002**;9:704–710.
- Bloch F**. Theory of Line Narrowing by Double-Frequency Irradiation. *Phys. Rev.* **1958**;111(3):841.
- Bloch F**, Hansen WW, Packard M. The Nuclear Induction Experiment. *Phys. Rev.* **1946**;70(7-8):474.
- Blount P**, Moe PC. Bacterial mechanosensitive channels: integrating physiology, structure and function. *Trends Microbiol.* **1999**;7:420–424.
- Blount P**, Schroeder MJ, Nagle SK, Kung C, Sukharev SI. Single residue substitutions that change the gating properties of a mechanosensitive channel in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **1996 a**;93:11652–11657.
- Blount P**, Sukharev SI, Moe PC, et al. Membrane topology and multimeric structure of a mechanosensitive channel protein of *Escherichia coli*. *EMBO J.* **1996 b**;15:4798–4805.
- Blount P**, Sukharev SI, Moe PC, Nagle SK, Kung C. Towards an understanding of the structural and functional properties of MscL, a mechanosensitive channel in bacteria. *Biol. Cell.* **1996 c**;87(1-2):1-8.
- Bordag N**, Keller S. Alpha-helical transmembrane peptides: a « divide and conquer » approach to membrane proteins. *Chem. Phys. Lipids.* **2010**;163(1):1-26.
- Breneman KD**, Brownell WE, Rabbitt RD. Hair cell bundles: flexoelectric motors of the inner ear. *PLoS ONE.* **2009**;4(4):e5201.
- Cady SD**, Schmidt-Rohr K, Wang J, et al. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature.* **2010**;463(7281):689-692.
- Cai ML**, Huang Y, Sakaguchi K, et al. An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J. Biomol. NMR.* **1998**;11(1):97-102.
- Canet D**. *La RMN: concepts et méthodes*. InterEditions; **1991**.
- Casado M**, Ascher P. Opposite modulation of NMDA receptors by lysophospholipids and arachidonic acid: common features with mechanosensitivity. *J. Physiol.* **1998**;513(2):317-330.
- Casales E**, Aranda A, Quetglas JI, et al. A novel system for the production of high levels of functional human therapeutic proteins in stable cells with a Semliki Forest virus noncytopathic vector. *N. Biotechnol.* **2010**;27(2):138-148.
- Castellani F**. Structure determination of immobilized proteins by solid-state NMR spectroscopy. **2003**.
- Castellani F**, van Rossum B, Diehl A, et al. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature.* **2002**;420:98–102.
- Catoire LJ**, Zoonens M, van Heijenoort C, et al. Solution NMR mapping of water-accessible residues in the transmembrane beta-barrel of OmpX. *Eur. Biophys. J.* **2010 a**;39(4):623-630.
- Catoire LJ**, Damian M, Giusti F, et al. Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts

- a highly constrained conformation when associated to human BLT2. *J. Am. Chem. Soc.* **2010** *b*;132(26):9049-9057.
- Chandran** V, Fronzes R, Duquerroy S, et al. Structure of the outer membrane complex of a type IV secretion system. *Nature.* **2009**;462(7276):1011-1015.
- Chang** G, Spencer RH, Lee AT, Barclay MT, Rees DC. Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science.* **1998**;282(5397):2220-2226.
- Chemin** J, Patel AJ, Duprat F, et al. Up- and down-regulation of the mechano-gated K-2P channel TREK-1 by PIP2 and other membrane phospholipids. *Pflugers Arch.* **2007**;455(1):97-103.
- Chiang** CS, Shirinian L, Sukharev SI. Capping transmembrane helices of MscL with aromatic residues changes channel response to membrane stretch. *Biochemistry.* **2005**;44:12589–12597.
- Christie** WW. Rapid separation and quantification of lipid classes by high performance liquid chromatography and mass (light-scattering) detection. *J. Lipid Res.* **1985**;26(4):507-512.
- Chroboczek** J. Incomplete polypeptides are formed in vitro by premature chain termination. *Eur. J. Biochem.* **1985**;149(3):565-569.
- Chung** S-H, Anderson OS, Krishnamurthy VV. *Biological Membrane Ion Channels: Dynamics, Structure, And Applications.* 1er éd. Springer-Verlag New York Inc.; **2006**.
- Cohen** BB. 2 Fractions required for cell-free protein synthesis with components from rabbit reticulocytes. *Biochem. J.* **1969**;115(3):523-527.
- Colbeau** A, Nachbaur J, Vignais PM. Enzymic characterization and lipid composition of rat liver subcellular membranes. *Biochim. Biophys. Acta.* **1971**;249(2):462-492.
- Cruickshank** CC, Minchin RF, Le Dain AC, Martinac B. Estimation of the pore size of the large-conductance mechanosensitive ion channel of *Escherichia coli*. *Biophys. J.* **1997**;73:1925–1931.
- Davies** PF. Flow-mediated endothelial mechanotransduction. *Physiol. Rev.* **1995**;75(3):519-560.
- Debret** G, Valadie H, Stadler AM, Etchebest C. New insights of membrane environment effects on MscL channel mechanics from theoretical approaches. *Proteins.* **2008**;71:1183–1196.
- Deisenhofer** J, Epp O, Miki K, Huber R, Michel H. X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodospirillum rubrum*. *J. Mol. Biol.* **1984**;180(2):385-398.
- Demel** RA, De Kruyff B. The function of sterols in membranes. *Biochim. Biophys. Acta.* **1976**;457(2):109-132.
- Dittmer** J, Thøgersen L, Underhaug J, et al. Incorporation of antimicrobial peptides into membranes: a combined liquid-state NMR and molecular dynamics study of alamethicin in DMPC/DHPC bicelles. *J Phys Chem B.* **2009**;113(19):6928-6937.
- Domanov** YA, Molotkovsky JG, Gorbenko GP. Coverage-dependent changes of cytochrome c transverse location in phospholipid membranes revealed by FRET. *Biochim. Biophys. Acta.* **2005**;1716(1):49-58.
- Dorwart** MR, Wray R, Brautigam CA, Jiang Y, Blount P. *S. aureus* MscL is a pentamer in vivo but of variable stoichiometries in vitro: implications for detergent-solubilized membrane proteins. *PLoS Biol.* **2010**;8(12):e1000555.
- Eddidin** M. Lipids on the frontier: a century of cell-membrane bilayers. *Nat. Rev. Mol. Cell Biol.* **2003**;4(5):414-418.
- Egawa** K, Shibata H, Yamashita S-I, et al. Overexpression and purification of rat peroxisomal membrane protein

- 22, PMP22, in *Pichia pastoris*. *Protein Expr. Purif.* **2009**;64(1):47-54.
- Egorova-Zachernyuk** TA, Hollander J, Fraser N, et al. Heteronuclear 2D-correlations in a uniformly [¹³C, ¹⁵N] labeled membrane-protein complex at ultra-high magnetic fields. *J. Biomol. NMR.* **2001**;19(3):243-253.
- Eifler** N, Duckely M, Sumanovski LT, et al. Functional expression of mammalian receptors and membrane channels in different cells. *J. Struct. Biol.* **2007**;159(2):179-193.
- Elmore** DE, Dougherty DA. Investigating lipid composition effects on the mechanosensitive channel of large conductance (MscL) using molecular dynamics simulations. *Biophys. J.* **2003**;85:1512–1524.
- Etzkorn** M, Martell S, Andronesi OC, et al. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2007**;46(3):459-462.
- Ferguson** N, Becker J, Tidow H, et al. General structural motifs of amyloid protofilaments. *Proc. Natl. Acad. Sci. U.S.A.* **2006**;103(44):16248-16253.
- Fernández** C, Hilty C, Bonjour S, et al. Solution NMR studies of the integral membrane proteins OmpX and OmpA from *Escherichia coli*. *FEBS Lett.* **2001**;504(3):173-178.
- Ferndahl** C, Bonander N, Logez C, et al. Increasing cell biomass in *Saccharomyces cerevisiae* increases recombinant protein yield: the use of a respiratory strain as a microbial cell factory. *Microb. Cell Fact.* **2010**;9:47.
- Fiaux** J, Bertelsen EB, Horwich AL, Wüthrich K. Uniform and residue-specific ¹⁵N-labeling of proteins on a highly deuterated background. *J. Biomol. NMR.* **2004**;29(3):289-297.
- Fink** M, Duprat F, Lesage F, et al. Cloning, functional expression and brain localization of a novel unconventional outward rectifier K⁺ channel. *EMBO J.* **1996**;15(24):6854-6862.
- Franks** WT, Wylie BJ, Schmidt HL, et al. Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105:4621–4626.
- Ganapathy** S, van Gammeren AJ, Hulsbergen FB, de Groot JJM. Probing secondary, tertiary, and quaternary structure along with protein-cofactor interactions for a helical transmembrane protein complex through ¹H spin diffusion with MAS NMR spectroscopy. *J. Am. Chem. Soc.* **2007**;129(6):1504-1505.
- Gandhi** CS, Walton TA, Rees DC. OCAM: a new tool for studying the oligomeric diversity of MscL channels. *Protein Sci.* **2011**;20(2):313-326.
- Gill** SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **1989**;182(2):319-326.
- Goñi** FM. Non-permanent proteins in membranes: when proteins come as visitors (Review). *Mol. Membr. Biol.* **2002**;19(4):237-245.
- Goñi** FM, Alonso A. Biophysics of sphingolipids I. Membrane properties of sphingosine, ceramides and other simple sphingolipids. *Biochim. Biophys. Acta.* **2006**;1758(12):1902-1921.
- Gorzelle** BM, Hoffman AK, Keyes MH, et al. Amphipols can support the activity of a membrane enzyme. *J. Am. Chem. Soc.* **2002**;124(39):11594-11595.
- Grashoff** C, Hoffman BD, Brenner MD, et al. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature.* **2010**;466(7303):263-266.
- Griffin** RG. Dipolar recoupling in MAS spectra of biological solids. *Nat. Struct. Biol.* **1998**;5 Suppl:508-512.
- Gu** L, Liu W, Martinac B. Electromechanical coupling model of gating the large mechanosensitive ion channel

- (MscL) of Escherichia coli by mechanical force. *Biophys. J.* **1998**;74:2889–2902.
- Guignard L**, Ozawa K, Pursglove SE, Otting G, Dixon NE. NMR analysis of in vitro-synthesized proteins without purification: a high-throughput approach. *FEBS Lett.* **2002**;524(1-3):159-162.
- Guilvout I**, Chami M, Berrier C, et al. In vitro multimerization and membrane insertion of bacterial outer membrane secretin PulD. *J. Mol. Biol.* **2008**;382(1):13-23.
- Gullingsrud J**, Schulten K. Gating of MscL studied by steered molecular dynamics. *Biophys. J.* **2003**;85:2087–2099.
- Hamill OP**. *Mechanosensitive ion channels*. Academic Press - Elsevier; **2007**.
- Hamill OP**, Marty A, Neher E, Sakmann B, Sigworth FJ. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Arch.* **1981**;391(2):85-100.
- Hamill OP**, McBride DW. Induced membrane hypo/hyper-mechanosensitivity: a limitation of patch-clamp recording. *Annu. Rev. Physiol.* **1997**;59:621-631.
- Hartmann SR**, Hahn EL. Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* **1962**;128(5):2042.
- Hase CC**, Minchin RF, Kloda A, Martinac B. Cross-linking studies and membrane localization and assembly of radiolabelled large mechanosensitive ion channel (MscL) of Escherichia coli. *Biochem. Biophys. Res. Commun.* **1997**;232:777–782.
- Heise H**, Hoyer W, Becker S, et al. Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2005**;102(44):15871-15876.
- Helenius A**, Simons K. Solubilization of membranes by detergents. *Biochim. Biophys. Acta.* **1975**;415(1):29-79.
- Helms V**. Attraction within the membrane. Forces behind transmembrane protein folding and supramolecular complex assembly. *EMBO Rep.* **2002**;3(12):1133-1138.
- Hiller M**, Higman VA, Jehle S, et al. [2,3-(13)C]-labeling of aromatic residues—getting a head start in the magic-angle-spinning NMR assignment of membrane proteins. *J. Am. Chem. Soc.* **2008 a**;130(2):408-409.
- Hiller S**, Garces RG, Malia TJ, et al. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science.* **2008 b**;321(5893):1206-1210.
- Holden MTG**, Feil EJ, Lindsay JA, et al. Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2004**;101(26):9786-9791.
- Hong M**. Determination of multiple phi-torsion angles in proteins by selective and extensive C-13 labeling and two-dimensional solid-state NMR. *J. Magn. Reson.* **1999**;139(2):389-401.
- Hong M**, Jakes K. Selective and extensive 13C labeling of a membrane protein for solid-state NMR investigations. *J. Biomol. NMR.* **1999**;14(1):71-74.
- Hovijitra NT**, Wu JJ, Peaker B, Swartz JR. Cell-Free Synthesis of Functional Aquaporin Z in Synthetic Liposomes. *Biotech. Bioeng.* **2009**;104(1):40-49.
- Hunte C**, Richers S. Lipids and membrane protein structures. *Curr. Opin. Struct. Biol.* **2008**;18(4):406-411.
- Hunter AR**, Farrell PJ, Jackson RJ, Hunt T. Role of polyamines in cell-free protein-synthesis in wheat-germ system. *Eur. J. Biochem.* **1977**;75(1):149-157.
- Huster D**, Yao X, Jakes K, Hong M. Conformational changes of colicin Ia channel-forming domain upon membrane binding: a solid-state NMR study. *Biochim. Biophys. Acta.* **2002**;1561(2):159-170.

- Ikai A.** Thermostability and aliphatic index of globular proteins. *J. Biochem.* **1980**;88(6):1895-1898.
- Ingber DE.** Cellular mechanotransduction: putting all the pieces together again. *FASEB J.* **2006**;20(7):811-827.
- Inoue R, Jensen LJ, Jian Z, et al.** Synergistic activation of vascular TRPC6 channel by receptor and mechanical stimulation via phospholipase C/diacylglycerol and phospholipase A2/omega-hydroxylase/20-HETE pathways. *Circ. Res.* **2009**;104(12):1399-1409.
- Iscla I, Wray R, Blount P.** On the structure of the N-terminal domain of the MscL channel: helical bundle or membrane interface. *Biophys. J.* **2008**;95:2283–2291.
- Ishihara G, Goto M, Saeki M, et al.** Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr. Purif.* **2005**;41(1):27-37.
- Ishima R, Louis JM, Torchia DA.** Transverse C-13 relaxation of CHD2 methyl isotopomers to detect slow conformational changes of protein side chains. *J. Am. Chem. Soc.* **1999**;121(49):11589-11590.
- Iwata K, Fujiwara T, Matsuki Y, et al.** 3D structure of amyloid protofilaments of beta2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2006**;103(48):18119-18124.
- Jaalouk DE, Lammerding J.** Mechanotransduction gone awry. *Nat. Rev. Mol. Cell Biol.* **2009**;10(1):63-73.
- Jacob F, Monod J.** Genetic regulatory mechanisms in synthesis of proteins. *J. Mol. Biol.* **1961**;3(3):318-&.
- Jansson M, Li YC, Jendeborg L, et al.** High-level production of uniformly ¹⁵N- and ¹³C-enriched fusion proteins in *Escherichia coli*. *J. Biomol. NMR.* **1996**;7(2):131-141.
- Johnston J, Cook GA, Tomich JM, Sansom MSP.** Conformation and Environment of Channel-Forming Peptides: A Simulation Study. *Biophys. J.* **2006**;90(6):1855-1864.
- Jones DT.** Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **1999**;292(2):195-202.
- Kaiser L, Graveland-Bikker J, Steuerwald D, et al.** Efficient cell-free production of olfactory receptors: Detergent optimization, structure, and ligand binding analyses. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105(41):15726-15731.
- Kamaraju K, Sukharev SI.** The membrane lateral pressure-perturbing capacity of parabens and their effects on the mechanosensitive channel directly correlate with hydrophobicity. *Biochemistry.* **2008**;47(40):10540-10550.
- Keeler J.** *Understanding NMR Spectroscopy.* John Wiley & Sons; **2010**.
- Keller T, Schwarz D, Bernhard F, et al.** Cell free expression and functional reconstitution of eukaryotic drug transporters. *Biochemistry.* **2008**;47(15):4552-4564.
- Kendrew JC, Dickerson RE, Strandberg BE, et al.** Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å. resolution. *Nature.* **1960**;185(4711):422-427.
- Ketchum RR, Hu W, Cross TA.** High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science.* **1993**;261(5127):1457-1460.
- Kigawa T, Muto Y, Yokoyama S.** Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J. Biomol. NMR.* **1995**;6(2):129-134.
- Kigawa T, Yabuki T, Yoshida Y, et al.** Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.* **1999**;442(1):15-19.
- Kim DM, Choi CY.** A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol. Prog.* **1996**;12(5):645-649.

- Kim** DM, Kigawa T, Choi CY, Yokoyama S. A highly efficient cell-free protein synthesis system from *Escherichia coli*. *Eur. J. Biochem.* **1996**;239(3):881-886.
- Kim** HJ, Howell SC, Van Horn WD, Jeon YH, Sanders CR. Recent Advances in the Application of Solution NMR Spectroscopy to Multi-Span Integral Membrane Proteins. *Prog Nucl Magn Reson Spectrosc.* **2009**;55(4):335-360.
- Klammt** C, Lohr F, Schafer B, et al. High level cell-free expression and specific labeling of integral membrane proteins. *Eur. J. Biochem.* **2004**;271(3):568-580.
- Klammt** C, Schwarz D, Eifler N, et al. Cell-free production of G protein-coupled receptors for functional and structural studies. *J. Struct. Biol.* **2007**;159(2):194-205.
- Klammt** C, Schwarz D, Fendler K, et al. Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J.* **2005**;272(23):6024-6038.
- Kloda** A, Ghazi A, Martinac B. C-terminal charged cluster of MscL, RKKEE, functions as a pH sensor. *Biophys. J.* **2006**;90:1992-1998.
- Kong** Y, Shen Y, Warth TE, Ma J. Conformational pathways in the gating of *Escherichia coli* mechanosensitive channel. *Proc. Natl. Acad. Sci. U.S.A.* **2002**;99:5999-6004.
- Kücken** M, Newell AC. Fingerprint formation. *J. Theor. Biol.* **2005**;235(1):71-83.
- Kyte** J, Doolittle RF. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **1982**;157(1):105-132.
- Laitko** U, Juranka PF, Morris CE. Membrane stretch slows the concerted step prior to opening in a Kv channel. *J. Gen. Physiol.* **2006**;127(6):687-701.
- Lange** A, Giller K, Hornig S, et al. Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature.* **2006**;440(7086):959-962.
- Lau** T-L, Partridge AW, Ginsberg MH, Ulmer TS. Structure of the integrin beta3 transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry.* **2008**;47(13):4008-4016.
- le Maire** M, Champeil P, Moller JV. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta.* **2000**;1508(1-2):86-111.
- Lee** AL, Urbauer JL, Wand AJ. Improved labeling strategy for ¹³C relaxation measurements of methyl groups in proteins. *J. Biomol. NMR.* **1997**;9(4):437-440.
- Lee** AG. Lipid-protein interactions in biological membranes: a structural perspective. *Biochim. Biophys. Acta.* **2003**;1612(1):1-40.
- Lee** AG. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta.* **2004**;1666(1-2):62-87.
- Lehner** I, Basting D, Meyer B, et al. The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J. Biol. Chem.* **2008**;283(6):3281-3288.
- LeMaster** DM, Kushlan DM. Dynamical mapping of *E. coli* thioredoxin via C-13 NMR relaxation analysis. *J. Am. Chem. Soc.* **1996**;118(39):9255-9264.
- Levina** N, Töttemeyer S, Stokes NR, et al. Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J.* **1999**;18(7):1730-1737.
- Li** Y, Berthold DA, Gennis RB, Rienstra CM. Chemical shift assignment of the transmembrane helices of DsbB, a

- 20-kDa integral membrane enzyme, by 3D magic-angle spinning NMR spectroscopy. *Protein Sci.* **2008**;17(2):199-204.
- Lian LY**, Middleton DA. Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**;39(3):171-190.
- Linares DM**, Geertsma ER, Poolman B. Evolved *Lactococcus lactis* strains for enhanced expression of recombinant membrane proteins. *J. Mol. Biol.* **2010**;401(1):45-55.
- Liu Z**, Gandhi CS, Rees DC. Structure of a tetrameric MscL in an expanded intermediate state. *Nature.* **2009**;461:120-124.
- Loquet A**, Bardiaux B, Gardiennet C, et al. 3D structure determination of the Crh protein from highly ambiguous solid-state NMR restraints. *J. Am. Chem. Soc.* **2008**;130(11):3579-3589.
- Lorch M**, Faham S, Kaiser C, et al. How to prepare membrane proteins for solid-state NMR: A case study on the alpha-helical integral membrane protein diacylglycerol kinase from *E. coli*. *ChemBiochem.* **2005**;6(9):1693-1700.
- Lowe IJ**. Free Induction Decays of Rotating Solids. *Phys. Rev. Lett.* **1959**;2(7):285.
- Lundström P**, Teilum K, Carstensen T, et al. Fractional ¹³C enrichment of isolated carbons using [1-¹³C]- or [2-¹³C]-glucose facilitates the accurate measurement of dynamics at backbone C α and side-chain methyl positions in proteins. *J. Biomol. NMR.* **2007**;38(3):199-212.
- Lyukmanova EN**, Shenkarev ZO, Paramonov AS, et al. Lipid-protein nanoscale bilayers: a versatile medium for NMR investigations of membrane proteins and membrane-active peptides. *J. Am. Chem. Soc.* **2008**;130(7):2140-2141.
- MacDonald JJ**, Sprecher H. Phospholipid fatty acid remodeling in mammalian cells. *Biochim. Biophys. Acta.* **1991**;1084(2):105-121.
- Macedo-Ribeiro S**, Bode W, Huber R, et al. Crystal structures of the membrane-binding C2 domain of human coagulation factor V. *Nature.* **1999**;402(6760):434-439.
- Maingret F**, Patel AJ, Lesage F, Lazdunski M, Honoré E. Lysophospholipids open the two-pore domain mechanogated K(+) channels TREK-1 and TRAAK. *J. Biol. Chem.* **2000**;275(14):10128-10133.
- Malmivuo J**, Plonsey R. *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press; **1995**.
- Mani R**, Cady SD, Tang M, et al. Membrane-dependent oligomeric structure and pore formation of a beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2006**;103(44):16242-16247.
- Marcu K**, Dudock B. Characterization of a highly efficient protein synthesizing system derived from commercial wheat germ. *Nucleic Acids Res.* **1974**;1(11):1385-1397.
- Markin VS**, Sachs F. Thermodynamics of mechanosensitivity. *Phys Biol.* **2004**;1(1-2):110-124.
- Markin VS**, Martinac B. Mechanosensitive ion channels as reporters of bilayer expansion. A theoretical model. *Biophys. J.* **1991**;60(5):1120-1127.
- Marley J**, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR.* **2001**;20(1):71-75.
- Marsh D**. Energetics of hydrophobic matching in lipid-protein interactions. *Biophys. J.* **2008**;94(10):3996-4013.
- Martinac B**, Buechner M, Delcour AH, Adler J, Kung C. Pressure-sensitive ion channel in *Escherichia coli*. *Proc.*

- Natl. Acad. Sci. U.S.A.* **1987**;84(8):2297-2301.
- Martinac B**, Hamill OP. Gramicidin A channels switch between stretch activation and stretch inactivation depending on bilayer thickness. *Proc. Natl. Acad. Sci. U.S.A.* **2002**;99(7):4308-4312.
- Martinac B**. Mechanosensitive ion channels: molecules of mechanotransduction. *J. Cell. Sci.* **2004**;117(Pt 12):2449-2460.
- Mascioni A**, Karim C, Barany G, Thomas DD, Veglia G. Structure and orientation of sarcolipin in lipid environments. *Biochemistry.* **2002**;41(2):475-482.
- Maurer JA**, Elmore DE, Clayton D, et al. Confirming the revised C-terminal domain of the MscL crystal structure. *Biophys. J.* **2008**;94:4662-4667.
- Maxwell ES**. Stimulation of amino acid incorporation into protein by natural and synthetic polyribonucleotides in a mammalian cell-free system. *Proc. Natl. Acad. Sci. U.S.A.* **1962**;48:1639-1643.
- Meyer GR**, Gullingsrud J, Schulten K, Martinac B. Molecular dynamics study of MscL interactions with a curved lipid bilayer. *Biophys. J.* **2006**;91:1630-1637.
- Moe PC**, Levin G, Blount P. Correlating a protein structure with function of a bacterial mechanosensitive channel. *J. Biol. Chem.* **2000**;275:31121-31127.
- Moe PC**, Blount P, Kung C. Functional and structural conservation in the mechanosensitive channel MscL implicates elements crucial for mechanosensation. *Mol. Microbiol.* **1998**;28:583-592.
- Morris CE**, Juranka PF. Nav channel mechanosensitivity: activation and inactivation accelerate reversibly with stretch. *Biophys. J.* **2007**;93(3):822-833.
- Muchmore DC**, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Meth. Enzymol.* **1989**;177:44-73.
- Neher E**, Sakmann B. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature.* **1976**;260(5554):799-802.
- Newstead S**, Hobbs J, Jordan D, Carpenter EP, Iwata S. Insights into outer membrane protein crystallization. *Mol. Membr. Biol.* **2008**;25(8):631-638.
- Nguyen T**, Clare B, Guo W, Martinac B. The effects of parabens on the mechanosensitive channels of *E. coli*. *Eur. Biophys. J.* **2005**;34:389-395.
- Nirenberg M**, Matthaei JH. Dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **1961**;47(10):1588-&.
- Oakley AJ**, Martinac B, Wilce MC. Structure and function of the bacterial mechanosensitive channel of large conductance. *Protein Sci.* **1999**;8:1915-1921.
- Opella SJ**, Marassi FM, Gesell JJ, et al. Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat. Struct. Biol.* **1999**;6(4):374-379.
- Opella SJ**, Zeri AC, Park SH. Structure, dynamics, and assembly of filamentous bacteriophages by nuclear magnetic resonance spectroscopy. *Annu. Rev. Phys. Chem.* **2008**;59:635-657.
- Orr AW**, Helmke BP, Blackman BR, Schwartz MA. Mechanisms of mechanotransduction. *Dev. Cell.* **2006**;10(1):11-20.
- Ou X**, Blount P, Hoffman RJ, Kung C. One face of a transmembrane helix is crucial in mechanosensitive channel gating. *Proc. Natl. Acad. Sci. U.S.A.* **1998**;95:11471-11475.

- Ozawa** K, Wu P, Dixon N, Otting G. N-15-Labelled proteins by cell-free protein synthesis - Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *FEBS J.* **2006**;273(18):4154-4159.
- Ozawa** K, Headlam MJ, Schaeffer PM, et al. Optimization of an Escherichia coli system for cell-free synthesis of selectively N-15-labelled proteins for rapid analysis by NMR spectroscopy. *Eur. J. Biochem.* **2004**;271(20):4084-4093.
- Paliy** O, Gunasekera TS. Growth of E. coli BL21 in minimal media with different gluconeogenic carbon sources and salt contents. *Appl. Microbiol. Biotechnol.* **2007**;73(5):1169-1172.
- Pandit** A, Wawrzyniak PK, van Gammeren AJ, et al. Nuclear magnetic resonance secondary shifts of a light-harvesting 2 complex reveal local backbone perturbations induced by its higher-order interactions. *Biochemistry.* **2010**;49(3):478-486.
- Park** K-H, Berrier C, Lebaupain F, et al. Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem. J.* **2007**;403:183-187.
- Park** K-H, Berrier C, Martinac B, Ghazi A. Purification and functional reconstitution of N- and C-halves of the MscL channel. *Biophys. J.* **2004**;86:2129-2136.
- Park** K-H, Billon-Denis E, Dahmane T, et al. In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *N. Biotechnol.* **2011**;28(3):255-261.
- Park** SH, Mrse AA, Nevzorov AA, et al. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein « u » (Vpu) from HIV-1. *J. Mol. Biol.* **2003**;333(2):409-424.
- Parks** LW, Casey WM. Physiological implications of sterol biosynthesis in yeast. *Annu. Rev. Microbiol.* **1995**;49:95-116.
- Patel** AB, Crocker E, Reeves PJ, et al. Changes in interhelical hydrogen bonding upon rhodopsin activation. *J. Mol. Biol.* **2005**;347(4):803-812.
- Patel** AJ, Honore E, Lesage F, et al. Inhalational anesthetics activate two-pore-domain background K⁺ channels. *Nat. Neurosci.* **1999**;2(5):422-426.
- Pebay-Peyroula** E, Rummel G, Rosenbusch JP, Landau EM. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science.* **1997**;277(5332):1676-1681.
- Perozo** E, Cortes DM, Somporapisut P, Kloda A, Martinac B. Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature.* **2002**;418(6901):942-948.
- Perozo** E, Kloda A, Cortes DM, Martinac B. Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nat. Struct. Biol.* **2002**;9:696-703.
- Perozo** E, Rees DC. Structure and mechanism in prokaryotic mechanosensitive channels. *Curr. Opin. Struct. Biol.* **2003**;13:432-442.
- Petkova** AT, Yau W-M, Tycko R. Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry.* **2006**;45(2):498-512.
- Pfleger** N, Wörner AC, Yang J, et al. Solid-state NMR and functional studies on proteorhodopsin. *Biochim. Biophys. Acta.* **2009**;1787(6):697-705.
- Pines** A, Gibby MG, Waugh JS. Proton-enhanced nuclear induction spectroscopy ¹³C chemical shielding anisotropy in some organic solids. *Chem. Phys. Lett.* **1972**;15(3):373-376.
- Pocanschi** CL, Dahmane T, Gohon Y, et al. Amphipathic polymers: tools to fold integral membrane proteins to their active form. *Biochemistry.* **2006**;45(47):13954-13961.

- Popot** J-L. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu. Rev. Biochem.* **2010**;79:737-775.
- Powl** AM, East JM, Lee AG. Different effects of lipid chain length on the two sides of a membrane and the lipid annulus of MscL. *Biophys. J.* **2007**;93:113–122.
- Powl** AM, East JM, Lee AG. Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. *Biochemistry.* **2008 a**;47:12175–12184.
- Powl** AM, East JM, Lee AG. Anionic phospholipids affect the rate and extent of flux through the mechanosensitive channel of large conductance MscL. *Biochemistry.* **2008 b**;47:4317–4328.
- Powl** AM, Carney J, Marius P, East JM, Lee AG. Lipid interactions with bacterial channels: fluorescence studies. *Biochem. Soc. Trans.* **2005**;33:905–909.
- Pozza** A, Pérez-Victoria J, Di Pietro A. Overexpression of homogeneous and active ABCG2 in insect cells. *Protein Expr. Purif.* **2009**;63(2):75-83.
- Privé** GG. Detergents for the stabilization and crystallization of membrane proteins. *Methods.* **2007**;41(4):388-397.
- Purcell** EM, Torrey HC, Pound RV. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **1946**;69(1-2):37.
- Raghava** GPS. APSSP2 : A combination method for protein secondary structure prediction based on neural network and example based learning. *CASP5.* **2002**:A-132.
- Roberts** BE, Paterson BM. Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proc. Natl. Acad. Sci. U.S.A.* **1973**;70(8):2330-2334.
- Róg** T, Pasenkiewicz-Gierula M, Vattulainen I, Karttunen M. Ordering effects of cholesterol and its analogues. *Biochim. Biophys. Acta.* **2009**;1788(1):97-121.
- Rost** B, Sander CR. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc. Natl. Acad. Sci. U.S.A.* **1993**;90(16):7558-7562.
- Rost** B, Yachdav G, Liu J. The PredictProtein server. *Nucleic Acids Res.* **2004**;32(Web Server issue):W321-326.
- Saint** N, Lacapère J-J, Gu LQ, et al. A hexameric transmembrane pore revealed by two-dimensional crystallization of the large mechanosensitive ion channel (MscL) of *Escherichia coli*. *J. Biol. Chem.* **1998**;273(24):14667-14670.
- Saitô** H, Naito A. NMR studies on fully hydrated membrane proteins, with emphasis on bacteriorhodopsin as a typical and prototype membrane protein. *Biochim. Biophys. Acta.* **2007**;1768(12):3145-3161.
- Sambrook** J, Russell DW. *Molecular cloning: a laboratory manual*. CSHL Press; **2001**.
- Sampson** J, Mathews MB, Osborn M, Borghetti AF. Hemoglobin messenger ribonucleic acid translation in cell-free systems from rat and mouse liver and Landschutz ascites cells. *Biochemistry.* **1972**;11(19):3636-3640.
- Sanders** CR, Sönnichsen F. Solution NMR of membrane proteins: practice and challenges. *Magn Reson Chem.* **2006**;44 Spec No:S24-40.
- Schneider** B, Junge F, Shirokov VA, et al. Membrane protein expression in cell-free systems. *Methods Mol. Biol.* **2010**;601:165-186.
- Schneider** R, Ader C, Lange A, et al. Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers. *J. Am. Chem. Soc.* **2008**;130(23):7427-7435.

- Schubert** M, Manolikas T, Rogowski M, Meier BH. Solid-state NMR spectroscopy of 10% C-13 labeled ubiquitin: spectral simplification and stereospecific assignment of isopropyl groups. *J. Biomol. NMR*. **2006**;35(3):167-173.
- Seidel** K, Andronesi OC, Krebs J, et al. Structural characterization of Ca²⁺-ATPase-bound phospholamban in lipid bilayers by solid-state nuclear magnetic resonance (NMR) Spectroscopy. *Biochemistry*. **2008**;47(15):4369-4376.
- Shi** L, Lake EMR, Ahmed MAM, Brown LS, Ladizhansky V. Solid-state NMR study of proteorhodopsin in the lipid environment: secondary structure and dynamics. *Biochim. Biophys. Acta*. **2009**;1788(12):2563-2574.
- Shon** KJ, Kim Y, Colnago LA, Opella SJ. NMR studies of the structure and dynamics of membrane-bound bacteriophage Pf1 coat protein. *Science*. **1991**;252(5010):1303-1305.
- Sillero** A, Ribeiro JM. Isoelectric points of proteins: theoretical determination. *Anal. Biochem*. **1989**;179(2):319-325.
- Singer** SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. *Science*. **1972**;175(23):720-731.
- Sperling** LJ, Berthold DA, Sasser TL, Jeisy-Scott V, Rienstra CM. Assignment strategies for large proteins by magic-angle spinning NMR: the 21-kDa disulfide-bond-forming enzyme DsbA. *J. Mol. Biol*. **2010**;399(2):268-282.
- Spirin** AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*. **1988**;242(4882):1162-1164.
- Stejskal** EO, Schaefer J. Removal of artifacts from cross-polarization NMR experiments. *J. Magn. Reson*. **1975**;18(3):560-563.
- Subbarao** GV, van den Berg B. Crystal structure of the monomeric porin OmpG. *J. Mol. Biol*. **2006**;360(4):750-759.
- Sukharev** S, Corey DP. Mechanosensitive channels: multiplicity of families and gating paradigms. *Sci. STKE*. **2004**;2004:re4.
- Sukharev** S, Betanzos M, Chiang CS, Guy HR, Sukharev S. The gating mechanism of the large mechanosensitive channel MscL. *Nature*. **2001**;409:720-724.
- Sukharev** S, Durell SR, Guy HR, Sukharev S. Structural models of the MscL gating mechanism. *Biophys. J*. **2001**;81:917-936.
- Sukharev** SI, Schroeder MJ, McCaslin DR, Sukharev SI. Stoichiometry of the large conductance bacterial mechanosensitive channel of E. coli. A biochemical study. *J. Membr. Biol*. **1999**;171:183-193.
- Sukharev** SI, Blount P, Martinac B, Blattner FR, Kung C. A large-conductance mechanosensitive channel in E. coli encoded by mscL alone. *Nature*. **1994 a**;368:265-268.
- Sukharev** SI, Martinac B, Blount P, Kung C. Functional Reconstitution as an Assay for Biochemical Isolation of Channel Proteins: Application to the Molecular Identification of a Bacterial Mechanosensitive Channel. *Methods*. **1994 b**;6(1):51-59.
- Sukharev** SI, Martinac B, Arshavsky VY, Kung C. Two types of mechanosensitive channels in the Escherichia coli cell envelope: solubilization and functional reconstitution. *Biophys. J*. **1993**;65(1):177-183.
- Surewicz** WK, Mantsch HH, Chapman D. Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry*. **1993**;32(2):389-394.
- Suryanarayana** T, Subramanian AR. An essential function of ribosomal protein S1 in messenger ribonucleic acid translation. *Biochemistry*. **1983**;22(11):2715-2719.
- Suzuki** H, Kerr R, Bianchi L, et al. In vivo imaging of C-elegans mechanosensory neurons demonstrates a specific role for the MEC-4 channel in the process of gentle touch sensation. *Neuron*. **2003**;39(6):1005-1017.

- Swartz** J. Developing cell-free biology for industrial applications. *J. Ind. Microbiol. Biotechnol.* **2006**;33(7):476-485.
- Takegoshi** K, Nakamura S, Terao T. ¹³C-¹H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **2001**;344(5-6):631-637.
- Tanford** C, Reynolds JA. Characterization of membrane proteins in detergent solutions. *Biochim. Biophys. Acta.* **1976**;457(2):133-170.
- Tang** Y, Yoo J, Yethiraj A, Cui Q, Chen X. Mechanosensitive channels: insights from continuum-based simulations. *Cell Biochem. Biophys.* **2008**;52:1-18.
- Tate** C. Biochemistry. Membrane protein gymnastics. *Science.* **2010**;328(5986):1644-1645.
- Teilum** K, Brath U, Lundström P, Akke M. Biosynthetic ¹³C labeling of aromatic side chains in proteins for NMR relaxation measurements. *J. Am. Chem. Soc.* **2006**;128(8):2506-2507.
- Tissieres** A, Watson JD, Schlessinger D, Hollingworth BR. Ribonucleoprotein particles from Escherichia coli. *J. Mol. Biol.* **1959**;1(3):221-233.
- Topper** JN, Gimbrone MA. Blood flow and vascular gene expression: fluid shear stress as a modulator of endothelial phenotype. *Mol. Med. Today.* **1999**;5(1):40-46.
- Traaseth** NJ, Shi L, Verardi R, et al. Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. *Proc. Natl. Acad. Sci. U.S.A.* **2009**;106(25):10165-10170.
- Triba** MN, Devaux PF, Warschawski DE. Effects of lipid chain length and unsaturation on bicelles stability. A phosphorus NMR study. *Biophys. J.* **2006**;91(4):1357-1367.
- Ujwal** R, Cascio D, Colletier J-P, et al. The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105(46):17742-17747.
- Valadié** H, Lacapère J-J, Sanejouand YH, Etchebest C. Dynamical properties of the MscL of Escherichia coli: a normal mode analysis. *J. Mol. Biol.* **2003**;332:657-674.
- van den Bogaart** G, Krasnikov V, Poolman B. Dual-color fluorescence-burst analysis to probe protein efflux through the mechanosensitive channel MscL. *Biophys. J.* **2007**;92:1233-1240.
- van Gammeren** AJ, Hulsbergen FB, Hollander JG, de Groot HJM. Biosynthetic site-specific (¹³C) labeling of the light-harvesting 2 protein complex: a model for solid state NMR structure determination of transmembrane proteins. *J. Biomol. NMR.* **2004**;30(3):267-274.
- van Horn** WD, Kim H-J, Ellis CD, et al. Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science.* **2009**;324(5935):1726-1729.
- van Meer** G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**;9(2):112-124.
- van Melckebeke** H, Wasmer C, Lange A, et al. Atomic-resolution three-dimensional structure of HET-s(218-289) amyloid fibrils by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2010**;132(39):13765-13775.
- Varga** K, Aslimovska L, Watts A. Advances towards resonance assignments for uniformly-¹³C, ¹⁵N enriched bacteriorhodopsin at 18.8 T in purple membranes. *J. Biomol. NMR.* **2008**;41(1):1-4.
- Venters** RA, Calderone TL, Spicer LD, Fierke CA. Uniform ¹³C isotope labeling of proteins with sodium acetate for NMR studies: application to human carbonic anhydrase II. *Biochemistry.* **1991**;30(18):4491-4494.
- Vinarov** DA, Lytle BL, Peterson FC, et al. Cell-free protein production and labeling protocol for NMR-based

- structural proteomics. *Nat. Methods*. **2004**;1(2):149-153.
- Vosegaard T**, Kamihira-Ishijima M, Watts A, Nielsen NC. Helix conformations in 7TM membrane proteins determined using oriented-sample solid-state NMR with multiple residue-specific ¹⁵N labeling. *Biophys. J.* **2008**;94(1):241-250.
- Vuister GW**, Kim SJ, Wu C, Bax A. 2D and 3D NMR-study of phenylalanine residues in proteins by reverse isotopic labeling. *J. Am. Chem. Soc.* **1994**;116(20):9206-9210.
- Wang J**, Kim S, Kovacs F, Cross TA. Structure of the transmembrane region of the M2 protein H(+) channel. *Protein Sci.* **2001**;10(11):2241-2250.
- Warschawski DE**. Membrane Proteins Of Known Structure Determined By NMR. Available at: <http://www.drorlist.com/nmr/MPNMR.html>
- White S**. Membrane Proteins of Known Structure. *Membrane Proteins Of Known 3D Structure*. Available at: http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html
- Williamson MP**, Havel TF, Wüthrich K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ¹H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **1985**;182(2):295-315.
- Wozniak MA**, Chen CS. Mechanotransduction in development: a growing role for contractility. *Nat. Rev. Mol. Cell Biol.* **2009**;10(1):34-43.
- Wuthier RE**. Two-dimensional chromatography on silica gel-loaded paper for the microanalysis of polar lipids. *J. Lipid Res.* **1966**;7(4):544-550.
- Yamaguchi S**, Tuzi S, Bowie JU, Saitô H. Secondary structure and backbone dynamics of Escherichia coli diacylglycerol kinase, as revealed by site-directed solid-state ¹³C NMR. *Biochim. Biophys. Acta.* **2004**;1698(1):97-105.
- Yeagle P**. The Structure of biological membranes. CRC Press; **2005**.
- Yefimov S**, van der Giessen E, Onck PR, Marrink SJ. Mechanosensitive membrane channels in action. *Biophys. J.* **2008**;94:2994–3002.
- Yoshimura K**, Batiza A, Kung C. Chemically charging the pore constriction opens the mechanosensitive channel MscL. *Biophys. J.* **2001**;80:2198–2206.
- Yoshimura K**, Batiza A, Schroeder M, Blount P, Kung C. Hydrophilicity of a single residue within MscL correlates with increased channel mechanosensitivity. *Biophys. J.* **1999**;77:1960–1972.
- Yoshimura K**, Usukura J, Sokabe M. Gating-associated conformational changes in the mechanosensitive channel MscL. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105:4033–4038.
- Yushmanov VE**, Mandal PK, Liu Z, Tang P, Xu Y. NMR structure and backbone dynamics of the extended second transmembrane domain of the human neuronal glycine receptor alpha1 subunit. *Biochemistry.* **2003**;42(13):3989-3995.
- Zeeman P**. The Effect of Magnetisation on the Nature of Light Emitted by a Substance. *Nature.* **1897**;55(1424):347-347.
- Zhang Y**, Doherty T, Li J, et al. Resonance assignment and three-dimensional structure determination of a human alpha-defensin, HNP-1, by solid-state NMR. *J. Mol. Biol.* **2010**;397(2):408-422.
- Zhao G**, London E. An amino acid « transmembrane tendency » scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity. *Protein Sci.* **2006**;15(8):1987-2001.

Références

Zoonens M, Catoire LJ, Giusti F, Popot J-L. NMR study of a membrane protein in detergent-free aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.* **2005**;102(25):8893-8898.

Zubay G. In vitro synthesis of protein in microbial systems. *Annu. Rev. Genet.* **1973**;7:267-287.

Matériels et Méthodes expérimentales

B Matériels et Méthodes expérimentales

Dans cette partie seront exposés successivement le matériel utilisé puis les protocoles expérimentaux nécessaires pour la production d'un échantillon de la protéine membranaire *Ec-MscL*, sa purification et sa reconstitution en liposomes, afin de l'étudier *in situ* par RMN en phase solide. Les différentes techniques employées pour analyser l'échantillon qualitativement et quantitativement, sont également détaillées.

Le gène codant pour *Ec-MscL* est cloné dans un plasmide de type *pIVEX-2.3*. La production de la protéine recombinante est réalisée avec le système d'ARN polymérase T7 (Studier & Moffatt, 1986) dans des bactéries *E. coli* de souche *BL21 (DE3)*. Le système T7 est également utilisé pour la production de la protéine *in vitro*. Une étiquette de 6 histidines est rajoutée au niveau C-terminal de la protéine permettant la purification de la protéine en une seule étape sur une colonne d'affinité au nickel. La protéine est reconstituée en vésicules multilamellaires géantes (*MLVs pour MultiLamellar Vesicles*) et les protéoliposomes sont transférés dans le rotor utilisé en RMN. Il s'agit ici des rotors de 4 mm de diamètre dont la capacité est de 50 μ L. Les rotors contenant l'échantillon sont conservés à -20°C jusqu'à l'acquisition des données au niveau du spectromètre.

La dernière partie est consacrée aux expériences de mise au point effectuées au niveau du spectromètre, ainsi que les méthodes utilisées pour les mesures de corrélations C-C et C-N au niveau de la protéine marquée. Enfin, les méthodes utilisées pour l'analyse et le traitement des données obtenues au niveau du spectromètre sont expliquées.

I Matériels

Une liste exhaustive du matériel utilisé pour les expériences décrites ultérieurement est citée ici. Il s'agit d'une notice permettant de regrouper par type d'expérience les produits et les outils nécessaires. Les marques des produits usuels ne sont pas précisées.

a Surexpression bactérienne

i. Transformation des bactéries

Bactéries électrocompétentes *BL21 (DE3)* conservées à -80°C (*STRATAGENE*)
Plasmide *pIVEX 2.3 -mscL-(H₆)* conservé à -20°C (*ROCHE*)
Milieux *Lysogeny Broth* lyophilisés *LB - LB-Agar*
Ampicilline sodique - Glycérol à pureté analytique
Électroporateur - Cuve d'électroporation conservée à -20°C
Équipement de préparation d'*ADN MIDI-PREP*
Boîtes de *PÉTRI* - Erlenmeyers (*200 mL, 500 mL, 1 L*) - Tubes *FALCON (15 et 50 mL)*
Tubes *EPPENDORF (1,5 mL)* - Incubateur - Machine d'autoclave

ii. Vérification des transformants

Milieux *Lysogeny Broth* lyophilisés *LB, LB-Agar*
Ampicilline sodique - Isopropyl- β -*D*-1-thiogalactopyranoside *IPTG*
Incubateur - Centrifugeuse de paillasse
Bec *BUNSEN* - Éthanol dénaturé - Râteau

iii. Surexpression de la protéine *MscL* en bactéries

Pré-culture en milieu riche

Milieu *Lysogeny Broth* lyophilisés *LB* - Erlenmeyers - Incubateur
Ampicilline sodique - Machine d'autoclave

Culture en milieu minimum *M9*

Chlorure d'ammonium $^{14}\text{NH}_4\text{Cl}$ - Glucose $^{12}\text{C}_6\text{H}_{12}\text{O}_6$

Chlorure d'ammonium marqué au ^{15}N $^{15}\text{NH}_4\text{Cl}$ - Glucose marqué au ^{13}C $^{13}\text{C}_6\text{H}_{12}\text{O}_6$ (*ISOTEC*)

Phosphate de potassium monobasique KH_2PO_4 - Phosphate de potassium dibasique K_2HPO_4 -

Phosphate de sodium dibasique Na_2HPO_4 - Sulfate de potassium K_2SO_4 - Chlorure de calcium $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ - Sulfate de magnésium MgSO_4 (Pureté analytique)

Thiamine hydrochloride - Nicotinamide - Ampicilline sodique (Qualité appropriée pour la culture

cellulaire) - Extraits de levure

Acide éthylène diamine tétraacétique *EDTA* (Pureté analytique)

Isopropyl- β -D-1-thiogalactopyranoside *IPTG* (Pureté analytique)

Sulfate ferreux $FeSO_4 \cdot 7H_2O$ - Chlorure de manganèse $MnCl_2 \cdot 4H_2O$ - Chlorure de cobalt

$CoCl_2 \cdot 6H_2O$ - Sulfate de zinc $ZnSO_4 \cdot 7H_2O$ - Chlorure de cuivre $CuCl_2 \cdot 2H_2O$ - Acide borique

H_3BO_3 - Molybdate d'ammonium $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (Pureté analytique)

Centrifugeuse - Tubes de centrifugation - Rotors *JA-10.5000*, *JA-12.500*

Erlenmeyer 5 L - Filtres 0,45 μm - Seringues

Incubateur - Machine d'autoclave - Azote liquide

Composant	Pour 1 L	
À autoclaver		
KH_2PO_4	13 g	Dans 970 mL À pH 6,8
K_2HPO_4	10 g	
Na_2HPO_4	9 g	
K_2SO_4	2,4 g	
$^{15}NH_4Cl$	1 g	
À filtrer (et à rajouter séparément)		
$CaCl_2 \cdot 2 H_2O$	600 mg	
$MgSO_4$	10 mL à 1 M	
Thiamine	30 mg	
Nicotinamide	30 mg	
Ampicilline	100 mg	
Extraits de levure	10 μL d'une solution à 10%	
$^{13}C_6$ -Glucose	1 à 4 g	
Solution <i>TE</i> (voir tableau 2)	10 mL	

Tableau 1 : Composition du milieu minimum *M9*.
(Sambrook et Russel, 2001 ; Cai *et al.*, 1998)

Composition de 100 mL (100x) de la solution d'oligo-éléments (<i>TE</i> pour <i>Trace Elements</i>)	
<i>EDTA</i> ($2Na^+$)	500 mg
$FeSO_4 \cdot 7 H_2O$	600 mg
$MnCl_2 \cdot 4 H_2O$	120 mg
$CoCl_2 \cdot 6 H_2O$	80 mg
$ZnSO_4 \cdot 7 H_2O$	70 mg
$CuCl_2 \cdot 2 H_2O$	30 mg
H_3BO_3	2 mg
$(NH_4)_6Mo_7 \cdot 4 H_2O$	25 mg

Tableau 2 : Composition d'une solution d'oligoéléments *TE*.

iv. Broyage des cellules

Presse de *FRENCH*

Spectrophotomètre - Cuves de spectrométrie

Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation
 Phosphate de sodium monobasique NaH_2PO_4 - Chlorure de sodium $NaCl$ - Sulfate de magnésium $MgSO_4$ - Saccharose - Fluorure de phenylméthanesulfonyle $PMSF$ (PIERCE)
 Fluorure du 4-(2-aminoéthyl)-benzènesulfonyl hydrochloride $AEBSF$ (PIERCE)
 $ADNase$ de type I de pancréas bovin - β -Mercaptoéthanol β -MetOH - Dithiothréitol DTT

Composant	Concentration
NaH_2PO_4	50 mM
$MgSO_4$	2 mM
$NaCl$	10 mM
Saccharose	5%
À rajouter 30 min avant l'utilisation du tampon	
$AEBSF$	1 mM
$ADNase$	10 μ g/mL
β -MetOH	2 mM
DTT	1 mM

Tableau 3 : Composition d'une solution de tampon FP (pour FRENCH Press)

v. Solubilisation des membranes

Phosphate de sodium dibasique Na_2HPO_4 - Phosphate de sodium monobasique NaH_2PO_4 - Chlorure de sodium $NaCl$ - Imidazole (Pureté analytique)
 Fluorure du 4-(2-aminoéthyl)-benzènesulfonyl hydrochloride $AEBSF$
 Triton-x-100 - Filtres 0,45 μ m - Seringues
 Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation

Composant	Concentration
Tampon Phosphate (pH 7,5)	50 mM
$NaCl$	300 mM
Imidazole	10 mM
Et après resuspension :	
$AEBSF$	1 mM
Triton-x-100	0,2 – 4,0 %
200 mM de tampon phosphate à pH 7,5 est obtenu en mélangeant 4/5 de Na_2HPO_4 à 0,2 M et 1/5 de NaH_2PO_4 0,2 M (Dawson <i>et al.</i> , 1989)	

Tableau 4 : Composition du tampon d'équilibration EB (pour Equilibration buffer)

vi. Purification de la protéine $MscL$

Phosphate de sodium dibasique Na_2HPO_4 - Chlorure de sodium $NaCl$
 Colonne Ni -NTA $HiTrap$ conservée dans 20% d'éthanol à 4°C (AMERSHAM BIOSCIENCES)
 Système de chromatographie liquide pour la purification des protéines FPLC (GE HEALTHCARE - AKTA) muni d'un collectionneur de fractions - Tubes pour collectionneur de fractions
 Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation

Filtres $0,45 \mu\text{m}$ - Seringues - Membranes PVDF
 2-amino-2-hydroxyméthyl-1,3-propanediol Tris - Triton-x-100
 Chlorure de sodium NaCl - Tween-20 - Lait en poudre
 Anticorps monoclonaux anti-histidine couplés à une activité peroxydase (SIGMA)
 Tablettes DAB SIGMAFAST pour la révélation (SIGMA)

Composant	Concentration		
Na_2HPO_4	50 mM		
NaCl	300 mM		
	A1	A2	B
Triton-X-100	4.0 % (p/v)	0.2 % (p/v)	0.2 % (p/v)
Imidazole	10 mM	10 mM	500 mM

Tableau 5 : Tampons de purification utilisés en FPLC

vii. Reconstitution de la protéine *MscL* en liposomes

Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique HEPES
 Chlorure de potassium KCl - Triton-x-100
 Cassette de dialyse SLIDE-A-LYZER, 10K MWCO 3–12 mL (PIERCE)
 Bêchers 2 L - Bain d'ultrasons
 Kits de dosage au BCA - Albumine de sérum bovin BSA
 Dioléoyl-phosphatidylcholine DOPC (AVANTI POLAR LIPIDS) - Chloroforme
 Extrudeur - Seringues (1 mL) - Membranes en polycarbonate $0,1 \mu\text{m}$ (AVANTI POLAR LIPIDS)
 Bio-billes en polystyrène (SM-2 BIO-RAD)
 Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation
 Fluorure du 4-(2-aminoéthyl)-benzènesulfonyl hydrochloride AEBSF

Tampon de reconstitution RB : HEPES 10 mM, KCl 100 mM à pH 7,4-7,6

Tampon de dialyse RB-T = RB + triton-x-100 à 0,2%

b Synthèse *in vitro*

Ce chapitre est composé de deux parties pour séparer la synthèse de *MscL in vitro* en micelles de détergent puis sa reconstitution en liposomes de la synthèse de la protéine directement en liposomes.

Acide aminé	PM_i (Da)	n	Poids (mg)	Volume (mL)
<i>Ala</i> (A)	89	15	1,7	1,1
<i>Arg</i> (R)	174	6	1,3	0,45
<i>Asn</i> (N)	132	6	1,0	0,45
<i>Asp</i> (D)	133	7	1,2	0,52
<i>Cys</i> (C)	121	0	0	0,10
<i>Gln</i> (Q)	146	4	0,73	0,30
<i>Glu</i> (E)	147	8	1,5	0,60
<i>Gly</i> (G)	75	13	1,2	0,97
<i>His</i> (H)	155	7	1,4	0,52
<i>Ile</i> (I)	131	16	2,6	1,2
<i>Leu</i> (L)	131	13	2,1	1,2
<i>Lys</i> (K)	146	9	1,6	0,67
<i>Met</i> (M)	149	5	0,93	0,37
<i>Phe</i> (F)	131	10	1,6	0,74
<i>Pro</i> (P)	115	7	1,0	0,52
<i>Ser</i> (S)	105	4	0,53	0,30
<i>Thr</i> (T)	119	3	0,45	0,22
<i>Trp</i> (W)	204	0	0	0,10
<i>Tyr</i> (Y)	181	1	0,23	0,10
<i>Val</i> (V)	117	13	1,9	0,97

Tableau 6 : Composition en acides aminés de la protéine *MscL*. Calcul des quantités d'acides aminés nécessaires pour la *SIV*, en surévaluant à 20 mg la quantité de protéines obtenues dans 10 mL. n étant le nombre d'apparition de chaque acide aminé et PM_i le poids moléculaire. Les valeurs de la colonne Poids sont à utiliser lorsqu'il s'agit d'un acide aminé marqué aux ^{13}C et ^{15}N . Celles de la colonne volume lorsqu'il s'agit d'une solution d'acide aminé fournie dans le kit *AMINO ACID SAMPLER*.

a Synthèse *in vitro* de MscL en micelles de détergent et reconstitution en liposomes

i. Synthèse *in vitro* en micelles de détergent

Kits *RTS-9000* conservés à -80°C - Kits *AMINO ACID SAMPLER* : Acides aminés et dithiothréitol
DTT conservés à -20°C (*ROCHE*; *5PRIME/VWR*)
Acides aminés marqués au $^{13}\text{C}/^{15}\text{N}$ conservés à -20°C (*CAMBRIDGE ISOTOPE LABORATORIES*)
Agitateur thermostaté *THERMOMIXER COMFORT* (*EPPENDORF*)
Plasmide *pIVEX 2.3 -mscL-(H₆)* conservé à -20°C (*ROCHE*)
Eau ultrapure - *Triton-X-100* - Pipettes *PASTEUR* autoclavées
Tubes *FALCON* (50 mL) - Tubes *EPPENDORF* (1,5 mL) - Bain d'ultrasons
Pipettes (0 - 10 μL , 10 - 200 μL , 200 - 1000 μL et 5 mL) autoclavées
Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation

ii. Purification de la protéine

Phosphate de sodium dibasique Na_2HPO_4 - Chlorure de sodium *NaCl*
Colonne *Ni-NTA HiTrap* conservée dans 20% d'éthanol à 4°C (*AMERSHAM BIOSCIENCES*)
Système de chromatographie liquide pour la purification des protéines *FPLC* (*GE HEALTHCARE* - *AKTA*) muni d'un collectionneur de fractions - Tubes pour collectionneur de fractions
Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation
Filtres 0,45 μm - Seringues - Membranes *PVDF*
2-amino-2-hydroxyméthyl-1,3-propanediol *Tris* (*SIGMA*) - *Triton-x-100*
Chlorure de sodium *NaCl* - *Tween-20* - Lait en poudre
Anticorps monoclonaux anti-histidine couplés à une activité peroxydase (*SIGMA*)
Tablettes *DAB SIGMAFAST* pour la révélation (*SIGMA*)

iii. Reconstitution de la protéine en liposomes

Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique *HEPES*
Chlorure de potassium *KCl* - *Triton-x-100*
Cassette de dialyse *SLIDE-A-LYZER*, 10K *MWCO* 3-12 mL (*PIERCE*)
Béchers 2 L - Bain d'ultrasons
Kits de dosage au *BCA* - Albumine de sérum bovin *BSA*
Dioléoyl-phosphatidylcholine *DOPC* (*AVANTI POLAR LIPIDS*) - Chloroforme
Extrudeur - Seringues (1 mL) - Membranes en polycarbonate 0,1 μm (*AVANTI POLAR LIPIDS*)
Bio-billes en polystyrène (*SM-2 BIO-RAD*)
Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation
Fluorure du 4-(2-aminoéthyl)-benzènesulfonyl hydrochloride *AEBSF* (*PIERCE*)
Diméthyl- β -cyclodextrine (*SIGMA*) - Extraits lipidiques d'*E. coli* (*AVANTI POLAR LIPIDS*)
1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine *DPPC* - Asolectine (*AVANTI POLAR LIPIDS*)
1-(9Z-octadecenoyl)-*sn*-glycero-3-phosphocholine *L₉so-PC* (*AVANTI POLAR LIPIDS*)

b Synthèse *in vitro* de MscL en liposomes

Kits *RTS-9000* conservés à -80°C - Kits *AMINO ACID SAMPLER* : Acides aminés et dithiothréitol
DTT conservés à -20°C (*ROCHE*; *5PRIME/VWR*)
Acides aminés marqués au $^{13}\text{C}/^{15}\text{N}$ conservés à -20°C (*CAMBRIDGE ISOTOPE LABORATORIES*)
Agitateur thermostaté *THERMOMIXER COMFORT* (*EPPENDORF*)
Plasmide *pIVEX 2.3 -mscL(H₆)* conservé à -20°C (*ROCHE*)
Eau ultrapure - *Triton-X-100* - Pipettes *PASTEUR* autoclavées
Tubes *FALCON* (50 mL) - Tubes *EPPENDORF* (1,5 mL) - Bain d'ultrasons
Pipettes (0 - 10 μL , 10 - 200 μL , 200 - 1000 μL et 5 mL) autoclavées
Centrifugeuse - Ultracentrifugeuse - Rotors - Tubes de centrifugation
Dioléoyl-phosphatidylcholine *DOPC* (*AVANTI POLAR LIPIDS*) - Chloroforme
Extrudeur - Seringues (1 mL) - Membranes en polycarbonate 0,1 μm (*AVANTI POLAR LIPIDS*)
Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique *HEPES*
Chlorure de potassium *KCl*

c Analyse de l'échantillon obtenu

i. Électrophorèse sur gel de polyacrylamide - *SDS-PAGE*

Équipement standard pour la migration des gels - Gels *SDS-PAGE* 8-20% pré-coulés
Tris-(hydroxy-méthyl)-amino-méthane *TRIS*
Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique *HEPES*
Sodium dodécyl sulfate *SDS* - Colorant *COOMASSIE BRILLIANT BLUE R-250*
Éthanol - acide acétique - glycérol - bleu de bromophénol - dithiothréitol *DTT*
Tampon de migration *SRB* (pour *SDS Running Buffer*): *Tris* 100 mM, *HEPES* 100 mM, *SDS* 3 mM, le *pH* doit être ajusté à 8
Tampon de coloration des gels *SB* (pour *Staining Buffer*) composé de 0,5% de colorant *COOMASSIE*, 50% d'éthanol et 10% d'acide acétique
Tampon de décoloration des gels *DSB* (pour *DeStaining Buffer*) composé de 20% d'éthanol et 10% d'acide acétique
Tampon de charge *SLB* (pour *SDS Loading Buffer*) 2x composé de 8% *SDS*, de 40% glycérol, 0,4% de bleu de bromophénol, 400 mM *DTT*, dans une solution de *Tris* à 100 mM *pH* 8

ii. Immunodétection *WESTERN BLOT* / *DOT BLOT*

WESTERN BLOT

Équipements standard pour immunodétection par luminescence et de développement de photo
Kit commercial de détection par *WESTERN BLOT* (réactifs de luminescence *luminol* et H_2O_2)
Membrane de nitrocellulose - Papiers-filtres *WHATMAN* - Lait en poudre
Tris - Glycine - Isopropanol - *PONCEAU S* - *SDS* - Acide acétique - *NaCl* - *Tween-20*
Anticorps anti-histidine de souris - Anticorps anti-souris couplés à la peroxydase
Tampon de transfert de *TOWBIN TTB* (pour *TOWBIN Transfer Buffer*) composé de *Tris* à 25 mM, de glycine à 192 mM, d'isopropanol à 20%, et de *SDS* à 0,1%. La solution est conservée à 4°C sans

isopropanol, celui-ci est rajouté avant utilisation

Solution de *PONCEAU S* à 0,1% dans 5% d'acide acétique

Tampon *TBS-T* (pour *Tris-Buffered Saline - Tween-20*) composé de *Tris* à 50 mM, de *NaCl* à 150 mM et de 0,05 % *Tween-20* à pH 7,4

Tampon *TBS-TM* (pour *Tris-Buffered Saline - Tween-20 and Milk*) composé de *TBS-T* contenant 5% de lait en poudre

Solution d'anticorps *anti-His* diluée au 1/5 000^e dans du *TBS-T*

Solution d'anticorps anti-souris diluée au 1/10 000^e dans du *TBS-TM*

DOT BLOT

Membrane *PVDF - NaCl - Tris - Tween-20*

Anticorps monoclonaux *anti-His* couplés à la peroxydase (*SIGMA*)

Tablettes de diaminobenzidine *DAB* – tablettes de peroxyde d'hydrogène H_2O_2 et urée (*SIGMA*)

Tampon *TBS-T* (pour *Tris-Buffered Saline - Tween-20*) composé de *Tris* à 50 mM, de *NaCl* à 150 mM et de 0,05 % *Tween-20* à pH 7,4

Solution d'anticorps *anti-His*-peroxydase diluée au 1/2 000^e dans du *TBS-T*

Solutions de *DAB* et d'urée et H_2O_2 préparées séparément et mélangées avant utilisation

iii. Quantification de la protéine - dosage au *BCA*

Kit de *BCA* (*PIERCE*) - spectrophotomètre - solution de *BSA* (*Bovine Serum Albumine*)

Acétone (conservé à -20°C)

À partir d'une solution-mère de *BSA* de 2 mg/mL, huit échantillons à des concentrations entre 0 et 100 µg/mL sont préparés afin de les utiliser pour étalonner le dosage au *BCA*. La solution utilisée pour les dilutions doit être identique à celle dans laquelle se trouve la protéine Mscl.

iv. Réticulation covalente ou *cross-linking*

Disuccinimidyl-subérate (*DSS*) - Formaldéhyde (*FA*) à 37%

Dimeéthyl-sulfoxyde (*DMSO*) - Solution de *Tris-HCl* (1 M *Tris* à pH 8)

(Ainsi que le matériel cité dans la section Électrophorèse sur gel de polyacrylamide – *SDS-PAGE*)

v. Quantification du phosphore - Méthode de *ROUSER*

Bloc chauffant - spectrophotomètre - Étuve - Tubes en verre

Acide nitrique HNO_3 - Acide perchlorique $HClO_4$

Molybdate d'ammonium $(NH_4)_6Mo_7O_{24}.4H_2O$

Phosphate de potassium monobasique KH_2PO_4

Solutions d'acide nitrique (1%) et d'acide perchlorique concentrée (70%) conservées sous la hotte

Solution stock de molybdate d'ammonium (25 g/L) conservée à 4°C (solution stable)

Solution d'acide ascorbique (100 g/L) conservée à 4°C (très instable, à ne pas conserver au-delà d'une semaine) - Solution stock de phosphate inorganique (KH_2PO_4 à 0,1 g/L) conservée à 4°C

Les tubes en verre sont rincés avec la solution d'acide nitrique et séchés à l'étuve avant utilisation

À partir de la solution-mère de KH_2PO_4 , cinq échantillons contenant de 0 à 5 mg de KH_2PO_4 par tube sont préparés

vi. Quantification de l'eau

Tubes *EPPENDORF* - Pompe à vide

vii. Caractérisation de la densité des protéoliposomes - Gradient de saccharose

Tubes de centrifugation - Saccharose - *HEPES* - *KCl*

Tampon *RB* - Solutions de saccharose à 15, 30 et 45% dans le tampon *RB*

d Spectroscopie de résonance magnétique nucléaire

Spectromètre de *RMN* haute résolution dans les solides (16,4 T - *AVANCE II* - *BRUKER*)

Sonde de mesure *CP-MAS* à triple détection ^1H - ^{13}C - ^{15}N (*BRUKER*)

Rotor de 4 mm de diamètre (50 μL de volume) en oxyde de zirconium ZrO_2 - bouchon en chloro-trifluoro-éthylène (*Kel-F*) (*CORTECNET*)

Bromure de potassium *KBr* - Adamantane - H_2O - un ou plusieurs acides aminés en poudre pour les réglages

II Méthodes expérimentales

a Surexpression bactérienne

i. Transformation des bactéries

Le vecteur d'expression *pIVEX 2.3* utilisé dans ces travaux contient le gène *mscL*, conférant une étiquette poly-histidine (*hexa-His*) en C-terminal à la protéine d'intérêt *MscL*. (Rogé & Betton 2005). La souche d'expression *E. coli BL21 (DE3)* employée dans cette étude comporte le système d'expression du promoteur *T7* leur permettant de synthétiser la protéine d'intérêt de manière contrôlée. Le plasmide est amplifié et purifié suivant le protocole standard *MIDI-PREP*. Il est gardé à -20°C . Les souches d'expression sont conservées à -80°C dans 20% de glycérol.

Le plasmide est introduit dans les bactéries *BL21* par électroporation. Pour cela, un tube contenant les bactéries compétentes est retiré du congélateur et placé sur la glace. Le plasmide (10 à 50 ng) est mis en contact avec les bactéries. Ensuite le mélange est placé dans une cuve d'électroporation préalablement refroidie à 4°C . Une impulsion de 2,5 kV est appliquée au niveau de l'électroporateur. Les bactéries transformées sont immédiatement retirées de la cuve et rajoutées à 900 μL de milieu *LB* à température ambiante, dans un tube de 15 mL. Le tube est placé dans l'incubateur pendant 1 heure à 37°C et en fixant la rotation du plateau à 180 rpm. Ceci permet de régénérer et multiplier les bactéries. Les bactéries ayant incorporé le plasmide referment leur paroi et expriment le gène de résistance à l'ampicilline.

Après 1 h de croissance, les bactéries sont étalées sur un milieu de culture solide et sélectif (*LB-Agar* + ampicilline) préalablement coulé dans une boîte de *PÉTRI*. La sélectivité du milieu est assurée par la présence d'ampicilline à 100 $\mu\text{g.mL}^{-1}$. La boîte contenant les bactéries étalées est incubée pendant la nuit à 37°C .

ii. Vérification des transformants

Pour confirmer la présence du plasmide dans les bactéries ayant formé des colonies, un test d'induction de l'expression de *MscL* est réalisé. Sur la boîte de *PÉTRI*, plusieurs colonies sont distinguables (les bactéries *E. coli* forment des colonies jaune opaque dont le centre est légèrement plus foncé, de 1 à 2 mm de diamètre). Chaque colonie choisie est utilisée pour ensemercer 5 mL de milieu liquide sélectif (*LB* + ampicilline) dans un tube *FALCON* de 50 mL. Les milieux sont incubés à 37°C pendant 12 h.

Pour le test d'induction, deux tubes *FALCON* de 50 mL contenant 14 mL de milieu liquide sélectif sont préparés. Dans chacun des deux tubes, 1 mL de la culture précédente est utilisé. Après 3 h d'incubation à 37°C, l'induction est réalisée dans un des deux tubes, par addition d'*IPTG* à une concentration finale de 1 mM. L'autre tube sert de référence ou témoin négatif d'induction. Différentes températures et temps d'induction ont été testés, mais ne seront pas détaillés ici. À 37°C, les cultures sont analysées après 3 h d'induction.

À la fin de la période d'induction, la densité optique des références ainsi que des cultures induites est mesurée. Pour chacun des tubes, un calcul est effectué de manière à prélever un volume contenant 100 µg de bactéries, en utilisant la règle suivante : pour une *DO* équivalente à 2, à la longueur d'onde de 600 nm ($DO_{600} = 2$), 1 mL de culture contient 1 mg de bactéries. Ces prélèvements sont centrifugés (5 minutes à 16 000 g) et les culots sont repris dans 100 µL de tampon de dépôt. Les différents échantillons sont alors analysés par *SDS-PAGE* ; l'induction et la surexpression de *MscL* sont qualitativement estimées par *WESTERN BLOT* en utilisant des anticorps primaires de souris dirigés contre l'étiquette histidine, et des anticorps secondaires dirigés contre les anticorps de souris couplés à une activité peroxydase.

iii. Surexpression de la protéine *MscL* en bactéries

Pré-culture en milieu riche

À partir des résultats de gel *SDS-PAGE* et de *WESTERN BLOT*, une souche bactérienne exprimant *MscL* est sélectionnée. La colonie correspondante au niveau de la boîte de

PÉTRI est utilisée pour réaliser l'inoculum. Cette colonie sert à ensemercer 5 mL de milieu riche sélectif. Après une incubation à 37°C sous agitation pendant 8 h, la pré-culture est réalisée. Pour cela, 100 à 500 mL de milieu liquide sélectif sont ensemencés avec 1 à 5 mL de l'inoculum (dilution au 1/100^{ème}) dans une fiole *ERLENMEYER* de 1 à 2 L (en règle générale, le volume de la fiole utilisée doit correspondre à un volume de 4 à 5 fois supérieur au volume du milieu de croissance des bactéries aérobies); la pré-culture est réalisée à 37°C pendant 12 h avec une agitation de 160 rpm.

Culture en milieu minimum

La surexpression de la protéine recombinante avec un marquage uniforme nécessite l'utilisation d'un milieu de culture minimum *M9* contenant une seule source de carbone et une seule source d'azote, les deux uniformément marquées. Typiquement ce sont le ¹³C₆-glucose et le chlorure d'ammonium ¹⁵NH₄Cl qui sont utilisés pour effectuer le marquage dans *E. coli*. L'incorporation de chaque isotope est ainsi homogène au sein de la protéine. La difficulté principale de cette méthode consiste à maintenir des niveaux de croissance élevés des bactéries en milieu minimum. Le milieu minimum est supplémenté d'oligoéléments servant de cofacteurs ou d'activateurs enzymatiques, et un extrait de levures est également ajouté. Cet extrait est riche en nucléotides et en acides aminés, et est essentiel au démarrage de la croissance bactérienne pendant la phase de latence. La croissance bactérienne est réalisée dans des grandes fioles de 5 L.

La pré-culture précédemment réalisée pendant 12 h est centrifugée à 6 000 g pendant 10 minutes à 4°C. Le culot obtenu est resuspendu en milieu minimum sans glucose, afin de le débarrasser des sources de carbone ¹²C, et la solution resuspendue est utilisée pour inoculer le milieu minimum avec la quantité de bactéries nécessaire pour démarrer la croissance à une *DO*₆₀₀ de 0,1. La variation de la *DO*₆₀₀ en fonction du temps est suivi. Un prélèvement est effectué toutes les 40 minutes, et la courbe est tracée afin de suivre l'évolution de la masse bactérienne pendant la phase exponentielle.

L'induction du promoteur *T7* à l'*IPTG* se fait couramment lorsque la *DO*₆₀₀ est entre 0,6 et 0,8. Il s'agit généralement du moment où les bactéries sont au début de la phase

stationnaire. Plusieurs paramètres sont importants lors de l'induction du promoteur *T7* : la concentration en *IPTG*, la température et la durée d'induction. Ces paramètres sont modifiés en fonction de la toxicité de la protéine et de sa stabilité. Pour la surexpression de *MscL*, l'induction est réalisée avec 1 mM d'*IPTG*, pendant 3 h . À la fin, la biomasse est récupérée par sédimentation à basse vitesse (centrifugation à $6\ 000\text{ g}$) pendant 15 minutes . Le culot bactérien obtenu est resuspendu dans 100 mL de tampon *M9* sans glucose puis centrifugé pendant 15 minutes . Le culot final est récupéré, congelé dans de l'azote liquide, pesé puis conservé à -20°C jusqu'à son utilisation lors du broyage des cellules.

iv. Broyage des cellules

Le culot bactérien est mis dans la glace et resuspendu dans 40 mL de tampon *FP* contenant des inhibiteurs de protéases. La resuspension est effectuée sur la glace pour limiter la dégradation des protéines. Le *PMSF* était couramment utilisé, mais il est préférable d'utiliser l'*AEBSF*, moins toxique et plus stable. Néanmoins, la manipulation est effectuée sous la hotte à cause du *DTT* et du β -*Met-OH*.

Les bactéries sont cassées en utilisant la technique de pression de *FRENCH* (Milner *et al.*, 1950). Pour cela, la suspension bactérienne est versée dans le cylindre après avoir fixé le piston dans le cylindre jusqu'au trait indiquant 40 mL . Le socle est ensuite posé et le robinet vissé. L'état de la bille est vérifié avant chaque utilisation. Le montage est inversé et posé avec précaution sur l'appareil. Ce dernier est allumé en position *down*, puis la pression est augmentée jusqu'à atteindre $1\ 000\text{ psi}$ (703 Kg.cm^{-3} ou $69\times 10^6\text{ Pa}$). Lorsque la pression est stabilisée, l'homogénat est récupéré, goutte à goutte, dans un tube *Falcon* posé dans un bécher contenant de la glace. Un seul passage suffit normalement pour les bactéries, mais en cas de doute, la procédure est répétée. Un moyen simple de savoir si les membranes sont cassées est la mesure de la densité optique à 600 nm avant et après le broyage ; les vésicules inversées produites après broyage des cellules absorbent moins de lumière que les bactéries entières. Une chute de l'absorbance de 80% est un témoin du bon déroulement de l'expérience. L'homogénat obtenu est centrifugé à $5\ 000\text{ g}$ pendant 15 minutes afin de sédimenter les fragments cellulaires et les bactéries non cassées. Le

surnageant contenant les vésicules inversées est récupéré et divisé en 4 tubes d'ultracentrifugation de 10,4 mL et ultracentrifugé à 300 000 g. Le culot obtenu est couvert de tampon d'équilibration *EB* sans détergent, et conservé à 4°C.

v. Solubilisation des membranes

La quantité de détergent nécessaire pour solubiliser les membranes est calculée en fonction du poids de bactéries en fin de culture. Approximativement, 10 g de bactéries contiennent 1 g de protéines membranaires, qui nécessitent 1% de détergent (*p/v*) pour être solubilisé. La quantité de *triton-x-100* nécessaire est solubilisée dans 20 mL de tampon *EB* avant d'être rajoutée aux membranes.

Le culot obtenu après ultracentrifugation est resuspendu dans 20 mL de tampon *EB* à 4°C à l'aide d'une seringue. Après resuspension totale des membranes, le détergent est rajouté et la solution finale est mélangée à température ambiante pendant 30 minutes, pour solubiliser les vésicules. Après solubilisation, une centrifugation (à 5 000 g pendant 10 minutes à 18°C) est effectuée pour sédimenter les fragments de membranes non solubilisées. Le surnageant est récupéré et purifié sur colonne d'affinité. Une filtration est éventuellement réalisée au lieu de la centrifugation. Un filtre de 0,45 µm de diamètre de pore est utilisé, puis la solution est injectée dans le système de purification.

vi. Purification de la protéine *MscL*

La protéine *MscL* porte une étiquette 6-*His* et de ce fait est purifiée sur colonne d'affinité *Ni-NTA*. Le principe de cette purification repose sur l'affinité existant entre les six histidines contiguës de la protéine et l'ion Ni^{2+} relié à la matrice de la colonne via un groupement d'acide nitriloacétique *NTA*. En effet, le *NTA* interagit, par l'intermédiaire de 4 atomes avec l'ion Ni^{2+} . Ce dernier possède alors encore deux sites d'interaction libres. Lorsque la protéine contenant 6 résidus histidines en N ou C-terminal passe dans la colonne, des interactions entre les histidines de la protéine et l'ion Ni^{2+} se créent. La protéine se fixe alors spécifiquement sur la matrice. L'élution de la protéine est ensuite effectuée en augmentant la concentration de l'imidazole ayant une meilleure affinité pour

le N_7^{2+} que l'étiquette *6-His*, et va donc déplacer la protéine.

La purification est effectuée en utilisant un système de chromatographie liquide *FPLC*. Ce système permet le suivi de différents paramètres au cours de la purification (*pH*, conductance). Il est également équipé d'un détecteur *UV* et d'un collecteur de fraction. Le logiciel commandant l'appareil permet un contrôle direct ainsi que l'écriture de programmes pour automatiser les procédures. Le système possède plusieurs entrées sur lesquelles sont reliées les bouteilles contenant les tampons *A1*, *A2* et *B*.

La solution contenant les membranes solubilisées est introduite dans la colonne, préalablement équilibrée dans le tampon *A1*. Après lavage des protéines non retenues (100% tampon *A1*), la concentration en imidazole est augmentée afin de décrocher le matériel fixé non spécifiquement (80% *A1*, 20% *B*). L'étape suivante (60% *A2*, 40% *B*) permet de doubler la concentration en imidazole, et de réduire à 0,2 % la concentration en *triton-x-100*. Pour l'élution de la protéine, la concentration en imidazole est augmentée à 500 *mM*, afin de déplacer entièrement le matériel fixé sur la colonne.

La présence de la protéine *MscL* dans les fractions obtenues est révélée par la technique de *DOT BLOT*. La technique repose sur le même principe que le *WESTERN BLOT*. L'échantillon est déposé tel quel sur une membrane de *PVDF*. Pour réaliser le *DOT BLOT*, 1-2 μL de chaque fraction obtenue après purification sont déposés sur la membrane. Celle-ci est ensuite séchée puis les sites d'interactions non spécifiques sont bloqués en incubant la membrane dans une solution de *TBS-TM*. La membrane est ensuite incubée dans une solution contenant un anticorps anti-*His* couplé à une activité peroxydase. La révélation de l'activité peroxydase se fait en incubant la membrane dans une tablette *DAB* solubilisée dans 5 *mL* d'eau, jusqu'à coloration des dépôts. La révélation se fait par l'action du composé *DAB* en présence de peroxyde de carbamide. Un dépôt brun signale la présence de l'activité peroxydase et donc de l'étiquette histidine.

vii. Reconstitution de la protéine en liposomes

L'analyse par *DOT BLOT* révèle en moyenne 5 fractions contenant la protéine

d'intérêt. Les fractions (2 mL chacune) contenant la protéine *MscL* sont rassemblées dans un tube de 50 mL. Elles contiennent une grande quantité d'imidazole (500 mM), qu'il faut réduire avant de quantifier la protéine (le dosage au *BCA* supporte au maximum 25 mM d'imidazole) (Smith *et al.*, 1985).

Dialyse - Élimination de l'imidazole

Une dialyse est donc effectuée. Pour cela, les fractions de protéines purifiées (10 mL) sont injectées dans une cassette de dialyse. La cassette est placée dans 1 L de tampon RB-T. La membrane de dialyse laisse passer les molécules de taille inférieure à 10 kDa de l'intérieur vers l'extérieur. À l'intérieur de la cassette se trouvent les micelles de *triton-x-100* contenant les protéines membranaires *MscL* dans le tampon B. À la fin de la première dialyse, à l'intérieur de la cassette, la concentration en *HEPES* est de 9,9 mM approximativement, celle du *KCl* de 99 mM. Des traces de Na_2HPO_4 (0,5 mM) de *NaCl* (3 mM) et d'imidazole (5 mM) subsistent. Une deuxième dialyse est effectuée. La cassette de dialyse est mise à nouveau dans 1 L de tampon RB-T. À la fin de la seconde dialyse, il reste à l'intérieur de la cassette 5 μ M de Na_2HPO_4 , 30 μ M de *NaCl* et 50 μ M d'imidazole. La première dialyse est effectuée toute la nuit en chambre froide. La deuxième est effectuée pendant 2 heures également en chambre froide. Les dialyses sont réalisées dans un grand bécher en effectuant une agitation modérée de la solution. À la fin des deux dialyses, la quantité d'imidazole est considérée négligeable et ne pourra pas interférer avec la quantification de la protéine.

Quantification de la protéine

La protéine est récupérée de l'intérieur de la cassette de dialyse à l'aide d'une seringue. L'échantillon est prêt pour l'étape de quantification. Le dosage au *BCA* (Brown *et al.*, 1989) est utilisé ici. Le réactif est constitué d'acide bicinchonique (*BCA*) et d'ions cuivriques en milieu alcalin. Les ions cuivreux générés après réduction par les protéines forment avec le *BCA* un complexe stable et intensément coloré ($\lambda_{Max} = 562 \text{ nm}$). Une gamme d'étalonnage de 0 à 100 μ g de protéines par tube est réalisée à partir d'une solution de sérum albumine bovine (*BSA*) à 1 mg.mL^{-1} . Les dilutions sont effectuées dans un

tampon *RB-T*. Pour doser la protéine *MscL*, 3 échantillons sont préparés, dilués au $1/10^{ème}$, au $1/100^{ème}$ et au $1/500^{ème}$. Les dilutions de la protéine *MscL* sont faites dans le tampon *RB-T*. Le réactif de *BCA* est rajouté selon le protocole défini par le fabricant (*PIERCE*). Les solutions sont incubées à $37^{\circ}C$ pendant 1 heure. Ensuite la densité optique des échantillons est lue à 562 nm . La courbe d'étalonnage est tracée, et le coefficient de proportionnalité entre la quantité de protéines et l'absorbance mesurée est calculé.

Élimination du détergent et reconstitution en vésicules lipidiques

Pour la protéine surexprimée en bactéries, le rendement était de 1 mg de protéine uniformément marquée pour 1 litre de culture. Au total, 3 mg de la protéine membranaire *MscL* sont solubilisés dans $0,2\%$ *triton-x-100*, pour un volume total de 10 mL . Pour la reconstitution de la protéine *MscL* en liposomes, la *DOPC* a été choisie. Le rapport protéine/lipide souhaité étant de $1/2$, 6 mg de lipides sont pesés. Les lipides sont hydratés pendant 30 minutes. Pour augmenter l'efficacité de la solubilisation des lipides dans 1 mL de tampon *RB*, la suspension peut être soumise à des cycles de gel - dégel.

Une fois l'échantillon de lipides entièrement hydraté, il est chargé dans une des seringues étanches et celle-ci est placée sur le support de l'extrudeur. La deuxième seringue vide est placée de l'autre côté sur le support. Le piston de la seringue contenant les lipides est poussé délicatement jusqu'à ce que la solution soit complètement transférée à l'autre seringue, puis les lipides sont transférés à la seringue originale. L'extrusion est répétée au moins 10 fois. L'étape finale devrait ramener la solution de vésicules lipidiques dans la seringue qui était vide au départ. Cela permet de réduire les chances de contamination avec de plus grandes particules restées dans la seringue d'origine. La seringue remplie est retirée délicatement de l'extrudeur. Les liposomes obtenus sont des vésicules multi-lamellaires de taille inférieure à la taille des pores de la membrane utilisée ($0,1\text{ }\mu\text{m}$).

Une fois les liposomes prêts, ils sont rajoutés aux 10 mL de protéines en micelles de détergent. Le volume de l'échantillon est maintenant de 11 mL , il contient 3 mg de protéines *MscL*, 20 mg de *triton-x-100*, et 6 mg de lipides. L'échantillon est mis dans un tube

de 15 mL, 100 mg de *BIO-BEADS* sont rajoutés (celles-ci sont préalablement préparées, le poids indiqué correspond à un volume sec). Le tube est placé sur un agitateur rotatif en chambre froide pendant 12 h, puis 100 mg de *BIO-BEADS* supplémentaires sont rajoutés pour 2 h.

Le tube est ensuite récupéré et posé verticalement de manière à sédimenter les *BIO-BEADS*. À l'aide d'une pipette *PASTEUR*, le surnageant est récupéré et ultracentrifugé à 9 000 g pendant 30 minutes. À la fin de cette étape, le surnageant est récupéré. Le culot est resuspendu dans 1 mL de tampon *RB*, et la suspension obtenue est ultracentrifugée. Cette étape est répétée afin de se débarrasser des traces de détergent. Le culot obtenu après les trois ultracentrifugations est conservé à -20°C jusqu'au transfert dans les rotors pour leur analyse par *RMN*.

b Synthèse *in vitro*

Dans cette partie la synthèse *in vitro* de la protéine membranaire *MscL* est décrite. Cette technique a été employée pour synthétiser plusieurs échantillons. La stratégie adoptée concernant le choix des acides aminés marqués au ^{13}C et au ^{15}N dans chaque échantillon est d'abord décrite, suivie des différents protocoles mis en place. Le protocole expérimental développé par la société *ROCHE* puis par *5PRIME* permet de synthétiser la protéine dans un volume de 10 mL dans les kits *RTS-9000* utilisant le système en échange continu *CECF* (pour *Continuous-Exchange Cell Free*). La *SIV* permet un marquage sélectif au ^{13}C et ^{15}N sur certains acides aminés par l'utilisation de l'échantillonneur d'acides aminés (*AMINO ACID SAMPLER*). Le rendement attendu pour la synthèse de *MscL* dans les kits *RTS-9000* est de 10 mg pour 10 mL. La quantité d'acides aminés à rajouter est calculée -en excès- pour 20 mg ou 1,25 μmol de *MscL* ($PM = 16\ 105\ \text{g.mol}^{-1}$). La quantité finale de chaque acide aminé à rajouter est $(n \times PM_i \times 1,25 \cdot 10^3\ \text{mg})$, n étant le nombre d'apparitions de chaque acide aminé dans la protéine et PM_i le poids moléculaire de chaque acide aminé (voir tableau 6 pour plus de détails). Pour une protéine membranaire, il existe différents protocoles en *SIV*: la protéine est récupérée sous forme de culot, elle est synthétisée en présence de détergent, ou alors elle est synthétisée et incorporée directement en liposomes. Le premier protocole n'a pas été testé dans ce travail. Les deux autres sont abordés successivement.

i. Synthèse *in vitro* de *MscL* en micelles de détergent et reconstitution en liposomes

Dans cette partie est présentée la synthèse *in vitro* de la protéine membranaire *MscL* en micelles de *triton-x-100*. Cette approche a été largement étudiée. Différents détergents et tensioactifs ont été utilisés (Berrier *et al.*, 2004 ; Park *et al.*, 2007 ; Park *et al.*, 2011). Parmi les détergents pouvant être utilisés en *SIV* pour *MscL*, le *triton-x-100* présente différents avantages : une *CMC* basse, un faible coût et une adsorption efficace aux *BIO-BEADS*. En utilisant l'échantillonneur d'acides aminés, la protéine est marquée sélectivement, au niveau des seize isoleucines et des trois thréonines. L'analyse de

l'échantillon marqué au niveau des *Ile* et des *Thr* (*Ile, Thr*)- ^{13}C , ^{15}N -*MscL* et sa comparaison à un échantillon uniformément marqué ont fait l'objet d'un article (Abdine *et al.*, 2010). Ce protocole de synthèse a été mis au point dans l'équipe d'*A. GHAZI* sur des équipements de *SIV* à petite échelle (*RTS-100* et *RTS-500*) (Berrier *et al.*, 2004) et a été étendu pour synthétiser une plus grande quantité de *MscL* pour son étude en RMN (*RTS-9000*). Un deuxième échantillon a été synthétisé *in vitro* en micelles de détergent où le marquage sélectif a été effectué au niveau des acides aminés *Ala, Gly, Lys, Val* et *Tyr* (*(Ala, Gly, Lys, Val, Tyr)*- ^{13}C , ^{15}N -*MscL*).

- Synthèse *in vitro* en micelles de détergent

Dans les extraits utilisés pour la *SIV* se trouvent les ribosomes et les différents *ARNs* sensibles aux *ARNases*. Pour cela tout le matériel utilisé doit être autoclavé ou garanti par le fabricant sans nucléases. Pour la préparation des acides aminés marqués, des tubes *EPENDORF* autoclavés sont nécessaires. Les acides aminés *Ile* et *Thr* sont solubilisés dans le tampon de reconstitution présent dans le kit de *SIV*. Le détergent utilisé pour la synthèse est préparé en avance, pour cela 40 g de *triton-x-100* sont solubilisés dans 200 mL d'eau milli-Q (pour une concentration en *triton-x-100* de 20%). Les solutions nécessaires à la purification de la protéine sur colonne *Ni-NTA* sont également préparées. (Voir le tableau 5 pour la composition des tampons). Pour la dialyse de l'échantillon après sa purification, le tampon de dialyse *RB-T* est préparé. Les *BIO-BEADS* nécessaires pour l'étape de reconstruction de la protéine en liposomes sont préparées. Pour cela 5 g de billes sont lavés dans 25 mL de méthanol, puis quatre fois dans 25 mL d'eau milli-Q. Elles sont conservées à 4°C dans l'eau jusqu'à leur utilisation.

Le contenu du kit *RTS-9000* est dégelé sur la glace, sauf le tampon de reconstitution qui peut être placé directement à température ambiante. Les flacons des acides aminés marqués aux ^{13}C et ^{15}N *Ile* et *Thr* sont mis à dégeler sur la glace. Pour les acides aminés non marqués, l'échantillonneur d'acides aminés contenant 20 tubes d'acides aminés et 4 tubes de *DTT* est mis à température ambiante. Les acides aminés *Ile* et *Thr* non marqués sont écartés des autres tubes. Ils sont remplacés par deux tubes *EPENDORF* dans lesquels

sont pesés 2,6 et 0,45 mg d'*Ile* et de *Thr* respectivement. Le plasmide est décongelé à température ambiante. Les acides aminés marqués *Ile* et *Thr* sont solubilisés dans 1,2 et 0,3 mL de tampon de reconstitution respectivement. Pour les acides aminés non marqués, les tubes présentant un aspect trouble sont mis dans le bain d'ultrasons pour 5 minutes. Le mélange d'acides aminés est préparé dans un tube *FALCON* de 50 mL en y ajoutant pour chaque acide aminé (sauf *Ile* et *Thr*) le volume indiqué dans le tableau 6 à partir du kit des acides aminés non marqués. Les acides aminés *Tyr* et *Leu* sont rajoutés en dernier pour éviter leur précipitation. Le contenu des deux tubes contenant les acides aminés marqués est rajouté au mélange d'acides aminés non marqués. Le *DTT* est ensuite rajouté au mélange (3 mL de *DTT* à 40 mM). Le tube contenant tous les acides aminés et le *DTT* est remis sur la glace. Le lysat de bactéries *E. coli* lyophilisé est reconstitué dans 5,2 mL de tampon de reconstitution. Le mélange réactionnel lyophilisé est reconstitué dans 2,2 mL de tampon de reconstitution. Le mélange d'alimentation lyophilisé est reconstitué dans 80 mL de tampon de reconstitution. Les lyophilisats sont solubilisés dans leur fiole d'origine. Pour solubiliser efficacement les lysats, les fioles sont agitées doucement.

La synthèse *in vitro* est effectuée dans une cuve réactionnelle fournie dans le kit *RTS-9000*. Celle-ci est composée d'une chambre réactionnelle de 11 mL dans une chambre d'alimentation de 110 mL. Une membrane semi-perméable sépare les deux compartiments et laisse passer les substrats de la chambre d'alimentation vers la chambre réactionnelle et les sous-produits de la synthèse de protéine dans l'autre direction. La chambre réactionnelle contient le lysat des bactéries (5,2 mL), le mélange réactionnel (2,2 mL), le plasmide (300 µL à 0,5 µg.µL⁻¹), le détergent (220 µL de la solution à 20%), une partie du mélange d'acides aminés pour démarrer la synthèse (2,7 mL) et le *DTT* (300 µL). Les différents composants sont mélangés dans la fiole contenant initialement le lysat bactérien puis le mélange est injecté dans la chambre réactionnelle. La chambre d'alimentation contient le mélange d'alimentation (80 mL), les acides aminés et le *DTT* (26 mL) ainsi que le détergent (2,2 mL de la solution à 20%). Les acides aminés, le *DTT* et le détergent sont rajoutés à la fiole du mélange d'alimentation puis le mélange est injecté dans la chambre d'alimentation. La chambre réactionnelle est toujours remplie en premier, et pour les deux compartiments, les bulles d'air sont enlevées en tapotant légèrement le

dispositif. Un échantillon du mélange réactionnel est prélevé pour les tests d'analyse ultérieurs. Le dispositif est refermé et la présence de bulles est vérifiée à nouveau en le retournant. Le volume de la chambre réactionnelle et celui de la chambre d'alimentation sont généralement de 11 mL et 110 mL respectivement. Pour des raisons inconnues, ces valeurs sont variables d'un kit à l'autre. Dans le cas où le volume des chambres serait supérieur à celui des solutions préparées, du tampon de reconstitution contenant du *triton-x-100* à $0,2\%$ est utilisé pour compléter. Une fois prêt, le dispositif est placé dans l'appareil *THERMOMIXER CONFORT* et celui-ci est refermé. La température est réglée à 30°C et la vitesse de rotation à 800 rpm . La réaction est déclenchée pour 22 heures.

- Purification de la protéine

Après 22 heures, l'appareil est arrêté et le dispositif retiré. Le mélange réactionnel est extrait et placé dans un tube *FALCON* de 50 mL . Pour éviter une éventuelle action de dégradation des protéines par des protéases, $80\text{ }\mu\text{L}$ d'*AEBSF* sont ajoutés (à partir de 8 mL de solution-mère à 50 mM) et le tube est agité pendant 15 minutes. L'échantillon est ensuite complété jusqu'à 50 mL avec du tampon *A1* (Voir tableau 5) et centrifugé à 6000 g pendant 15 minutes à 4°C et après la centrifugation le surnageant est dilué deux fois. Pour éviter toute obstruction des capillaires pendant la purification par *FPLC*, l'échantillon est centrifugé, ou filtré en utilisant des filtres de $0,45\text{ }\mu\text{m}$. Le culot issu de la centrifugation est resuspendu en utilisant le tampon *A1*, et un aliquote est prélevé pour le tester en *SDS-PAGE*. Le protocole de purification de l'échantillon de *(Ile, Thr)-¹³C,¹⁵N-MscL* obtenu en *SIV* en micelles de détergent, suit les mêmes étapes que celui présenté précédemment. La seule différence est que la purification est effectuée en deux fois. Les fractions obtenues après purification sont analysées par *DOT BLOT* pour détecter la présence de *MscL*. Les fractions issues des deux purifications et contenant *MscL* sont réunies dans deux tubes, et une dialyse est effectuée pour réduire la concentration en imidazole. Ainsi, les fractions sont transférées dans deux cassettes de dialyse avec une limite d'exclusion de 10 kDa et chaque échantillon est placé dans 1 L de tampon *RB-T* à 4°C , pendant 12 heures.

- Reconstitution de la protéine en liposomes

Pour l'échantillon (*Ile, Thr*)-¹³C,¹⁵N-*MscL*, après 12 heures, les cassettes sont transférées dans un nouveau bain de dialyse pour 2 heures à 4°C. Les échantillons obtenus après dialyse, ainsi que les échantillons prélevés au cours de la *SIV* et pendant la purification, sont analysés par électrophorèse sur gel de polyacrylamide pour tester le déroulement de la *SIV* et de la purification. Après migration, les protéines sont transférées sur membrane de nitrocellulose. Avant de reconstituer la protéine en liposomes, une quantification par la méthode *BCA* est effectuée. Pour l'échantillon *MscL* synthétisé *in vitro* en micelles de *triton-x-100*, la quantité totale de protéine obtenue dans les 20 mL (après purification et dialyse) est de 9,6 mg. Le ratio protéine/lipide souhaité pour cet échantillon étant de 1/2 (*m/m*), la quantité de lipides à rajouter est de 19,2 mg. Les lipides sont solubilisés dans 10 mL de tampon *RB-T* contenant du *triton-x-100* à 0,8%. Les lipides solubilisés sont rajoutés aux micelles de *triton-x-100* contenant la protéine pour former des micelles ternaires. Le volume final est de 30 mL. La quantité de détergent présente est de $(20 \times 0,2 \times 10^{-2}) + (10 \times 0,8 \times 10^{-2}) = 0,12$ g. Le mélange est séparé en quatre tubes. Le premier lot de *BIO-BEADS* est rajouté après 15 minutes. La quantité de *BIO-BEADS* rajoutée correspond à 10 fois la quantité de détergent présente (4×300 mg). Les tubes sont mis sur un agitateur rotatif en chambre froide pendant toute la nuit. Le lendemain, le deuxième lot de *BIO-BEADS* est rajouté (4×300 mg) pour 2 heures à 4°C. Après deux heures, les quatre tubes sont placés verticalement pour faire décanter les *BIO-BEADS*. Après environ 5 minutes, le surnageant est retiré (30 mL) et centrifugé à 100 000 g pendant 30 minutes à 4°C afin de récupérer les liposomes contenant *MscL*. À la fin de la première ultracentrifugation, les culots obtenus sont solubilisés dans 15 mL de tampon *RB*. L'étape est répétée progressivement pour avoir la totalité de l'échantillon dans un seul tube. Le tube contenant le culot final est conservé à -20°C jusqu'à son analyse par RMN.

Pour l'échantillon (*Ala, Gly, Lys, Val, Tyr*)-¹³C,¹⁵N-*MscL*, la diméthyl- β -cyclodextrine a été utilisée pour l'extraction du détergent et la reconstitution de la protéine en liposomes. Cet échantillon a servi pour la comparaison de l'effet des lipides sur la protéine. Différentes compositions ont été testées : un mélange *DOPC/DPPC* ; extraits de lipides

d'*E. coli* ; extraits d'asolectine ; et un mélange *DOPC/LPC*. La quantité de protéine obtenue après *SIV*, purification et dialyse était de 10,5 mg dans 13,2 mL de tampon *RB-T*. Pour 4 échantillons, cela correspondait à 2,6 mg de *MscL* pour chacun, dans 3,3 mL de *RB-T*, et donc 6,6 mg de *triton-x-100*.

Les quantités de lipides utilisées sont différentes d'un échantillon à un autre. Pour le mélange *DOPC/DPPC*, les lipides sont solubilisés ensemble, et les vésicules sont formées à l'aide de l'extrudeur à partir de 2,6 mg de *DOPC* et 2,6 mg de *DPPC*. De même, des vésicules sont formées à partir de : 1,3 mg d'extraits d'*E. coli* ; de 5,4 mg d'extraits d'asolectine et de 13,5 mg de *DOPC*. Les vésicules sont rajoutées aux micelles de *triton-x-100* contenant la protéine *MscL*, et le mélange est réalisé pendant 30 min à température ambiante. Ensuite pour chaque tube, contenant donc 6,6 mg de *triton-x-100*, 6,6 mg de cyclodextrine sont rajoutés en quatre fois, à partir d'une solution-mère à 250 mg/mL (4 * 6,6 µL/tube). Les ajouts sont espacés de 15 min, et après le dernier lot de cyclodextrine, les échantillons sont dilués dans le tampon *RB* puis centrifugés à 300 000 g pendant 30 min pour récupérer les protéoliposomes formés. Ceux-ci sont ensuite placés sur un gradient de saccharose pour les débarrasser des molécules de cyclodextrine. Les échantillons reconstitués en *DOPC/DPPC*, extraits d'*E. coli* et extraits d'asolectine sont récupérés, repris dans du tampon *RB* puis centrifugés. Le culot est ensuite transféré dans un rotor et conservé à -20°C jusqu'aux expériences de RMN. La *LPC* (3,4 mg) est rajoutée à l'échantillon reconstitué en *DOPC* après l'avoir récupéré sur le gradient de saccharose. L'échantillon est mélangé délicatement, dilué au tampon *RB* puis centrifugé. Le culot est ensuite transféré dans un rotor et conservé à -20°C.

ii. Synthèse *in vitro* de *MscL* en liposomes

Dans cette partie est présentée la synthèse *in vitro* de la protéine membranaire *MscL* directement en liposomes. Ce protocole permet de réduire les étapes nécessaires pour obtenir un échantillon de protéine membranaire pour son étude par RMN en phase solide. Le protocole décrit a été mis au point d'abord en utilisant une *RTS-500*, et a été étendu pour la synthèse de trois échantillons (*Ala, Lys, Val, Tyr*)-¹³C, ¹⁵N-*MscL*, (*Phe, Ile,*

Met, Pro, Arg)-¹³C, ¹⁵N-MscL et (*Asp, Gly, Leu, Ser*)-¹³C, ¹⁵N-MscL à l'aide des kits RTS-9000. Les liposomes utilisés ici sont composés de DOPC. D'autres lipides ou les nano-disques peuvent être utilisés (Berrier *et al.*, 2010 ; Katzen *et al.*, 2008). Tout le matériel utilisé doit être autoclavé ou garanti par le fabricant sans nucléases. En fonction du marquage souhaité et comme précisé précédemment, certains acides aminés nécessitent un autre tampon que le tampon de reconstitution présent dans les kits RTS. Les volumes des acides aminés non marqués et les masses d'acides aminés marqués à prélever sont calculés en avance (Voir le [tableau 6](#)). Une solution de tampon RB est préparée (500 mL).

Pour l'échantillon (*Ala, Lys, Val, Tyr*)-¹³C, ¹⁵N-MscL, un tampon HEPES à pH 13 est nécessaire pour la solubilisation de la tyrosine (HEPES 60 mM, KOH 0.7 M).

Pour l'échantillon (*Phe, Ile, Met, Pro, Arg*)-¹³C, ¹⁵N-MscL, un tampon HEPES à pH 7,45 est préparé pour la solubilisation de la phénylalanine.

Le contenu du kit RTS-9000 est dégelé sur la glace, sauf le tampon de reconstitution qui peut être placé directement à température ambiante. Les flacons des acides aminés marqués aux ¹³C et ¹⁵N sont mis à dégeler sur la glace.

Pour l'échantillon (*Ala, Lys, Val, Tyr*)-¹³C, ¹⁵N-MscL, 1,7 mg d'*Ala*, 1,6 mg de *Lys*, et 1,9 mg de *Val* sont solubilisés dans 1,1 0,67 et 0,97 mL de tampon de reconstitution respectivement ; pour la *Tyr*, 0,23 mg sont solubilisés dans 0,1 mL de tampon HEPES à pH 13.

Pour l'échantillon (*Phe, Ile, Met, Pro, Arg*)-¹³C, ¹⁵N-MscL, 2,6 mg d'*Ile*, 1 mg de *Pro* et 1,3 mg d'*Arg* sont solubilisés dans 1,2 0,52 et 0,45 mL de tampon de reconstitution respectivement ; pour la *Phe*, 1,6 mg sont solubilisés dans 0,74 mL de tampon HEPES à pH 7,45 et le tube est placé dans le bain d'ultrasons pendant 5 minutes ; enfin pour la *Met*, 0,93 mg sont solubilisés dans 0,37 mL de tampon de reconstitution contenant 4 mM de DTT.

Pour l'échantillon (*Asp, Gly, Leu, Ser*)-¹³C, ¹⁵N-MscL, 1,2 mg d'*Asp*, 1,2 mg de *Gly*, 2,1 mg de *Leu* et 0,53 mg de *Ser* sont solubilisés dans 0,52 0,97 1,2 et 0,3 mL de tampon de

reconstitution respectivement.

Pour les acides aminés non marqués, l'échantillonneur d'acides aminés contenant 20 tubes d'acides aminés et 4 tubes de *DTT* est mis à température ambiante. Les acides aminés à marquer sont écartés de l'échantillonneur. Ils sont remplacés par les tubes *EPPENDORF* dans lesquels sont solubilisés les acides aminés selon l'échantillon à préparer. Le plasmide est décongelé à température ambiante. Pour les acides aminés non marqués, les tubes présentant un aspect trouble sont mis dans le bain d'ultrasons pour 5 minutes.

Avant de mélanger les acides aminés, les liposomes sont préparés. Le rendement de la *SIV* pour la protéine *MscL* dans les kits *RTS-9000* est (par expérience) estimé à 10 mg de protéines. Le ratio protéine/lipide souhaité est de 1/2, 20 mg de lipides sont donc nécessaires. Deux stratégies pour obtenir la bonne quantité de lipides ont été utilisées. Dans la première, les lipides sont pesés en excès, puis solubilisés en chloroforme dans un tube en verre. Le chloroforme est évaporé sous un flux d'argon. À l'issue de l'évaporation du solvant organique, un film de lipides se forme. Les lipides sont solubilisés dans le tampon de reconstitution. Une fois l'échantillon entièrement hydraté, le volume contenant 20 mg de lipides est chargé dans une des seringues étanches de l'extrudeur et la seringue y est placée soigneusement. Les seringues, la membrane et les filtres sont préalablement préparés. La deuxième stratégie consiste à peser la quantité exacte souhaitée dans un tube en verre et solubiliser les lipides directement dans 1 mL de tampon de reconstitution en utilisant la spatule pour écraser les lipides contre la paroi du tube facilitant ainsi la solubilisation complète, avant de passer l'échantillon dans l'extrudeur. Le protocole de préparation des vésicules est le même que celui décrit précédemment. Les vésicules sont gardées dans la seringue sur la glace jusqu'à leur utilisation. Une fois les vésicules prêtes, le mélange d'acides aminés est préparé dans un tube *FALCON* de 50 mL en y ajoutant pour chaque acide aminé, sauf ceux qui sont marqués aux ^{13}C et ^{15}N , le volume indiqué dans le tableau 6 à partir de l'échantillonneur des acides aminés non marqués. Les acides aminés *Tyr* et *Leu* sont rajoutés en dernier pour éviter leur précipitation. Le contenu des tubes contenant les acides aminés marqués est rajouté au mélange d'acides aminés non marqués. Le *DTT* est ensuite rajouté au mélange (3 mL de *DTT* à 40 mM). Le tube contenant tous

les acides aminés et le *DTT* est remis sur la glace. Le lysat de bactéries *E. coli* lyophilisé est reconstitué dans 5,2 mL de tampon de reconstitution dans la fiole d'origine. Le mélange réactionnel lyophilisé est reconstitué dans 2,2 mL de tampon de reconstitution dans la fiole d'origine. Le mélange d'alimentation lyophilisé est reconstitué dans 80 mL de tampon de reconstitution dans la fiole d'origine. Pour bien solubiliser les lysats, les fioles sont agitées doucement. La *SIV* est effectuée dans la cuve réactionnelle. La chambre réactionnelle contient le lysat des bactéries (5,2 mL), le mélange réactionnel (2,2 mL), le plasmide (300 μL à 0,5 $\mu\text{g}\cdot\mu\text{L}^{-1}$), les liposomes (1 μL de la solution à 20 $\text{mg}\cdot\text{mL}^{-1}$), une partie du mélange d'acides aminés pour démarrer la synthèse (2,7 mL) et le *DTT* (300 μL). Les différents composants sont mélangés dans la fiole contenant initialement le lysat bactérien puis le mélange est injecté dans la chambre réactionnelle. La chambre d'alimentation contient le mélange d'alimentation (80 mL), les acides aminés (26 mL), ainsi que le *DTT* (3 mL d'une solution à 40 mM). Les acides aminés et le *DTT* sont rajoutés à la fiole du mélange d'alimentation, puis le mélange est injecté dans la chambre d'alimentation. La chambre réactionnelle est toujours remplie en premier, et pour les deux compartiments, les bulles d'air sont enlevées en tapotant légèrement le dispositif. Un échantillon du mélange réactionnel est prélevé pour les différents tests d'analyse. Le dispositif est refermé, et la présence de bulles est vérifiée à nouveau en le retournant. Une fois prêt, le dispositif est placé dans l'appareil *THERMOMIXER CONFORT* et l'appareil refermé. La température est réglée à 30°C, et la vitesse de rotation à 800 rpm. La réaction est déclenchée pour 22 heures.

Après 22 heures de synthèse, l'appareil est arrêté et le dispositif en est retiré. Le mélange réactionnel est extrait, et placé dans un tube *FALCON* de 15 mL. Pour éviter une éventuelle action de dégradation des protéines par des protéases, 80 μL d'*AEBSF* sont ajoutés (à partir de 8 mL de solution-mère à 50 mM) et le tube est agité pendant 15 minutes. La solution est ensuite centrifugée puis le surnageant est ultracentrifugé à 100 000 g pendant 20 minutes à 4°C pour récupérer les protéoliposomes. À la fin de l'ultracentrifugation, les culots sont resuspendus dans 3 mL de tampon *RB*, placés dans un seul tube et une nouvelle ultracentrifugation est effectuée. Le culot final est récupéré et conservé à -20°C.

c Analyse de l'échantillon obtenu

Après avoir surexprimé, purifié et reconstitué la protéine dans un environnement natif, différents tests sont réalisés. Ces tests sont effectués afin de suivre le bon déroulement des différentes étapes (*SDS-PAGE*) et la présence de la protéine d'intérêt dans les différentes fractions obtenues (*WESTERN* et *DOT BLOTS*). La protéine *MscL* ne contient pas de tryptophanes au niveau de la séquence d'acides aminés, la concentration de la protéine est alors déterminée par le test du biuret (dosage au *BCA*). Le pentamère se dénaturant facilement en *SDS*, les expériences de *cross-linking* permettent de déterminer l'état d'oligomérisation de la protéine, en liant covalamment les sous-unités, en vue d'une analyse postérieure en *SDS-PAGE*. Une fois la reconstitution finie, l'échantillon est composé de la protéine *MscL* intégrée dans des vésicules géantes unilamellaires de *DOPC*. Le rapport protéine/lipides est mesuré en déterminant la concentration finale en protéine (dosage au *BCA*) et celle des lipides (dosage des phosphates - méthode de *ROUSER*). L'eau est quantifiée pour déterminer le taux d'hydratation des protéoliposomes obtenus.

- Électrophorèse sur gel de polyacrylamide - *SDS-PAGE*

Le principe du gel d'électrophorèse en conditions dénaturantes est de traiter un mélange de protéines avec : un agent réducteur, le β -*Me*OH ou le *DTT* qui réduisent les ponts disulfure et un détergent anionique, le *SDS* qui dénature les protéines et leur confère une enveloppe de charges négatives ; la plupart des protéines se retrouvent sous une forme monomérique. Les protéines présentes sont ensuite séparées par électrophorèse sur gel de polyacrylamide. La migration s'effectue sous l'action d'un champ électrique, dans les mailles formées par la polymérisation de l'acrylamide. Une protéine migre d'autant moins que sa masse molaire est élevée (Laemmli, 1970). À la fin, les protéines sont révélées par une coloration (le bleu de *COOMASSIE* ou le nitrate d'argent). La masse molaire des protéines est déterminée à l'aide de marqueurs qui sont des protéines de masses molaires connues. Les gels utilisés dans cette étude sont des gels d'acrylamide de gradient avec un continuum de concentrations en acrylamide à 4-20% sur la longueur du gel.

Vérification des transformants

Pour le protocole de synthèse de *MscL* en bactéries, afin de choisir le meilleur clone d'*E. coli* pour la surexpression de la protéine, les différents transformants sont vérifiés sur gel. La densité optique DO_{600} des cultures de bactéries induites et non induites est mesurée, et la même quantité de cellules est prélevée de chaque culture et centrifugée. Le culot obtenu est solubilisé dans le tampon *SLB*. La présence d'acides nucléiques pouvant perturber la migration sur gel, les échantillons sont mis dans un bain d'ultrasons pendant 10 minutes. Si cela s'avère être insuffisant, l'ajout d'*ADNase* est nécessaire. Après migration, les protéines sont transférées sur membrane de nitrocellulose (Voir chapitre suivant).

Suivi de la surexpression de la protéine en bactéries

Au cours de la surexpression de la protéine en milieu minimum, un échantillon est prélevé avant l'induction par l'*IPTG*, et un autre après 3 heures d'induction. D'autres échantillons sont prélevés après la presse de *FRENCH* dont un, directement à la fin du broyage des cellules, et deux autres après l'ultracentrifugation (le surnageant et le culot resuspendu). Pour la purification de la protéine, les différentes fractions en sortie de la colonne sont testées sur gel. Les différentes étapes concernant la reconstitution de la protéine en liposomes peuvent être suivies aussi par *SDS-PAGE* afin de s'assurer qu'il n'y a pas de pertes de protéine. Le volume prélevé des différents échantillons de la surexpression à la reconstitution est calculé de manière à avoir approximativement un suivi qualitatif et proportionnel (ainsi, afin d'évaluer leur contenu en protéines, les culots sont resuspendus avec le même volume que celui d'avant la centrifugation). L'électrophorèse sur gel de polyacrylamide technique permet donc de déterminer la masse molaire d'un monomère de la protéine et renseigne sur le rendement de purification; elle permet également la séparation des protéines présentes dans un échantillon avant de transférer ces protéines sur membrane de nitrocellulose, pour une détection plus spécifique de la protéine d'intérêt.

Suivi de la synthèse de la protéine *in vitro*

Pour la synthèse de *MscL* en micelles de détergent, des échantillons de mélange réactionnel avant et après la *SIV* sont prélevés, et sont analysés par *SDS-PAGE* avec d'autres échantillons prélevés au cours de la purification, de la dialyse et de la reconstitution de la protéine. Pour la synthèse de *MscL* en liposomes, des échantillons sont également prélevés avant et après la *SIV*, ainsi qu'après les étapes de centrifugation, afin de les analyser par *SDS-PAGE*.

- Immunodétection *WESTERN BLOT* / *DOT BLOT*

Les deux techniques utilisent le principe d'immunodétection, le but étant de révéler spécifiquement la protéine d'intérêt. Pour cela des anticorps reconnaissant cette protéine sont utilisés, et plus particulièrement, les anticorps sont dirigés contre l'étiquette histidine en C-terminal. Ces anticorps sont détectés directement dans le cas du *DOT BLOT*, ou indirectement après la fixation d'anticorps secondaires porteurs d'une fonction détectable pour le *WESTERN BLOT*. Dans le cas du *DOT BLOT*, les protéines sont transférées directement du milieu liquide sur la membrane, le transfert se réalisant par diffusion simple. Avant la détection, le tampon est laissé s'évaporer. Pour le *WESTERN BLOT*, les protéines sont séparées par *SDS-PAGE* puis transférées sur une membrane de nitrocellulose sous l'action d'un champ électrique. Les protéines sont adsorbées à la membrane. Après transfert des protéines sur la membrane, une étape de blocage des sites d'interactions non spécifiques est nécessaire. Ceci est réalisé en incubant la membrane dans une solution diluée de protéines de lait en présence de détergent.

WESTERN BLOT

La membrane de nitrocellulose est mise en présence d'anticorps primaires. Ces anticorps sont dirigés contre l'étiquette histidine, et ont été exprimés chez la souris. Après un cycle de lavage permettant de se débarrasser des molécules adsorbées non spécifiquement, une deuxième étape consiste à mettre la membrane en présence d'anticorps secondaires exprimés chez le lapin et dirigés contre les anticorps de souris. Les anticorps secondaires sont couplés à l'enzyme peroxydase du radis noir (*HRP*). Après un

cycle de lavage pour laver la membrane, la révélation de l'activité peroxydase est effectuée en appliquant sur la membrane du *luminol* et de l'eau oxygénée. La dégradation de l'eau oxygénée par la peroxydase libère des radicaux libres oxygénés. Sous l'action des radicaux libres, le *luminol* produit des photons détectés sur un film photographique. Les protéines apparaissent alors sous la forme de bandes sombres. La luminescence est proportionnelle à la quantité de protéines. Après chaque *SDS-PAGE*, un transfert sur membrane de nitrocellulose est effectué, afin de s'assurer que la bande visible vers 15 kDa est bien le monomère de la protéine d'intérêt *MscL* et non pas un contaminant.

DOT BLOT

À la fin de la purification sur colonne d'affinité, les différentes fractions obtenues sont déposées directement sur une membrane de *PVDF*. La détection se fait en une étape, en utilisant les anticorps *anti-His* couplés à la peroxydase. Ceci est généralement nécessaire pour savoir dans quelles fractions se trouve la protéine pure. La révélation du *DOT BLOT* se fait par l'action d'un composé chimique (*DAB*) en présence de peroxyde de carbamide. En effet, ce dernier libère du dioxygène lorsqu'il interagit avec la peroxydase. L'action du O_2 sur le *DAB* donne un précipité brun révélant la présence de la protéine dans les fractions de *FPLC*.

- Quantification de la protéine - dosage au *BCA*

Il est nécessaire de connaître la concentration en protéine avant l'étape de reconstitution pour connaître la quantité de lipides à rajouter pour avoir le ratio protéine/lipide souhaité. La protéine *MscL* ne contient pas de tryptophane et contient uniquement une seule tyrosine, ce qui la rend indétectable par une simple mesure d'absorption aux *UV*. Une méthode colorimétrique est donc utilisée, basée sur la réduction des ions cuivriques Cu^{2+} par les protéines en ions cuivreux Cu^+ en milieu alcalin (en présence de tartrate de sodium et de potassium $KNaC_4H_4O_6 \cdot 4H_2O$ qui chélate les ions cuivriques et évite leur précipitation en milieu alcalin, et d'iodure de potassium *KI* pour éviter de réduire le cuivre). Les ions cuivreux obtenus réagissent avec l'acide bicinchonique ou *BCA*. Un ion cuivreux est chélaté par deux molécules de *BCA*

permettant de former un complexe pourpre ayant une absorption optique maximale à 562 nm. La technique présente plusieurs avantages, parmi lesquels sa linéarité accrue par rapport à la technique de BRADFORD, et sa capacité à mieux supporter la présence de détergent. Dans le cas du triton-X-100 utilisé pour solubiliser *MscL*, le dosage est perturbé lorsque la concentration en détergent dépasse les 4%. D'autres substances peuvent interférer, notamment l'imidazole, d'où l'intérêt majeur de la dialyse de l'échantillon avant de quantifier la protéine. En fin de dialyse, la protéine *MscL* est dans un tampon RB-T à pH 7,6 ; la concentration en protéines étant inconnue, un aliquote est prélevé et dilué au 1/10^{ème}, 1/100^{ème} et au 1/500^{ème}. La courbe-étalon est réalisée dans les mêmes conditions (RB-T à pH 7,6) en utilisant la protéine BSA. Huit échantillons sont préparés en doublon avec des concentrations finales situées entre 0 et 100 µg.mL⁻¹. Le tampon de réaction est préparé en mélangeant la solution de BCA (50 parts) avec celle de KNaC₄H₄O₆.4H₂O (48 parts) en rajoutant la solution contenant les ions Cu²⁺. Un volume de 100 µL de chaque échantillon est prélevé et mis dans un tube EPPENDORF de 1,5 mL, et 1 mL de tampon de réaction est rajouté dans chaque tube. Les tubes sont ensuite fermés, et placés dans le bain-marie à 37°C. Après une heure, les tubes sont placés dans la glace pour interrompre la réaction chimique. La densité optique des échantillons est mesurée à 562 nm. La variation de l'absorbance des échantillons de BSA en fonction de la concentration est tracée ; elle est utilisée pour déterminer la concentration de la protéine *MscL* des 3 échantillons testés. Pour la protéine *MscL* synthétisée dans 3 L de milieu minimum M9, la concentration finale après purification était de 1 mg par litre de culture avec une incertitude de 5-10 %. Cette incertitude est due en partie à la technique car les réactifs ne réagissent pas de la même manière avec toutes les protéines. Elle est due aussi à l'utilisation d'une protéine globulaire et soluble pour étalonner le dosage, alors que c'est la concentration d'une protéine membranaire qui est recherchée. Pour éviter d'utiliser le détergent et afin d'exposer toute la protéine à doser aux réactifs du BCA, une précipitation à l'acétone a été effectuée, puis la protéine a été quantifiée. Aucune variation n'a été décelée.

Pour les échantillons synthétisés *in vitro*, la quantité de *MscL* après 22 heures d'expression est de 7-10 mg. Lorsque la protéine est synthétisée en présence de détergent,

la quantification est effectuée après purification et dialyse de l'échantillon. Le rendement pour l'échantillon $(Ile, Thr)^{-13}C,^{15}N$ -MscL était de 9,6 mg, mais ceci prend en compte le taux de synthèse, mais également les pertes subies pendant les différentes étapes qui suivent. Ce fut le cas également pour l'échantillon $(Ala, Gly, Lys, Val, Tyr)^{-13}C,^{15}N$ -MscL où le rendement final après purification et dialyse était de 10,5 mg. Dans le cas des échantillons synthétisés directement en liposomes, le rendement était plus élevé compte tenu du fait que la quantification était réalisée sur l'échantillon directement à la fin de la SIV. Pour les 3 échantillons $(Ala, Lys, Val, Tyr)^{-13}C,^{15}N$ -MscL, $(Phe, Ile, Met, Pro, Arg)^{-13}C,^{15}N$ -MscL et $(Asp, Gly, Leu, Ser)^{-13}C,^{15}N$ -MscL, le rendement était de 9 mg en moyenne.

- Réticulation covalente ou *cross-linking*

Le principe de la réticulation chimique est de créer un lien covalent entre deux groupements fonctionnels d'une molécule biologique par l'intermédiaire d'un agent réticulant. Il existe plus d'une centaine d'agents réticulants décrits dans la littérature et disponibles commercialement (Hermanson, 1996 ; Wong 1993). Un agent de réticulation est caractérisé par ses groupements fonctionnels, qui vont réaliser le lien covalent, par son bras espaceur, qui par sa taille permettra d'associer des molécules plus ou moins éloignées et enfin par sa nature chimique lui conférant des propriétés plus ou moins solubles dans l'eau. La réticulation des protéines est un moyen de connaître l'état d'oligomérisation. Concernant MscL, après la surexpression de la protéine, la technique permet de savoir si la protéine est exprimée et organisée en pentamère. Après le broyage des cellules, l'homogénéisation et la solubilisation des membranes, la réticulation informe sur la stabilité de l'oligomère en micelles de détergent. Après reconstitution de la protéine en membrane, l'état final de la protéine *in situ* est vérifié par la même technique.

Parmi les différents agents réticulants, deux ont été utilisés pour tester l'oligomérisation de MscL: le formaldéhyde FA et le disuccinimidyl subérate DSS. Les deux molécules agissent préférentiellement au niveau des fonctions amines des chaînes latérales des lysines. Au niveau de MscL, il existe une lysine périplasmique, et sept cytoplasmiques. Le formaldéhyde CH_2O est hydrophile, et considérant la taille de la

molécule, la réaction effectuée entre la molécule et deux groupements amines tertiaires, la réticulation ne peut se réaliser que sur une courte distance. Le *EA* interagit d'abord avec une molécule d'eau formant une molécule de méthylène glycol $HOCH_2OH$. Ensuite, la réaction de réticulation s'effectue entre une lysine et une fonction amine proche, formant entre les deux azotes un pont méthylène, avec libération d'une molécule d'eau. Cette liaison est rompue à des températures supérieures à $70^\circ C$. À la différence du *EA*, le *DSS* est plutôt hydrophobe et donc capable de pénétrer la membrane des vésicules. Son poids moléculaire est de $368,35 Da$. Le *DSS* est un agent réticulant contenant une fonction ester de *N*-hydroxy-succinimide (*NHS*) de chaque côté et un bras espaceur de 8 atomes de carbone ($\sim 11 \text{ \AA}$). Un ester de *NHS* réagit avec une amine primaire à *pH* 7-9 pour former une liaison amide stable, avec libération d'un groupement *N*-hydroxy-succinimide. Il permet de lier des fonctions amines éloignées de 11 \AA . Au niveau d'une protéine, le *DSS* agit avec les fonctions amines des lysines, mais aussi avec la fonction *N*-terminale. Considérant son caractère hydrophobe, le *DSS* est d'abord dissous dans un solvant organique (le *DMSO* par exemple) puis rajouté au tampon utilisé pour l'expérience de réticulation. Il existe un analogue hydrophile du *DSS*, il s'agit du *BS3*, les deux molécules agissent de manière identique vis-à-vis des fonctions amines primaires, possèdent la même longueur de bras espaceur, mais le *DSS* agirait à l'intérieur des liposomes et le *BS3* à l'extérieur.

Le tampon utilisé pour l'expérience de *cross-linking* ne doit pas contenir des fonctions amines primaires. Le *EA* ou le *DSS* sont rajoutés dans la solution contenant la protéine purifiée et solubilisée en détergent ou reconstituée en liposomes. L'incubation est réalisée sans agitation à température ambiante. L'action de réticulation est arrêtée par ajout du tampon *Tris* à *pH* 8. Pour leur analyse *SDS-PAGE*, le tampon *SLB* est rajouté aux échantillons. L'échantillon contenant le *EA* ne doit pas être chauffé au-delà de $60^\circ C$, au risque de rompre les liaisons réalisées par la molécule. Celui qui contient le *DSS* supporte une température de $90^\circ C$. Les échantillons sont chauffés pendant 5 minutes dans un bain-marie, puis une fois froids, sont déposés sur le gel de polyacrylamide.

- Quantification du phosphore - Méthode de ROUSER

Il s'agit donc d'une méthode permettant de quantifier le phosphore minéral. Pour un échantillon biologique, la quantification du phosphore total est effectuée après minéralisation de l'extrait dans du $HClO_4$ à 70% à $180^\circ C$ pendant 30 minutes ou jusqu'à ce que la couleur jaune du mélange disparaisse. Après refroidissement des échantillons, le phosphate inorganique obtenu est complexé avec une solution de molybdate d'ammonium à 2,5% (p/v) en présence d'acide ascorbique à 10% (p/v). Le complexe n'étant pas stable, sa réduction par l'acide ascorbique est nécessaire afin d'effectuer le dosage colorimétrique. Le complexe phosphomolybdique-molybdeux de couleur bleu a un maximum d'absorption pour une longueur d'onde de 830 nm . Une gamme étalon est réalisée à partir d'une solution de KH_2PO_4 à $100\ \mu g.mL^{-1}$, les dilutions couvrant la gamme $2-100\ \mu g.mL^{-1}$. La densité optique des échantillons est réalisée à 800 nm . Cette méthode permet de doser des échantillons contenant au moins $2\ \mu g$ de phosphore (ce qui correspond à $50\ \mu g$ de phospholipides). D'autres méthodes permettent ce dosage dans des échantillons contenant moins de $2\ \mu g$ de phosphore inorganique. Il s'agit d'une amélioration de la méthode de ROUSER (Zhou, 1992).

- Quantification de l'eau

Pour estimer la quantité d'eau dans les protéoliposomes, un échantillon est pesé avant et après 16 heures de mise sous vide.

- Caractérisation de la densité des protéoliposomes - Gradient de saccharose

L'analyse de la migration d'un échantillon de protéoliposomes, sous l'effet de la centrifugation, sur un gradient discontinu de saccharose permet d'estimer sa densité. Elle permet également de séparer les protéoliposomes des agrégats de protéine ainsi que des vésicules lipidiques vides. Un gradient discontinu est formé en juxtaposant une par-dessus l'autre quatre solutions ayant des différences suffisantes de densités. De bas en haut, des solutions de moins en moins denses en saccharose sont donc empilées. Deux stratégies

ont été testées : les protéoliposomes sont déposés en bas du gradient dans la solution de saccharose à 45%, suivis d'une couche à 30%, une autre à 10% et une dernière sans saccharose ; les protéoliposomes peuvent être également déposés en haut du gradient dans la fraction sans saccharose, sur trois solutions à 10%, 30% et 45% respectivement. Les différentes solutions de saccharose sont solubilisées dans le tampon *RB*. Les gradients sont réalisés dans des tubes de centrifugation pouvant s'adapter à deux types de rotors utilisables : à godets mobiles ou verticaux. Les rotors à angle fixe ne sont normalement pas utilisables pour ce type d'application. Après ultracentrifugation à 90 000 *g* pendant 1 h 30 minutes et à 18°C, les protéoliposomes se retrouvent à l'interface entre 30 et 45%, les liposomes vides en haut du gradient, et les agrégats de protéines sous forme de culot. Pour récupérer les protéines agrégées autrement, une solution de saccharose plus dense (60%) peut être rajoutée en bas du tube, les protéines se retrouveraient ainsi à l'interface entre 45 et 60%. À la fin de l'ultracentrifugation, des petites fractions sont récupérées délicatement pour éviter le mélange des différentes solutions. Ces fractions sont analysées par *SDS-PAGE* et *WESTERN BLOT*, afin de détecter la présence de *MscL*. Les fractions contenant les protéoliposomes sont ensuite mélangées et diluées afin de réduire la densité, pour pouvoir centrifuger l'échantillon et le débarrasser du saccharose.

d Spectroscopie de résonance magnétique nucléaire

Dans ce chapitre, les techniques utilisées pour réaliser l'étude structurale de la protéine membranaire *MscL* par résonance magnétique nucléaire en phase solide sont présentées. Dans la première partie, la méthodologie employée pour la mise en place d'une expérience *RMN* au niveau du spectromètre permettant l'acquisition des données est décrite. L'analyse des données obtenues sera ensuite exposée dans la deuxième partie.

À l'issue de la reconstitution de la protéine *MscL* dans des vésicules de *DOPC*, l'échantillon est transféré dans un rotor de 4 mm. Le volume du rotor est de 50 μL dont la moitié est occupée par l'eau. Les expériences *RMN* ont été enregistrées sur un spectromètre *BRUKER AVANCE* opérant à la fréquence du proton à 700 MHz, équipé d'une sonde *MAS* permettant la détection des noyaux ^1H , ^{13}C et ^{15}N . Le logiciel *TOPSPIN* (*BRUKER*) permet le suivi des expériences et l'acquisition des spectres. Deux méthodes ont été mises au point et réalisées sur les échantillons uniformément et sélectivement marqués : la première utilise une séquence de type *DARR* pour les corrélations homo-nucléaires entre atomes de carbones et met en évidence des interactions dipolaires à courte et longue portées ; la seconde consiste à déterminer la position des acides aminés ainsi identifiés dans la séquence de la protéine. Il s'agit de la méthode *specific-CP* identifiant des interactions ^{15}N - ^{13}C .

i. Mise au point d'une expérience de *RMN*

La polarisation croisée combinée à la rotation à l'angle magique est la méthode utilisée dans cette étude. La mise au point de cette expérience est expliquée dans ce paragraphe. Les premières étapes consistent à mettre en place et vérifier les différents paramètres et constantes utilisés. L'orientation pour laquelle l'interaction dipolaire est réduite à zéro correspond à l'angle magique ($54,7^\circ$) par rapport au champ statique B_0 . Pour régler l'angle magique, la poudre de bromure de potassium est utilisée (Pour plus de détails voir Frye *et al.*, 1982). Pour corriger les inhomogénéités du champ magnétique les

shims (aimants annexes réglables) doivent être réglés sur un pic isolé et intense de proton. Pour cela un rotor rempli d'eau est utilisé, et les *shims* sont réglés de manière à obtenir le pic le plus fin possible. L'adamantane $C_{10}H_{16}$ est ensuite utilisé pour le référencement des déplacements chimiques du ^{13}C . L'adamantane peut également être utilisé pour régler les *shims* (pour des détails supplémentaires, voir Morcombe & Zilm, 2003).

DARR (^{13}C - ^{13}C)

Les expériences décrites ici suivent le plan de la séquence d'impulsion *DARR* (Voir Takegoshi *et al.*, 2001). La mise au point de la séquence est effectuée sur un échantillon lyophilisé de lysine marqué aux ^{13}C et ^{15}N . Le rotor est inséré et avant de commencer les mesures, la rotation à l'angle magique et la température de l'expérience sont réglées. Une fois ces deux paramètres stables, le réglage des différents paramètres est effectué. L'expérience commence par une impulsion 90° sur le 1H , suivie par une période de transfert d'aimantation (*CP*) entre les noyaux 1H et les noyaux ^{13}C . Après la polarisation croisée, l'aimantation évolue pendant le temps t_1 et est placée le long de l'axe *Z* par une impulsion de 90° sur le ^{13}C . Pendant le temps de mélange, une faible puissance est appliquée sur le canal du 1H afin de permettre le recouplage. Finalement, après la deuxième impulsion 90° sur le ^{13}C , et pendant la détection sur le canal du ^{13}C une impulsion intense de découplage est appliquée sur le 1H pour permettre un retour à l'équilibre avant de recommencer la séquence. Les expériences ont été effectuées pour différents temps de mélange entre 10 et 500 *ms*. Toutes les expériences sont effectuées en rotation à l'angle magique à une fréquence de 11 *kHz*. À cette fréquence et pour une température affichée de 233 *K*, la température effective à l'intérieur du rotor était de 253 *K*. L'effet de la température a été mesuré par la suite sur l'échantillon (*Phe, Ile, Met, Pro, Arg*)- ^{13}C , ^{15}N -*MscL*. L'effet de l'environnement lipidique a été testé sur l'échantillon (*Ala, Gly, Lys, Val, Tyr*)- ^{13}C , ^{15}N -*MscL*. Les paramètres d'acquisition sont indiqués dans les légendes des spectres.

Specific-CP (^{15}N - ^{13}C)

Cette méthode permet de mesurer les interactions entre les noyaux ^{15}N et les noyaux ^{13}C . L'expérience consiste tout d'abord à irradier les protons. Cette aimantation est transférée aux azotes par polarisation croisée. Ensuite, pendant que les protons sont découplés, l'aimantation des ^{15}N évolue. L'aimantation est transférée sélectivement soit aux $\text{C}\alpha$, soit aux C' . Pendant le temps d'évolution de l'aimantation, les protons sont découplés. Cette expérience a été mise au point et des spectres ont été obtenus sur les échantillons U - ^{13}C , ^{15}N - MscL , (*Ile*, *Thr*)- ^{13}C , ^{15}N - MscL , et (*Ala*, *Lys*, *Val*, *Tyr*)- ^{13}C , ^{15}N - MscL .

ii. Analyse des spectres et traitement des données

La transformation des données obtenues au niveau du spectromètre est effectuée en utilisant le programme *bruk2pipe*, permettant de créer le fichier contenant tous les signaux de précession libre ou signal temporel initial (*FID* pour *Free Induction Decay*) en lisant les informations accumulées au niveau du spectromètre à l'aide du logiciel *TOPSPIN*, sauvegardées dans le fichier *ser*. Le traitement des spectres est réalisé avec le logiciel *NMRPipe* (Delaglio *et al.*, 1995). Les différentes fonctions d'apodization, de phasage et la transformation de *FOURIER* sont appliquées successivement selon un schéma de *pipelines*, les dimensions sont transformées les unes après les autres.

Un exemple de macro utilisée pour le traitement d'un spectre bidimensionnel est présenté à la page suivante :

nmrPipe -in test.fid \ Cette première ligne permet la lecture du fichier contenant le *fid*.

| nmrPipe -fn ZF -zf 2 \ L'interpolation ou *zero-filling* permet d'obtenir plus de points dans le spectre final que dans le *FID* en rajoutant des zéros à la fin de la série temporelle avant l'application de la transformée de *FOURIER* ; ici -zf 2 permet de doubler la taille du *FID*.

| nmrPipe -fn LP -fb -after -ord 8 \ La prédiction linéaire permet un prolongement du *FID* au-delà des points effectivement mesurés réalisant une opération proche du *zero-filling* mais au lieu de rajouter des zéros, modélise des points théoriques à partir des points du *FID*. Les points *n* sont calculés à partir du point *n-1* et *n+1* en utilisant le paramètre *-fb*. Ils sont placés après les points du *FID* en utilisant le paramètre *-after*. L'ordre de la prédiction est de 8 et définit le nombre de points successifs, dans chaque groupe, utilisés pour construire le modèle.

| nmrPipe -fn EM -lb 10 -c 1.0 \ L'apodization est l'opération qui permet de multiplier le signal temporel par une fonction du temps. La fonction est choisie de manière à augmenter le rapport signal/bruit apparent sans dégrader sensiblement la résolution ; ici une fenêtre exponentielle est utilisée avec un coefficient d'amortissement de 10 Hz. Le paramètre *-c 1.0* signifie que la totalité du premier point est prise en compte.

| nmrPipe -fn FT -auto -verb \ La transformée de *FOURIER* permet de convertir le signal temporel en fréquences. Les paramètres de la transformée sont choisis automatiquement. Le paramètre *-verb* permet de suivre les opérations explicitement.

| nmrPipe -fn PS -p0 36.0 -p1 180.0 -di \ L'opération de phasage, par combinaison des parties réelles et imaginaires du spectre, se fait en mode interactif par correction de la phase d'ordre zéro *p₀* et de celle de premier ordre *p₁* dans la fenêtre du programme *nmrDraw*, puis en corrigeant les valeurs au niveau de la *macro*.

| nmrPipe -fn EXT -x1 -20ppm -xn 220ppm -sw \ L'extraction de la région contenant les informations intéressantes permet de réduire la taille des fichiers générés. Dans la dimension directe, seules les fréquences entre -20 et 220 ppm seront affichées.

| nmrPipe -fn TP \ Cette fonction est essentielle pour le traitement des données bidimensionnelles car elle permet le passage de la dimension directe à la dimension indirecte pour effectuer sur cette dernière les opérations permettant de transformer le signal temporel en fréquences.

| nmrPipe -fn ZF -zf 2 \ L'interpolation est effectuée à nouveau sur la dimension indirecte.

| nmrPipe -fn LP -fb -after -ord 8 \ L'opération de prédiction linéaire peut aussi être appliquée au niveau de la dimension indirecte, mais elle est à utiliser avec précaution, car généralement cette dimension contient beaucoup moins de points que la dimension directe.

| nmrPipe -fn SP -off 0.4 -end 0.95 -pow 0.9 -c 0.5 \ La fonction d'apodization utilisée dans la dimension indirecte est généralement moins abrupte afin de garder le maximum d'informations pour la transformée de *FOURIER*.

| nmrPipe -fn FT -auto -verb \ Cette fonction permet la transformée de *FOURIER* des *FID* de la dimension indirecte.

| nmrPipe -fn PS -p0 -41.9 -p1 87.0 -di \ Le phasage se fait en mode interactif dans la fenêtre du programme *nmrDraw*.

| nmrPipe -fn EXT -x1 -20ppm -xn 220ppm -sw \ Dans la dimension indirecte, seules les fréquences situées entre -20 et 200 ppm seront extraites.

| nmrPipe -fn TP \ Cette fonction permet le passage de la dimension temporelle à la dimension des fréquences.

| nmrPipe -fn POLY -auto \ Cette opération permet de multiplier le spectre obtenu par une fonction polynomiale afin de corriger la ligne de base.

-ov -out test.ft2 La dernière ligne permet de créer un fichier en sortie contenant le spectre obtenu après cette série de fonctions.

Après transformation, les spectres sont convertis au format *UCSF* par le programme *pipe2ucsf* afin d'être visualisé à l'aide du programme *SPARKY* (Goddard & Kneller, *Sparky 3*). Ce logiciel permet le traitement et l'analyse des données *RMN* obtenues. La ligne de base et les niveaux de contour des pics y sont définis. La ligne de base est choisie en vue d'optimiser le rapport signal/bruit observé sur le spectre. Les niveaux de contour sont définis en guise d'une meilleure résolution des pics de corrélation. Après avoir trouvé les meilleurs paramètres, les spectres sont imprimés dans un fichier *POSTSCRIPT*, et les figures sont créées ensuite à partir de ces fichiers à l'aide du logiciel *ADOBE ILLUSTRATOR* (*ADOBE SYSTEMS INCORPORATED*). L'attribution des pics est facilitée par les menus conviviaux du programme *SPARKY*. Pour les tracés des intensités de certains pics en fonction du temps de mélange, les pics sont sélectionnés, et intégrés à l'aide d'une fonction d'ajustement gaussien.

III Références

- Abdine** A, Verhoeven MV, Park K-H, et al. Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* **2010**;204(1):155-159.
- Baldus** M, Petkova AT, Herzfeld J, Griffin RG. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Physics.* **1998**;95(6):1197-1207.
- Berrier** C, Park K-H, Abes S, et al. Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry.* **2004**;43(39):12585-12591.
- Brown** RE, Jarvis KL, Hyland KJ. Protein measurement using bicinchoninic acid: elimination of interfering substances. *Anal. Biochem.* **1989**;180(1):136-139.
- Cai** ML, Huang Y, Sakaguchi K, et al. An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J. Biomol. NMR.* **1998**;11(1):97-102.
- Dawson** RMC, Elliott DC, Elliott WH. Data for Biochemical Research. Oxford University Press; **1989**.
- Delaglio** F, Grzesiek S, Vuister GW, et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* **1995**;6(3):277-293.
- Frye** JS, Maciel GE. Setting the magic angle using a quadrupolar nuclide. *J. Magn. Reson.* **1982**;48(1):125-131.
- Goddard** TD, Kneller DG. Sparky - NMR Assignment Program. University of California, San Francisco Available at: <http://www.cgl.ucsf.edu/home/sparky/>.
- Hermanson** GT. Bioconjugate techniques. Academic Press; **1996**.
- Katzen** F, Fletcher JE, Yang J-P, et al. Cell-free protein expression of membrane proteins using nanolipoprotein particles. *BioTechniques.* **2008**;45(4):469.
- Laemmli** UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **1970**;227(5259):680-685.
- Milner** HW, Lawrence NS, French CS. Colloidal dispersion of chloroplast material. *Science.* **1950**;111(2893):633-634.
- Morcombe** C, Zilm KW. Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson.* **2003**;162(2):479-486.
- Park** K-H, Berrier C, Lebaupain F, et al. Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem. J.* **2007**;403(1):183-187.
- Park** K-H, Billon-Denis E, Dahmane T, et al. In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *N Biotechnol.* **2011**.
- Sambrook** J, Russell DW. Molecular cloning: a laboratory manual. CSHL Press; **2001**.
- Smith** PK, Krohn RI, Hermanson GT, et al. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **1985**;150(1):76-85.
- Takegoshi** K, Nakamura S, Terao T. C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **2001**;344(5-6):631-637.
- Wong** SS, Wong IJ. Chemical crosslinking and the stabilization of proteins and enzymes. *Enzyme Microb. Technol.* **1992**;14(11):866-874.

Zhou X, Arthur G. Improved procedures for the determination of lipid phosphorus by malachite green. *J. Lipid Res.* **1992**;33(8):1233-1236.

Résultats

C Résultats

Les résultats obtenus sont exposés dans cette partie. Les bilans de synthèse des différents échantillons de *Ec-MscL* suivant les protocoles expérimentaux précédemment décrits sont d'abord mentionnés, ainsi que les différentes expériences de caractérisation des échantillons obtenus, puis finalement les résultats obtenus au niveau du spectromètre sont détaillés. Les différents échantillons de *Ec-MscL* synthétisés *in vivo* ou *in vitro*, marqués uniformément ou sélectivement sur certains résidus, ont été testés par la technique de *PATCH CLAMP* dans le laboratoire d'A. GHAZI. Les tests ont montré que le marquage ne perturbait pas l'activité mécanosensible de la protéine.

Avant d'aborder les résultats concernant la protéine *Ec-MscL*, les spectres uni- et bidimensionnels obtenus sur les vésicules de *DOPC* sont présentés. Une série de spectres acquis à différentes températures ont permis de calibrer la température. Le premier chapitre consacré à *Ec-MscL* décrit les expériences effectuées sur l'échantillon uniformément marqué $U\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$. Les corrélations $^{13}\text{C}\text{-}^{13}\text{C}$ mais aussi $^{13}\text{C}\text{-}^{15}\text{N}$ ont été mesurées. Les résultats obtenus sur les échantillons sélectivement marqués sont décrits par la suite. Le deuxième chapitre comporte les résultats obtenus sur l'échantillon $(Ile, Thr)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$ et la comparaison des informations spectrales entre $(Ile, Thr)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$ et $U\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$, qui a fait récemment l'objet d'une communication (Abdine *et al.*, 2010). Dans le troisième chapitre la stratégie adoptée pour le marquage des acides aminés, et les résultats obtenus sur l'échantillon $(Ala, Lys, Val, Tyr)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$ sont détaillés. Le quatrième chapitre traite de la stratégie de marquage et des résultats obtenus sur les échantillons $(Phe, Ile, Met, Pro, Arg)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$ et $(Asp, Gly, Leu, Ser)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$. Le dernier chapitre sera consacré à l'effet de la composition lipidique, mesuré sur $(Ala, Gly, Lys, Val, Tyr)\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$.

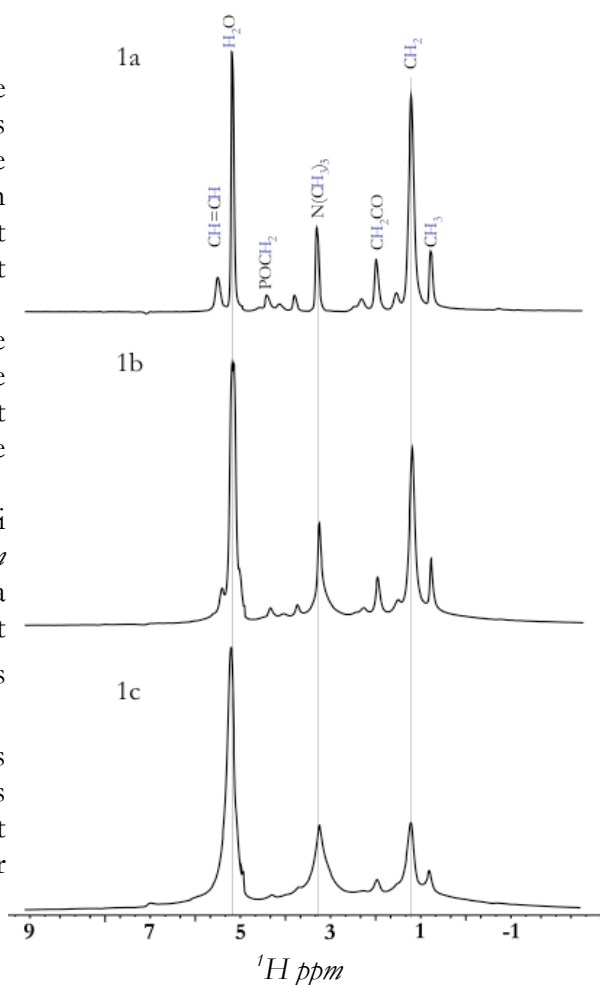
Les expériences RMN ont été effectuées sur un spectromètre *Bruker Avance II* opérant à 16,44 T (équivalent à une fréquence de 700 MHz pour le ^1H) et équipé d'une sonde 4 mm CPMAS pour une triple résonance H/C/N. La vitesse de rotation est fixée à $11\,000 \pm 3\text{ Hz}$. Les expériences sont réalisées à 258 K (sauf exceptions précisées). Les expériences de corrélations ^{13}C - ^{13}C ont été réalisées en utilisant la séquence *DARR*, mais d'autres méthodes de recouplage dipolaire ont été testées, comme le *PDSD* (Castellani *et al.*, 2002 ; Szeverenyi *et al.*, 1982), le B^2R^3 (Duma *et al.*, 2008) ou le *RFDR* (Bayro *et al.*, 2008 ; Griffiths & Griffin, 1993).

Figure 1 : Spectres MAS du ^1H des vésicules de *DOPC* obtenus à différentes températures. Les spectres sont obtenus en appliquant une impulsion de 90° sur le canal du proton, et en enregistrant le *FID* obtenu. Les spectres sont obtenus à 243 (1a), 233 (1b) et 223 K (1c). Huit scans sont enregistrés.

Le traitement des *FID* est effectué en utilisant le programme *MestReNova*. Il consiste en une transformée de *Fourier* suivie d'un phasage et d'une correction automatique de la ligne de base.

Le pic le plus intense présent à 5,2 ppm est celui des protons de l'eau. Le pic à 3,15 ppm correspond aux protons des méthyles de la choline. Les protons des groupements $-\text{CH}_2-$ et $-\text{CH}_3$ des chaînes d'acides gras présentent des pics à 1,2 et 0,8 ppm.

D'autres pics sont indiqués sur la figure 1a. Les raies des protons des chaînes hydrocarbonées sont fines en phase fluide, et s'élargissent lorsque la température diminue, jusqu'à devenir non détectable en phase gel.



Avant de procéder à l'acquisition des données spectroscopiques, l'effet de la rotation à l'angle magique sur la température de l'échantillon, et l'effet des impulsions utilisées lors des séquences *DARR* ont été testés. Ces expériences ont été effectuées sur des vésicules de *DOPC*. Dans les conditions physico-chimiques de préparation des vésicules, la

température de transition de la *DOPC* est située à 253 K . Pour observer l'effet de la rotation à l'angle magique sur la température, le rotor contenant les lipides est inséré dans la sonde à température ambiante, puis la rotation à l'angle magique est fixée à $11\ 000\text{ Hz}$. La température cible est établie, et après 1 h , les spectres unidimensionnels du ^1H sont enregistrés. Les données spectrales ont été traitées identiquement en utilisant le programme *MESTRENOVA* (*MESTRELAB RESEARCH*). Après la transformée de *FOURIER*, une correction de phase automatique est effectuée, suivie d'une correction de la ligne de base selon une polynomiale de *BERNSTEIN*.

La figure 1 montre ces spectres à trois différentes températures (243 , 233 et 223 K). Les résonances ont été résolues et attribuées, en se basant sur des travaux antérieurs (Halladay *et al.*, 1990 ; Zhou *et al.*, 1999). Lorsque les phospholipides sont en phase fluide, leur mobilité rend les interactions dipolaires proton-proton homogènes, et les spectres sont alors bien résolus. Le spectre 1a obtenu à 243 K indique clairement une phase fluide. À 233 K , les fortes interactions dipolaires proton-proton sont à l'origine de l'élargissement des raies de résonance et ne peuvent être éliminées par la rotation à l'angle magique. Le spectre 1c obtenu à 223 K indique clairement une phase gel. Les raies de résonance des protons de la chaîne hydrocarbonée sont bien résolues.

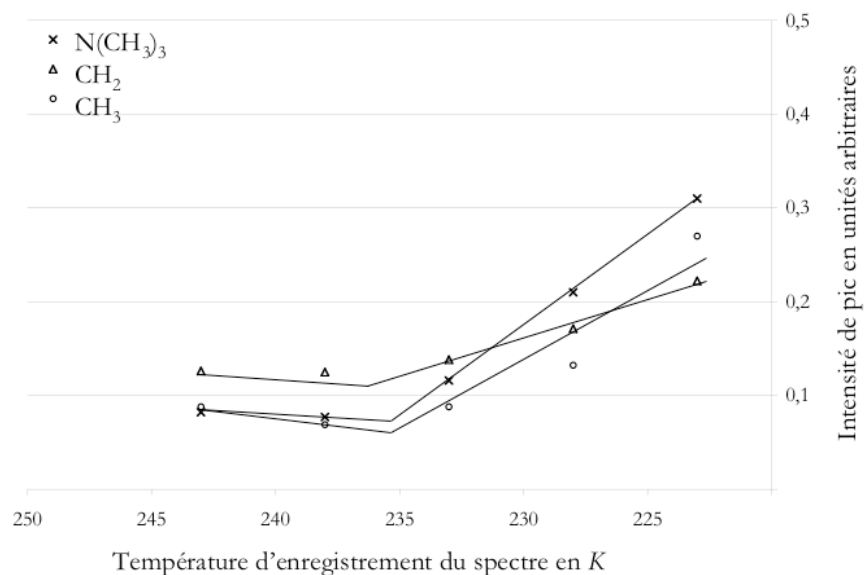


Figure 2 : Evolution de la largeur de raie de protons de la *DOPC* en fonction de la température. La largeur des raies des signaux est mesurée sur les spectres obtenus à 243 , 238 , 233 , 228 et 223 K . Les signaux sont ceux qui sont mentionnés dans la figure 1. Le tracé de la largeur de raie en fonction de la température montre une augmentation des valeurs à partir de 240 - 235 K . La température de transition de la *DOPC* est donc située entre ces deux valeurs.

Les largeurs de raies des protons de la tête polaire $N(CH_3)_3$, de ceux des groupements CH_2 et CH_3 des chaînes d'acides gras en fonction de la température sont présentées dans la [figure 2](#).

Les valeurs obtenues à cinq différentes températures montrent qu'à 243 K , les lipides sont déjà en phase gel. Ceci indique que la température de transition est située entre 238 et 243 K . Théoriquement celle-ci est de 253 K , donc pour une température x affichée au niveau du spectromètre, pour un échantillon en rotation à l'angle magique à une fréquence de $11\ 000\text{ Hz}$, la température effective est située entre $x+10$ et $x+15$.

Des spectres $^{13}\text{C-MAS}$ uni- et bidimensionnels ont été obtenus sur les vésicules de *DOPC*. L'acquisition de ces spectres est réalisée dans les mêmes conditions que pour les expériences concernant les différents échantillons de *Ec-MscL*. Ces expériences permettent d'obtenir les déplacements chimiques des différents carbones des molécules de *DOPC* pour éviter de les confondre, au niveau des spectres de *Ec-MscL*, avec des carbones de la protéine. En effet, dans un échantillon typique de *Ec-MscL*, pour 3 mg de protéines, il faut 6 mg de lipides. La formule chimique de la *DOPC* est $C_{44}H_{84}NO_8P$ et la masse moléculaire est de $786,1\text{ Da}$. Dans 6 mg de *DOPC*, il y a $7,6 \times 10^6$ moles et donc $4,6 \times 10^{18}$ molécules de *DOPC*. Celles-ci contiennent 2×10^{20} atomes de carbone. La *DOPC* utilisée contient l'isotope ^{13}C du carbone en une abondance naturelle. Sur 2×10^{20} noyaux de carbone, cela correspond à $2 \times 10^{18}\text{ }^{13}\text{C}$. Ceci est largement suffisant pour obtenir un signal. La [figure 3](#) montre un spectre $1D$ obtenu sur les vésicules de *DOPC*. Le spectre est obtenu dans des conditions identiques à celles des expériences sur *Ec-MscL*. De même un spectre $2D$ est obtenu sur cet échantillon afin d'obtenir l'empreinte des lipides dans les spectres $2D$ obtenus sur *Ec-MscL* et exposés ultérieurement (voir [Figure 4](#)). Les spectres $^{13}\text{C-MAS}$ uni- et bidimensionnels obtenus sur les liposomes de *DOPC* permettent d'obtenir les déplacements chimiques et les corrélations des lipides dans les conditions expérimentales utilisées. La [figure 3](#) présente une bonne résolution au niveau du spectre $1D$. La [figure 4](#) montre les régions dans lesquelles des corrélations dues aux ^{13}C présents à l'état naturel dans les lipides pourraient apparaître au niveau des spectres $2D$ lors de l'étude de la protéine.

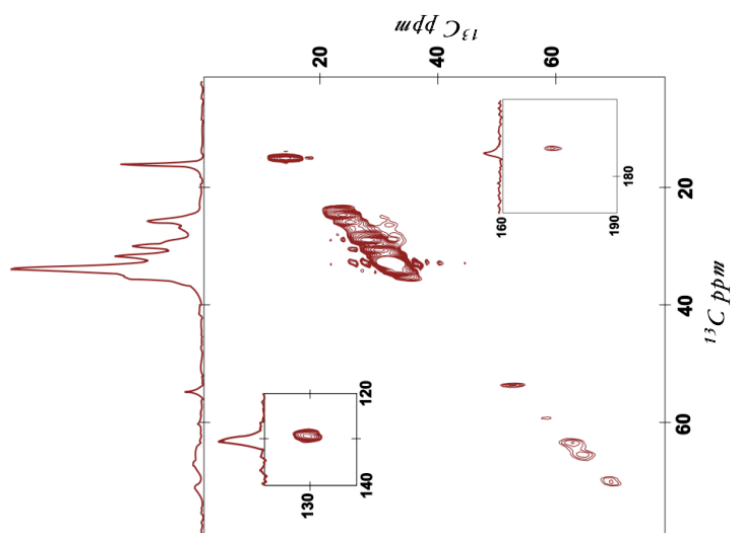


Figure 4 : Spectre ^{13}C -MAS bidimensionnel des vésicules de DOPC. L'expérience est réalisée sur des vésicules multilamellaires de DOPC dans un tampon HEPES-KCl ($10\text{ mM} - 100\text{ mM}$) à $\text{pH } 7,5$. Le spectre ^{13}C - ^{13}C est enregistré à $11\ 000\text{ Hz}$ et à 258 K , avec un temps de mélange de 50 ms . La figure principale contient la région $0\text{-}80\text{ ppm}$. Le signal du carbone de la double liaison de la chaîne aliphatique est représenté dans le carré en haut à gauche. Celui des fonctions carbonyles est situé dans le carré en bas à droite.

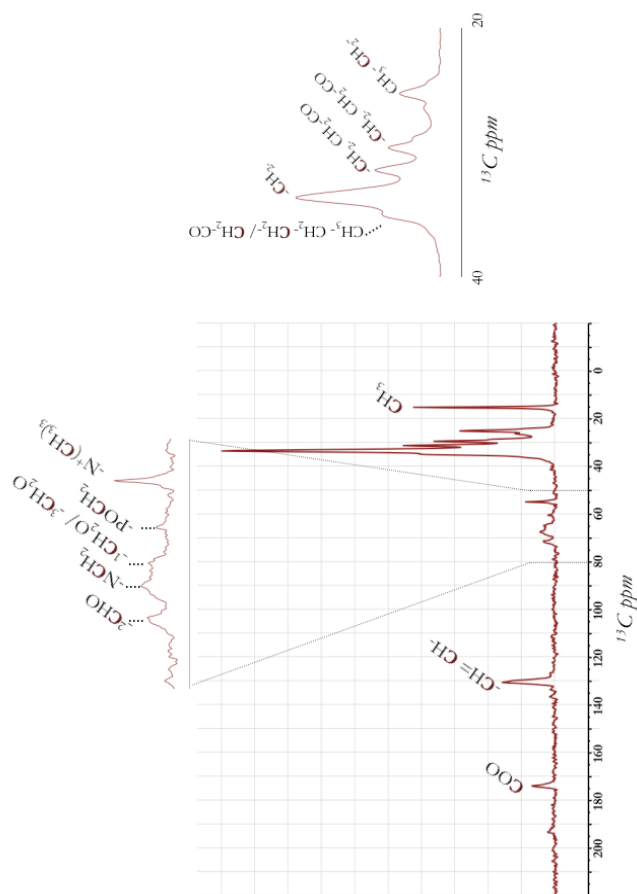


Figure 3 : Spectre ^{13}C -MAS unidimensionnel des vésicules de DOPC. L'expérience est réalisée sur des vésicules multilamellaires de DOPC dans un tampon HEPES-KCl ($10\text{ mM} - 100\text{ mM}$) à $\text{pH } 7,5$. Le spectre est enregistré à $11\ 000\text{ Hz}$ et à 258 K . La figure principale contient le spectre entier (256 scans) où quelques pics sont annotés. Les régions les plus encombrées sont agrandies. À droite figure l'agrandissement de la région entre 20 et 40 ppm , et en haut celui de la région 50 à 80 ppm . Les attributions sont basées sur des expériences de la littérature scientifique (Forbes *et al.*, 1988 ; Neumann *et al.*, 1985 ; Warschawski, 1995).

I Analyse de l'échantillon uniformément marqué

La protéine *Ec-MscL* étudiée comporte 136 acides aminés, dont la séquence est présentée dans la [figure 1.1](#). Le plasmide *pIVEX* a été optimisé pour la *SIV*, mais il est également utilisé en surexpression bactérienne, et permet d'obtenir un bon rendement. La surexpression est effectuée dans la souche *BL21 (DE3)* d'*E. coli* dans un milieu minimum où les seules sources de carbone et d'azote sont respectivement du glucose marqué au ^{13}C et du chlorure d'ammonium marqué au ^{15}N . La protéine contient une étiquette histidine en C-terminal, et sa purification se fait donc sur colonne de *Ni-NTA*. La séquence complète du His-tag est *PGGGSHHHHHH* (PM 1,2 *kDa*). Après purification, la protéine est reconstituée en liposomes de *DOPC*, puis analysée par RMN. La séquence *DARR* a été principalement utilisée. Différents temps de mélanges allant de 1 à 500 *ms* ont été testés. Les expériences *DARR* sur l'échantillon $U\text{-}^{13}\text{C},^{15}\text{N}\text{-MscL}$ en variant le temps de mélange ont toutes été effectuées à 248 K. D'autres séquences d'impulsions (B^2R^3 et *PDSD*) ont été testées sur l'échantillon à 248 K. Des mesures ont été effectuées pour comparer les séquences *DARR* et *PDSD* à plus haute température (268 K).

a Production et purification de *U-MscL*

10 20 30 40 50 60 70
 | | | | | | |
 MSIIKEFREFAMRGNVVDLAVGVIIIGAAFGKIVSSLVADIIMPPGLLLIGGIDFKQFAVTLRDAQGDIPA
 VVMHYGVFIQNVDFLIVAFAlFMAIKLINKLNRRKKEEPAAPAPTKEEVLLTEIRDLLKEQNNRSPGGG
 SHHHHHH

Figure 1.1 : Séquence de *MscL* d'*E. coli* dans le plasmide *pIVEX*. La protéine *Ec-MscL* est composée de 147 acides aminés. Son poids moléculaire est de 16,2 *kDa*. Sur cette séquence, les acides aminés hydrophobes *A*, *L*, *I*, *V*, *M* et *W* sont en rouges. Ces derniers sont au nombre de 62 sur 147 acides aminés soit 42,2 % de la protéine (*NPS@*: *Network Protein Sequence Analysis*. Combet *et al.*, 2000).

Le protocole de production et de purification de la protéine a été conçu et optimisé avec les impératifs de la biologie structurale en général, et de la RMN en particulier : une production à grande échelle dans un milieu permettant une maîtrise totale des sources de carbone et d'azote. Après le test des différentes souches ayant intégré le plasmide, une

préculture est réalisée à partir d'un clone dans un milieu de culture riche, en présence d'ampicilline. Le lendemain, la préculture est centrifugée et les bactéries sont suspendues dans une solution du tampon *M9* avant l'inoculation du milieu de culture. La culture en milieu minimum *M9* contenant du glucose et du NH_4Cl est débutée à une DO_{600} de 0,1 à 30°C. La croissance bactérienne est suivie au cours du temps, et lorsque les bactéries arrivent en fin de phase exponentielle de croissance, généralement autour d'une DO_{600} de 0,8, l'induction est réalisée pendant 3 heures. Tous ces paramètres ont fait l'objet d'une optimisation précédemment dans le laboratoire.

Pour la préculture, d'autres variantes ont été testées au cours de ce travail: une préculture en milieu riche pendant 4 heures, suivie d'un transfert des bactéries en milieu minimum à 37°C pendant la nuit ; une préculture directement en milieu minimum à 37°C pendant la nuit. Dans les deux cas, la densité atteinte était plus faible que lorsque la préculture est réalisée toute la nuit en milieu riche, et la vitesse initiale après inoculation était plus grande. Cependant, la phase stationnaire était atteinte rapidement, et l'induction était réalisée à des valeurs inférieures à 0,8. Les densités atteintes après induction étaient également inférieures au cas de préculture réalisée entièrement en milieu riche.

Différentes températures de croissance en milieu *M9* ont été testées également pendant ce travail. En effet, la surexpression d'*Ec-MscL* en bactéries *BL21* a été réalisée à 23, 30 et 37°C ; la croissance bactérienne est d'autant plus rapide que la température est élevée. À 23°C, il faut 14 h pour atteindre une DO_{600} de 0,8, 8-10 h à 30°C et moins de 8 h à 37°C. Cependant, le rendement de surexpression est meilleur lorsque la croissance bactérienne est réalisée à 30°C. Différentes concentrations d'*IPTG* ainsi que différents temps d'induction ont été testés. Aucune variation significative n'a été observée. En ce qui concerne les produits marqués, le chlorure d'ammonium $^{15}NH_4Cl$ est rajouté dans le milieu minimum *M9* au début de l'expérience. Le glucose marqué au ^{13}C est rajouté au moment de l'induction à l'*IPTG*, pour diminuer le coût de l'expérience.

À la fin de la surexpression, les bactéries sont récupérées par centrifugation, et conservées à -20°C, jusqu'au broyage des cellules. Généralement, un passage en presse de

FRENCH suffit pour casser la majorité des bactéries. La mesure de DO_{600} est le seul moyen permettant d'estimer quantitativement le broyage. Une chute de 80% de la DO_{600} témoigne du bon déroulement de la presse de *FRENCH*. Si la baisse est beaucoup plus faible, un deuxième passage est nécessaire. Cette technique permet de casser les bactéries, et de former des vésicules inversées contenant les protéines membranaires. La solubilisation de ces vésicules en *triton-x-100* à température ambiante ($\sim 20^{\circ}\text{C}$) est rapide et efficace. Une demi-heure est généralement suffisante. Si la solution n'est pas homogène, une centrifugation ou une filtration est réalisée, cela afin d'éviter le risque d'obstruction des capillaires par des particules ou fragments membranaires en suspension. La purification de *Ec-MscL* est réalisée sur colonne de *Ni-NTA*. Les différentes étapes de purification sont suivies en temps réel au niveau de l'écran relié au système de purification. À la fin de la purification, les fractions éluées sont analysées par *Dot Blot*, pour y détecter la présence de la protéine. Les fractions contenant *Ec-MscL* sont rassemblées. Une dialyse est effectuée pour réduire la concentration d'imidazole.

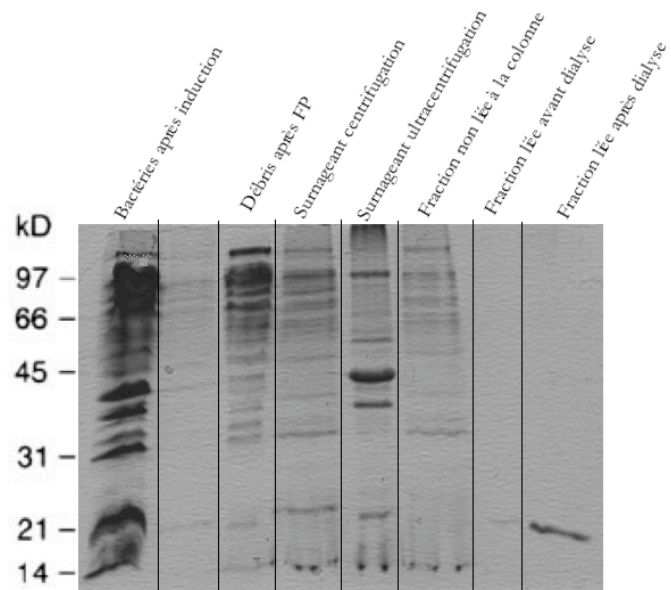
b Reconstitution de *U-MscL* en liposomes de *DOPC*

Après la purification et la dialyse, la protéine *Ec-MscL* se trouve donc en micelles de *triton-x-100*. La seule étape restante au niveau de la préparation de l'échantillon pour la RMN en phase solide est la reconstitution de la protéine en liposomes. Pour cela il est nécessaire de connaître la concentration de la protéine. La quantification se fait par dosage au *BCA*. En parallèle, les divers échantillons obtenus au cours de la surexpression et de la purification sont analysés, qualitativement, par *SDS-PAGE* puis *WESTERN BLOT*. À partir de 3 L de *M9*, et après purification et dialyse, le dosage indique 3 mg de *Ec-MscL*. L'échantillon obtenu est pur, comme en témoigne le profil de migration sur *SDS-PAGE* (voir [figure 1.2](#)).

Pour reconstituer 3 mg de *Ec-MscL* en liposomes, 6 mg de *DOPC* sont nécessaires. Les vésicules multilamellaires sont rajoutées aux micelles de *triton-x-100* contenant la protéine *Ec-MscL*. Le détergent est ensuite extrait en utilisant les *BIO-BEADS*. Les liposomes contenant la protéine *Ec-MscL* sont récupérés par ultracentrifugation et gardés à -20°C pour les expériences de RMN. Après quantification de l'eau, des lipides et des protéines, il

s'est avéré que le complexe était composé de 66% d'eau, de 23% de lipides et de 11% de protéines.

Figure 1.2 : Suivi de la purification de la protéine par *SDS-PAGE*. Les différentes fractions obtenues au cours de la surexpression, la presse de *French*, les centrifugations, la purification et la dialyse sont déposées sur un gel *SDS-PAGE* (les quantités de protéine déposées ne sont pas constantes). Le profil de migration permet le suivi des différentes étapes permettant d'obtenir la protéine *Ec-MscL* par surexpression en bactéries. La bande figurant dans le dernier puits présente l'échantillon de *Ec-MscL* qui est par la suite reconstitué en liposomes puis analysé par *RMN*.



c Analyse de *U-MscL* par *RMN*

La protéine *Ec-MscL* contenant l'étiquette histidine et uniformément marquée est composée de 147 résidus. Chaque molécule contient 738 atomes ^{13}C et 198 atomes ^{15}N . Le poids moléculaire après le marquage isotopique est de 17,075 *kDa*. Les protéoliposomes contenant la protéine uniformément marquée sont transférés dans le rotor adéquat et conservés à -20°C jusqu'aux expériences *RMN*. Après la mise au point des différentes valeurs des séquences d'impulsions utilisées, le rotor est placé à température ambiante. Il est ensuite inséré dans la sonde *MAS*, la température est abaissée (valeur indiquée 233 *K*) et la rotation à l'angle magique est augmentée par paliers (jusqu'à 11 000 *Hz*). Une fois ces deux paramètres stabilisés, la sonde est accordée finement aux fréquences des noyaux étudiés.

Le premier spectre *2D* a été réalisé avec un temps de mélange de 1,5 *ms*. Quelques corrélations $C'-C\alpha$ et $C\alpha-C\beta$ sont visibles sur le spectre (**Figure 1.3**). Dans la région 20-40

ppm , des pics sont présents. Ce sont probablement des corrélations entre des carbones de la protéine et ceux de la chaîne latérale des acides gras de la *DOPC*. Certaines des corrélations ^{13}C observées sont isolées et ont une largeur de raie comprise entre 1,5 et 2 ppm . Le spectre 1D observé sur la même figure présente des signaux peu dispersés. La largeur de raie du pic correspondant aux C' est de 5 ppm . Ce pic contient les déplacements chimiques de 147 C' . La région des $C\alpha$ (45-65 ppm) présente plus de dispersion. Les signaux des $C\beta$ et des carbones aliphatiques des chaînes latérales occupent toute la zone entre 5 et 45 ppm .

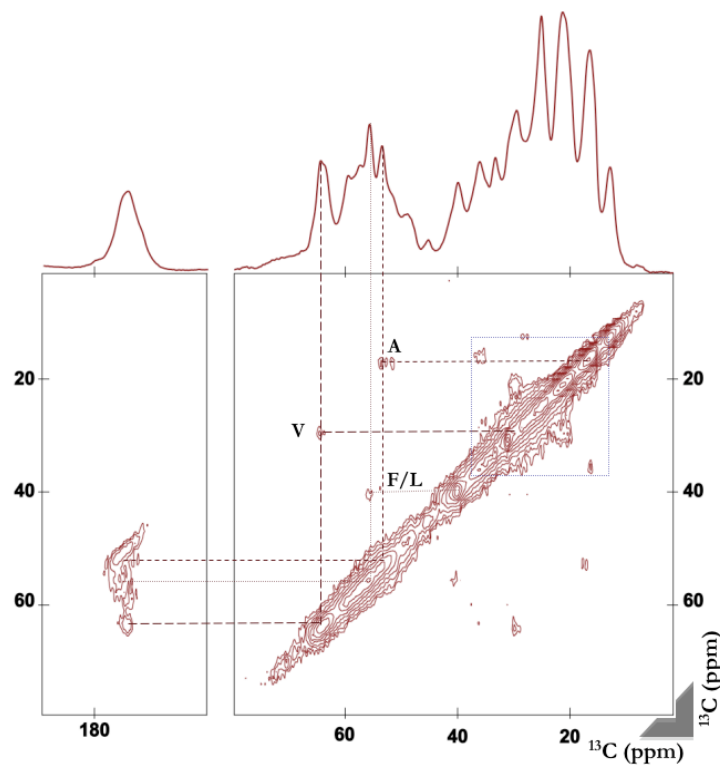


Figure 1.3 : Spectre ^{13}C - ^{13}C obtenu sur l'échantillon $u\text{-}^{13}\text{C}$ - ^{15}N -*MscL* en utilisant la séquence *DARR* et en appliquant un temps de mélange de 1,5 *ms*. Quelques pics sont identifiés : il s'agit des corrélations $C\alpha$ - $C\beta$ des résidus *Ala* (57-17 ppm), des résidus *Phe* ou *Leu* (65-40 ppm) et des résidus *Val* (65-30 ppm). Les corrélations CO - $C\alpha$ des mêmes résidus sont également visibles à 176 ppm . Les pics figurant dans le petit encadré sont probablement des corrélations entre des $C\alpha$ et des $\text{-CH}_2\text{-}$ des chaînes latérales de la *DOPC*.

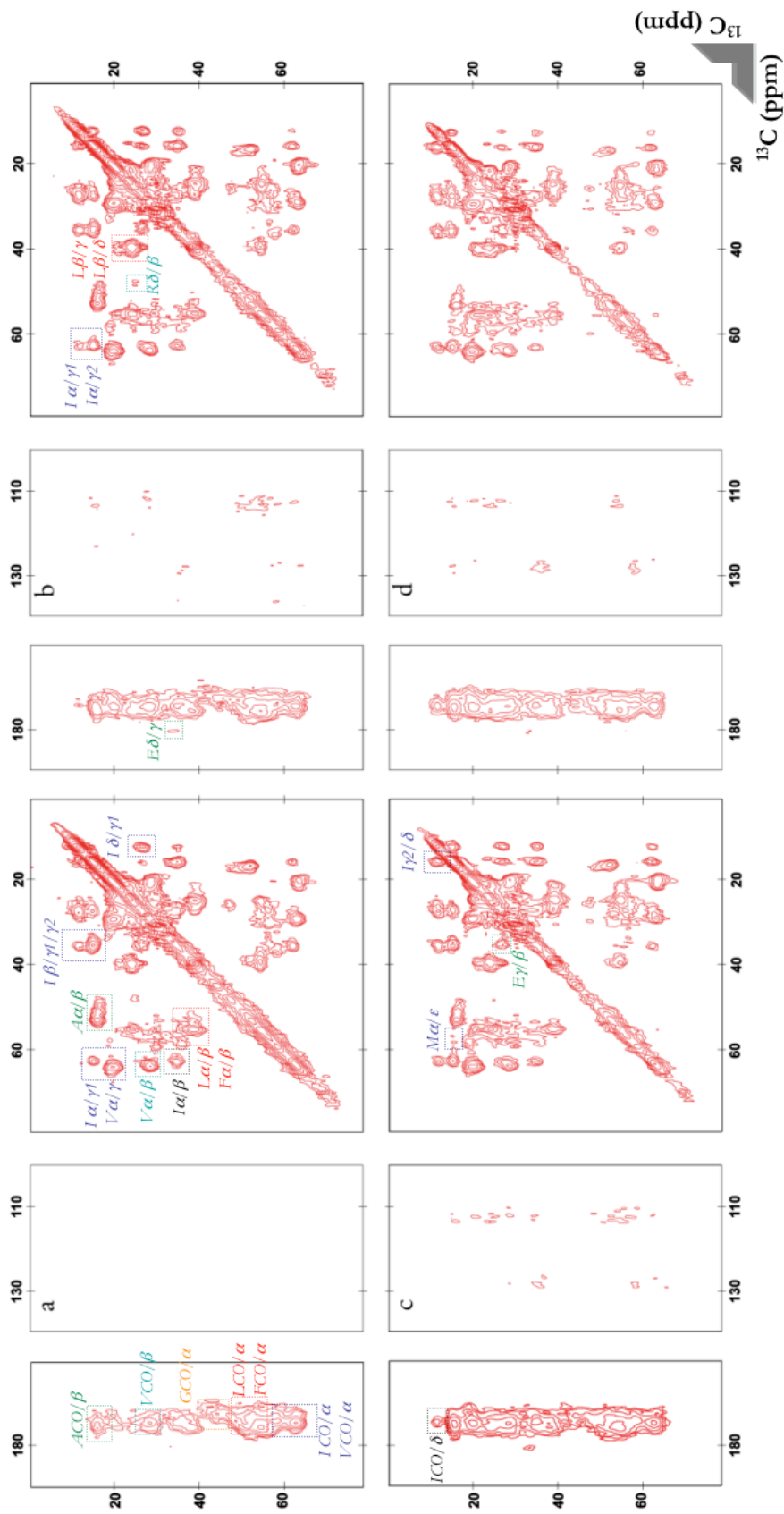


Figure 1.4 : Spectres ^{13}C - ^{13}C obtenus sur l'échantillon $\mu\text{-}^{13}\text{C}$ - ^{15}N -MscL en utilisant la séquence DARR et en appliquant des temps de mélange de 10 (a), 25(b), 50 (c) et 100 ms (d).

Pour obtenir des informations supplémentaires sur la protéine, des spectres $2D^{13}C-^{13}C-DARR$ à différents temps de mélange ont été enregistrés. À partir de 10 ms de temps de mélange (Figure 1.4a), beaucoup de corrélations $C'-C\alpha$ et $C\alpha-C\beta$ sont observables. Dans la région des C' , les pics ne sont pas dispersés. Dans la région correspondant aux $C\alpha-C\beta$, les corrélations visibles sur le spectre obtenu avec $1,5\text{ ms}$ de temps de mélange sont plus intenses. D'autres corrélations sont présentes. Les corrélations $C\alpha-C\beta$ des résidus les plus hydrophobes et les plus représentés sont toutes présentes (*Ala* 55-19 ppm, *Ile* 62-38 ppm, *Val* 65-30 ppm, *Leu* ou *Phe* 56-40 ppm). Des corrélations sont présentes aussi au niveau 54/58-20/40 ppm, mais le recouvrement des fréquences de résonance ne permet pas de les identifier. D'autres corrélations dans la région des $C\alpha$ sont identifiables. Les pics présents à 65-20 ppm et 62-18 ppm sont des corrélations $C\alpha-C\gamma$ des *Val* et *Ile* respectivement. Dans la région correspondant aux carbones aliphatiques des chaînes latérales, plusieurs pics sont présents. Les seuls identifiables sont les pics de corrélations des chaînes latérales des *Ile* et des *Val*. En augmentant le temps de mélange à 25 ms , les corrélations sont plus intenses, les signaux des carbones liés par plus d'une liaison sont plus marqués (Figure 1.4b). De nouveaux signaux apparaissent à partir de 50 ms , les pics sont plus fins dans la région 0-80 ppm (Figure 1.4c). Les signaux sont moins intenses pour un temps de mélange de 100 ms (Figure 1.4d). Sur tous ces spectres, aucune des corrélations observées ne correspond au signal d'un seul noyau. L'évolution du signal en fonction du temps de mélange est mesurée sur quelques pics fins. Le maximum d'intensité est localisé entre 25 et 50 ms de temps de mélange. Pour un temps de mélange supérieur à 50 ms , le signal décroît (Figure 1.5). Sur un spectre enregistré avec un temps de mélange de 500 ms , le signal disparaît complètement.

La séquence *DARR* a été comparée à d'autres séquences d'impulsion : *PDSD* et B^2R^3 (Figure 1.6). Le signal obtenu en utilisant la séquence *PDSD* est nettement moins intense que celui des deux autres séquences. La séquence *PDSD* (pour *PROTON DRIVEN SPIN DIFFUSION*) contient une impulsion sur les protons puis la magnétisation est transférée aux noyaux ^{13}C par polarisation croisée ; ensuite, pendant le temps t_1 une séquence de découplage est utilisée pour arrêter le couplage $^1H-^{13}C$ et une impulsion 90° est utilisée pour ramener l'aimantation des noyaux ^{13}C sur l'axe z . Par la suite, pendant le temps de

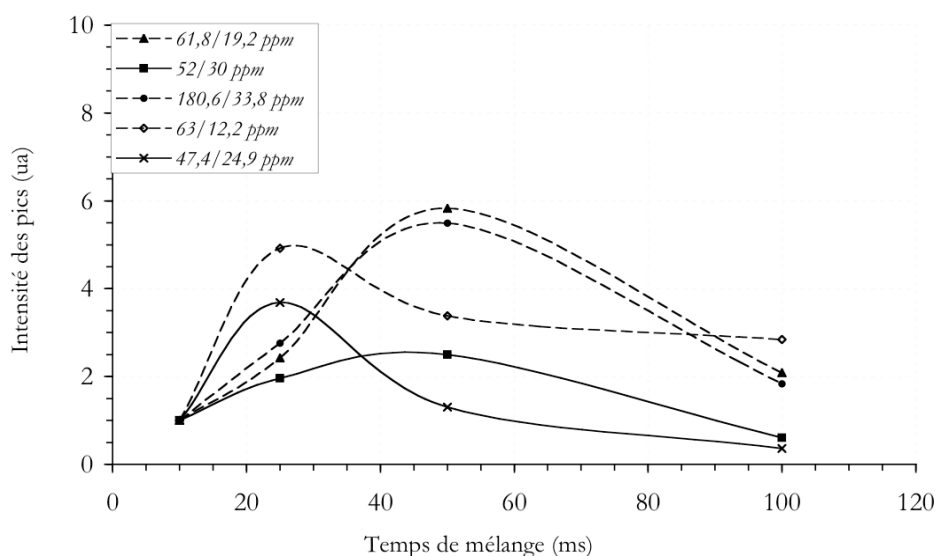


Figure 1.5 : Evolution de l'intensité des corrélations ^{13}C - ^{13}C en fonction du temps de mélange DARR. Cinq pics ont été isolés au niveau des spectres obtenus à 10, 25, 50 et 100 ms. Deux correspondent à ces corrélations entre deux carbones reliés par une seule liaison, les trois autres pics à deux carbones reliés par deux liaisons. Une fenêtre gaussienne est utilisée pour intégrer l'aire sous les pics. Dans chaque série, les intensités sont ramenées à celle du premier point. La figure montre que pour toutes les corrélations, les maxima sont atteints pour un temps de mélange entre 25 et 50 ms. Le pic 61,8/19,2 représente une corrélation $\text{C}\alpha$ - $\text{C}\gamma$ d'un résidu *Ile* ou *Val*, ce qui correspond à une distance de 3 Å. Le pic 52/30 ppm correspond à une corrélation $\text{C}\alpha$ - $\text{C}\beta$ d'une *His*, *Arg* ou *Lys*, la distance entre les deux carbones étant de 1,5 Å. Le pic 180,6-33,8 ppm représente une corrélation C' - $\text{C}\gamma$ d'une *Glu* ou *Gln* pour une distance de 2,5-3 Å. Le pic 47,4/24,9 ppm représente une corrélation $\text{C}\delta$ - $\text{C}\beta$ d'une *Arg* pour une distance de 1,5 Å. Finalement le pic 63/12,2 ppm représenterait une corrélation $\text{C}\alpha$ - $\text{C}\delta$ d'une *Ile* ou une corrélation inter-résiduelle entre le $\text{C}\alpha$ d'une *Ile*, *Val* ou *Pro* et le $\text{C}\delta$ d'une *Ile*.

mélange t_{mix} , la diffusion de spin a lieu entre les noyaux ^{13}C , tandis que pour la séquence d'impulsion DARR, une faible irradiation de radiofréquence facilite la diffusion de spin. Dans la séquence B^2R^3 (pour *BROADBAND VARIANT OF ROTARY RESONANCE RECOUPLING*), pendant le temps de mélange t_{mix} , aucune impulsion n'est appliquée au niveau des noyaux ^{13}C mais deux impulsions de radiofréquence sont appliquées au niveau des noyaux ^1H , d'une amplitude proche de la fréquence de rotation, et de phase variable. En superposant la région des corrélations entre les $\text{C}\alpha$ et les carbones des chaînes latérales, le spectre obtenu avec la séquence B^2R^3 paraît contenir plus d'informations. Cependant, en comparant les intensités des pics isolés, il s'est avéré que la résolution des pics isolés obtenus sur le spectre DARR était de $1,3 \pm 0,2$ ppm. Au niveau des spectres B^2R^3 la résolution était de $1,5 \pm 0,3$ ppm.

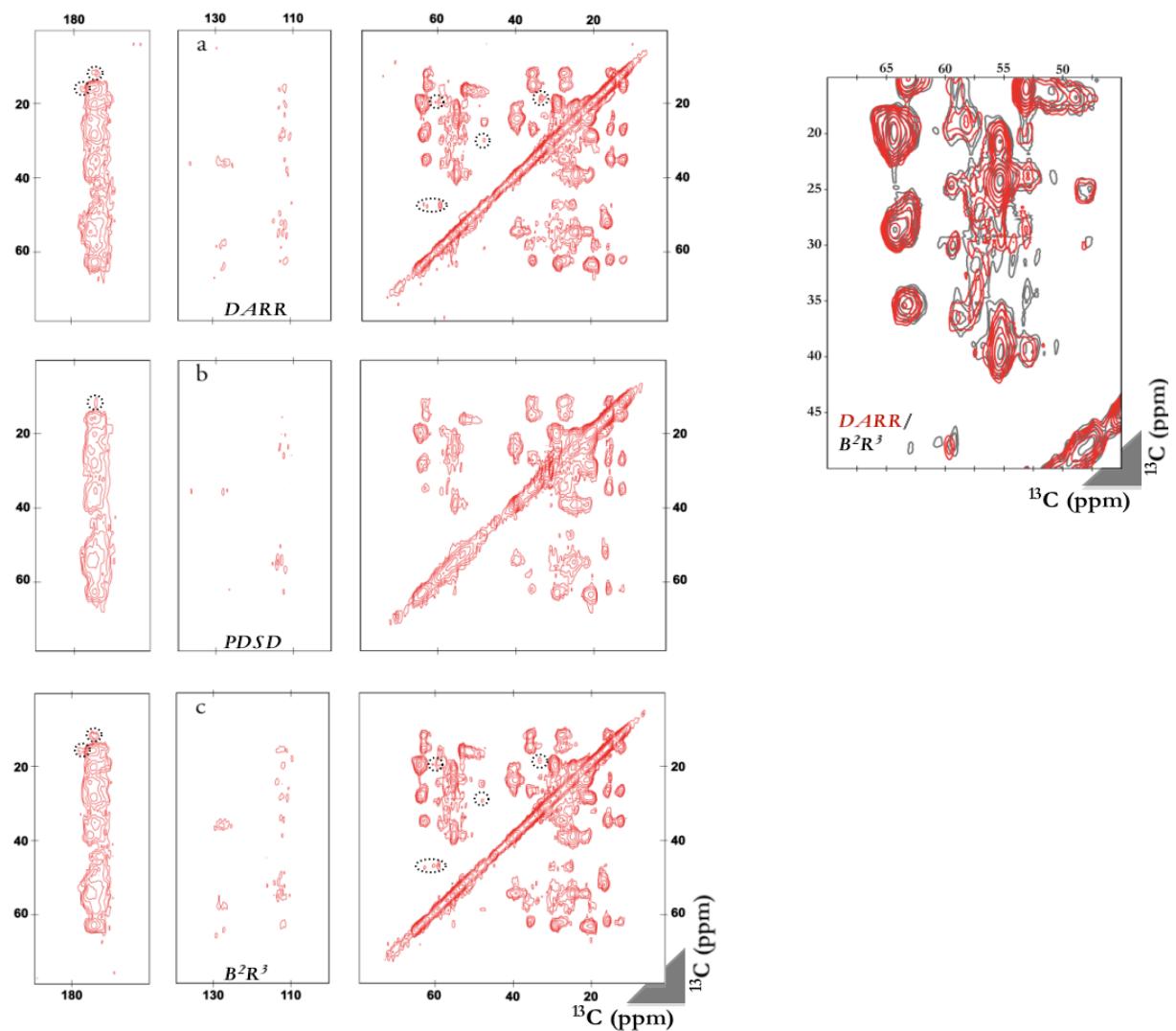


Figure 1.6 : Spectres ^{13}C - ^{13}C obtenus sur l'échantillon u - ^{13}C - ^{15}N -MscL en utilisant les séquences d'impulsion *DARR* (a), *PSD* (b) et B^2R^3 (c) avec un temps de mélange de 50 ms. Certaines des corrélations sont isolées et ont une largeur de raie entre 1 et 1,8 ppm. La région des corrélations des $C\alpha$ avec les carbones des chaînes latérales obtenues avec les séquences *DARR* et B^2R^3 sont superposées ci-contre. La figure montre que le spectre B^2R^3 contient plus de signal et les corrélations sont plus intenses pour le B^2R^3 mais de meilleure résolution avec la séquence *DARR*.

En augmentant le temps d'acquisition de 14 à 56 h, des signaux supplémentaires apparaissent au niveau du spectre *DARR* avec un temps de mélange de 50 ms (Figure 1.7). Les corrélations $C\alpha$ - $C\beta$ des résidus *Thr* et *Ser* sont visibles. Le recouvrement des pics empêche de voir les autres corrélations $C\alpha$ - $C\beta$ situées entre 50-60 ppm pour les $C\alpha$ et 30-40 ppm pour les $C\beta$. Dans la région des carbonyles, les corrélations C - $C\alpha$, C' - $C\beta$ et C' - $C\delta$

des *Glu* sont visibles également. Les corrélations dans la région des carbones aromatiques sont beaucoup plus intenses. La mesure de la largeur des rares pics isolés montre une résolution comprise entre 1,2 et 1,5 ppm.

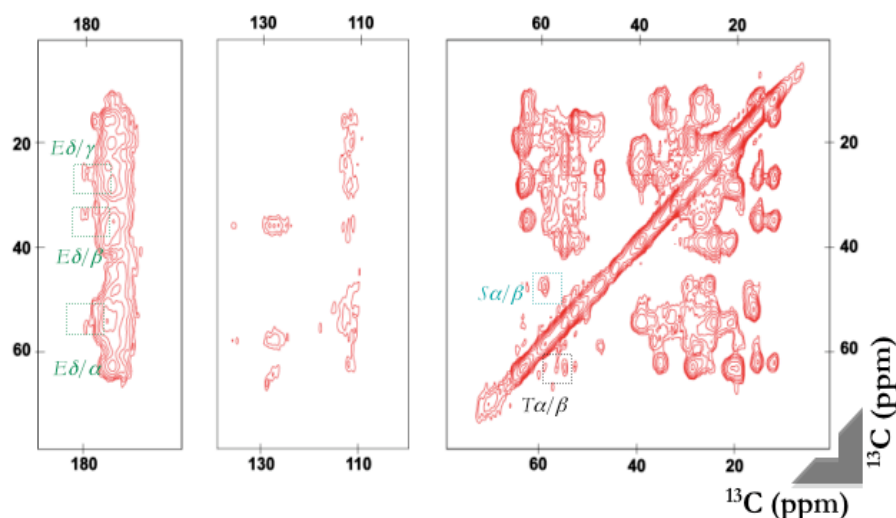


Figure 1.7 : Spectre ^{13}C - ^{13}C obtenu sur l'échantillon u - ^{13}C - ^{15}N -MscL en utilisant la séquence d'impulsion DARR, en appliquant un temps de mélange de 50 ms et par enregistrement de 256 scans. Des corrélations $C\alpha$ - $C\beta$ des résidus *Ser* et *Thr* sont visibles, ainsi que les corrélations entre le $C\delta$ des *Glu* et les autres carbones de la chaîne latérale et le $C\alpha$. Les signaux sont intenses, mais le recouvrement des signaux empêche les attributions ses systèmes de spins.

La diminution de la température permet d'une part une amélioration de la résolution spectrale en augmentant la différence entre les deux niveaux d'énergies de *BOLTZMANN* et en réduisant la dynamique moléculaire, mais provoque d'autre part l'augmentation du temps de relaxation spin-réseau $T1$. De plus, lors du refroidissement de l'échantillon de protéines, plusieurs conformations peuvent être gelées. Tout cela provoquerait un élargissement de la largeur de raie diminuant ainsi la résolution spectrale. Pour vérifier cette hypothèse, un test a été effectué en procédant à une acquisition d'un spectre $2D$ ^{13}C - ^{13}C -DARR à 268 K sur le même échantillon utilisé auparavant. Le temps de mélange utilisé était de 20 ms afin d'obtenir surtout les corrélations entre carbones liés par une seule liaison covalente. La résolution est améliorée comme en témoigne la largeur de pic des corrélations $C\gamma$ - $C\alpha$ et $C\gamma$ - $C\beta$ d'un système de spin isolé d'une thréonine (**Figure 1.8**). Les mêmes déplacements chimiques sont présents sur le spectre DARR obtenu avec un temps de mélange de 25 ms, à 248 K, mais le pic $C\gamma$ - $C\beta$ est très peu intense, et peut être confondu avec le bruit.

Néanmoins, après l'enregistrement du spectre *DARR* à 268 K, un spectre unidimensionnel ^1H a été enregistré, et montre une dégradation des phospholipides des protéoliposomes.

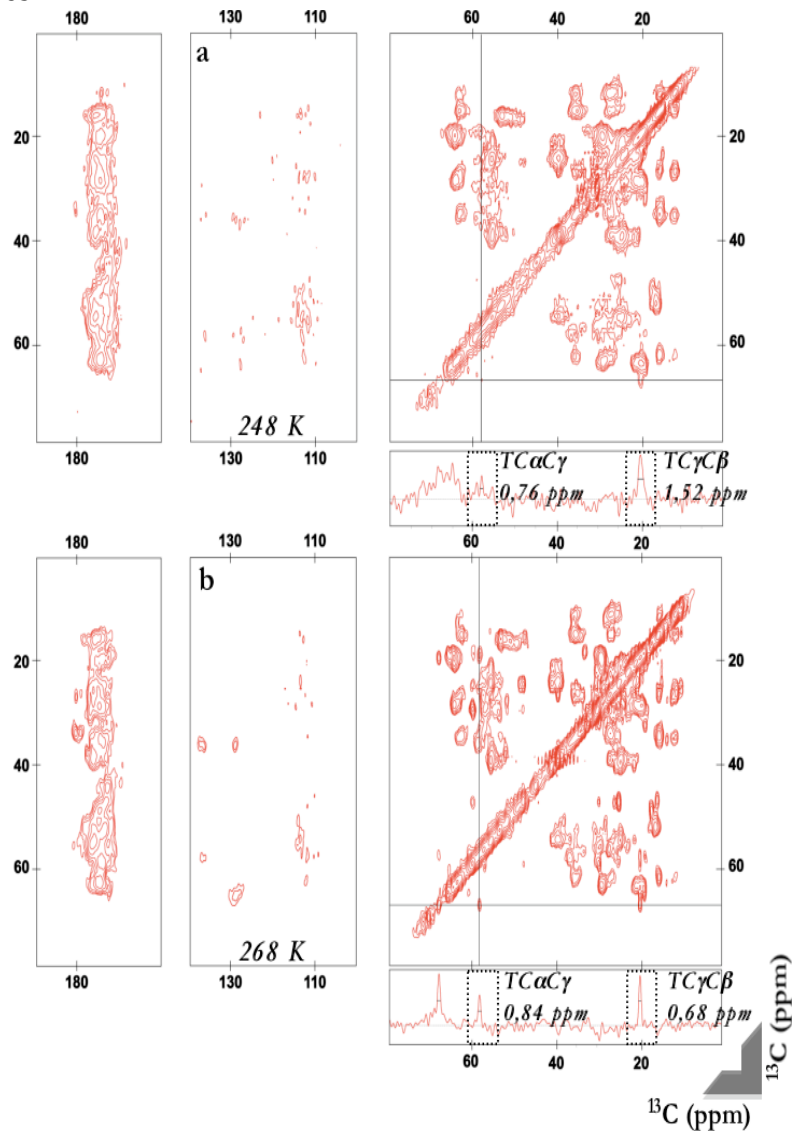


Figure 1.8 : Spectres ^{13}C - ^{13}C obtenu sur l'échantillon u - ^{13}C - ^{15}N -MscL en utilisant la séquence d'impulsion *DARR*, en appliquant un temps de mélange de 25 ms à 248 K (a) et un temps de mélange de 20 ms à 268 K (b). Le spectre obtenu à 248 K est le même que celui qui est présenté dans la figure 9b, mais en diminuant le niveau du signal jusqu'à apparition du système de spin d'une *Thr*. Les largeurs de raies des corrélations *TCα-Cγ* et *TCγ-Cβ* ont été déterminées au niveau de la ligne extraite et les valeurs figurent à côté des pics encadrés. Sur le spectre obtenu à 268 K, la résolution est améliorée. Sur la ligne extraite au niveau du système de spin de la même *Thr*, les pics *TCα-Cγ* et *TCγ-Cβ* sont plus fins.

Afin d'obtenir des informations séquentielles, les corrélations *N-C'* sont indispensables. Les signaux *N-Cα* servent à compléter les informations au niveau de

chaque acide aminé, et les signaux $N-C'$ correspondent aux liaisons amides entre deux acides aminés voisins. La séquence d'impulsion utilisée pour obtenir ces corrélations est la séquence *specific-CP*. La séquence contient deux polarisations croisées : le transfert d'aimantation des protons aux azotes, et le transfert de cohérence des azotes vers les $C\alpha$ ou les C' . Ce transfert est très sensible aux paramètres expérimentaux. Les spectres *specific-CP* obtenus sur l'échantillon uniformément marqué contiennent les informations sur 146 liaisons $N-C'$, et 147 $N-C\alpha$, ainsi que les chaînes latérales des résidus *Asn*, *Gln*, *His* et *Arg* (Figure 1.9). Les spectres obtenus présentent une faible résolution. Un seul pic représente les 146 corrélations $N-C'$ (Figure 1.9a), tandis que les corrélations $N-C\alpha$ forment trois groupes (Figure 1.9b). La résolution n'a pas pu être estimée en raison des difficultés pour identifier des pics isolés.

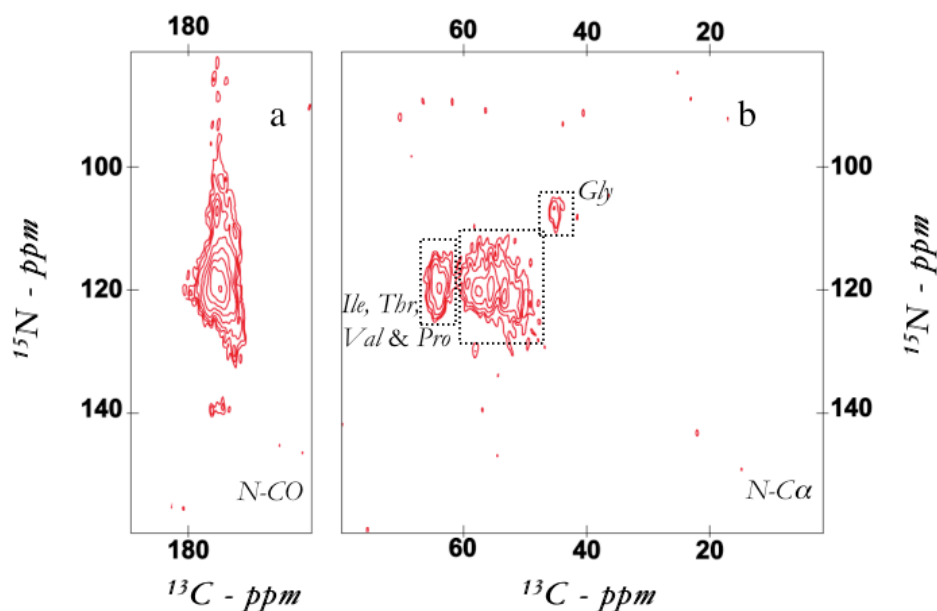


Figure 1.9 : Spectre $^{15}\text{N}-^{13}\text{C}$ obtenu sur l'échantillon $u\text{-}^{13}\text{C}\text{-}^{15}\text{N}\text{-MscL}$ en utilisant la séquence d'impulsion *specific-CP*. Le spectre est obtenu en deux fois en fixant la porteuse, soit sur la fréquence des carbonyles pour obtenir les corrélations $N-C'$, soit sur celle des carbones quaternaires pour avoir les corrélations $N-C\alpha$. Au niveau des corrélations $N-C'$, le pic contient les informations de 145 liaisons amides inter-résiduelles, ainsi que les 10 liaisons amides des résidus *Asn* et *Gln*. Au niveau des corrélations $N-C\alpha$, trois groupes sont distincts, le premier contient probablement les corrélations $N-C\alpha$ des 39 résidus *Ile*, *Thr*, *Val* et *Pro* ; le deuxième contient celles des autres résidus sauf les 12 *Gly*, dont les corrélations sont distinctes.

II Analyse de l'échantillon sélectivement marqué *IT-Ec-MscL* - Apport de la *SIV* pour l'étude des protéines membranaires par *RMN* en phase solide.

L'utilisation d'un marquage isotopique uniforme sur une protéine membranaire comme *MscL* montre la difficulté d'obtenir des informations structurales, malgré une résolution pouvant atteindre $0,6 \text{ ppm}$ pour certains pics isolés. Différentes stratégies sont envisageables à ce stade pour améliorer la lisibilité des spectres, permettre une identification des systèmes de spins, et procéder à l'attribution. Plusieurs stratégies consistent à utiliser des métabolites marqués au ^{13}C , autre que le glucose, procédant ainsi à un marquage sélectif sur certains acides aminés, voire certains carbones au sein des acides aminés. Pour des protéines membranaires où le taux de surexpression n'excède que très rarement le mg par litre de culture bactérienne, ceci paraît difficile à mettre en place, à cause de la chute de productivité observée lorsque les bactéries sont cultivées avec une autre source de carbone que le glucose. L'ajout d'acides aminés marqués dans un milieu minimum contenant une source de carbone non marquée est une autre alternative, mais la méthode nécessite de grandes quantités d'acides aminés marqués, et des souches bactériennes auxotrophes pour les acides aminés supplémentés, afin d'éviter les problèmes de redistribution des isotopes marqués ou "*scrambling*" intervenant dans la biosynthèse des acides aminés. Les avancées effectuées dans le domaine de la transcription et la traduction *in vitro* ont permis depuis une dizaine d'années la mise en place de systèmes assurant l'expression de protéines, en dehors de la cellule. La commercialisation de ces systèmes de synthèse *in vitro* a permis de surpasser les inconvénients de la surexpression de certaines protéines dans des systèmes recombinants. Dans ce système ouvert, toute la machinerie de synthèse de protéine est consacrée à la transcription de l'oligonucléotide rajouté par l'utilisateur, et la traduction du messenger correspondant en protéine. De plus, le système présente à l'utilisateur un contrôle total des acides aminés utilisés. Le plasmide *pIVEX* a été optimisé pour la *SIV*. Le gène *mscL* et son étiquette histidine ont été clonés dans ce plasmide permettant d'atteindre de très bons rendements. Dans un premier temps, pour démontrer l'efficacité de la *SIV* dans l'amélioration des jeux de données obtenus par

RMN en phase solide sur la protéine membranaire *MscL* reconstituée en liposomes, un marquage sélectif au niveau des *Ile* et des *Thr* a été effectué. Les isoleucines sont les acides aminés les plus représentés dans la protéine *MscL*, et les thréonines sont parmi les moins présents. Il s'agit, pour les *Ile*, d'un témoin de résolution, et d'un témoin de sensibilité pour les *Thr*. La SIV est réalisée en présence de micelles de *triton-x-100*. La purification de la protéine est faite sur colonne de *Ni-NTA*. Après purification, la protéine est reconstituée en liposomes de *DOPC*, puis analysée par RMN. La séquence *DARR* a été principalement utilisée. Différents temps de mélanges allant de 25 à 500 ms ont été testés. Les expériences *DARR* sur l'échantillon (*Ile, Thr*)-¹³C,¹⁵N-*MscL* en variant le temps de mélange ont toutes été effectuées à 248 K. D'autres séquences d'impulsions (*B²R³* et *PDSD*) ont été testées sur l'échantillon à 248 K.

a Production et purification de (*Ile, Thr*)-*MscL*

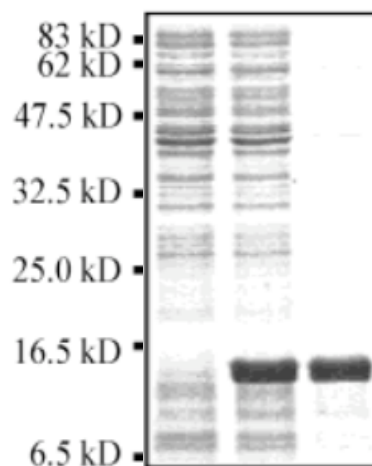
Le protocole de synthèse *in vitro* de la protéine membranaire *MscL* en micelles de *triton-x-100* a été mis en place dans le laboratoire d'A. GHAZI, et a été transposé pour réaliser la synthèse dans les kits de 10 mL. Pour les acides aminés, les quantités nécessaires pour la synthèse de 10 mg de protéines ont été rajoutées. Le détergent est rajouté dans le compartiment réactionnel mais aussi dans la chambre d'alimentation. La synthèse est réalisée pendant 22 heures à 30°C. Ensuite, le mélange réactionnel est récupéré et purifié par *FPLC*. La capacité de liaison de la colonne utilisée doit être vérifiée afin de savoir si l'échantillon est injecté en une ou plusieurs fois. Ici, la purification a été effectuée en deux fois. Le protocole de purification est le même que celui qui est décrit précédemment. À la fin de la purification, les fractions élues sont analysées par *Dot Blot*. Les fractions contenant *Ec-MscL* sont rassemblées. Une dialyse est effectuée pour réduire la concentration d'imidazole.

b Reconstitution de (*Ile, Thr*)-*MscL* en liposomes de *DOPC*

La protéine est quantifiée par dosage au *BCA*. Après la synthèse, la purification et la dialyse, 9,6 mg de *Ec-MscL* ont été obtenus. Pour reconstituer cet échantillon en liposomes,

19,2 mg de DOPC sont nécessaires. Le protocole de reconstitution est décrit dans la section Matériel et Méthodes expérimentales. En parallèle, les échantillons prélevés avant et après la SIV, ainsi que ceux qui sont prélevés au cours de la purification et après la dialyse sont analysés par SDS-PAGE (Figure 2.1). Les protéines séparées sur gel sont ensuite transférées sur membrane de nitrocellulose pour détecter la présence d'*Ec*-MscL par WESTERN BLOT. La protéine reconstituée en liposomes est conservée à -20°C jusqu'à son étude par RMN.

Figure 2.1 : Suivi de la synthèse de la protéine MscL par SIV en triton-x-100 par SDS-PAGE. L'analyse du mélange réactionnel avant (1ère colonne) et après SIV (2ème colonne) montre une bande représentant la protéine exprimée. Après purification sur colonne de Ni-NTA (3ème colonne) la protéine est reconstituée en liposomes, et conservée à -20°C.



c Analyse de (Ile, Thr)-MscL par RMN

La protéine *Ec*-MscL contenant l'étiquette histidine est composée de 147 résidus dont 16 Ile et 3 Thr. Chaque molécule marquée aux ^{13}C et ^{15}N sur les Ile et Thr contient donc 108 atomes ^{13}C et 19 atomes ^{15}N . Le poids moléculaire après le marquage isotopique est de 16,270 kDa. Les protéoliposomes contenant la protéine sélectivement marquée sont transférés dans le rotor adéquat et conservés à -20°C jusqu'aux expériences RMN. Après la mise au point des différentes valeurs des séquences d'impulsions utilisées, le rotor est placé à température ambiante. Il est ensuite inséré dans la sonde MAS, la température est réglée à 248 K et la rotation à l'angle magique est augmentée par paliers (jusqu'à 11 000 Hz). Une fois ces deux paramètres stabilisés, la sonde est accordée aux fréquences des noyaux étudiés.

Le premier spectre 2D ^{13}C - ^{13}C -DARR enregistré sur (Ile, Thr)- ^{13}C , ^{15}N -MscL est réalisé à 5 ms de temps de mélange. Sur ce spectre apparaissent plusieurs pics (Figure 2.2). Les

corrélations entre noyaux séparés par une seule liaison sont toutes présentes. Quelques corrélations entre carbones séparés par plusieurs liaisons sont visibles, notamment la corrélation $C'-C\beta$, $C'-C\gamma_2$ et $C\alpha-C\gamma_2$. L'analyse des largeurs de raies ^{13}C obtenues au niveau des corrélations des *Thr* montre une résolution située entre 1,4 et 1,6 ppm.

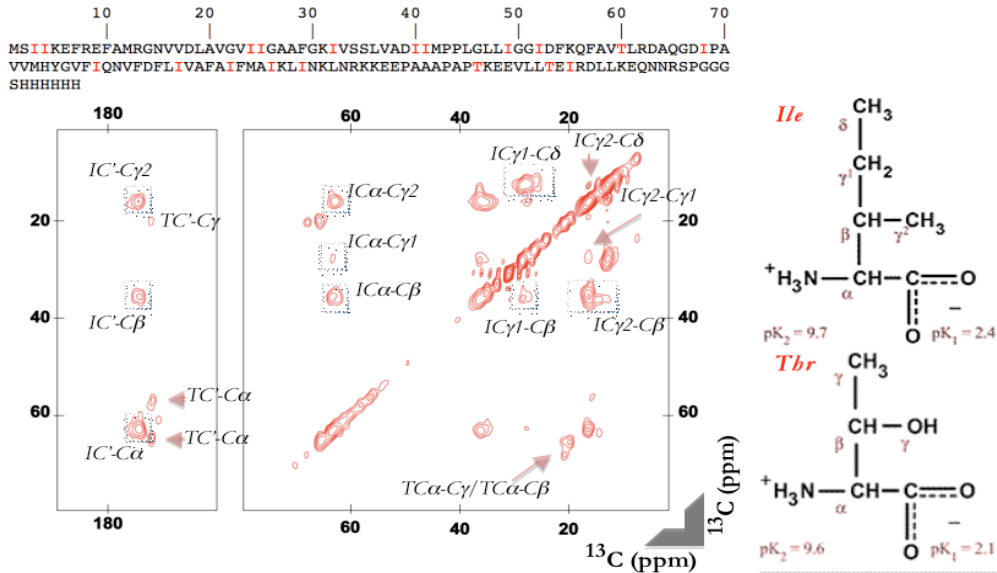


Figure 2.2 : Spectre 2D ^{13}C - ^{13}C -DARR obtenu sur l'échantillon (Ile, Thr)- ^{13}C , ^{15}N -MscL avec un temps de mélange de 5 ms à 248 K. Sur cette figure apparaissent des corrélations entre noyaux proches. En haut du spectre, la séquence d'*Ec*-MscL est présentée, les acides aminés Ile et Thr sont marqués en rouge. À droite du spectre, les formules chimiques de ces deux acides aminés sont représentées, ainsi que les nomenclatures des atomes de carbones.

En appliquant un temps de mélange plus élevé, les corrélations entre noyaux séparés par une seule liaison s'intensifient. D'autres corrélations apparaissent. Le spectre obtenu avec un temps de mélange de 25 ms montre une réduction importante de l'encombrement dans toutes les régions du spectre (Figure 2.3a). Les pics sont distincts et présentent des largeurs comprises entre 0,6 et 1,1 ppm. Pour les Ile, des corrélations entre les C' et les carbones $C\alpha$, $C\beta$, $C\gamma_1$ et $C\gamma_2$ sont visibles. Les corrélations entre les $C\alpha$ et tous les carbones de la chaîne latérale sont présentes sur le spectre. Un seul pic est présent pour chacune de ces corrélations. Les déplacements chimiques des $C\alpha$ et $C\beta$ sont de 63,2 et 34,9 ppm respectivement. Les valeurs de ces déplacements chimiques dans une pelote aléatoire sont de 61,6 et 38,6 ppm (source : BMRB ; Ulrich *et al.*, 2007).

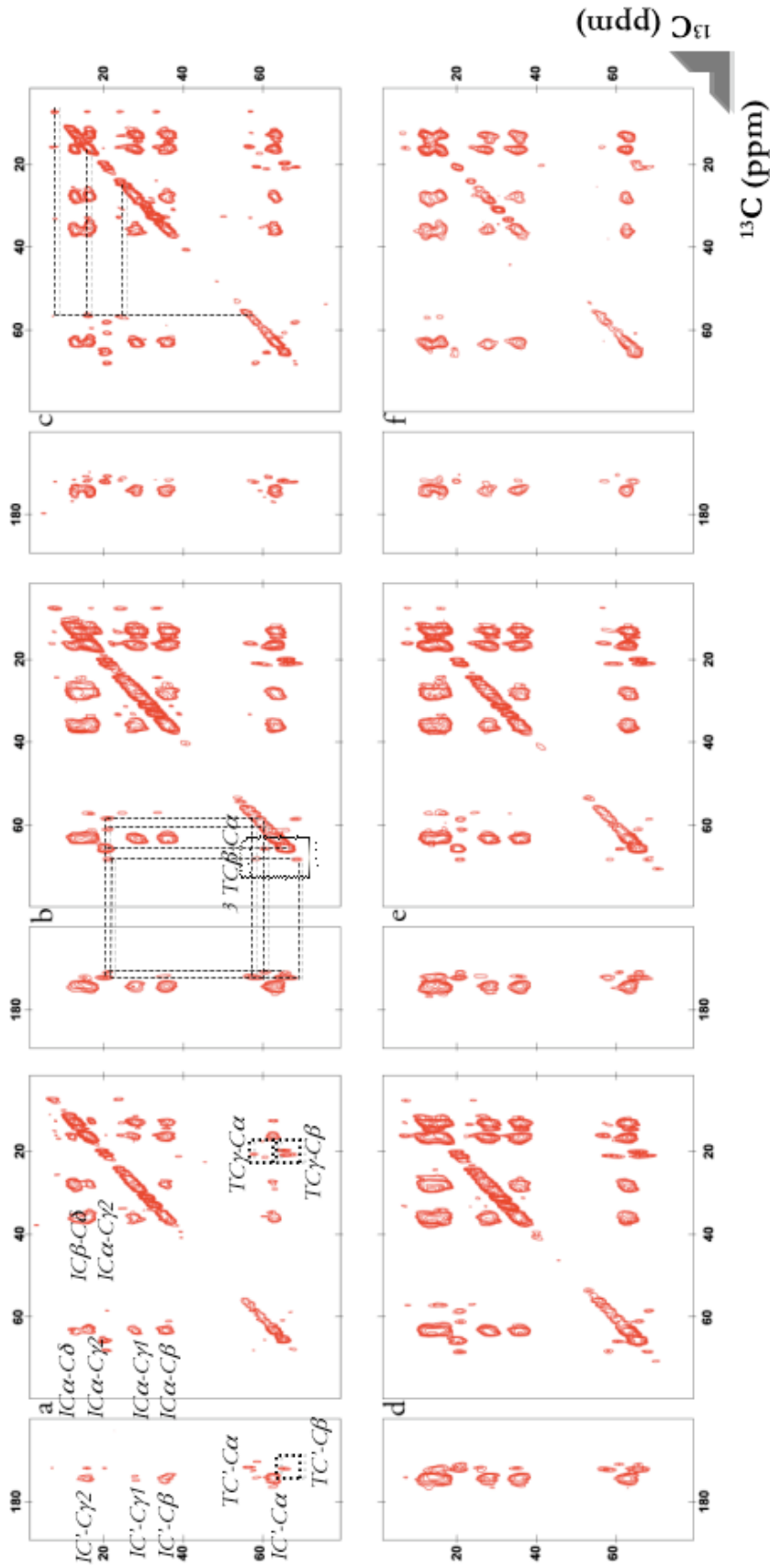


Figure 2.3 : Spectres 2D ^{13}C - ^{13}C -DARR obtenus sur l'échantillon (Ile, Thr)- ^{13}C , ^{15}N -MscL avec différents temps de mélange. Des temps de mélange de 25 (a), 50 (b), 100 (c), 200 (d), 500 (e) et 1000 ms (f) ont été utilisés pour obtenir des spectres DARR à 258 K avec une rotation de 11 000 Hz. Pour un temps de mélange de 25 ms (a), différents pics sont identifiés. Il s'agit essentiellement de corrélations des Ile. Les pics des trois Thr sont visibles mais sont beaucoup plus intenses pour des temps de mélange de 50 (b) et 100 ms (c). Le rectangle présent sur le spectre obtenu à 50 ms de temps de mélange montre les corrélations caractéristiques des Thr, qui permettent de trouver les autres corrélations de ces résidus dans la région des C γ , mais aussi les corrélations au niveau de la chaîne latérale (comme indiquent les traits). À 50 et 100 ms de temps de mélange apparaissent les corrélations d'une Ile probablement située en dehors des hélices. La résolution obtenue de pics isolés sur tous ces spectres est aux alentours de 0,8 ppm.

Cela signifie que les corrélations présentes sur le spectre sont celles d'une ou plusieurs *Ile* situées dans des régions de la protéine où la structure secondaire est en hélices α (Wishart *et al.*, 2001). Les systèmes de spin des β *Thr* sont visibles sur le spectre. Pour un temps de mélange de 50 ms , 4 pics caractéristiques des *Thr* sont visibles dans la région $58\text{-}68\text{ ppm}$ (Figure 2.3b). Ces pics correspondent à des corrélations $C\alpha\text{-}C\beta$ et $C\beta\text{-}C\alpha$ de résidus présents dans des régions où la structure secondaire est en feuillet β . La largeur de raie de ces pics est de $0,8\text{-}0,9\text{ ppm}$. Les corrélations $C\alpha\text{-}C\beta$ et $C\beta\text{-}C\alpha$ de la troisième *Thr* se trouvent au niveau de la diagonale vers 65 ppm . Cette valeur est caractéristique d'une *Thr* présente dans une région où la structure secondaire est en hélices α . En partant de ces pics, la recherche des autres corrélations montre 4 pics distincts pour les $C\beta\text{-}C\gamma$ et $C\alpha\text{-}C\gamma$ avec des valeurs de déplacement chimique autour de $19\text{-}20\text{ ppm}$ pour les *C γ* . Les déplacements chimiques des *C'* complétant ces systèmes de spin se trouvent autour de 172 ppm . Pour un temps de mélange de 100 ms , la résolution est améliorée, et la largeur de raie de quelques pics isolés est située entre $0,8$ et $0,9\text{ ppm}$. Les pics de corrélations $C\beta\text{-}C\gamma1$ et $C\beta\text{-}C\gamma2$ d'une *Ile* isolée sont plus fins (Figure 2.3c). À partir de 200 ms , les pics sont plus larges et aucune nouvelle information n'apparaît dans les spectres (Figures 2.3d, e et f).

Cependant les spectres montrent que le signal ne disparaît pas à des longs temps de mélange. Même pour des temps de mélange de $2\ 000\text{ ms}$, des signaux sont encore visibles sur un spectre où la séquence *PDSD* est appliquée. L'intensité des pics isolés a été extraite sur les spectres obtenus en appliquant la séquence d'impulsion *DARR*. Le tracé de l'intensité des pics représentant des corrélations de noyaux reliés par une seule liaison en fonction du temps de mélange montre des valeurs maximales obtenues pour 50 à 100 ms de temps de mélange. Le signal est fortement réduit pour un temps de mélange de $1\ 000\text{ ms}$ (Figure 2.4). Pour les pics représentant les corrélations de noyaux séparés par deux ou trois liaisons, les valeurs maximales sont plutôt vers 200 ms (Figure 2.5).

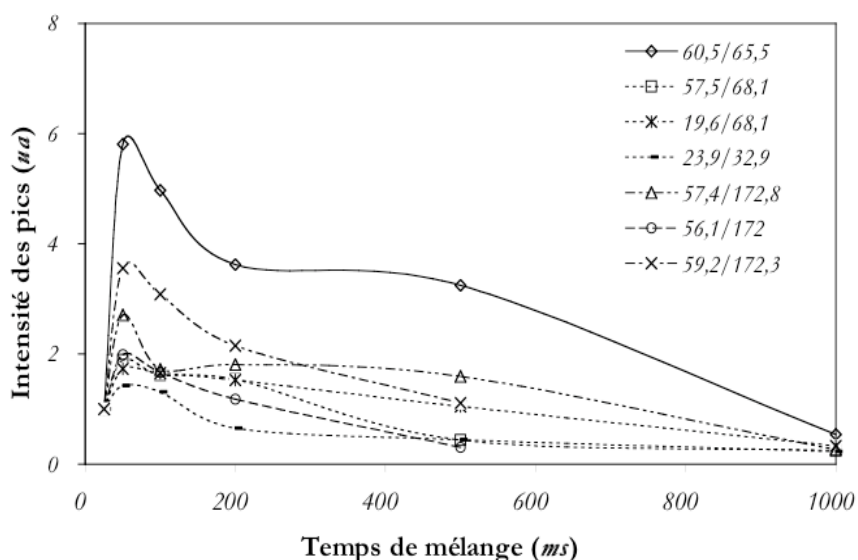


Figure 2.4 : Évolution du signal de certaines corrélations isolées entre atomes séparés par une seule liaison. Les valeurs sont extraites par intégration gaussienne sur les pics au niveau des différents spectres obtenus en appliquant la séquence d'impulsion *DARR* à différents temps de mélange entre 25 et 1 000 ms. Sur sept corrélations isolées, cinq appartiennent probablement à des résidus *Thr*. L'intensité des signaux est ramenée à celle de la corrélation identifiée sur le spectre obtenu à 25 ms. Les pics 60,5/65,5 et 57,5/68,1 représentent des corrélations $C\alpha-C\beta$ de deux *Thr*. Le pic 19,6/68,1 représente une corrélation $C\gamma-C\beta$ d'une *Thr*. Les pics 57,4/172,8 - 56,1/172 - 59,2/172,3 représentent des corrélations $C\alpha-C'$ de résidus *Ile* ou *Thr*. Le pic 23,9/32,9 correspond à une corrélation $C\gamma-C\beta$ d'une *Ile*. Toutes ces corrélations représentent des noyaux séparés par une liaison atomique C-C d'une valeur de 1,5 Å.

D'autres séquences d'impulsions ont été testées sur cet échantillon, mais aucune différence remarquable n'a été décelée. Afin d'obtenir les corrélations $N-C'$ et $N-C\alpha$, la séquence *specific-CP* a été appliquée. Le spectre montre trois pics au niveau des corrélations $N-C'$. Les pics présentent une résolution de 0,7 ppm dans la dimension du ^{13}C ; celle-ci est inférieure ou égale à 8 ppm dans la dimension de l'azote ^{15}N . Au niveau des corrélations $N-C\alpha$, le spectre montre un pic majeur, et trois autres de très faible rapport signal/bruit. La résolution dans la dimension du ^{13}C est de 0,8 à 1 ppm, tandis que celle dans la dimension du ^{15}N est inférieure ou égale à 11 ppm (figure 2.6).

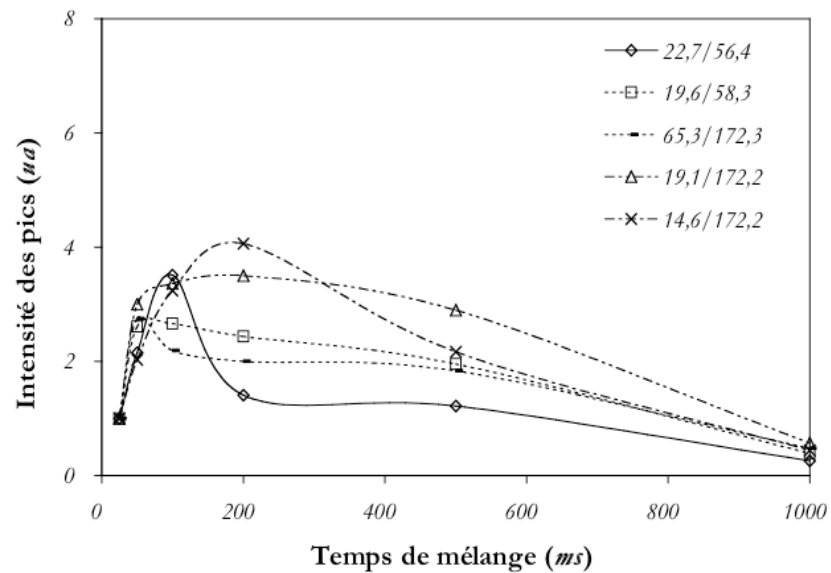


Figure 2.5 : Évolution du signal de certaines corrélations isolées entre atomes séparés par deux ou trois liaisons. Les valeurs sont extraites par intégration gaussienne sur les pics au niveau des différents spectres obtenus en appliquant la séquence d'impulsion *DARR* à différents temps de mélange entre 25 et 1 000 ms. Le pic 22,7/56,4 correspond à une corrélation $C\gamma-C\alpha$ d'une *Ile*. Le pic 19,6/58,3 correspond à une corrélation $C\gamma-C\alpha$ d'une *Thr*. Le pic 65,3/172,3 correspond à une corrélation $C\beta-C'$ d'une *Thr*. Ces corrélations sont établies entre noyaux séparés par deux liaisons, à une distance de 3 Å environ. Les pics 19,1/172,2 et 14,6/172,2 représentent des corrélations $C\gamma-C'$ et $C\delta-C'$ d'une *Thr*, et d'une *Ile* à 4,5 et 6 Å respectivement.

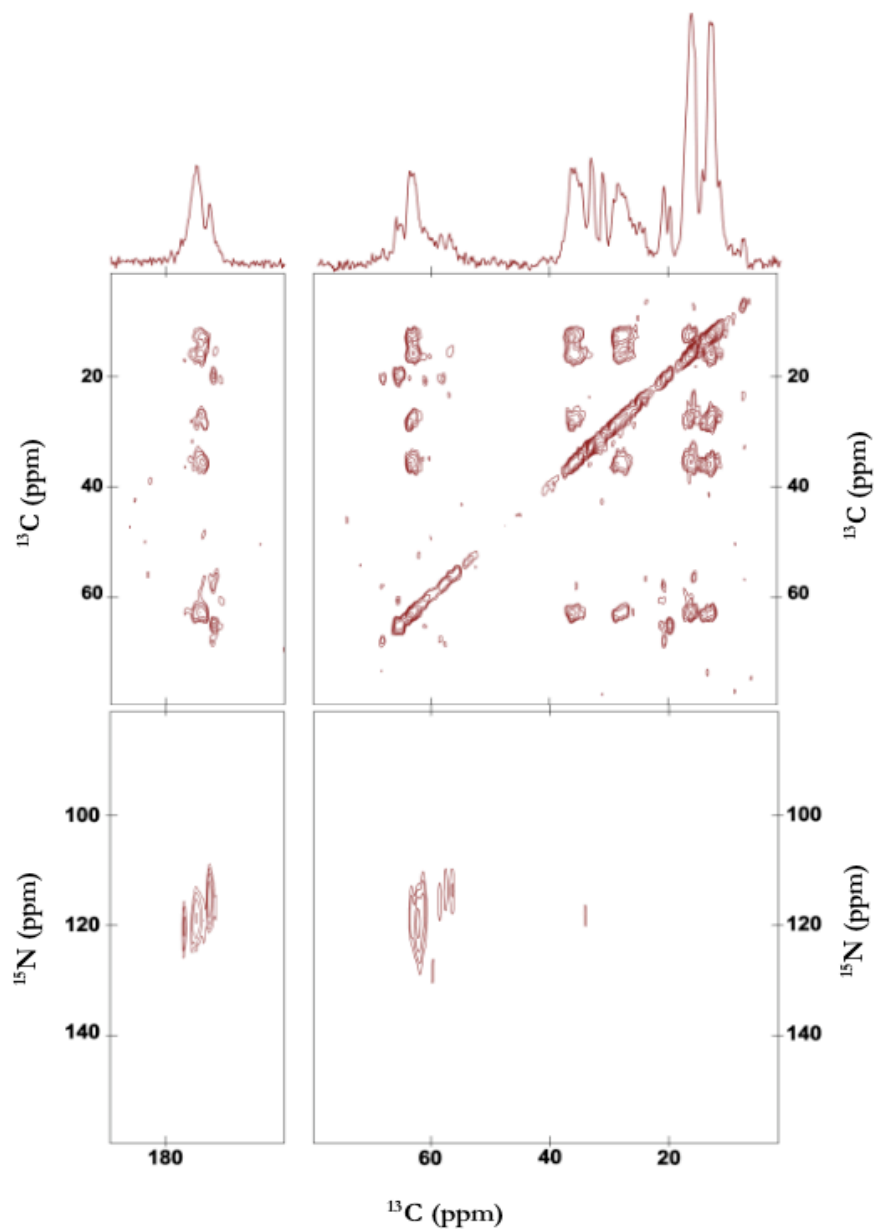


Figure 2.6 : Spectre ^{15}N - ^{13}C obtenu sur l'échantillon (Ile, Thr)- ^{13}C - ^{15}N -MscL utilisant la séquence d'impulsion *specific-CP*. De haut en bas, sont présentés un spectre $1D$ ^{13}C , un spectre $2D$ B^2R^3 ^{13}C - ^{13}C , et le spectre de corrélations $2D$ ^{15}N - ^{13}C . Ce dernier est obtenu en deux temps en fixant la porteuse, soit sur la fréquence des carbonyles pour obtenir les corrélations N - C' ou sur les carbones quaternaires pour avoir les corrélations N - $C\alpha$. La séquence d'impulsion B^2R^3 ne montre aucune différence significative au niveau de la résolution.

L'analyse des spectres a permis d'obtenir les déplacements chimiques des carbones des 3 *Thr*. Sur le spectre obtenu avec un temps de mélange de 200 ms, les systèmes de spin sont annotés (Figure 2.7). Pour chaque valeur de déplacement chimique, la moyenne a été calculée ainsi que l'écart à la moyenne, en se basant sur le nombre de fois où des corrélations impliquant ces carbones ont été observées (Tableau 2.1). Les trois *Thr* ont été annotées *Ta*, *Tb* et *Tc*. Pour les *Ile*, huit différents systèmes de spin ont été obtenus (Tableau 2.2). Certains de ces systèmes sont représentés dans la figure 2.8. Finalement, pour des temps de mélange supérieurs à 200 ms, de nouveaux pics ont été identifiés, correspondant à des corrélations inter-résiduelles entre les carbones d'une *Ile* et ceux de la *Thr Tb* (Figure 2.9).

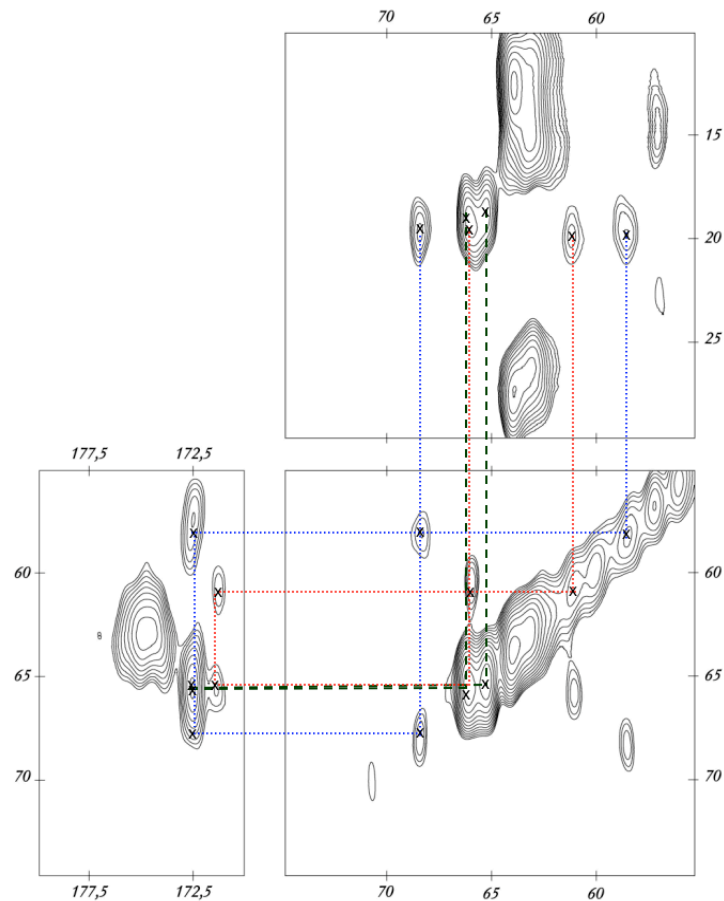


Figure 2.7 : Agrandissement des régions contenant les pics de corrélations entre les différents carbones des *Thr* de la protéine *Ec-MscL*. Ces informations sont issues du spectre 2D DARR obtenu à 263 K avec un temps de mélange de 200 ms. Chaque système de spin est relié par des traits de la même couleur. Les traits bleus relient les corrélations de la *Thr Ta*, les traits rouges, celles de *Tb* et les traits verts celles de *Tc*.

Déplacements chimiques des 3 Thr					
observations effectuées sur tous les spectres obtenus à différents temps de mélange					
<i>aa</i>	¹³ C	<i>C'</i>	<i>Cα</i>	<i>Cβ</i>	<i>Cγ</i>
Ta	déplacement chimique (ppm)	172,2	58,2	68,3	19,6
	Erreur (ppm)	0,2	0,2	0,2	0,2
	Nombre de corrélations observées	17	19	18	20
Tb	déplacement chimique (ppm)	170,9	60,7	65,7	19,9
	Erreur (ppm)	0,2	0,1	0,2	0,2
	Nombre de corrélations observées	15	22	23	20
Tc	déplacement chimique (ppm)	171,7	64,9	65,8	18,5
	Erreur (ppm)	0,2	0,1	0,1	0,1
	Nombre de corrélations observées	7	11	7	11

Tableau 2.1 : Déplacements chimiques des carbones des 3 Thr de la protéine *Ec-MscL*.

Déplacements chimiques de 7 Ile identifiées observations effectuées sur tous les spectres obtenus à différents temps de mélange						
<i>aa</i> \ ^{13}C	<i>C'</i>	<i>Cα</i>	<i>Cβ</i>	<i>Cγ1</i>	<i>Cγ2</i>	<i>Cδ</i>
<i>Ia</i>	172,2	55,9	-	24,1	14,3	9,9
	0,3	0,1	-	0,2	-	0,1
	5	5	-	3	1	3
<i>Ib</i>	172,5	57,2	36,4	23,1	15,2	11,2
	0,5	0,1	0,1	0,2	0,5	0,2
	2	5	2	5	3	5
<i>Ic</i>	171,5	56,6	32,4	23,1	14,7	6,2
	-	0,1	0,3	0,2	0,3	0,2
	1	13	7	11	13	7
<i>Id</i>	-	59,1	36,2	27,1	16,1	-
	-	0,1	0,1	0,4	0,3	-
	-	5	5	3	3	-
<i>Ie</i>	175,2	60,1	34,5	26,2	15,5	12,1
	0,1	0,3	0,2	-	0,2	-
	3	5	4	1	3	1
<i>If</i>	-	60,7	-	26,6	15,3	-
	-	0,1	-	0,2	0,2	-
	-	7	-	3	3	-
<i>Ig</i>	175,2	63,5	35,3	27,7	14,9	12,0
	0,3	0,1	0,1	0,1	0,1	0,1
	5	7	7	7	7	7
<i>Ih</i>	171,9	65,4	36,0	28,3	-	12,3
	0,2	0,4	0,2	0,2	-	0,2
	3	3	2	3	-	3

Tableau 2.2 : Déplacements chimiques des carbones de 8 Ile identifiées de la protéine *Ec-MscL*.

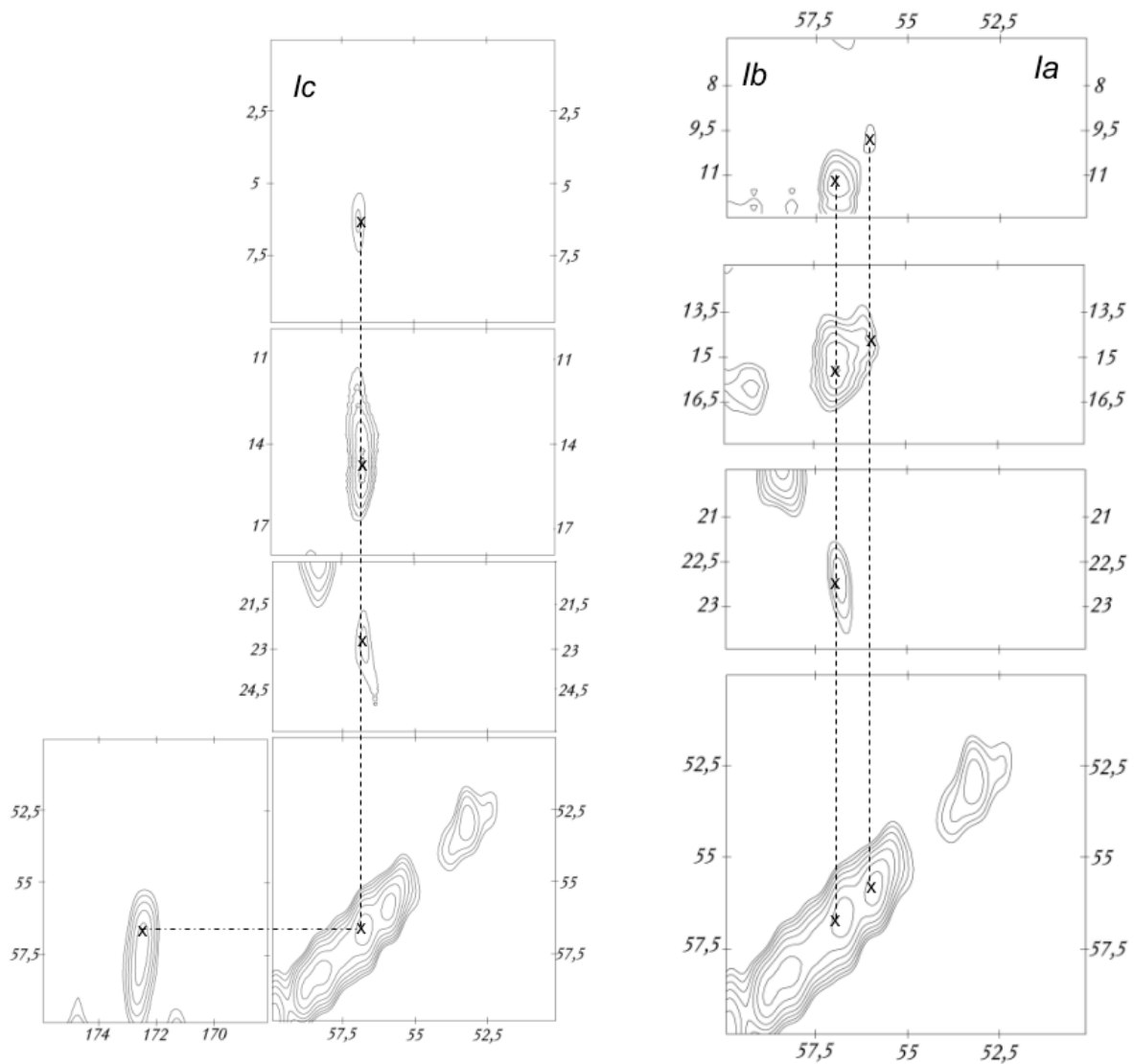
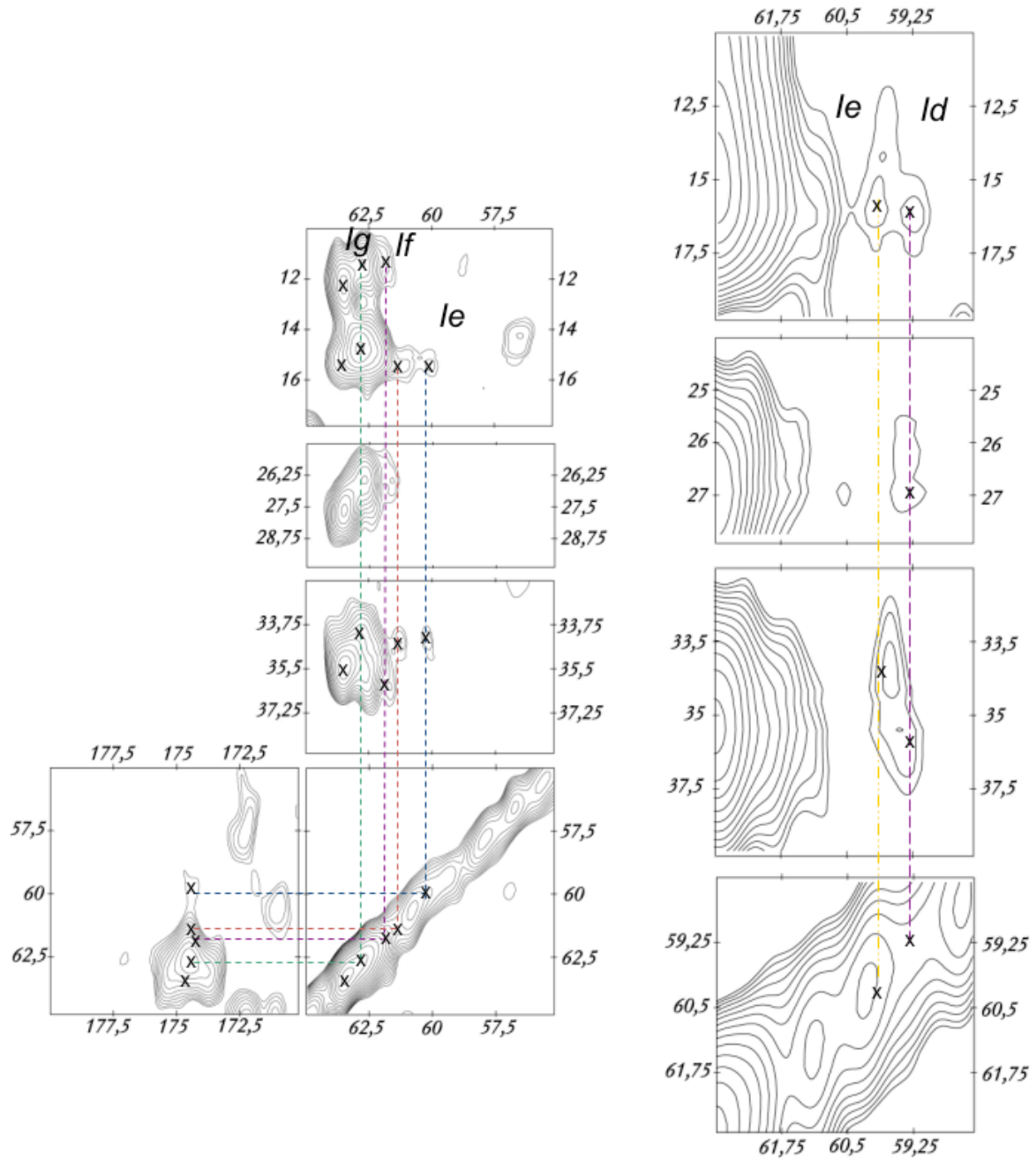


Figure 2.8 : Agrandissement des régions contenant les pics de corrélations entre les différents carbones de quelques *Ile* de la protéine *Ec-MscL*. Ces informations sont issues de différents spectres 2D DARR obtenus à 263 K avec différents temps de mélange. Trois systèmes de spin sont présentés ici. Il s'agit de ceux des résidus annotés *Ia*, *Ib* et *Ic* (suite à la page suivante).



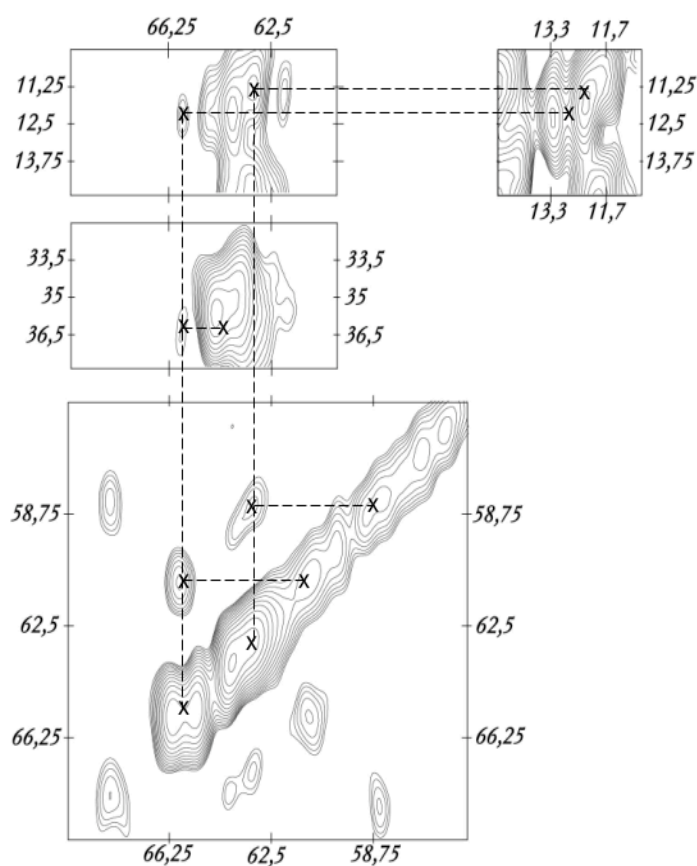


Figure 2.9 : Agrandissement des régions contenant les pics de corrélations inter-résiduelles entre une *Thr* et une *Ile*.

d Comparaison des résultats obtenus sur (Ile, Thr) -*MscL* par rapport à ceux obtenus sur *U-MscL*

Les spectres bidimensionnels obtenus sur l'échantillon (Ile, Thr) - $^{13}C, ^{15}N$ -*MscL* montrent moins d'encombrement au niveau des corrélations que ceux obtenus sur l'échantillon produit par surexpression bactérienne. De plus, les corrélations des *Thr* apparaissent alors qu'elles n'étaient pas présentes sur les spectres obtenus sur l'échantillon uniformément marqué (Figure 2.10). L'analyse des spectres obtenus à différents temps de mélange sur l'échantillon (Ile, Thr) - $^{13}C, ^{15}N$ -*MscL* a permis d'identifier les systèmes de spin des trois *Thr* (Tableau 1, Figure 2.8) et la moitié des *Ile* (Tableau 2). Le marquage sélectif permet aussi d'obtenir des spectres avec de très longs temps de mélange (jusqu'à 2 000 ms).

Cette partie de la thèse a fait l'objet d'un article qui est apparu dans la revue scientifique *JOURNAL OF MAGNETIC RESONANCE* en mai 2010. L'article est joint en annexe à la fin du manuscrit.

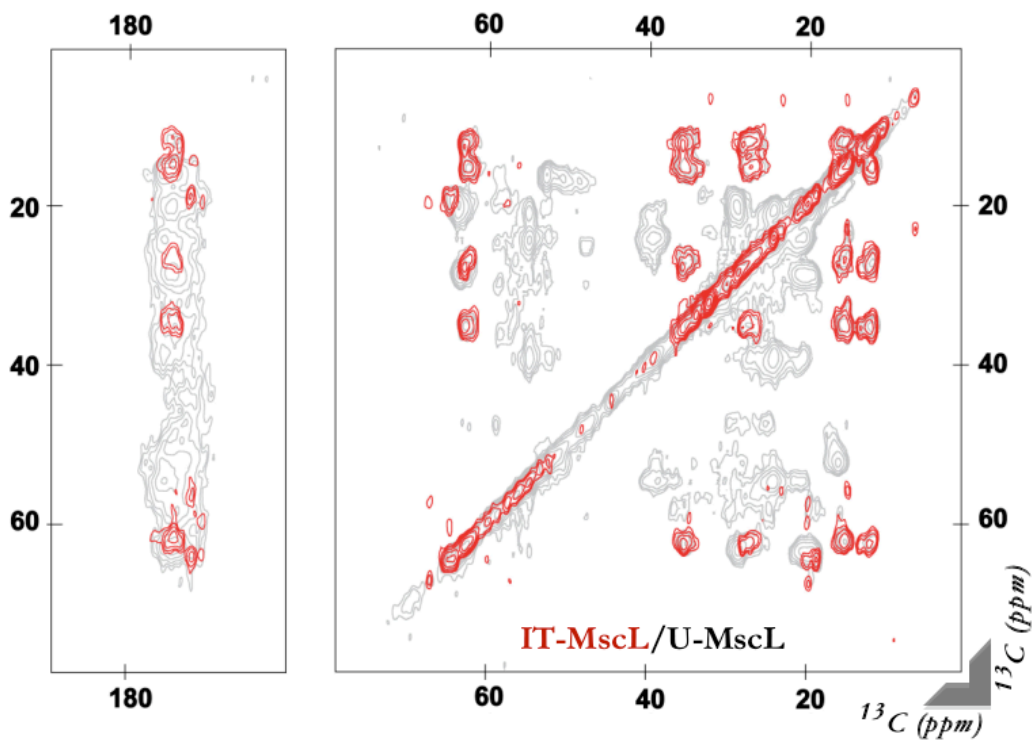


Figure 2.10 : Superposition des spectres $2D$ ^{13}C - ^{13}C obtenus sur les échantillons IT - ^{13}C - ^{15}N - $MscL$ et u - ^{13}C - ^{15}N - $MscL$. Les spectres montrent l'avantage du marquage sélectif au niveau de la réduction du recouvrement de pics de corrélations. Les signaux des trois *Thr* et d'une *Ile* apparaissent sur le spectre de l'échantillon sélectivement marqué, mais sont noyés dans le bruit au niveau du spectre de l'échantillon uniformément marqué.

III Analyse de l'échantillon sélectivement marqué *AKVY-Ec-MscL* - Apport des algorithmes de prédiction de déplacement chimiques

L'utilisation de la *SIV* pour un marquage isotopique sélectif de la protéine membranaire *MscL* a montré l'intérêt majeur dans la réduction du recouvrement des pics de corrélations sur des spectres obtenus sur l'échantillon par *RMN* en phase solide. Afin d'obtenir des informations sur l'ensemble de la protéine, il fallait une stratégie permettant de connaître les combinaisons d'acides aminés à marquer aux isotopes ^{13}C et ^{15}N conjointement. Les travaux effectués sur la protéine *Ec-MscL* ainsi que la structure obtenue par radiocristallographie de la protéine *Mt-MscL* ont permis à différents laboratoires d'établir des modèles structuraux de la protéine. Le modèle proposé par *S. SUKHAREV* et *H. GUY* montre la protéine sous trois représentations, où le canal est respectivement fermé, étendu ou ouvert (Sukharev *et al.*, 2001). Un autre modèle a été construit dans le laboratoire de *C. ETCHEBEST* et a permis d'observer le comportement du système en dynamique moléculaire et par modes normaux (Valadié *et al.* 2003 ; Valadié, 2003). Différents *snapshots* ont été obtenus au cours de ces études. Ces différents modèles, ainsi que les trois proposés par *SUKHAREV* et *GUY*, ont été utilisés comme fichiers d'entrée dans le programme *SPARTA*. Il s'agit d'un algorithme de prédiction de déplacements chimiques des noyaux de la chaîne principale d'une protéine, en se basant sur une banque de données expérimentales de déplacements chimiques (Shen & Bax, 2007).

En analysant les déplacements chimiques prédits en fonction de la nature de l'acide aminé, trois ensembles de quatre acides aminés chacun sont isolés. Au sein de chaque ensemble, les déplacements chimiques prédits de chacun des quatre acides aminés ne présentent aucun recouvrement. En utilisant cette séparation, le premier ensemble étudié est celui composé des acides aminés *Ala*, *Lys*, *Val* et *Tyr*. Par *SIV*, 10 mg de la protéine *Ec-MscL* marquée en ^{13}C et ^{15}N au niveau de ces quatre acides aminés sont obtenus. La *SIV* est réalisée directement en liposomes. La protéine est ensuite analysée par *RMN*. La séquence *DARR* a été utilisée. Des temps de mélanges de 10 à 500 ms ont été testés. Les

expériences *DARR* sur l'échantillon (*Ala, Lys, Val, Tyr*)- ^{13}C , ^{15}N -*MscL* ont toutes été effectuées à 248 K.

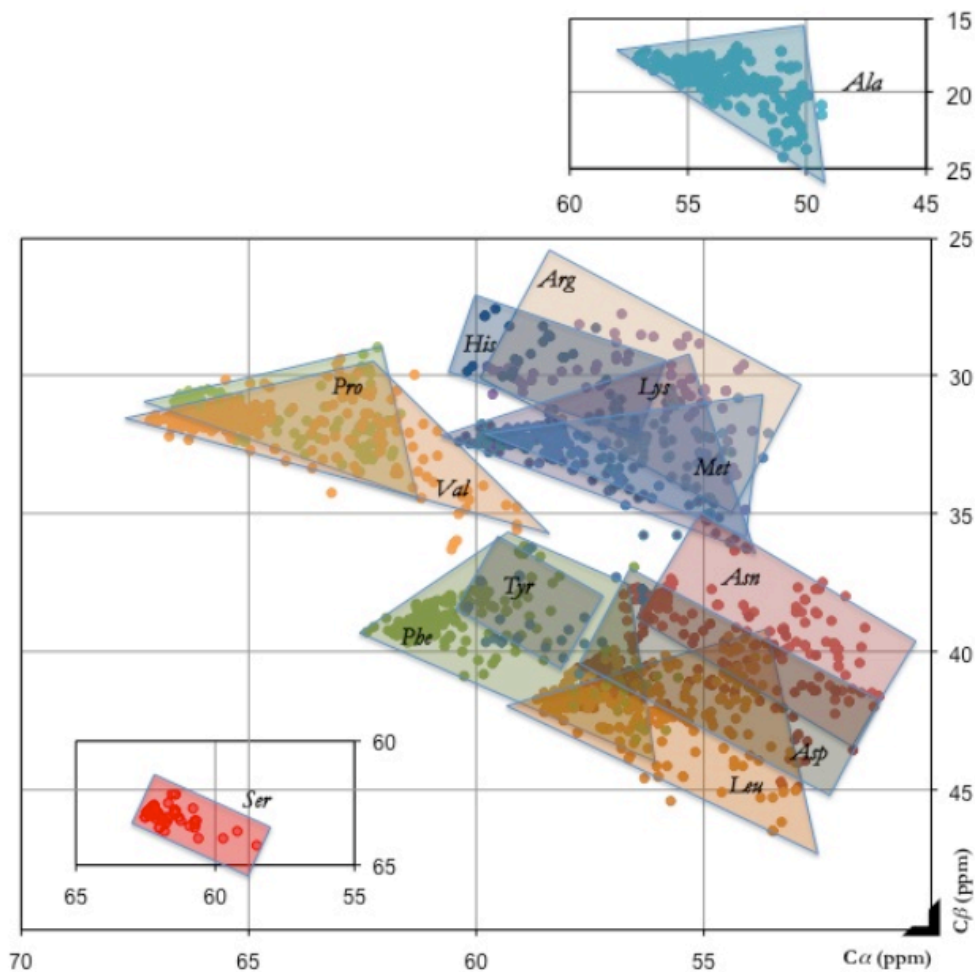
a Prédiction des déplacements chimiques

Dans *SPARTA*, les effets des angles de torsion sont combinés à la recherche d'homologie de séquence (Shen & Bax 2007). Cet algorithme utilise la structure locale d'un résidu pour prédire les déplacements chimiques des atomes de la chaîne principale, à partir d'une structure tridimensionnelle en entrée. Il recherche dans une base de données des triplets des résidus qui correspondent au triplet formé par le résidu et ses voisins. La base de données contient les déplacements chimiques des atomes ^{15}N , ^1HN , $^1\text{H}\alpha$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ et $^{13}\text{C}'$ de 200 protéines pour lesquelles une structure cristallographique à une résolution inférieure à 2,4 Å est disponible. La base de données peut être actualisée manuellement. L'importance des facteurs de pondération pour les angles de torsion et l'homologie de séquence est optimisée de façon empirique. Les cycles aromatiques et liaisons *H* sont pris en compte dans l'algorithme. La déviation de moyenne quadratique entre les déplacements prédits et les déplacements expérimentaux était de 2,36, 0,46, 0,25, 0,88, 0,97 et 1,01 ppm, respectivement pour les atomes ^{15}N , ^1HN , $^1\text{H}\alpha$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ et $^{13}\text{C}'$.

Afin d'identifier tous les acides aminés de la protéine, il fallait synthétiser plusieurs échantillons en *SIV* comprenant les 18 familles d'acides aminés présentes dans la protéine (*MscL* ne contient pas de *Trp* ni de *Cys*). Les kits de *SIV* étant coûteux, il fallait donc limiter le nombre d'échantillons synthétisés. Les acides aminés *Ile* et *Thr* étant déjà obtenus, il restait 16 acides aminés. Les acides aminés *Glu* et *Gln* ne peuvent être marqués par *SIV* dans les kits utilisés car ils sont présents dans le tampon, il restait donc 14 acides aminés.

Pour choisir les familles d'acides aminés à marquer par échantillon, l'algorithme *SPARTA* a été appliqué sur 15 modèles de *Ec-MscL*. Parmi ces modèles, trois sont issus du travail de *SUKHAREV* et *GUY* pour étudier le mécanisme d'ouverture de la protéine. Les douze autres modèles ont été obtenus dans le laboratoire de *C. ETCHEBEST*. Il s'agit de

snapshots de dynamique moléculaire à différents intervalles, à partir de quatre différents modèles issus de différents alignements de séquence (Valadié, 2003). Les déplacements chimiques ont été prédits sur ces quinze modèles. Les valeurs obtenues sur les carbones C' , $C\alpha$ et $C\beta$ ont été utilisées pour tracer les données 2D $C'-C\alpha$ et $C\alpha-C\beta$. La [figure 3.1](#) montre les corrélations $C\alpha-C\beta$ de 13 familles d'acides aminés. Chaque famille d'acides aminés est représentée par une couleur. Pour le premier échantillon, le choix a été porté sur les familles d'acides aminés *Ala*, *Lys*, *Val* et *Tyr*. La figure montre que ces résidus ne présentent aucun recouvrement au niveau des corrélations.

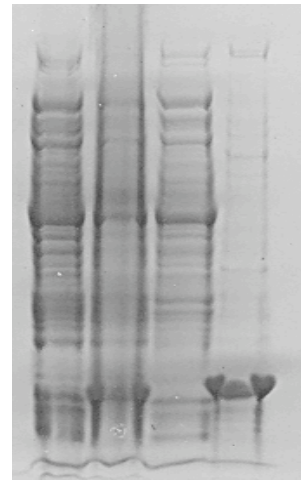


[Figure 3.1](#) : Prédiction des déplacements chimiques sur quinze modèles de *Ec-MscL*. Représentation de la zone des corrélations $C\alpha-C\beta$ des acides aminés présents dans la protéine, sauf *Ile* et *Thr* qui ont déjà été marqués. Les acides aminés *Glu* et *Gln* ne sont pas présents car impossibles à marquer dans les kits standards de *SIV*.

b Production de l'échantillon (Ala, Lys, Val, Tyr)-MscL en liposomes de DOPC

Le protocole de synthèse de la protéine *Ec-MscL* par *SIV* directement en liposomes a été mis en place dans le laboratoire d'A. GHAZI. L'insertion dans la bicouche lipidique serait spontanée et ne nécessite pas la présence de cofacteur ou de chaperon (Berrier *et al.*, 2010). L'échantillon (Ala, Lys, Val, Tyr)-¹³C,¹⁵N-MscL représente la première tentative de synthèse d'une protéine membranaire *in vitro*, directement en liposomes, et dans le but de son étude par RMN en phase solide. Les liposomes sont préparés avec le tampon présent dans le kit de *SIV*. Ils sont incorporés uniquement dans la chambre réactionnelle. La synthèse est réalisée pendant 22 heures à 30°C. L'échantillon est ensuite retiré, un inhibiteur de protéases est rajouté pour 15 minutes, puis les protéoliposomes sont récupérés par une étape de centrifugation. Après la synthèse et la centrifugation, 9 mg de *Ec-MscL* ont été obtenus. Des échantillons prélevés avant et après la *SIV* sont analysés par SDS-PAGE (Figure 3.2). La protéine est conservée à -20°C jusqu'à son étude par RMN.

Figure 3.2 : Suivi de la synthèse de la protéine *MscL* par *SIV* en liposomes de DOPC. L'analyse du mélange réactionnel avant (1ère colonne) et après (2ème colonne) *SIV* montre une bande représentant la protéine exprimée. Après une étape de centrifugation à basse vitesse, le surnageant est ultra-centrifugé. Le surnageant (3ème colonne) et le culot (4ème colonne) de l'ultra-centrifugation sont également analysés montrant la présence de la protéine dans le culot. Ce dernier est conservé à -20°C jusqu'à son analyse par RMN.



c Analyse de (Ala, Lys, Val, Tyr)-MscL par RMN

La protéine *Ec-MscL* contenant l'étiquette histidine est composée de 147 résidus dont 15 Ala, 9 Lys, 13 Val et 1 Tyr. Chaque molécule marquée aux ¹³C et ¹⁵N sur les Ala, Lys, Val et Tyr contient donc 173 atomes ¹³C et 47 atomes ¹⁵N. Le poids moléculaire après le marquage isotopique est de 16,363 kDa. Les protéoliposomes contenant la protéine sélectivement marquée sont transférés dans le rotor adéquat et conservés à -20°C

jusqu'aux expériences RMN. Après la mise au point des différentes valeurs des séquences d'impulsions utilisées, le rotor est placé à température ambiante. Il est ensuite inséré dans la sonde MAS, la température est réglée à 248 K et la rotation à l'angle magique est augmentée par paliers (jusqu'à 11 000 Hz). Une fois ces deux paramètres stabilisés, la sonde est accordée aux fréquences des noyaux étudiés.

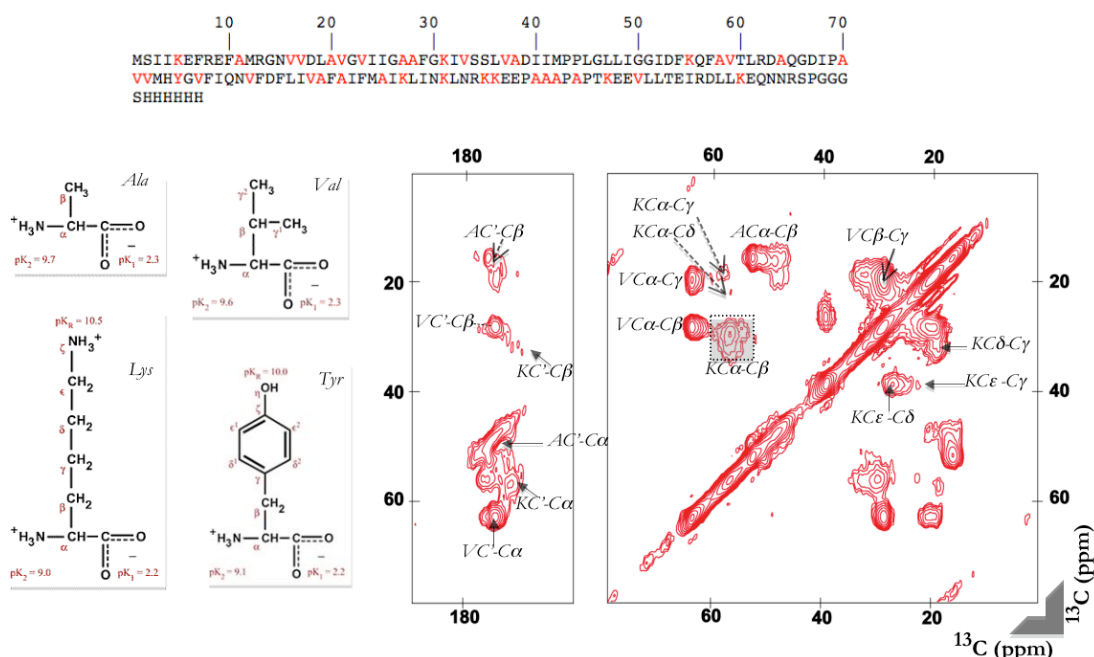


Figure 3.3 : Spectre 2D ^{13}C - ^{13}C obtenu sur l'échantillon (Ala, Lys, Val, Tyr)- ^{13}C , ^{15}N -MscL à 248 K avec un temps de mélange de 10 ms. Plusieurs corrélations des résidus Ala, Lys et Val sont mentionnées sur le spectre

Une série de sept expériences 2D ^{13}C - ^{13}C DARR a été effectuée sur l'échantillon (Ala, Lys, Val, Tyr)- ^{13}C , ^{15}N -MscL avec des temps de mélange de 10, 20, 50, 100, 200, 400 et 500 ms. Pour un temps de mélange de 10 ms, plusieurs pics sont identifiables (Figure 3.3). Un pic fin situé au niveau des corrélations Cα-Cε des Lys montre une résolution de 1,2 ppm. La région des corrélations C'-Cα montre des pics peu résolus. Au niveau des corrélations entre les C' et les carbones de la chaîne latérale, deux pics sont présents. Il s'agit très probablement des corrélations C'-Cβ des résidus Ala, Lys et Val. Concernant les corrélations Cα-Cβ, celles des résidus Ala, Lys et Val sont présentes. Les valeurs de déplacements chimiques des Cα et des Cβ correspondent à celles prédites par le programme SPARTA. Aucun signal spécifique à la Tyr n'est présent. Des corrélations entre

les $C\alpha$ et les autres carbones de la chaîne latérale des résidus *Lys* et *Val* sont présentes. D'autres corrélations sont visibles, notamment celles des carbones des chaînes latérales des *Lys*.

En augmentant le temps de mélange, les corrélations présentes pour 10 ms de temps de mélange sont plus intenses (Figure 3.4). D'autres corrélations apparaissent. Celles des C' avec les carbones des chaînes latérales sont plus larges. Des signaux apparaissent dans la région des carbones aromatiques ($100\text{-}140\text{ ppm}$) mais aucun signal $C\alpha\text{-}C\beta$ de la *Tyr* n'est visible (Figures 3.4a et 3.4b). Pour des temps de mélange supérieurs à 100 ms (Figures 3.4c et 3.4d), une corrélation intense apparaît. Un pic croisé entre un carbone résonnant autour de $56\text{-}57\text{ ppm}$ et un autre à 40 ppm et présentant une résolution de $1,5\text{ ppm}$ est visible sur le spectre. Pour un temps de mélange de 200 ms , un pic juxtaposé mais moins intense apparaît. Ce sont des corrélations entre les $C\alpha$ et les $C\epsilon$ des *Lys*. Un autre pic apparaît pour des temps de mélange supérieurs à 100 ms à $64,5/53\text{ ppm}$. Il s'agit très probablement d'une corrélation entre un $C\alpha$ d'une *Ala* et un $C\alpha$ d'une *Val*. Sur le spectre obtenu avec 50 ms de temps de mélange, et pour un faible rapport signal/bruit, ces deux corrélations sont présentes. Les déplacements chimiques des groupes d'acides aminés sont répertoriés dans le tableau 3.1. À partir des spectres obtenus, les intensités des corrélations $57,1/39,1$, $55,8/39,2$ et $64,4/53\text{ ppm}$ ont été mesurées. L'évolution de l'intensité de ces trois pics en fonction du temps de mélange est représentée dans la figure 3.5. Au niveau du pic intense des corrélations $C\alpha\text{-}C\gamma$ des *Val*, une excroissance apparaît pour des temps de mélange supérieurs à 200 ms . Il s'agit très probablement d'une corrélation entre le $C\alpha$ d'une *Val* et le $C\beta$ d'une *Ala*.

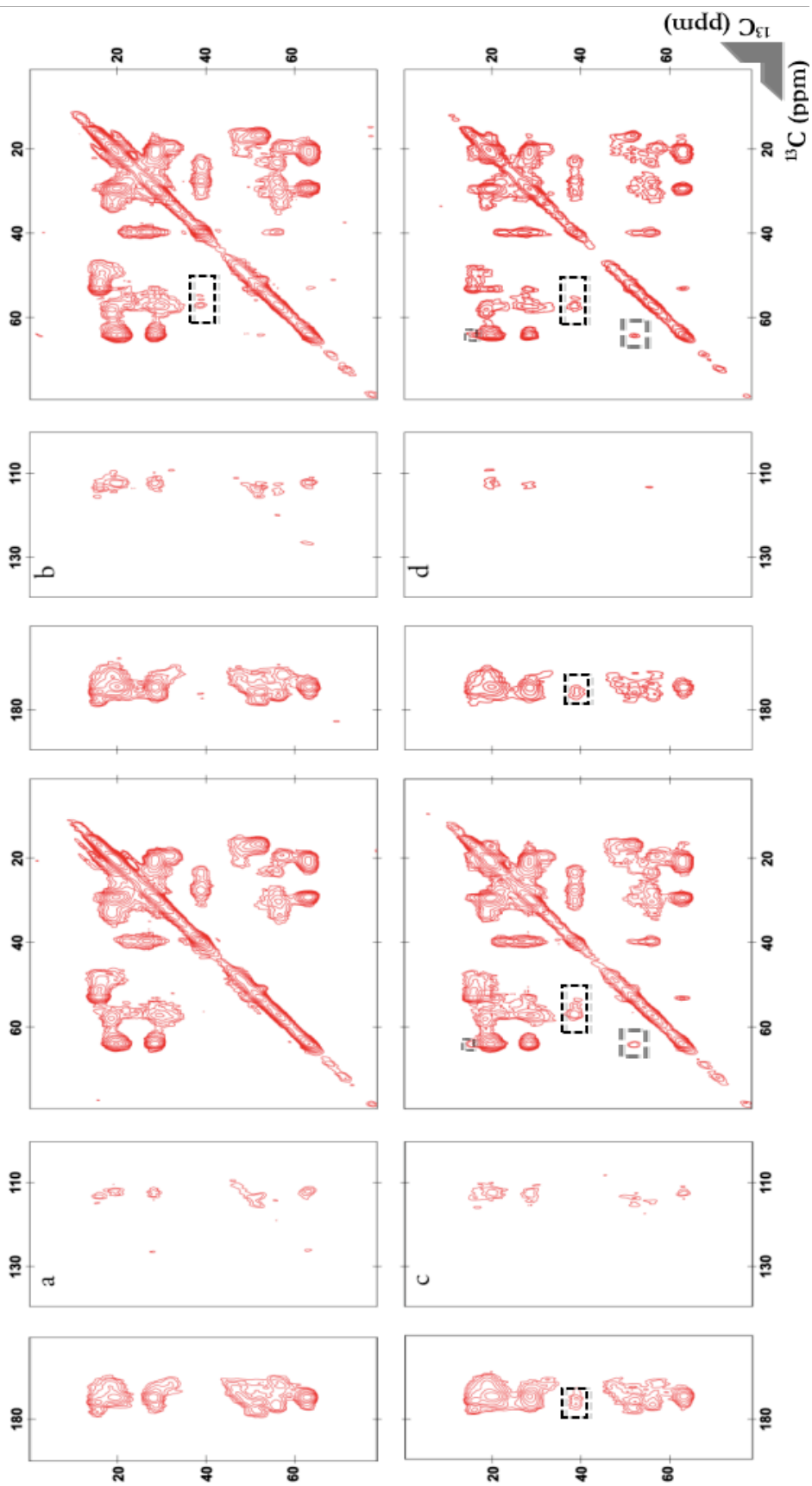


Figure 3.4 : Spectres $2D\ ^{13}C-^{13}C$ obtenus sur l'échantillon (Ala, Lys, Val, Tyr)- $MscL$ à $258\ K$ avec différents temps de mélange. Les spectres présentés ont été obtenus en appliquant la séquence $DARR$ avec des temps de mélange de $20, 50, 200$ et $500\ ms$. Les pics encadrés en noir correspondent à des corrélations intra-résiduelles $C\alpha-C\beta$ et $VC\alpha-AC\beta$ et $VC\alpha-AC\alpha$ d'une ou plusieurs Lys. Les pics encadrés en gris correspondent à des corrélations inter-résiduelles $VC\alpha-AC\alpha$ et $VC\alpha-AC\beta$ entre une Ala et une Val.

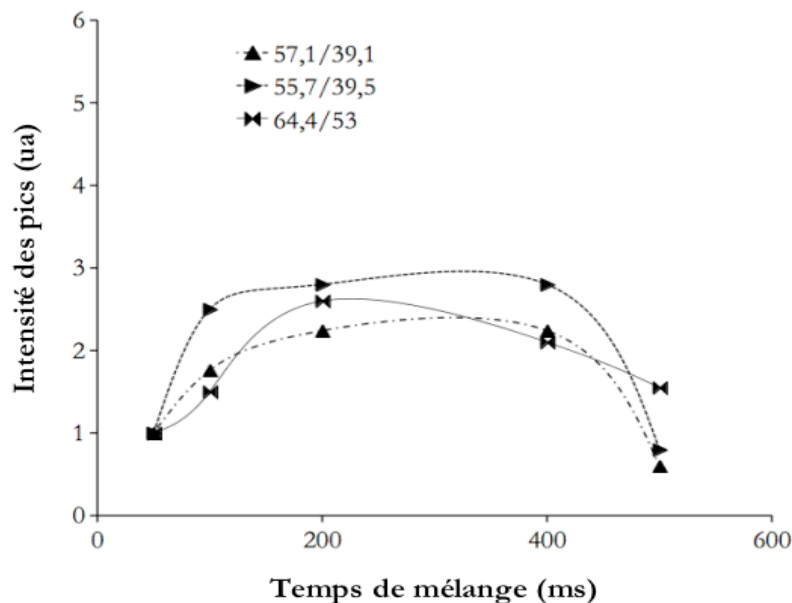


Figure 3.5 : Évolution du signal de trois corrélations isolées des différents spectres obtenus sur l'échantillon (Ala, Lys, Val, Tyr)-*Ec-MscL*. Les valeurs sont extraites par intégration gaussienne sur les pics au niveau des différents spectres obtenus en appliquant la séquence d'impulsion *DARR* à différents temps de mélange entre 50 et 500 ms. Les deux corrélations situées en 57,1/39,1 et 55,7/39,5 sont probablement des corrélations $C\alpha-C\epsilon$ de deux ou plusieurs Lys, et la corrélation située en 64,4/53 est une corrélation inter-résiduelle entre le $C\alpha$ d'une ou plusieurs Ala, et le $C\alpha$ d'une ou plusieurs Val. L'intensité des signaux est ramenée à celle de la corrélation identifiée sur le spectre obtenu à 50ms. La distance entre le $C\alpha$ et le $C\epsilon$ de la Lys est de 6 Å. Pour la corrélation inter-résiduelle Ala-Val, la distance est de 4,4,5 Å s'il s'agit de résidus liés à travers une liaison peptidique ; elle est de 5,4 Å s'il s'agit de deux résidus séparés par un pas d'hélice dans une hélice α ; elle est supérieure à 5,4 Å dans les autres cas de figure.

La séquence *specific-CP* a été appliquée sur cet échantillon. Cette stratégie de marquage implique la présence de onze doublets d'acides aminés marqués. Il s'agit des doublets $_{16}V-_{17}V$, $_{20}A-_{21}V$, $_{27}A-_{28}A$, $_{37}V-_{38}A$, $_{58}A-_{59}V$, $_{70}A-_{71}V$, $_{71}V-_{72}V$, $_{88}V-_{89}A$, $_{105}K-_{106}K$, $_{110}A-_{111}A$ et $_{111}A-_{112}A$. Sur la partie *N-C'*, un seul pic apparaît représentant les onze corrélations entre acides aminés voisins (Figure 3.6a). Sur la partie *N-C α* , trois pics apparaissent représentant les 47 corrélations des résidus marqués de la protéine (Figure 3.6b).

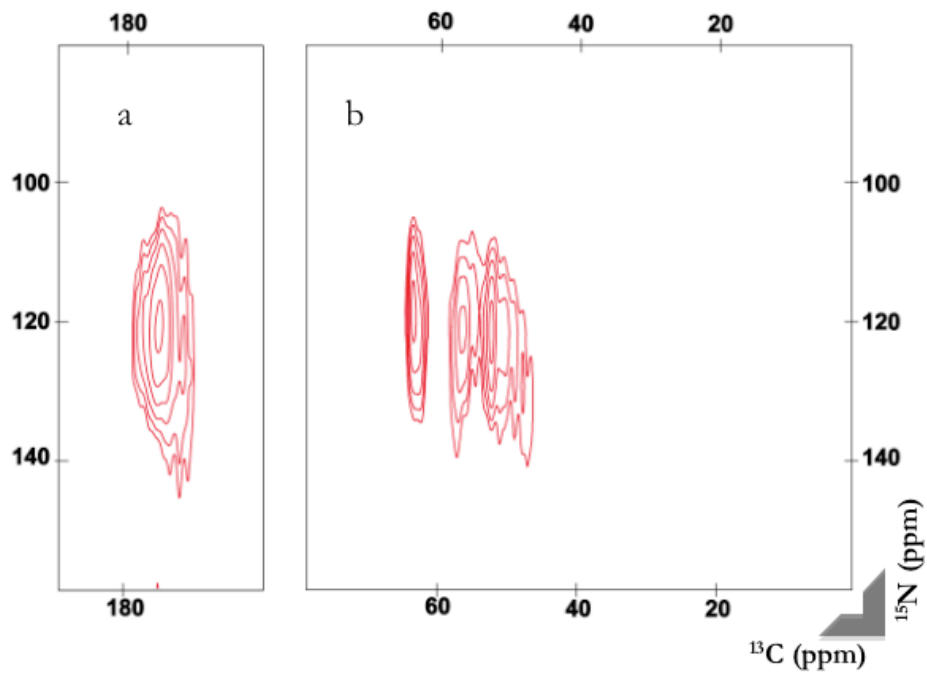


Figure 3.6 : Spectres $2D$ ^{15}N - ^{13}C obtenus sur l'échantillon (Ala, Lys, Val, Tyr) ^{13}C , ^{15}N -MscL à 248 K. Les spectres présentés ont été obtenus en appliquant la séquence *specific-CP* entre le ^{15}N d'un résidu n et les ^{13}C d'un résidu $n-1$ (a), et entre les ^{15}N d'un résidu et les ^{13}C du même résidu (b).

Déplacements chimiques des groupes d'acides aminés <i>Ala</i> , <i>Lys</i> et <i>Val</i> . Observations effectuées sur tous les spectres obtenus à différents temps de mélange							
<i>aa</i>	¹³ C	<i>C'</i>	<i>Cα</i>	<i>Cβ</i>	<i>Cγ</i>	<i>Cδ</i>	<i>Cε</i>
<i>Ala</i> *	déplacement chimique (ppm)	173,2	49,2	16,5			
	Erreur (ppm)	3,5	2	3,5			
<i>Ala</i> **	déplacement chimique (ppm)	176,1	53,0	16,1			
	Erreur (ppm)	2	3	2,5			
<i>Lys</i>	déplacement chimique (ppm)	174,5	57,3	29,8	24,6	29,3	39,4
	Erreur (ppm)	2	2,4	5	2,6	3	1,5
<i>Val</i>	déplacement chimique (ppm)	174,5	64,2	29,1	19,9		
	Erreur (ppm)	1,5	2,5	2,0	3		
<i>Val</i> ***	déplacement chimique (ppm)	-	58,5	-	19,8		
	Erreur (ppm)	-	1,5	-	1,5		

Tableau 3.1 : Déplacements chimiques des groupes d'acides aminés *Ala*, *Lys* et *Val*. Ces valeurs sont issues de tous les spectres 2D DARR obtenus sur l'échantillon (*Ala*, *Lys*, *Val*, *Tbr*)-*Ec-MscL* à différents temps de mélange. Pour les *Ala*, les déplacements chimiques sont séparés en deux groupes. Le premier (*) représente les résidus situés dans des régions de la protéine en dehors des hélices. Le deuxième (**) montre des déplacements chimiques caractéristiques des résidus situés au niveau d'une hélice α . Pour les *Val*, des corrélations *Cα-Cγ* sont identifiées au niveau des spectres 2D à des faibles temps de mélange (***), et sont confondues avec les corrélations dues aux chaînes latérales des *Lys* pour des temps de mélange supérieurs à 100 ms. Les déplacements chimiques des *Cα* de ces *Val* montrent des valeurs caractéristiques de résidus situés en dehors des hélices.

IV Analyse de l'échantillon sélectivement marqué *FIMPR-Ec-MscL* - Analyse de la séquence

La deuxième stratégie testée pour l'élaboration de nouveaux schémas de marquage par *SIV* est basée uniquement sur l'analyse de la séquence de la protéine. L'idée développée est basée sur le marquage double combinatoire ou *CDL* (*combinatorial dual-isotope labelling*) mis en place dans le laboratoire de *V. DÖTSCH* et *F. BERNHARD* (Sobhanifar *et al.*, 2010), mais également sur une méthode issue du laboratoire de *N.E. DIXON* et *G. OTTING* (Ozawa *et al.*, 2006). La technique *CDL* consiste à identifier des paires uniques d'acides aminés au sein de la séquence. En procédant ensuite au marquage des résidus impliqués, la recherche de corrélations sur un spectre donnera des informations précises et uniques, permettant une attribution simplifiée. La deuxième méthode de marquage combinatoire consiste à analyser la composition de la protéine en acides aminés, et classer ces derniers en fonction de leur fréquence. Cinq schémas de marquage sont proposés : les cinq acides aminés les plus fréquents sont marqués une fois, les cinq les moins fréquents sont marqués trois fois, et les dix intermédiaires sont marqués deux fois. Cela permet de simplifier les spectres et facilite ainsi l'attribution.

a Marquage basé sur l'analyse de la structure

Concernant la composition de *Ec-MscL*, les acides aminés les plus fréquents sont les résidus hydrophobes *Ile*, *Ala*, *Leu*, *Val* et *Gly*. Les acides aminés les moins fréquents sont *Met*, *Ser*, *Gln*, *Thr* et *Tyr*. En prenant les résidus *Ile* du premier groupe, et le résidu *Met* du deuxième, une paire unique peut être identifiée au niveau de l'hélice transmembranaire *TM1* de *Ec-MscL*. Il s'agit de la paire $_{41}I-_{42}M$. En cherchant parmi le groupe intermédiaire comprenant les résidus *Phe*, *Lys*, *Glu*, *Asp*, *His*, *Pro*, *Arg* et *Asn*, le choix du triplet *Phe*, *Pro* et *Arg* permet d'obtenir huit autres paires uniques. Il s'agit de $_{7}F-_{8}R$, $_{12}M-_{13}R$, $_{42}M-_{43}P$, $_{68}I-_{69}P$, $_{78}F-_{79}I$, $_{92}I-_{93}F$, $_{93}F-_{94}M$, $_{125}I-_{126}R$. De plus, ces paires sont réparties entre les différentes parties de *MscL* afin d'avoir au moins une paire par élément de structure secondaire (Figure 4.1).

MGI¹³IK E¹⁵FREF AMRGN VVDLA VGVII GAAFG KIVSS LVADI IMPEL GLLIG
 GIDFK QFAVT LRDAQ GDIPA VVMHY GVFIQ NVFDF LIVAF AIFMA IKLIN
 KLN¹³RK KEEPA AAPAP TKEEV LLTEI RDLLK EQNNR SPGGG SHHHH HH

Figure 4.1 : Analyse de la séquence de *Ec-MscL* pour établir une stratégie de marquage. La séquence de la protéine *Ec-MscL* contenant l'étiquette histidine est représentée en haut. Les parties soulignées correspondent aux deux hélices transmembranaires et à l'hélice cytoplasmique. Les résidus colorés sont ceux qui ont été marqués aux ^{13}C et ^{15}N . Pour trouver la stratégie de marquage, les acides aminés composant la protéine ont été classés en fonction de leur pourcentage de présence. Les acides aminés sont séparés par groupes de cinq. Les premiers acides aminés des trois premiers groupes (*Ile*, *Phe* et *Pro*) sont choisis. Les deux autres acides aminés (*Met* et *Arg*) ont été choisis en analysant la séquence à la recherche de paires uniques (surlignées en gris) dans les différentes parties de la protéine.

<i>Ile</i> (I)	16	10,80%
<i>Ala</i> (A)	15	10,20%
<i>Leu</i> (L)	13	8,80%
<i>Val</i> (V)	13	8,80%
<i>Gly</i> (G)	12	8,10%
<i>Phe</i> (F)	10	6,80%
<i>Lys</i> (K)	9	6,10%
<i>Glu</i> (E)	8	5,40%
<i>Asp</i> (D)	7	4,70%
<i>His</i> (H)	7	4,70%
<i>Pro</i> (P)	7	4,70%
<i>Arg</i> (R)	6	4,00%
<i>Asn</i> (N)	6	4,00%
<i>Met</i> (M)	5	3,40%
<i>Ser</i> (S)	5	3,40%
<i>Gln</i> (Q)	4	2,70%
<i>Thr</i> (T)	3	2,00%
<i>Tyr</i> (Y)	1	0,10%
<i>Cys</i> (C)	0	0,00%
<i>Trp</i> (W)	0	0,00%

Ainsi, un marquage aux isotopes ^{13}C et ^{15}N a été effectué par *SIV* sur les 10 *Phe*, 16 *Ile*, 5 *Met*, 7 *Pro* et 6 *Arg* de la protéine *Ec-MscL*. La *SIV* a été réalisée en présence de liposomes de *DOPC*, avec dans un premier temps un ratio protéine/lipide de 1/4. Dans un second temps, l'échantillon de protéoliposomes a été solubilisé en micelles de *triton-x-100*, et la protéine a été purifiée sur colonne de *Ni-NTA*. Après purification, la protéine a été reconstituée en liposomes de *DOPC* avec un ratio protéine/lipide de 1/2. Ces deux échantillons ont été étudiés en utilisant la séquence *DARR* avec des temps de mélange de 20 à 500 ms à 248 K.

b Production de l'échantillon (*Phe*, *Ile*, *Met*, *Pro*, *Arg*)-*MscL* en liposomes de *DOPC*

La protéine a été synthétisée *in vitro* en utilisant le même protocole que celui qui a servi à produire l'échantillon (*Ala*, *Lys*, *Val*, *Tyr*)- ^{13}C , ^{15}N -*MscL*. Pour un ratio de protéine/lipide de 1/4 et pour un rendement estimé à 10 mg de protéines, 40 mg de *DOPC* sont nécessaires. Après la *SIV* et la centrifugation, 9 mg de *Ec-MscL* sont obtenus.

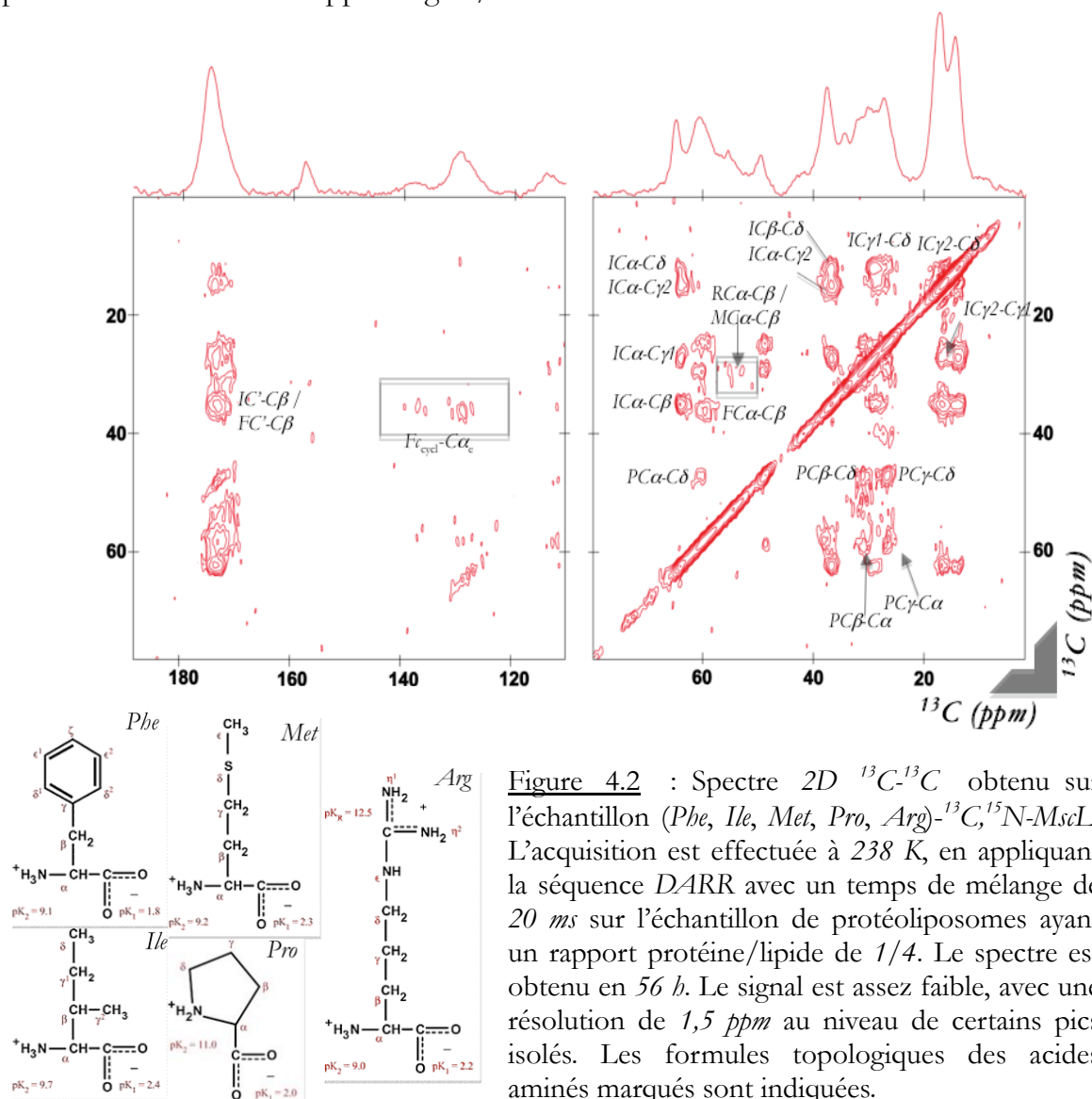
L'échantillon a été conservé à -20°C jusqu'à l'analyse d'une partie de l'échantillon au spectromètre de RMN. Après la première série d'expériences, l'autre partie de l'échantillon a été solubilisée en micelles de détergent. Un volume de 20 mL de tampon *A1* a été utilisé pour obtenir des micelles ternaires. La protéine a été purifiée sur une colonne d'affinité. Les fractions récupérées après purification ont été analysées par *DOT BLOT*. Les fractions contenant la protéine *Ec-MscL* ont été dialysées contre 2 L de tampon *RB-T*. Après la dialyse, la protéine est quantifiée par dosage au *BCA*, indiquant la présence de 5 mg de protéines. Pour reconstituer l'échantillon 10 mg de *DOPC* sont nécessaires. Les vésicules sont préparées à l'aide de l'extrudeur, puis rajoutées à l'échantillon de protéines en micelles de détergents. Le complexe ternaire est mélangé pendant 20 min , puis le détergent est retiré à l'aide des *BIO-BEADS*. Après extraction totale du détergent, les protéoliposomes sont centrifugés, et conservés à -20°C jusqu'à leur analyse au spectromètre de RMN.

c Analyse de (Phe, Ile, Met, Pro, Arg)-MscL par RMN

La protéine *Ec-MscL* contenant l'étiquette histidine est composée de 147 résidus dont 10 Phe, 16 Ile, 5 Met, 7 Pro et 6 Arg. Chaque molécule marquée aux ^{13}C et ^{15}N sur les Phe, Ile, Met, Pro et Arg contient donc 277 atomes ^{13}C et 62 atomes ^{15}N . Le poids moléculaire après le marquage isotopique est de $16,487\text{ kDa}$. Les protéoliposomes contenant la protéine sélectivement marquée sont transférés dans le rotor adéquat et conservés à -20°C jusqu'aux expériences RMN. Après la mise au point des différentes valeurs des séquences d'impulsions utilisées, le rotor est placé à température ambiante. Il est ensuite inséré dans la sonde *MAS*, et la rotation à l'angle magique est augmentée par paliers (jusqu'à $11\ 000\text{ Hz}$). La température souhaitée est fixée. Une fois ces deux paramètres stabilisés, la sonde est accordée aux fréquences des noyaux étudiés.

Sur l'échantillon de protéoliposomes avec un ratio protéine/lipide de $1/4$, l'effet de la température a été testé. Des spectres $2\text{D }^{13}\text{C}-^{13}\text{C}$ ont été obtenus à des températures allant de 233 à 268 K . L'analyse de la largeur de raies de certains pics isolés montre une résolution variant entre $1,5$ et $2,2\text{ ppm}$. Malgré un rapport signal sur bruit assez faible, la

meilleure résolution est obtenue pour une température de 238 K. L'effet de la variation du temps de mélange a ensuite été mesuré à cette température. La séquence *DARR* a été appliquée avec des temps de mélange variant entre 20 et 500 ms. Pour chaque valeur de temps de mélange, plusieurs spectres ont été enregistrés avec des temps d'acquisition de 14 h. Pour chaque temps de mélange, quatre *FID* sont additionnés pour obtenir un spectre avec un meilleur rapport signal/bruit.



Les corrélations identifiées entre les atomes de carbone de ces résidus sont indiquées sur le spectre 2D. La zone des corrélations entre les C' et les autres carbones des acides aminés montre très peu de séparation entre les pics. Des signaux issus de corrélations entre les carbones aromatiques du cycle phényl des *Phe* et les $\text{C}\alpha$ sont présents entre 120 et 140 ppm. Dans la région 0-80 ppm, différentes corrélations sont identifiées. Parmi lesquelles les corrélations $\text{C}\alpha$ - $\text{C}\beta$ des cinq familles de résidus sont mentionnées. Les corrélations entre les $\text{C}\alpha$ et le reste de la chaîne carbonée sont présentes pour les résidus *Ile* et *Pro* mais pas pour les résidus *Arg* et *Met*.

La figure 4.2 montre le spectre obtenu pour un temps de mélange de 20 ms . La résolution est assez faible. Sur certains pics isolés, elle est située autour de $1,9\text{ ppm}$ avec un minimum de $1,5\text{ ppm}$. La région des C' présente très peu de séparation entre les différents pics. Les corrélations entre les $C\alpha$ et les carbones des chaînes latérales sont identifiables.

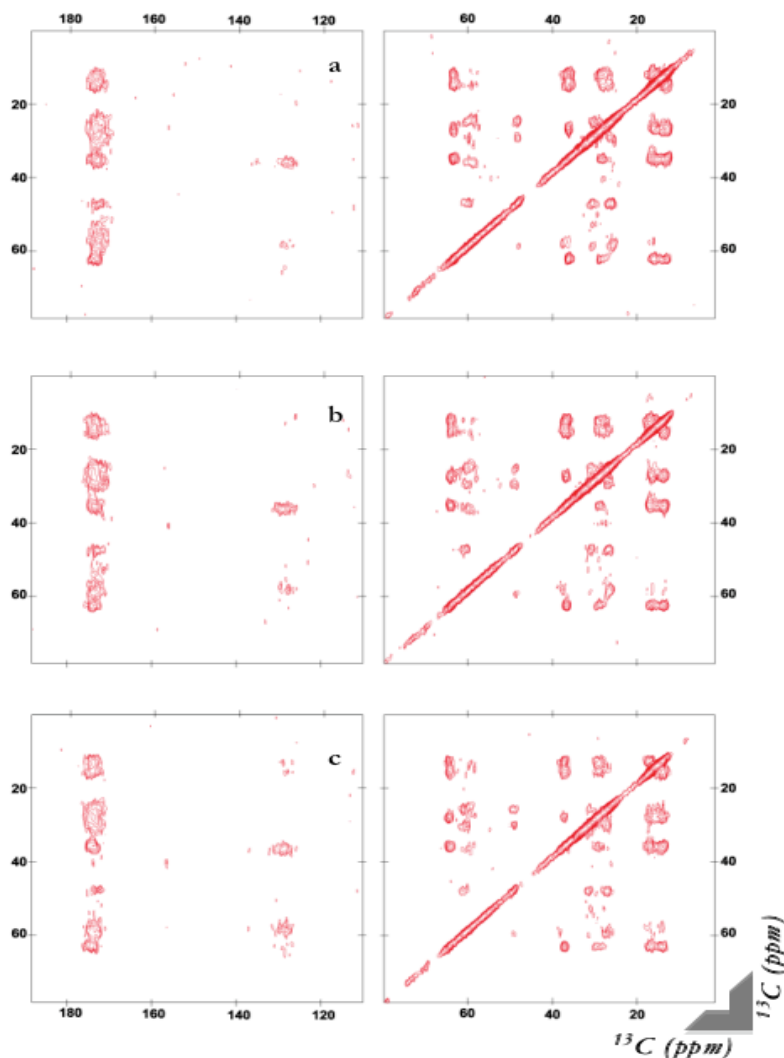


Figure 4.3 : Spectres $2D\ ^{13}C-^{13}C$ obtenus sur l'échantillon $(Phe, Ile, Met, Pro, Arg)-^{13}C,^{15}N$ -MscL à différents temps de mélange. L'acquisition est effectuée à 238 K , en appliquant la séquence *DARR* avec des temps de mélange de 50 (a), 100 (b), et 200 ms (c) sur l'échantillon de protéoliposomes ayant un rapport protéine/lipide de $1/4$. Chaque spectre est obtenu en 56 h . Le signal est assez faible, avec une résolution entre $1,4$ et 2 ppm au niveau de certains pics isolés.

Les pics présentent une meilleure résolution pour un temps de mélange de 50 ms (Figure 4.3a). Les signaux sont plus intenses pour un temps de mélange de 100 ms (Figure 4.3b). Pour des temps de mélange supérieurs à 100 ms , le signal diminue. Aucune

corrélation inter-résiduelle n'est présente sur les spectres (Figure 4.3c). Pour avoir un meilleur rapport signal/bruit, le reste de l'échantillon a été solubilisé en micelles de détergent, purifié puis reconstitué en liposomes avec un ratio protéine/lipide de 1/2. Les spectres de la figure 4.4 montrent une très bonne résolution. Les pics sont plus fins et la résolution est de 1,3 ppm en moyenne. Sur tous les spectres obtenus, une corrélation à 65/48,5 ppm apparaît. Il peut s'agir d'une corrélation $IC\alpha-PC\delta$. Au dessus de 50 ms de temps de mélange, une autre corrélation inter-résiduelle apparaît à 65/59,5 ppm. Il peut s'agir d'une corrélation entre le $C\alpha$ d'une *Ile* et d'un $C\alpha$ d'une *Phe* ou d'une *Pro*. Lorsque pour chaque temps de mélange (50, 100, 200 et 500 ms) quatre FIDs sont additionnés, les corrélations impliquant les carbones des chaînes latérales des *Arg* apparaissent (Figure 4.5). Les autres acides aminés présentent des pics croisés très intenses. La résolution ne permet pas néanmoins de distinguer deux acides aminés de la même famille. L'évolution de l'intensité de quelques signaux isolés en fonction du temps de mélange est indiquée dans la figure 4.6. Les déplacements chimiques observés sont regroupés dans le tableau 4.1.

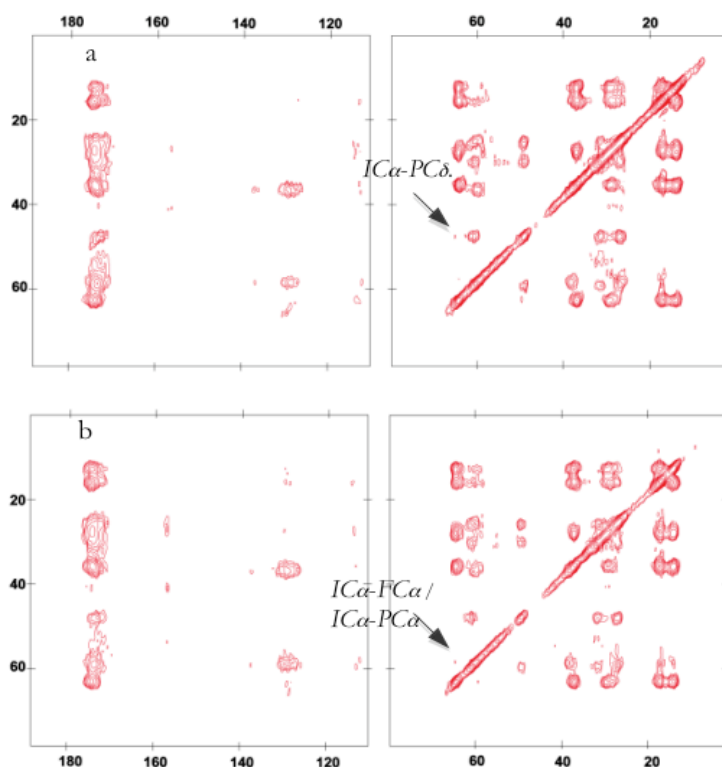


Figure 4.4 : Spectres 2D ^{13}C - ^{13}C obtenus sur l'échantillon (Phe, Ile, Met, Pro, Arg)- ^{13}C , ^{15}N -MscL à différents temps de mélange. L'acquisition est effectuée à 238 K, en appliquant la séquence DARR avec des temps de mélange de 50 (a) et 200 ms (b) sur l'échantillon ayant un rapport protéine/lipide de 1/2. Chaque spectre est obtenu en 14 h.

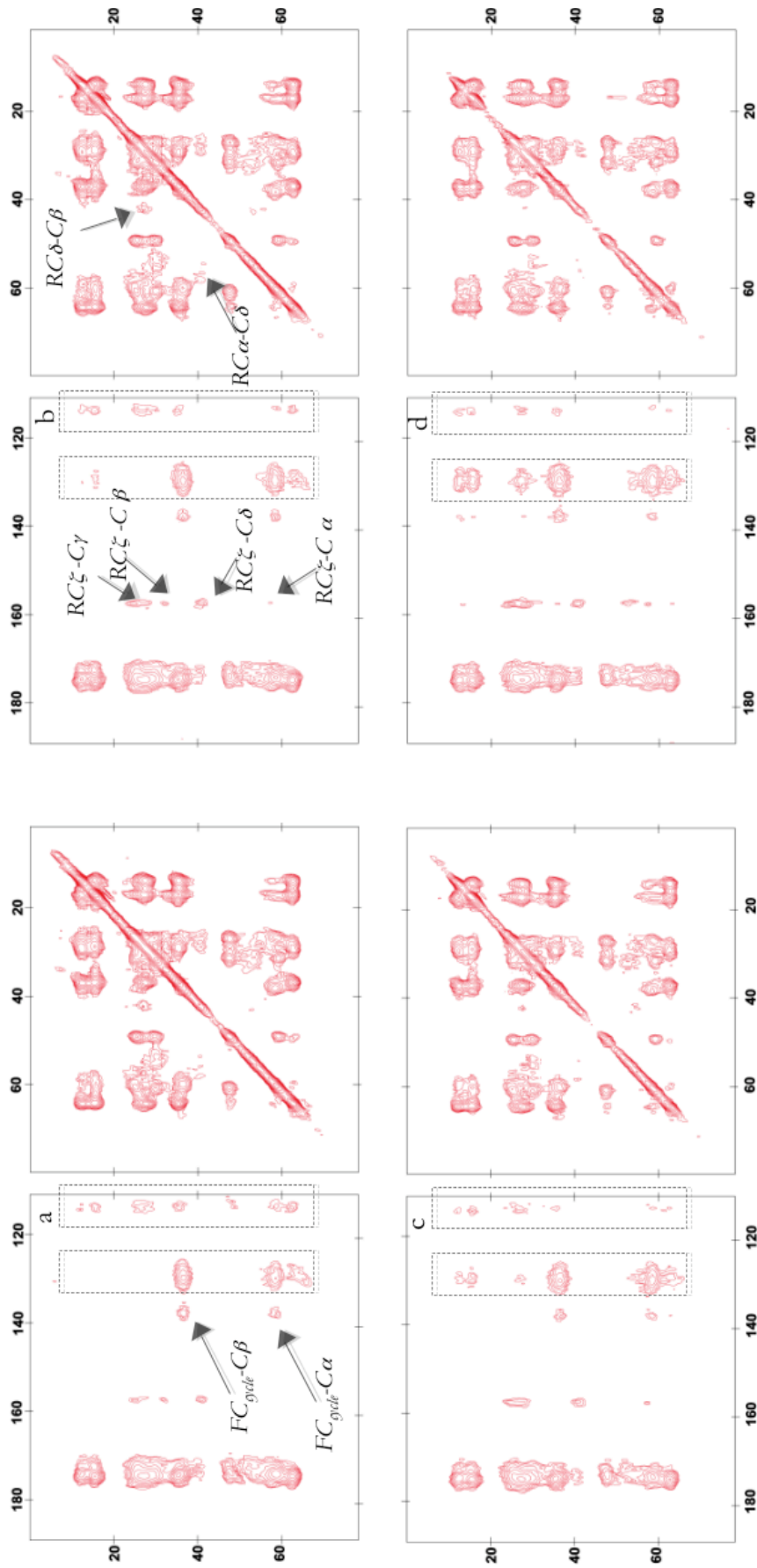


Figure 4.5 : Spectres $2D^{13}C-^{13}C$ obtenus sur l'échantillon (Phe, Ile, Met, Pro, Arg)- $^{13}C,^{15}N$ -MscL à différents temps de mélange. L'acquisition est effectuée à 238 K, en appliquant la séquence $DARR$ avec des temps de mélange de 50 (a), 100 (b), 200 (c) et 500 ms (d) sur l'échantillon de protéoliposomes ayant un rapport protéine/lipide de 1/2. Chaque spectre est obtenu en 56 h.

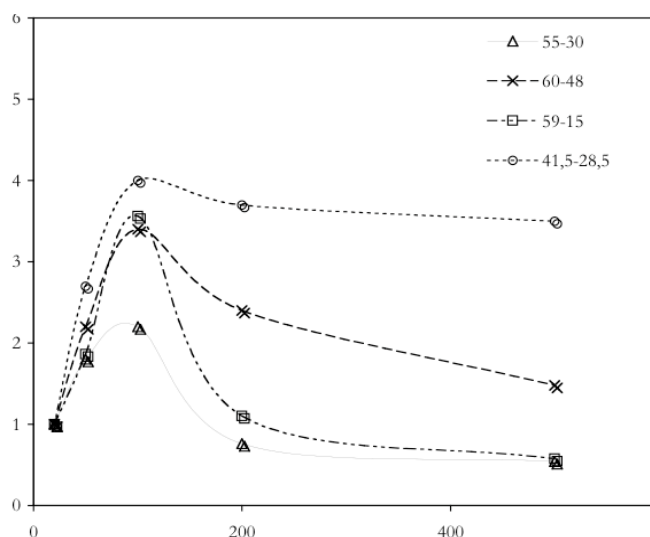


Figure 4.6 : Évolution du signal de quatre corrélations isolées des différents spectres obtenus sur l'échantillon (*Phe, Ile, Met, Pro, Arg*)-*Ec-MscL*. Les valeurs sont extraites par intégration gaussienne sur les pics au niveau des différents spectres obtenus en appliquant la séquence d'impulsion *DARR* à différents temps de mélange entre 50 et 500 *ms*. Le pic 55-30 représente une corrélation $C\alpha-C\beta$ d'une *Arg* ou d'une *Met*, ce qui correspond à une distance de 1,5 Å. Le pic 60-48 représente probablement une corrélation $IC\alpha-PC\delta$. Le pic 59-15 représente une corrélation $IC\alpha-C\delta$ ou $IC\alpha-C\gamma_2$ correspondant à une distance de 3 à 4,5 Å. Le pic 41,5-28,5 représente une corrélation $IC\beta-C\gamma_1$ ou $RC\delta-C\gamma$ correspondant dans les deux cas à une distance de 1,5 Å.

Déplacements chimiques des groupes d'acides aminés <i>Phe, Ile, Met, Pro</i> et <i>Arg</i> . Observations effectuées sur tous les spectres obtenus à différents temps de mélange								
<i>aa</i> \ ^{13}C	C'	$C\alpha$	$C\beta$	$C\gamma$	$C\delta$	$C\epsilon$	$C\zeta$	
<i>Phe</i> déplacement chimique (<i>ppm</i>)	174,8	59,6	37,5					
Erreur (<i>ppm</i>)	3,0	0,5	1,5					
<i>Ile</i> déplacement chimique (<i>ppm</i>)	175,4	64,2	36,2	28,4	15,8	12,7		
Erreur (<i>ppm</i>)	2,0	1,5	3,0	2,5	1,5	1,0		
<i>Met</i> déplacement chimique (<i>ppm</i>)	172,5	58,6	31,0	31,0		16		
Erreur (<i>ppm</i>)	3,0	1,5	2,0	2,0		1,5		
<i>Pro</i> déplacement chimique (<i>ppm</i>)	175,0	62,0	30,5	26,0	48,9			
Erreur (<i>ppm</i>)	3,0	2,0	2,5	1,5	1,5			
<i>Arg</i> déplacement chimique (<i>ppm</i>)	173,0	54,2	30,2	-	41,5		157,8	
Erreur (<i>ppm</i>)	2,0	2,5	2,0	-	0,2		0,2	

Tableau 4.1 : Déplacements chimiques des groupes d'acides aminés *Phe, Ile, Met, Pro* et *Arg*. Ces valeurs sont issues de tous les spectres 2D *DARR* obtenus sur l'échantillon (*Phe, Ile, Met, Pro, Arg*)-*Ec-MscL* à différents temps de mélange. Pour les résidus *Phe*, les déplacements chimiques des carbones du cycle aromatiques sont situés entre 128,5 et 138,6 *ppm*.

V Analyse de l'échantillon sélectivement marqué *DGLS-Ec-MscL* - une combinaison des approches prédictives

La dernière stratégie employée pour l'élaboration du schéma impliquant le marquage des résidus *Asp*, *Gly*, *Leu* et *Ser* en *SIV* utilise des informations issues des prédictions mais aussi de l'analyse de la séquence. Pour l'approche prédictive utilisée précédemment (voir le chapitre III), le spectre obtenu sur l'échantillon *AKVY-Ec-MscL* contient des informations à une bonne résolution. Pour certains pics isolés, la résolution était de 1,2 *ppm*. Cependant, le spectre n'est pas interprétable à cause de l'imbrication des déplacements chimiques de tous les acides aminés de la même famille dans une même zone. En considérant les valines par exemple, le spectre montre des corrélations $C\alpha-C\beta$ et $C\alpha-C'$ d'une dizaine de *Val* dans une petite surface n'excédant pas 2 *ppm* dans les deux dimensions (figure 23). Le recouvrement est moindre dans le cas des *Ala* et des *Lys*, mais aucun pic ne peut être identifié. Le pic isolé retrouvé en 59-40 *ppm* en ordonnée représente une corrélation $C\alpha-C\epsilon$ d'une ou plusieurs *Lys*. Une corrélation $C\alpha-C\alpha$ et une autre $C\alpha-C\beta$ entre *Val* et *Ala* sont visibles également, mais l'existence de cinq doublets *AV* ou *VA* rend l'identification impossible. La prédiction des déplacements chimiques a été effectuée sur un plus grand nombre de modèles basés sur quatre alignements de séquence différents, en séparant les déplacements chimiques des résidus au sein de la même famille d'acides aminés. En combinant les résultats de prédiction des déplacements chimiques, avec une analyse de la séquence pour trouver le maximum de paires isolées, le choix a été porté sur les résidus *Asp*, *Gly*, *Leu* et *Ser*.

a Stratégie de marquage

Pour prédire les déplacements chimiques, de nouveaux alignements de séquence ont été effectués. Comme précédemment, les séquences alignées sont la séquence complète (136 acides aminés) de *Ec-MscL* et la séquence tronquée (125 acides aminés) de *Mt-MscL*, correspondant à la structure du fichier *PDB* (2OAR). Les algorithmes utilisés sont les suivants : *ClustalW2* (Larkin *et al.*, 2007) ; *Tcoffee* (Poirot *et al.*, 2003) ; *MUSCLE* (Edgar, 2004)

et *AlignMe* (Khafizov *et al.*, 2010). Les paramètres utilisés sont les suivants : pour *ClustalW2*, un alignement par paires utilisant la matrice *GONNET* ; pour *Tcoffee*, le mode de calcul régulier ; pour *AlignMe*, les paramètres optimisés pour les protéines membranaires en hélices α , une matrice *blosum62* et l'échelle d'hydrophobicité de *KYTE & DOOLITTLE*. Les alignements varient peu, avec 5 à 7 *gaps* selon l'alignement (Figure 5.1).

<i>ClustalW2</i>	
MSIIKEFRFAMRGNVVDLAVGVIIGAAFVKIVSSLVADIMPPLGLLIGGIDFKQFAVT	60
--MLKGFKEFLARGNIVDLAVAVVIGTAF TALVTKFTDSITPLINRIG--VNAQSDVGI	56
LRDAQGDIPAVVMHYGVFIQNVDFLIVAFVAFMAIKLINKLNRRKKEPAAAPAPTKEEV	120
LRIGIGGGQTIDLN--VLLSAAINFFLIAFAVYFLVVLVLPYNTLRKKGE--VEQPGDTQVV	112
LLTEIRDLLKEQNNRS	136
LLTEIRDLLAQTN---	125

<i>Tcoffee</i>	
MSIIKEFRFAMRGNVVDLAVGVIIGAAFVKIVSSLVADIMPPLGLLIGGIDFKQFAVT	60
M--LKGFKEFLARGNIVDLAVAVVIGTAF TALVTKFTDSITPLINR-I-GVNAQSDVGI	56
LRDAQGDIPAVVMHYGVFIQNVDFLIVAFVAFMAIKLINKLNRRKKEPAAAPAPTKEEV	120
LRIGIGGGQ--TIDLNVLLSAAINFFLIAFAVYFLVVLVLPYNTLRKKGEV-EQPGDT-QVV	112
LLTEIRDLLKEQNNRS	136
LLTEIRDLLAQT---N	125

<i>AlignMe</i>	
MSIIKEFRFAMRGNVVDLAVGVIIGAAFVKIVSSLVADIMPPLGLLIGGIDFKQFAVT	60
M--LKGFKEFLARGNIVDLAVAVVIGTAF TALVTKFTDSITPLINRI--GVNAQSDVGI	56
LRDAQGDIPAVVMHYGVFIQNVDFLIVAFVAFVAFMAIKLINKLNRRKKEPAAAPAPTKEEV-	120
LRIGIGGGQTI--DLNVLLSAAINFFLIAFAVYFLVVLVLPYNTLRKKGEVEQ----PGDTQVV	112
LLTEIRDLLKEQNNRS	136
LLTEIRDLLAQTN---	125

<i>Muscle</i>	
MSIIKEFRFAMRGNVVDLAVGVIIGAAFVKIVSSLVADIMPPLGLLIGGIDFKQFAVT	60
--MLKGFKEFLARGNIVDLAVAVVIGTAF TALVTKFTDSITPLINRI--GVNAQSDVGI	56
LRDAQGDIPAVVMHYGVFIQNVDFLIVAFVAFVAFMAIKLINKLNRRKKEPAAAPAPTKEEV	120
LRIGIGG-GQTIDLNVLLSAAINFFLIAFAVYFLVVLVLPYNTLRKKGE-VEQPGDT-QVV-	112
LLTEIRDLLKEQNNRS	136
LLTEIRDLLAQTN---	125

Figure 5.1 : Alignements de séquence entre *Ec-MscL* et *Mt-MscL* en utilisant différents algorithmes. *ClustalW2*, *Tcoffee*, *AlignMe* & *Muscle* sont utilisés. La séquence d'*Ec-MscL* est complète (136 *aas*), alors que celle de *Mt-MscL* est tronquée et correspond à la séquence dont la structure est connue (Steinbacher *et al.*, 2007). Les alignements varient très peu entre eux. Les régions les plus stables sont d'une part, le début de l'hélice transmembranaire *TM1* (RGNXVDLAV) et, d'autre part, la région correspondant à l'hélice cytoplasmique située en C-ter (LLTEIRDLL).

À partir de chaque alignement de séquence, dix modèles par homologie ont été créés, vérifiés à l'aide de la série d'algorithmes *Protein Model Check* du serveur *WHAT IF*. Les protons sont rajoutés à la structure à l'aide de *Reduce* (Word *et al.*, 1999). Les déplacements chimiques sont prédits sur les quarante modèles. Les données obtenues sont séparées en

fonction de la nature de l'acide aminé, et les corrélations théoriques $C\alpha-C\beta$ sont tracées. L'analyse des déplacements chimiques prédits pour les résidus *Ala*, *Lys* et *Val* montre une très faible répartition (Figure 5.2). En effet, 11 *Ala* (sur 15) présentent des corrélations situées dans une région de 4 ppm en abscisse par 2 ppm en ordonnée (Figure 5.2a) ; les corrélations de 7 *Lys* (sur 9) sont situées dans une région de 2 ppm sur les deux axes (Figure 5.2b) ; et finalement pour les *Val*, les corrélations de 9 résidus sur 13 sont présentes dans une zone de 3 ppm en abscisses et 1 ppm en ordonnées (Figure 5.2c). Les déplacements chimiques prédits des résidus *Phe*, *Leu* et *Ile* présentent également peu de répartition, tandis que les résidus *Asp*, *Met*, *Asn*, *Ser* et *Thr* présentent des corrélations réparties sur plus de 4 ppm dans les deux dimensions. Les résidus *Met* et *Thr* ont déjà été marqués dans les échantillons précédents.

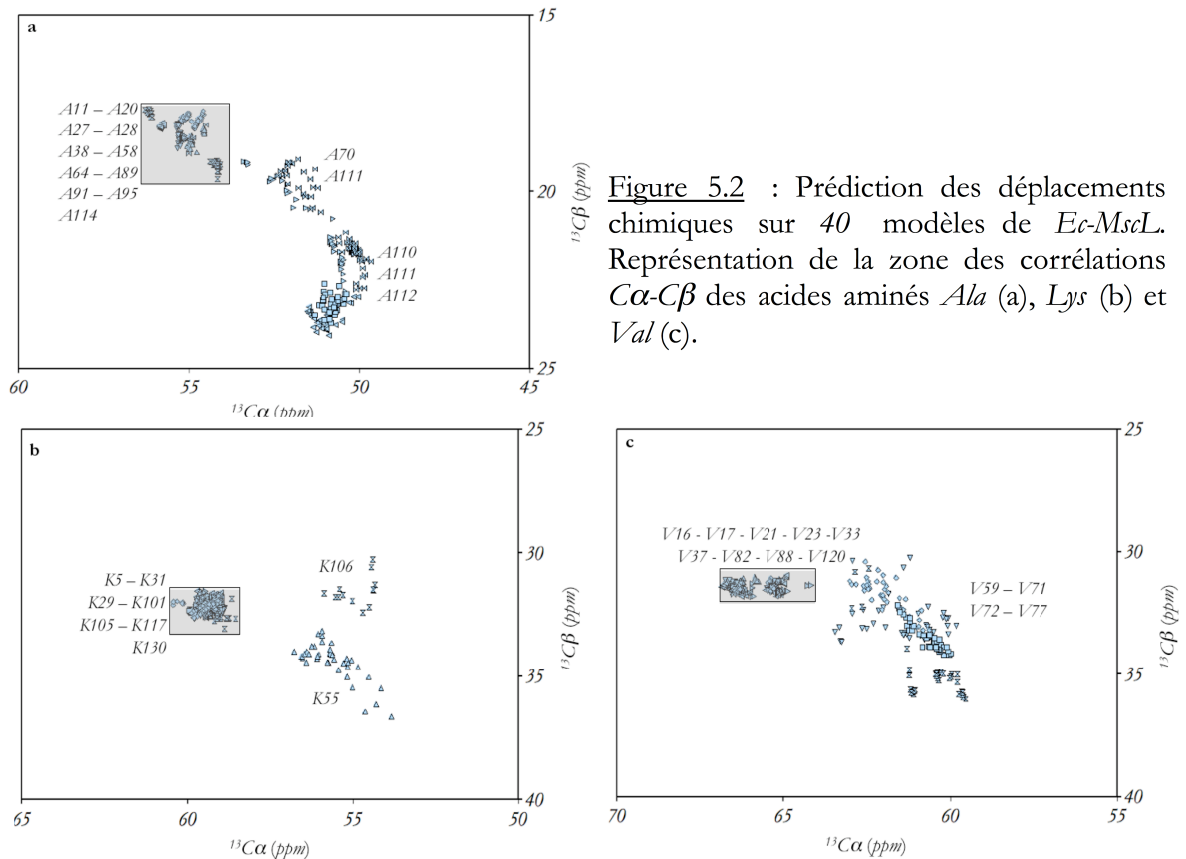


Figure 5.2 : Prédiction des déplacements chimiques sur 40 modèles de *Ec-MscL*. Représentation de la zone des corrélations $C\alpha-C\beta$ des acides aminés *Ala* (a), *Lys* (b) et *Val* (c).

Les résidus *Asp* et *Ser* ont été choisis (Figures 5.3 et 5.6) et les résidus *Gly* et *Leu* y ont été rajoutés car l'ensemble présentait plusieurs paires d'acides aminés voisins, situées dans différentes régions de la protéine. Ainsi, les paires marquées sont *18D-19L* au début

de l'hélice *TM1*, les paires $_{34}S$ - $_{35}S$ et $_{35}S$ - $_{36}L$ au sein de l'hélice *TM1*, les paires $_{45}L$ - $_{46}G$, $_{46}G$ - $_{47}L$, $_{47}L$ - $_{48}L$ à l'interface entre l'hélice *TM1* et la boucle périplasmique, les paires $_{50}G$ - $_{51}G$ et $_{66}G$ - $_{67}D$ au sein de la boucle périplasmique, les paires $_{121}L$ - $_{122}L$, $_{127}D$ - $_{128}L$ et $_{128}L$ - $_{129}L$ du côté C-terminal. De plus, les études de mutagenèse citées précédemment (Introduction - Chapitre III.C.1) accordent beaucoup d'intérêt aux résidus $_{14}G$ et $_{19}L$ susceptibles de former le site de constriction, ainsi qu'aux résidus $_{22}G$, $_{26}G$ et $_{30}G$ pour participer à ce site, mais forment également des motifs $G_{xxx}G$ d'oligomérisation des hélices transmembranaires (King & Dixon, 2010 ; Russ & Engelman, 2000).

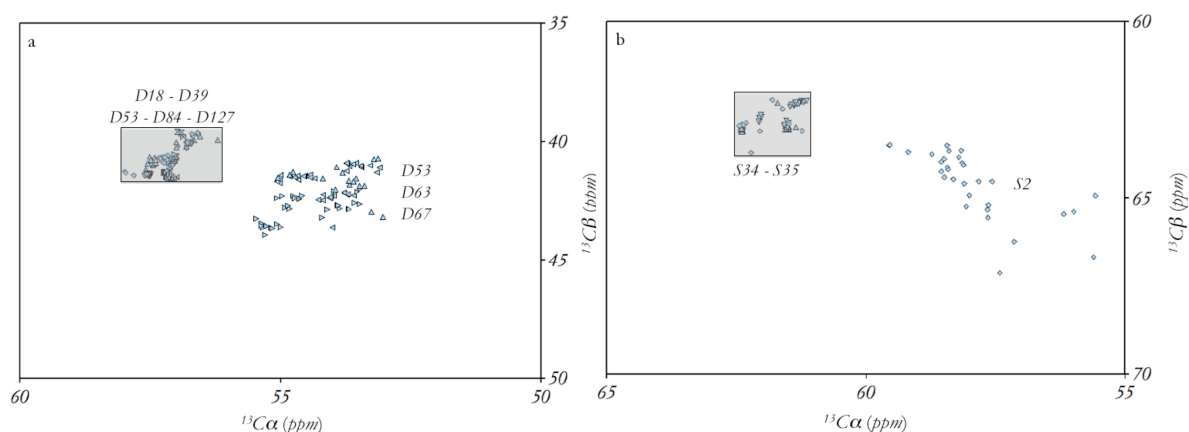


Figure 5.3 : Prédiction des déplacements chimiques sur quarante modèles de *Ec-MscL*. Représentation de la zone des corrélations $C\alpha$ - $C\beta$ des acides aminés *Asp* (a) et *Ser* (b).

Ainsi, un marquage aux isotopes ^{13}C et ^{15}N a été effectué par *SIV* sur les 7 *Asp*, 12 *Gly*, 13 *Leu* et 4 *Ser* de la protéine *Ec-MscL*. La *SIV* a été réalisée comme pour l'échantillon *FIMPR-MscL* en présence de liposomes de *DOPC*, avec dans un premier temps un ratio protéine/lipide de 1/4. Dans un second temps, l'échantillon de protéoliposomes a été solubilisé en micelles de *triton-x-100*, et la protéine a été purifiée sur colonne de *Ni-NTA*. Après purification, la protéine a été reconstituée en liposomes de *DOPC* avec un ratio protéine/lipide de 1/2. Cet échantillon a été étudié en utilisant la séquence *DARR* avec des temps de mélange de 20 à 500 ms à 258 K.

b Production de l'échantillon (*Asp, Gly, Leu, Ser*)-MscL en liposomes de DOPC

La protéine a été synthétisée *in vitro* en utilisant le même protocole que celui qui a servi à produire l'échantillon (*Ala, Lys, Val, Tyr*)- $^{13}\text{C},^{15}\text{N}$ -MscL. Pour un ratio de protéine/lipide de 1/4 et pour un rendement estimé à 10 mg de protéines, 40 mg de DOPC sont nécessaires. Après la *SIV* et la centrifugation, 9 mg de *Ec*-MscL sont obtenus. L'échantillon a été conservé à -20°C jusqu'à l'obtention des premiers spectres sur l'échantillon *FIMPR*-MscL avec le même ratio protéine/lipide, présentant une mauvaise résolution. L'échantillon a donc été solubilisé en micelles de détergent afin d'augmenter le ratio protéine/lipide. Un volume de 20 mL de tampon *A1* a été utilisé pour obtenir des micelles ternaires. La protéine a été purifiée sur une colonne d'affinité. Les fractions récupérées après purification ont été analysées par *DOT BLOT*. Les fractions contenant la protéine *Ec*-MscL ont été dialysées contre 2 L de tampon *RB-T*. Après la dialyse, la protéine est quantifiée par dosage au *BCA*, indiquant la présence de 5 mg de protéines. Pour reconstituer l'échantillon 10 mg de DOPC sont nécessaires. Les vésicules sont préparées à l'aide de l'extrudeur, puis rajoutées à l'échantillon de protéines en micelles de détergents. Le complexe ternaire est mélangé pendant 20 min, puis le détergent est retiré à l'aide des *BIO-BEADS*. Après extraction totale du détergent, les protéoliposomes sont centrifugés, et conservés à -20°C jusqu'à leur analyse au spectromètre de RMN.

c Analyse de (*Asp, Gly, Leu, Ser*)-MscL par RMN

La protéine *Ec*-MscL contenant l'étiquette histidine est composée de 147 résidus dont 7 *Asp*, 12 *Gly*, 13 *Leu*, et 4 *Ser*. Chaque molécule marquée aux ^{13}C et ^{15}N sur les *Asp*, *Gly*, *Leu*, et *Ser* contient donc 143 atomes ^{13}C et 36 atomes ^{15}N . Le poids moléculaire après le marquage isotopique est de 16,321 kDa. Les protéoliposomes contenant la protéine sélectivement marquée sont transférés dans le rotor adéquat et conservés à -20°C jusqu'aux expériences RMN. Après la mise au point des différentes valeurs des séquences d'impulsions utilisées, le rotor est placé à température ambiante. Il est ensuite inséré dans la sonde *MAS*, et la rotation à l'angle magique est augmentée par paliers (jusqu'à 11 000

$H\alpha$). La température souhaitée est fixée. Une fois ces deux paramètres stabilisés, la sonde est accordée aux fréquences des noyaux étudiés.

Une série d'expériences a été accomplie sur l'échantillon (*Asp, Gly, Leu, Ser*)- $^{13}C, ^{15}N$ -*MscL* à 248 K en utilisant la séquence d'impulsions *DARR*, avec des temps de mélange de 20 à 200 ms (Figure 5.4). L'analyse d'un pic isolé montre une résolution de 1,6 ppm pour le spectre obtenu à 50 ms de temps de mélange (Figure 5.4b). Au niveau des spectres obtenus, des pics correspondant aux corrélations des *Leu* sont présentes. Des corrélations $C'-C\alpha$ des résidus *Gly* sont également visibles sur les spectres.

La majorité des corrélations des résidus *Leu* sont présentes à 20 ms. L'intensité des pics augmente lorsque le temps de mélange s'élève jusqu'à 100 ms (Figure 5.4c), puis diminue entre 100 et 200 ms (Figure 5.4d). Aucune corrélation correspondant aux résidus *Ser* n'est présente. Les pics de faible intensité situés à proximité de celui des corrélations $C\alpha-C\beta$ des *Leu* (58-40 ppm) représentent des corrélations $C\alpha-C\beta$ de résidus *Leu* situés dans un environnement autre qu'une hélice α ou des corrélations $C\alpha-C\beta$ de résidus *Asp*. Aucune autre corrélation des résidus *Asp* n'est identifiée. Le suivi de l'intensité de certains pics isolés est représenté dans la figure 5.5.

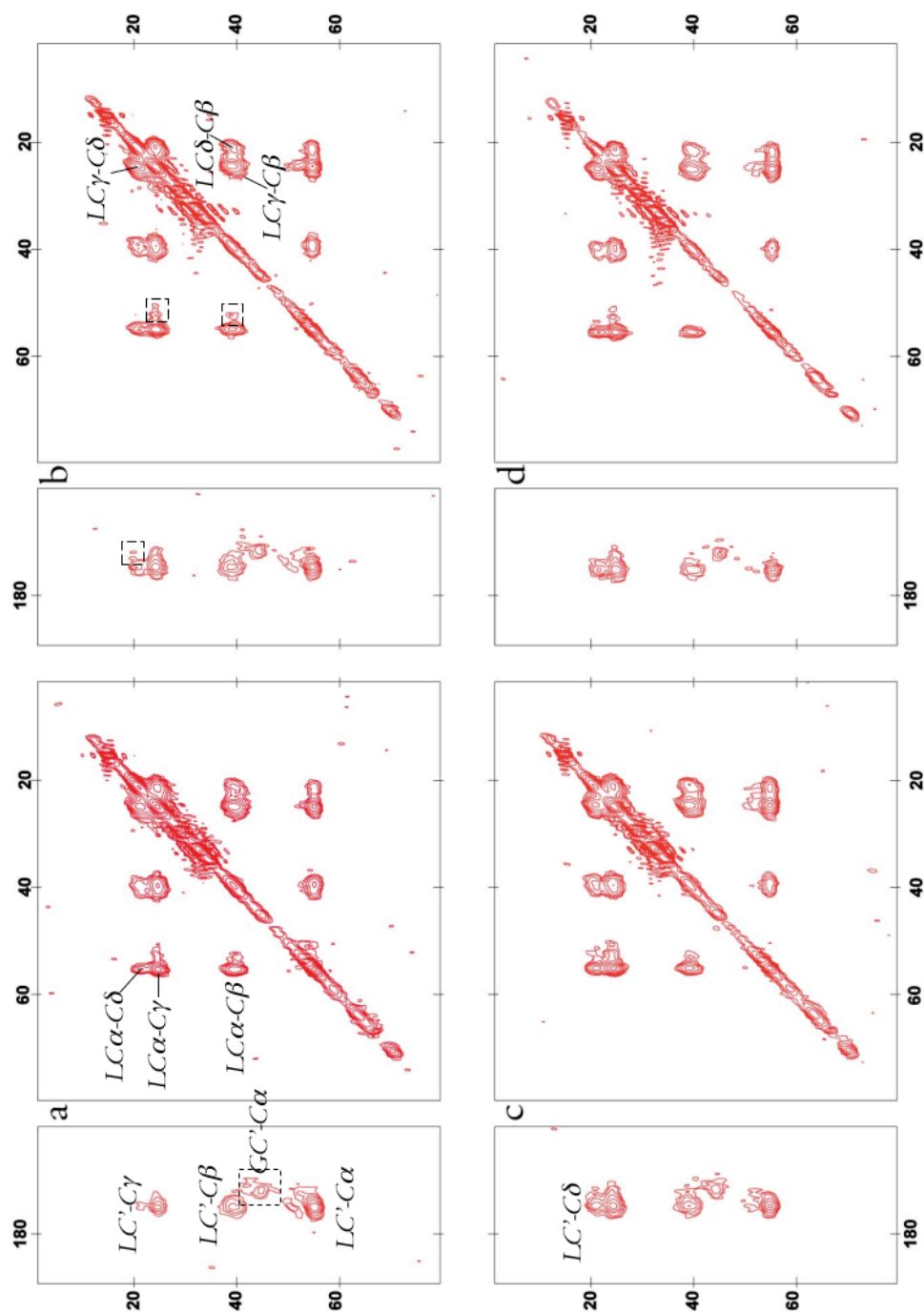


Figure 5.4 : Spectres $2D\ ^{13}C-^{13}C$ obtenus sur l'échantillon (Asp, Gly, Leu, Ser)- $^{13}C,^{15}N$ -MscL à 248 K avec différents temps de mélange. Les spectres présentés ont été obtenus en appliquant la séquence *DARR* avec des temps de mélange de 20 (a), 50 (b), 100 (c) et 200 ms (d). Aucune corrélation propre aux systèmes de spin des Asp ou Ser n'est présente. Les pics encadrés correspondent aux corrélations dont l'intensité au cours du temps de mélange est suivie.

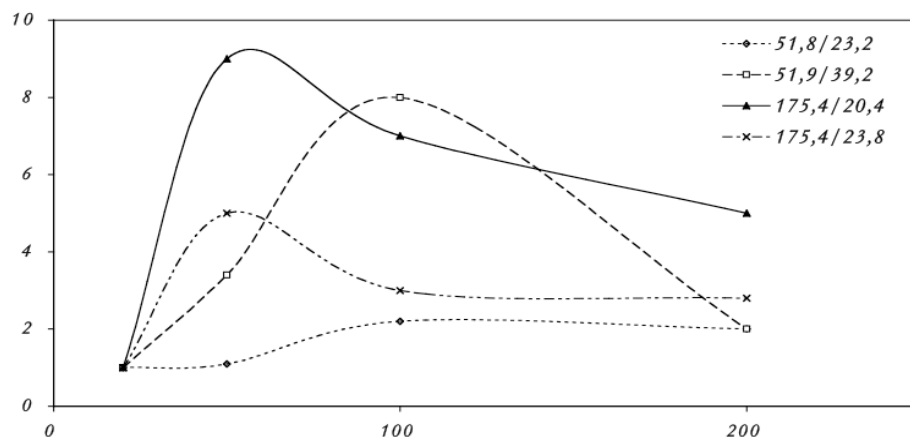


Figure 5.5 : Évolution du signal de quatre corrélations issues des différents spectres obtenus sur l'échantillon (*Asp*, *Gly*, *Leu*, *Ser*)-*Ec-MscL*. Le pic 51,8-23,2 représente une corrélation $LC\alpha-C\delta$ correspondant à une distance de 4,4,5 Å. Le pic 51,9/39,2 représente une corrélation $LC\alpha-C\beta$ ou $DC\alpha-C\beta$ correspondant à une distance de 1,5 Å. Les pics 175,5/20,7 et 175,4/23,8 représentent des corrélations $LC'-C\delta$.

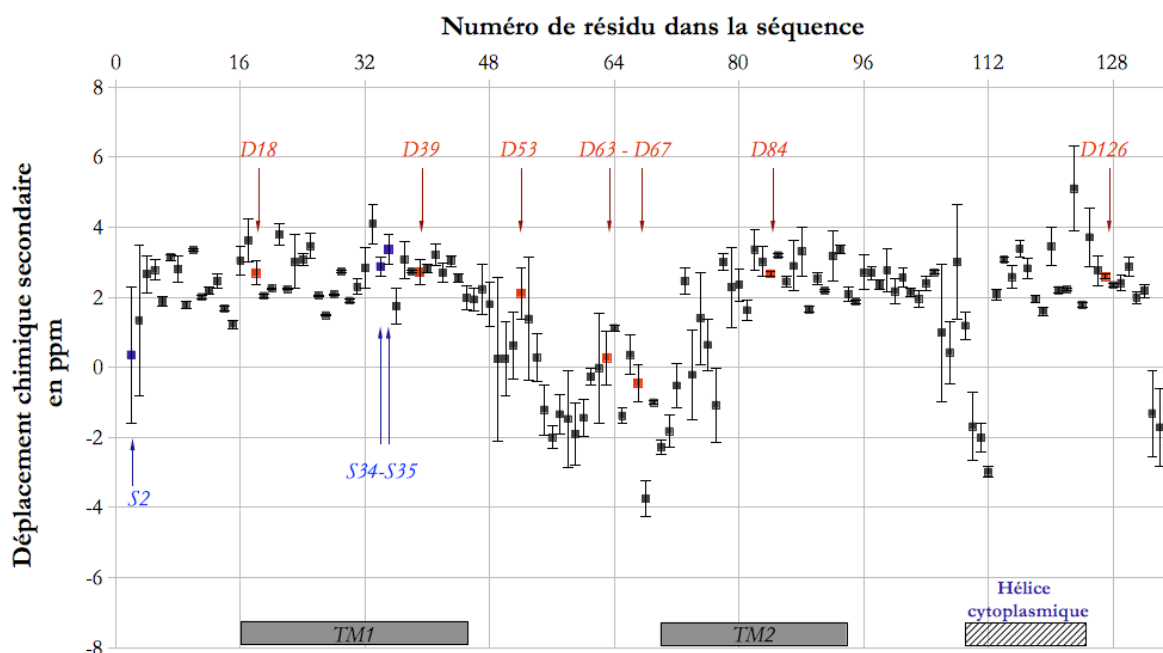


Figure 5.6 : Fluctuation des valeurs de déplacements chimiques prédites $C\alpha$ en fonction de la séquence. Les valeurs au niveau de l'axe des abscisses correspondent aux moyennes des déplacements chimiques prédits sur les quarante modèles, auxquelles sont soustraites les valeurs correspondantes de déplacement chimique du résidu en *random coil* (valeurs issues de la *BMRB*). Les écart-types à la moyenne des quarante valeurs obtenues pour chaque $C\alpha$ sont représentés. Les valeurs mises en évidence correspondent à celles des *Asp* (en rouge) et les *Ser* (en bleu). Les emplacements des hélices transmembranaires (rectangles gris) et de l'hélice cytoplasmique (rectangle hachuré) sont indiqués.

VI Rôle de l'environnement lipidique - Analyse de l'échantillon *AGKVY-Ec-MscL*

Après avoir testé différentes stratégies de marquage de la protéine *Ec-MscL*, avec différentes séquences d'impulsions, et en acquérant des spectres à différents temps de mélange et différentes températures, l'effet de la variation de l'environnement lipidique a été étudié. Pour cela, un nouvel échantillon a été synthétisé *in vitro*, avec un marquage au niveau des acides aminés *Ala*, *Gly*, *Lys*, *Val* et *Tyr*. La synthèse est effectuée en présence de micelles de détergent. L'échantillon obtenu a été séparé en quatre lots pour une reconstitution en liposomes d'asolectine (*AGKVY-aso*), de lipides extraits de la membrane interne d'*E. coli* (*AGKVY-Eco*), d'un mélange *DPPC/DOPC* (*AGKVY-DPDO*), et d'un mélange *DOPC/LPC* (*AGKVY-DOL*). Pour le premier lot, l'asolectine est utilisé car il s'agit du mélange de lipides utilisé pour les expériences de *PATCH CLAMP* (*AGKVY-aso*) ; le ratio protéine/lipide souhaité est de 1/1 afin de se rapprocher des conditions expérimentales utilisées pour la cristallisation bidimensionnelle (Saint *et al.*, 1998) pour le second lot (*AGKVY-Eco*) ; pour le troisième lot, le mélange est composé de *DPPC* à 50% et de *DOPC* à 50%, afin d'augmenter la rigidité de la bicouche ainsi formée (*AGKVY-DPDO*) ; et pour le dernier lot, l'addition de lysophospholipides permet en sens inverse, de diminuer la rigidité des liposomes, en favorisant la courbure de la membrane (Perozo *et al.*, 2002 ; *AGKVY-DOL*). Pour l'extraction du détergent, la méthyl- β -cyclodextrine *MCD* a été utilisée. Il s'agit d'un dérivé synthétique de la β -cyclodextrine, présentant une meilleure solubilité dans l'eau que celle-ci. La β -cyclodextrine est un heptamère cyclique de glucose. Chaque molécule présente une structure délimitant une cavité hydrophobe capable de piéger une molécule de détergent (Degrip *et al.*, 1998). Le protocole utilisé est celui décrit dans cette référence. L'addition des molécules de *MCD* sur le mélange ternaire protéine/lipide/détergent permet la formation de protéoliposomes. Les échantillons sont conservés à -20°C jusqu'à l'acquisition des données RMN. Celles-ci ont été réalisées à 258 K en utilisant la séquence *DARR* avec un temps de mélange de 50 ms.

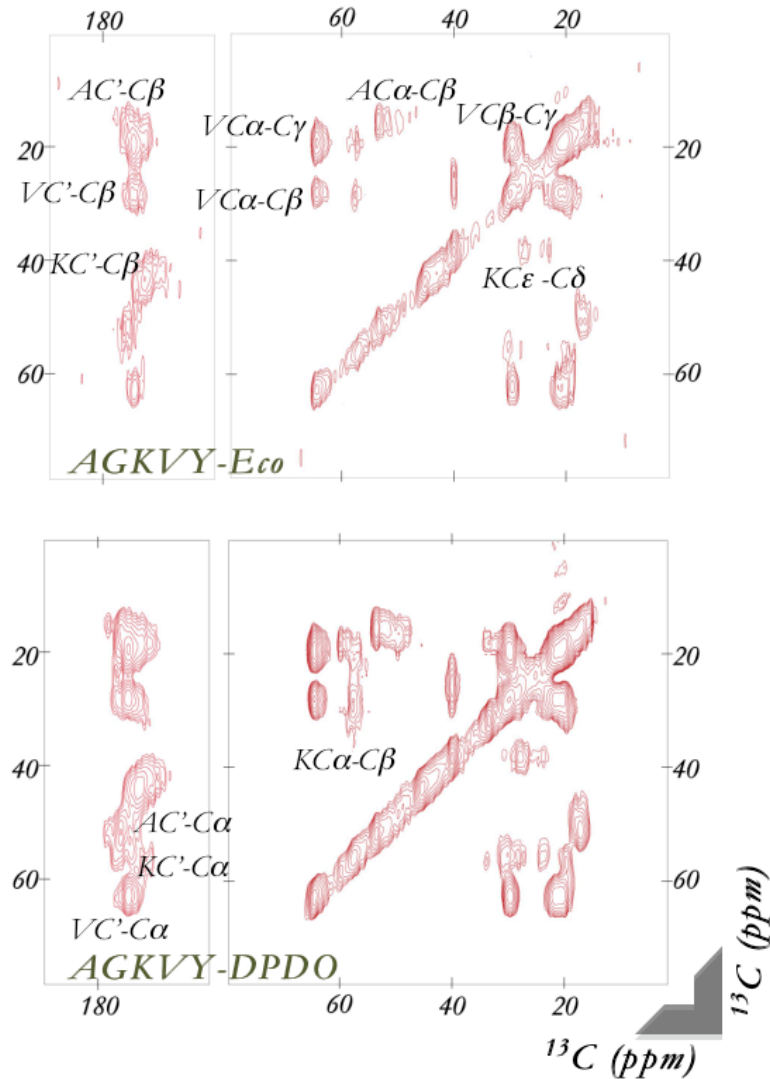
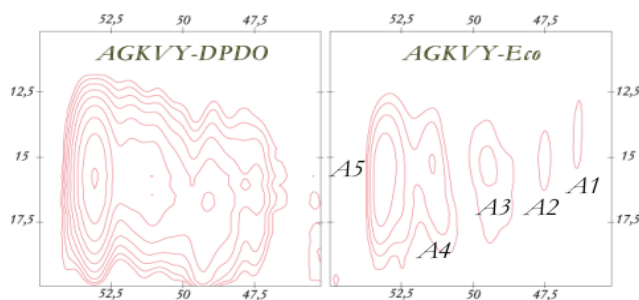


Figure 6.1 : Spectres 2D ^{13}C - ^{13}C obtenus sur l'échantillon (*Ala, Gly, Lys, Val, Tyr*)-*Ec-MscL* dans deux environnements lipidiques différents. En haut, le spectre DARR de l'échantillon *AGKVY-Eco* et en bas celui de *AGKVY-DPDO*. Les deux spectres sont obtenus avec un temps de mélange de 50 ms à une température de 258 K (14 h).

L'échantillon *AGKVY-DOL* n'était pas suffisamment solide pour être étudié, les lysophospholipides agissant comme tensioactifs déstabilisant les protéoliposomes. Une erreur s'est produite pendant la manipulation des protéoliposomes *AGKVY-aso* entraînant la perte d'une bonne partie de l'échantillon. Les spectres DARR obtenus sur les échantillons *AGKVY-Eco*, *AGKVY-aso* et *AGKVY-DPDO* montrent une résolution de 1,6 à 2,4 ppm. La résolution de 1,6 ppm est obtenue sur le spectre de l'échantillon *AGKVY-Eco*. Sur ce spectre apparaissent les corrélations des résidus *Ala* et *Val*, celles des

résidus *Lys* sont très peu intenses. La résolution est de $1,9 \text{ ppm}$ sur le spectre de l'échantillon *AGKVY-DPDO*, et toutes les corrélations présentes sur le spectre obtenu en liposomes de *DOPC* (Figure 3.4b) sont visibles également. Le spectre obtenu sur l'échantillon *AGKVY-*aso** (données non présentées) montre de faibles signaux, et une faible résolution. La comparaison des deux spectres de la [figure 6.1](#) montre des différences au niveau des corrélations des carbonyles avec les autres carbones, les corrélations des *Lys*, et surtout au niveau de la région des corrélations $C\alpha-C\beta$ des *Ala*. En effet, le spectre de l'échantillon *AGKVY-E*co** montre cinq pics distincts pour les corrélations $C\alpha-C\beta$ des *Ala* tandis que celles-ci ne forme qu'un seul large pic pour l'échantillon *AGKVY-DPDO* (Figure 6.2). La région des carbonyles montre également une meilleure résolution dans le cas de l'échantillon *AGKVY-E*co**. Le [tableau 6.1](#) contient les déplacements chimiques des carbones des cinq groupes d'*Ala* identifiés au niveau de ce spectre.

Un spectre 2D *DARR* obtenu dans les mêmes conditions sur l'échantillon précédemment étudié (*Ala*, *Lys*, *Val*, *Tyr*)-*E*c**-*MscL* synthétisé en liposomes de *DOPC* montre au niveau des corrélations $C\alpha-C\beta$ des *Ala* un seul large pic (Figure 6.4).



[Figure 6.2](#) : Agrandissement de la région des corrélations $C\alpha-C\beta$ des résidus *Ala*. Ces informations sont issues des spectres *DARR* obtenus sur les échantillons *AGKVY-DPDO* et *AGKVY-E*co** et montrent un signal plus intense lorsque la protéine est reconstituée en liposomes de *DPPC-DOPC*, mais une meilleure résolution pour une reconstitution en lipides extraits de la membrane d'*E. coli*.

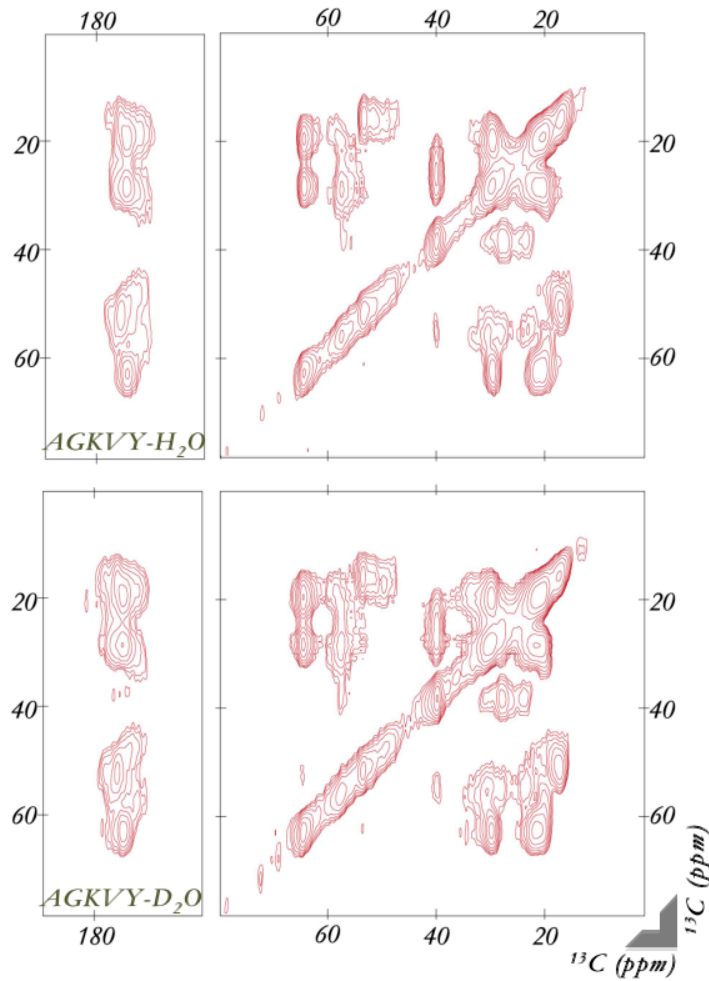


Figure 6.3 : Spectres 2D ^{13}C - ^{13}C obtenus sur l'échantillon (*Ala*, *Gly*, *Lys*, *Val*, *Tyr*)-*Ec*-*MscL* reconstitués en liposomes de *DOPC* dans de l'eau H_2O ou D_2O .

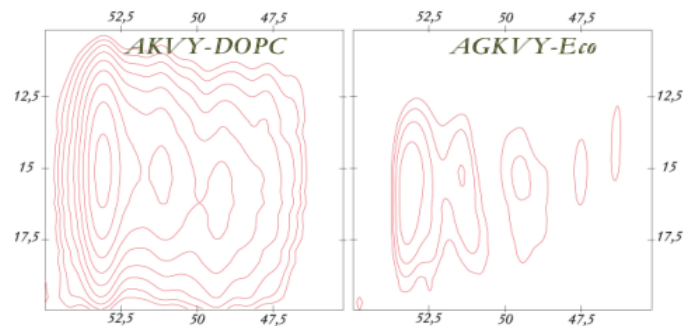


Figure 6.4 : Agrandissement de la région des corrélations $\text{C}\alpha$ - $\text{C}\beta$ des résidus *Ala*. Ces informations sont issues des spectres *DARR* obtenus sur les échantillons *AKVY-DOPC* (voir [Chapitre III](#)) et *AGKVY-Eco* et montrent un signal plus intense lorsque la protéine est reconstituée en liposomes de *DOPC*, mais une meilleure résolution pour une reconstitution en lipides extraits de la membrane d'*E. coli*.

Déplacements chimiques des acides aminés <i>Ala</i>				
Observations effectuées sur le spectre <i>DARR</i> (50 ms) sur l'échantillon <i>AGKVY-Eco</i>				
<i>aa</i>	^{13}C	<i>C'</i>	<i>Cα</i>	<i>Cβ</i>
<i>A1</i>	déplacement chimique (<i>ppm</i>)	178,1	46,2	14,1
	Erreur (<i>ppm</i>)	0,1	0,1	0,1
<i>A2</i>	déplacement chimique (<i>ppm</i>)	173,9	47,4	15,1
	Erreur (<i>ppm</i>)	0,5	0,1	0,1
<i>A3</i>	déplacement chimique (<i>ppm</i>)	174,9	49,3	15,8
	Erreur (<i>ppm</i>)	2,0	0,9	,02
<i>A4</i>	déplacement chimique (<i>ppm</i>)	174,5	51,4	15,4
	Erreur (<i>ppm</i>)	1,5	0,6	2,0
<i>A5</i>	déplacement chimique (<i>ppm</i>)	176,4	53,3	15,7
	Erreur (<i>ppm</i>)	1,0	0,3	3,0

Tableau 6.1 : Déplacements chimiques des carbones de cinq groupes d'*Ala* de la protéine *Ec-MscL*.

VII Références

- Abdine** A, Verhoeven MV, Park K-H, et al. Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* **2010**;204(1):155-159.
- Bayro** MJ, Ramachandran R, Caporini MA, Eddy MT, Griffin RG. Radio frequency-driven recoupling at high magic-angle spinning frequencies: homonuclear recoupling sans heteronuclear decoupling. *J Chem Phys.* **2008**;128(5):052321.
- Castellani** F, van Rossum B, Diehl A, et al. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature.* **2002**;420:98–102.
- Combet** C, Blanchet C, Geourjon C, Deléage G. NPS@: network protein sequence analysis. *Trends Biochem. Sci.* **2000**;25(3):147-150.
- Debret** G. Etude par modélisation moléculaire des propriétés mécaniques d'un système membranaire : le canal mécanosensible MscL au sein de bicouches lipidiques modèles. **2007**.
- Degrip** WJ, Vanoostrum J, Bovee-Geurts PH. Selective detergent-extraction from mixed detergent/lipid/protein micelles, using cyclodextrin inclusion compounds: a novel generic approach for the preparation of proteoliposomes. *Biochem J.* **1998**;330(Pt 2):667-674.
- Duma** L, Abergel D, Ferrage F, et al. Broadband dipolar recoupling for magnetization transfer in solid-state NMR correlation spectroscopy. *Chemphyschem.* **2008**;9(8):1104-1106.
- Edgar** RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics.* **2004**;5:113.
- Forbes** J, Husted C, Oldfield E. High-field, high-resolution proton « magic-angle » sample-spinning nuclear magnetic resonance spectroscopic studies of gel and liquid crystalline lipid bilayers and the effects of cholesterol. *J. Am. Chem. Soc.* **1988**;110(4):1059-1065.
- Griffiths** JM, Griffin RG. Nuclear magnetic resonance methods for measuring dipolar couplings in rotating solids. *Anal. Chim. Acta.* **1993**;283(3):1081-1101.
- Halladay** HN, Stark RE, Ali S, Bittman R. Magic-angle spinning NMR studies of molecular organization in multibilayers formed by 1-octadecanoyl-2-decanoyl-sn-glycero-3-phosphocholine. *Biophys. J.* **1990**;58(6):1449-1461.
- Khafizov** K, Staritzbichler R, Stamm M, Forrest LR. A study of the evolution of inverted-topology repeats from LeuT-fold transporters using AlignMe. *Biochemistry.* **2010**;49(50):10702-10713.
- King** G, Dixon AM. Evidence for role of transmembrane helix-helix interactions in the assembly of the Class II major histocompatibility complex. *Mol Biosyst.* **2010**;6(9):1650-1661.
- Larkin** MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* **2007**;23(21):2947-2948.
- Ozawa** K, Wu PSC, Dixon NE, Otting G. N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *FEBS J.* **2006**;273(18):4154-4159.
- Perozo** E, Cortes DM, Somporapisut P, Kloda A, Martinac B. Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature.* **2002**;418(6901):942-948.
- Poirot** O, O'Toole E, Notredame C. Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res.* **2003**;31(13):3503-3506.
- Roux** M, Neumann J-M, Bloom M, Devaux PF. 2H and 31P NMR study of pentyllysine interaction with headgroup

deuterated phosphatidylcholine and phosphatidylserine. *Eur Biophys J.* **1988**;16(5).

Russ WP, Engelman DM. The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **2000**;296(3):911-919.

Saint N, Lacapère J-J, Gu LQ, et al. A hexameric transmembrane pore revealed by two-dimensional crystallization of the large mechanosensitive ion channel (MscL) of *Escherichia coli*. *J. Biol. Chem.* **1998**;273(24):14667-14670.

Shen Y, Bax A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR.* **2007**;38:289–302.

Shen Y, Bax A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR.* **2010**;48(1):13-22.

Shen Y, Bax A. SPARTA Protein Backbone Chemical Shifts Prediction Program. Available at: <http://spin.niddk.nih.gov/bax/software/SPARTA/index.html>

Sobhanifar S, Reckel S, Junge F, et al. Cell-free expression and stable isotope labelling strategies for membrane proteins. *J. Biomol. NMR.* **2010**;46(1):33-43.

Steinbacher S, Bass R, Strop P, Rees DC. Structures of the Prokaryotic Mechanosensitive Channels MscL and MscS. Dans: *Mechanosensitive Ion Channels, Part A*. Vol Volume 58. Academic Press; **2007**:1-24.

Sukharev SI, Durell SR, Guy HR. Structural models of the MscL gating mechanism. *Biophys. J.* **2001**;81(2):917-936.

Szeverenyi NM, Sullivan MJ, Maciel GE. Observation of spin exchange by two-dimensional fourier transform ¹³C cross polarization-magic-angle spinning. *J. Magn. Reson.* **1982**;47(3):462-475.

Valadié H. Etude des propriétés dynamiques du canal mécanosensible MscL par des approches théoriques. 2003.

Valadié H, Lacapère J-J, Sanejouand Y-H, Etchebest C. Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis. *J. Mol. Biol.* **2003**;332(3):657-674.

Warschawski DE. Membranes biologiques, membranes modèles: Nouvelles approches utilisant la RMN solide en rotation à l'angle magique. **1995**.

Word JM, Lovell SC, Richardson JS, Richardson D. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **1999**;285(4):1735-1747.

Zhou Z, Sayer BG, Hughes DW, Stark RE, Eppand RM. Studies of Phospholipid Hydration by High-Resolution Magic-Angle Spinning Nuclear Magnetic Resonance. *Biophys. J.* **1999**;76(1):387-399.

Discussion

D Discussion

À la lumière des résultats présentés dans la section précédente et obtenus par l'application de différents protocoles et stratégies de marquage, les hypothèses de travail sont discutées dans ce chapitre afin de permettre une analyse plus profonde. Chaque point abordé dans le chapitre précédent est expliqué, puis une discussion générale s'ensuit, abordant l'intérêt global de la démarche suivie.

La plupart des expériences *RMN* sont réalisées à 248 K. Cette valeur est une estimation basée sur la transition de phase de la membrane de *DOPC*. L'écart observé entre la valeur affichée au niveau de la console de contrôle et la température à l'intérieur du rotor est provoqué surtout par les frottements dus à la rotation, et à une moindre échelle par les impulsions radiofréquence ainsi que les différents flux de gaz contrôlant la rotation à l'angle magique. La [figure 1](#) montrant trois spectres 1H de la *DOPC* à des températures différentes, présente clairement une phase fluide en 1a et une phase gel en 1c. La mesure des largeurs de raie de quelques signaux et son évolution en fonction de la température a permis de situer la température de transition. Celle-ci est de 253 K. La transition de phase est située ici autour de 238-243 K (point d'inflexion observé sur la [figure 2](#)). La valeur de température retenue pour l'acquisition des spectres est donc de 233 K au niveau de la console, ce qui correspond à 243-248 K dans le rotor. Les différentes expériences montrent qu'il s'agit d'un bon compromis entre une résolution suffisamment bonne au niveau des spectres de corrélations ^{13}C - ^{13}C et une stabilité des échantillons permettant leur utilisation à plusieurs reprises. La résolution était bonne pour les corrélations des résidus présents dans les régions très structurées de la protéine comme les hélices. Lors de l'acquisition des spectres sur les différents échantillons de *Ec-MscL*, les carbones des résidus présents dans les parties mobiles de la protéine présentaient peu ou pas de signaux. Ceci provient très probablement des mouvements de ces résidus. En effet, à faible température, la membrane est figée et les hélices adoptent une conformation stable. Les résidus sont animés d'un mouvement très lent, ce qui est traduit par des

signaux intenses. Le fait que la protéine soit figée, entraîne un ralentissement des mouvements des parties extra-membranaires. Un mouvement lent entraîne la présence de différentes conformations équiprobables, ce qui se traduit par la disparition des signaux des résidus correspondants. De plus, la polarisation croisée est plus efficace dans les parties rigides. De ce fait, l'augmentation de la température pourrait améliorer la résolution des résidus des boucles, en augmentant leur mouvement, ce qui pourrait se traduire par la présence de signaux sur les spectres. Les tests de température devraient être suivis quasi-systématiquement de tests fonctionnels pour voir si la protéine n'a pas été affectée, et une chromatographie sur couche mince des lipides, pour détecter une éventuelle dégradation.

Les expériences de corrélations ^{13}C - ^{13}C ont été réalisées en utilisant la séquence *DARR*. D'autres méthodes de recouplage dipolaire ont été testées, les différences n'étant pas significatives, alors la méthode *DARR* a été maintenue. La méthode *PDSD* a été largement employée pour obtenir des contraintes dans des systèmes biologiques microcristallins (Castellani *et al.*, 2002 ; Manolikas *et al.*, 2008). Au niveau de cette séquence, le transfert d'aimantation a lieu sans condition particulière, ce qui provoque un transfert de polarisation diffus et homogène. L'application de cette séquence sur des protéines synthétisées en présence de glycérol marqué en C^2 ou en C^1 et C^3 est plus favorable que pour des systèmes uniformément marqués, car le marquage se charge de diminuer les interactions multi-spins intenses (chaque acide aminé présente plusieurs motifs de marquage aléatoire). Dans le cas d'un marquage uniforme, la séquence *PDSD* pourrait être appliquée avec de très courts temps de mélange pour obtenir les corrélations entre carbones séparés par une seule liaison (ceci est valable également pour le marquage spécifique par *STV*, car tous les carbones d'un même acide aminé sont marqués). Le transfert de polarisation non contrôlé provoque la propagation et dissipation de l'énergie de l'aimantation, présentant de multiples signaux peu intenses pour de plus grandes distances (Szeverenyi *et al.*, 1982 ; Grommek *et al.*, 2006). La séquence *DARR* utilise un champ radiofréquence continu au niveau des protons ayant la même fréquence que la vitesse de rotation, en ce plaçant donc dans les conditions résonantes de recouplage, ce qui permet d'accélérer la diffusion de spin, et donc le transfert d'aimantation. Cette

séquence a prouvé son efficacité pour la détermination de structures de protéines (Bertini *et al.*, 2010 ; Franks *et al.*, 2008 ; Loquet *et al.*, 2008). La séquence B^2R^3 a été développée pour étendre l'application de la résonance rotationnelle pour des fréquences de rotation supérieures à 25 kHz (Duma *et al.*, 2008). Ceci est la raison pour laquelle aucune différence n'est perçue à 11 kHz (fréquence de rotation appliquée au cours de cette étude).

D'autres séquences pourraient être appliquées : la méthode *PARIS* (pour *Phase-Alternated Recoupling Irradiation Scheme*) développée par P. TEKELY à l'école normale supérieure (Weingarth *et al.*, 2009) permet de réduire le transfert d'aimantation à longue distance pour des temps de mélange courts. La séquence *CHHC* (Lange *et al.*, 2002) permet d'obtenir des contraintes entre carbones liés à un proton, ce qui pourrait réduire le signal au niveau des spectres bidimensionnels.

Les expériences de corrélations ^{15}N - ^{13}C ont été réalisées en utilisant la méthode *specific-CP* (*Spectrally Induced Filtering In Combination with Cross Polarization*). Les résultats obtenus sont assez décevants. La résolution au niveau de la dimension des azotes est très faible. Les spectres *specific-CP* ont tous été enregistrés à des basses températures. Des tests à des températures plus élevées sont envisageables. L'application d'autres séquences est également envisageable : la séquence *CANCO* (Li *et al.*, 2007) et *NHHC* (Lange *et al.*, 2002) permettraient l'exploitation du marquage à l'azote ^{15}N , pour aider dans la détermination des contraintes structurales hétéronucléaires.

I Analyse de l'échantillon uniformément marqué

La protéine *Ec-MscL* avec l'étiquette histidine comporte 147 acides aminés, son poids moléculaire est de 15 kDa. Malgré l'absence d'une limite de taille pour l'étude d'une protéine par RMN en phase solide, la technique ne s'est pas encore attaquée à des molécules de poids moléculaire supérieur à 11 kDa (Cady *et al.*, 2009), sauf lorsque les molécules présentent un arrangement fibrillaire ou cristallin. Le défi est ici l'étude d'une protéine membranaire uniformément marquée dans des conditions proches de l'environnement natif. La séquence codant pour la protéine a été clonée dans un plasmide *pIVEX* et la surexpression est effectuée dans la souche *BL21 (DE3)* d'*E. coli* dans un milieu minimum où les seules sources de carbone et d'azote sont respectivement du glucose marqué au ^{13}C et du chlorure d'ammonium marqué au ^{15}N . Après purification, la protéine est reconstituée en liposomes de *DOPC*, puis analysée par RMN.

a Production et purification de *U-MscL*

Le vecteur contenant la séquence codante pour la protéine *MscL* d'*E. coli* est le plasmide de type *pIVEX*. Ce plasmide a été initialement développé pour la synthèse *in vitro* (*pIVEX* pour *plasmid for In Vitro Expression*) mais a été depuis adapté à la surexpression bactérienne (Rogé & Betton, 2005). Cependant, au cours de la croissance bactérienne un suivi de l'expression de la protéine d'intérêt a été effectué sur plusieurs échantillons avant et après l'induction à l'*IPTG*. Après migration des échantillons sur gel de *SDS-PAGE* et transfert des protéines sur membrane de nitrocellulose, la révélation par la méthode *WESTERN BLOT* témoigne d'une expression de la protéine avant l'induction (données non présentées). Ceci révèle une fuite au niveau du promoteur *T7* du vecteur d'expression. Les taux de synthèse sont néanmoins négligeables par rapport au taux obtenu après 3 heures d'induction. Cependant, pour limiter la synthèse de la protéine d'intérêt avant l'induction, la souche *C43* pourrait réduire la fuite du promoteur (Miroux & Walker, 1996). La sélection de nouveaux mutants adaptés à la surexpression en milieu minimum pourrait être envisagée également, à partir de la souche *C43*.

Deux échantillons de protéine *Ec-MscL* uniformément marquée ont été analysés par spectroscopie de masse pour évaluer le taux d'incorporation des isotopes ^{13}C et ^{15}N . Le pic le plus intense montre des valeurs de $16\,886,37$ et $16\,893,76$ *kDa*. Cela signifie un taux d'incorporation de 80% si aucun clivage en N- ou C-terminal ne s'est effectué, ou un taux d'incorporation plus élevé si un clivage d'un ou plusieurs acides aminés a eu lieu.

b Reconstitution de *U-MscL* en liposomes de *DOPC*

Après la purification et la quantification, la protéine *Ec-MscL* est reconstituée en liposomes et les protéoliposomes ainsi obtenus sont purifiés sur gradient de saccharose afin de séparer les protéoliposomes des liposomes vides et des agrégats de protéines. À partir de 3 L de *M9*, et après purification et dialyse, le dosage indique 3 mg de *Ec-MscL*. L'échantillon obtenu est pur, comme en témoigne le profil de migration sur *SDS-PAGE* (Figure 1.2). La méthode de *cross-linking* de la protéine en détergent, ainsi qu'après la reconstitution en liposomes montre une organisation pentamérique. Les tests de caractérisation montrent que l'échantillon final contient 66% d'eau, 23% de lipides et 11% de protéines.

c Analyse de *U-MscL* par RMN

Des spectres $1D$ ^{13}C ont été enregistrés au début et à la fin de chaque série d'expériences à 243 - 248 K. Ces spectres permettent d'estimer la qualité de l'échantillon. Les signaux des lipides sont plus facilement détectables sur le spectre $1D$ que sur les spectres $2D$. Différents spectres $2D$ ^{13}C - ^{13}C -*DARR* ont été réalisés sur cet échantillon. Des temps de mélange de $1,5$ à 500 ms ont été testés. Aucune information significative n'est présente sur le spectre obtenu avec un temps de mélange de $1,5$ ms (Figure 1.3). À partir de 10 ms, le signal augmente, et des corrélations ^{13}C - ^{13}C apparaissent (Figure 1.4). Les systèmes de spin de quelques acides aminés sont identifiables, mais aucun spectre ne permet d'attribuer les déplacements chimiques à un résidu en particulier. La présence de quelques signaux isolés a permis d'évaluer les intensités des pics en fonction du temps de mélange. Les pics représentent des corrélations entre carbones séparés par une ou deux

liaisons. La [figure 1.5](#) montre que le signal baisse à partir de 20-30 ms dans le cas où les noyaux sont séparés par une seule liaison. Le point d'inflexion atteint les 50 ms dans le cas des noyaux séparés par deux liaisons. Ceci est probablement dû à la quantité de spins ^{13}C présents dans le système, ce qui provoque la dissipation du signal. Les spectres obtenus à des temps de mélange supérieurs à 100 ms en témoignent également. En effet, le signal pour un temps de mélange de 100 ms est inférieur à celui obtenu à 10 ms. Et sur un spectre enregistré avec un temps de mélange de 500 ms, le signal disparaît complètement. Les largeurs de raie des pics isolés prouvent que la protéine est correctement repliée. Elle est équivalente à des résolutions précédemment atteintes dans d'autres laboratoires, mais les pics sont bien plus larges que ceux obtenus sur un système microcristallin (Etzkorn *et al.*, 2007 ; Frericks *et al.*, 2006 ; Sperling *et al.*, 2010). Les protéoliposomes ne présentant pas d'ordre particulier, comparés aux systèmes fibrillaires ou microcristallins en général, ce résultat est donc logique, et ne peut probablement pas être amélioré.

Changer la séquence d'impulsion n'apporte pas d'informations supplémentaires ([Figure 1.6](#)). La résolution ne varie que très peu. En augmentant le temps d'acquisition de 14 à 56 h, des signaux supplémentaires apparaissent au niveau du spectre ([Figure 1.7](#)). Les corrélations $C\alpha-C\beta$ des résidus *Thr* et *Ser* minoritaires dans la protéines, sont visibles. Mais le recouvrement des pics empêche les attributions des déplacements chimiques. De plus, étant donné que le signal s'estompe rapidement, aucune information structurale ne peut être obtenue. Le spectre enregistré à une température supérieure (263-268 K) montre une meilleure résolution, notamment au niveau des corrélations des résidus hydrophiles. Cependant, le spectre $1D$ ^{13}C obtenu après cette expérience a montré une forte dégradation au niveau des lipides.

Les spectres de corrélations homonucléaires ne permettent pas d'attribuer des déplacements chimiques ou d'identifier les résidus. L'échantillon étudié est issu d'un gradient de saccharose, et représente donc des protéines membranaires correctement insérées dans les liposomes, comme en témoigne la largeur de raie de certains pics isolés. La différence de comportement des résidus hydrophobes et hydrophiles par rapport à la température d'acquisition des spectres paraît comme la conséquence du comportement de

la bicouche lipidique. En effet, les résidus hydrophobes, majoritaires dans la bicouche, présentent des signaux forts lorsque la bicouche est en phase gel (243-248 K), et voient leurs signaux diminuer lorsque la bicouche est en phase fluide (263-268 K). À l'opposé, les résidus hydrophiles, présents dans les régions cyto- et péri-plasmiques de la protéine, montrent des corrélations de faible intensité lorsque la bicouche est en phase gel. L'intensité augmente considérablement lorsque la bicouche est en phase fluide. Cela pourrait traduire un degré d'ordre acquis par ces résidus au cours du passage à des températures supérieures à la température de transition des lipides.

II Analyse de l'échantillon sélectivement marqué *IT-Ec-MscL* - Apport de la *SIV* pour l'étude des protéines membranaires par *RMN* en phase solide

Le marquage isotopique uniforme sur la protéine membranaire *MscL* montre une résolution pouvant atteindre $0,6 \text{ ppm}$ pour certains pics isolés, du même ordre que les valeurs atteintes sur des systèmes membranaires ($0,5-0,55 \text{ ppm}$ pour Varga *et al.*, 2008 ; $0,5-0,6 \text{ ppm}$ pour Shi *et al.*, 2009) ou même microcristallins ($0,5-1 \text{ ppm}$ pour Loquet *et al.*, 2008). Malgré cela, l'abondance du signal et le recouvrement des corrélations empêche l'obtention d'informations relatives à la structure de la protéine. Afin de diminuer le nombre de noyaux marqués, la technique de synthèse *in vitro* a été appliquée. Le plasmide *pIVEX* optimisé pour la *SIV* contient le gène *mscL*. Pour tester la *SIV* dans le but d'obtenir des informations structurales en *RMN* en phase solide sur la protéine membranaire *Ec-MscL*, un marquage sélectif a été effectué. Ceci est réalisé en introduisant dans le système des acides aminés marqués. Il s'agit dans ce cas des *Ile* et des *Thr*, les *Ile* étant les acides aminés les plus représentés, et les *Thr* parmi les moins présents. Il s'agit, pour les *Ile*, d'un témoin de résolution, et d'un témoin de sensibilité pour les *Thr*. La *SIV* est réalisée en présence de micelles de *triton-x-100*. Après purification, la protéine est reconstituée en liposomes de *DOPC*, puis analysée par *RMN*.

a Production et purification de (*Ile*, *Thr*)-*MscL*

Le protocole de synthèse *in vitro* de la protéine membranaire *MscL* en micelles de *triton-x-100* a été mis en place dans le laboratoire d'*A. GHAZI* (Berrier *et al.*, 2004), et a été transposé pour réaliser la synthèse dans les kits de 10 mL . À la fin de la synthèse, l'analyse du surnageant et du culot montre que la protéine *Ec-MscL* est majoritairement présente sous forme soluble, associée donc aux micelles de détergent.

b Reconstitution de (*Ile*, *Thr*)-*MscL* en liposomes de *DOPC*

Après la purification et la quantification, la protéine *Ec-MscL* sélectivement marquée

est reconstituée en liposomes de *DOPC*. À partir d'une seule synthèse, 9,6 mg de protéines sont obtenus dans 10 mL. La migration sur gel montre un échantillon pur, et son organisation est pentamérique comme en témoigne le profil sur gel *SDS-PAGE* (Figure 2.1).

c Analyse de (Ile, Thr)-MscL par RMN

Des spectres $1D\ ^{13}C$ ont été enregistrés au début et à la fin de chaque série d'expériences à 243-248 K. Ces spectres permettent d'estimer la qualité de l'échantillon. Différents spectres $2D\ ^{13}C\text{-}^{13}C\text{-DARR}$ ont été réalisés sur cet échantillon. Des temps de mélange de 5 à 500 ms ont été testés. Le spectre $2D\ ^{13}C\text{-}^{13}C\text{-DARR}$ réalisé à 5 ms de temps de mélange montre plusieurs pics (Figure 2.2). Les corrélations entre noyaux séparés par une seule liaison sont toutes présentes. La présence de corrélations entre carbones des *Ile* séparés par plusieurs liaisons ($C'\text{-}C\beta$, $C'\text{-}C\gamma 2$ et $C\alpha\text{-}C\gamma 2$) montre la présence de deux ou plusieurs *Ile* extrêmement proches dans l'espace. L'analyse des largeurs de raies ^{13}C obtenues au niveau des corrélations des *Thr* montre une résolution supérieure à 1 ppm, mais il s'agit surtout d'un artefact de la faiblesse du signal.

Pour des temps de mélange plus élevés, les corrélations entre noyaux séparés par une seule liaison s'intensifient et de nouvelles corrélations apparaissent (Figure 2.3). Les spectres obtenus montrent une réduction importante de l'encombrement dans toutes les régions du spectre. La résolution atteint 0,6 ppm pour certains pics isolés. Les corrélations présentes au niveau des spectres ont permis d'identifier trois systèmes de spin correspondant aux trois *Thr* présentes dans la protéine. En analysant les déplacements chimiques des $C\alpha$ et $C\beta$ et leur écart par rapport aux valeurs des déplacements chimiques théorique d'une *Thr* en *RANDOM COIL* (62,1 ppm pour le $C\alpha$, et 69,6 ppm pour le $C\beta$), lorsque le déplacement chimique du $C\alpha$ est supérieur à la valeur *RANDOM COIL*, et celui du $C\beta$ est inférieur à la valeur *RANDOM COIL*, le résidu se situe dans une hélice α . Lorsque la situation est inversée, le résidu est dans un feuillet β (Wishart *et al.*, 2001). Le résidu présentant une corrélation $C\alpha\text{-}C\beta$ au niveau de la diagonale (*Tc*) rentre dans le premier cas de figure. Les deux autres *Thr* présentent des déplacements chimiques $C\alpha/C\beta$ de 60,5/65,5 (*Tb*) et 57,5/68,1 (*Ta*). Il s'agit de valeurs correspondants à un repliement en *RANDOM COIL*. Les

trois *Thr* présentes dans la protéine sont : le résidu *T60* situé dans la boucle périplasmique, le résidu *T116* situé au milieu de l'hélice cytoplasmique, et le résidu *T123* situé du côté C-terminal de cette hélice. Le résidu présentant des valeurs de déplacements chimiques lui attribuant une conformation en hélice α est probablement le *T116* (*Tc*). Pour les deux autres résidus, les valeurs de déplacements chimiques ne permettent pas de faire une distinction. Or, le résidu *T126* situé au niveau C-terminal de l'hélice cytoplasmique est séparé d'une *Ile* par un seul résidu, alors que le résidu *T60* situé dans la boucle périplasmique est séparé par huit résidus d'une *Ile* en amont, et par sept résidus d'une autre *Ile* en aval. L'analyse des spectres obtenus avec des temps de mélange supérieurs à *200 ms* montre des corrélations inter-résiduelles entre les carbones d'une *Ile* et le $C\beta$ de la *Thr* résonnant à *65,7 ppm* (Figure 2.9). Cela suggérerait que la *Thr* présentant des déplacements chimiques $C\alpha/C\beta$ de *60,7/65,7* (*Tb*) serait le résidu *T123*.

Pour les *Ile*, des corrélations entre les C' et les carbones $C\alpha$, $C\beta$, $C\gamma1$ et $C\gamma2$ sont très intenses. Les déplacements chimiques des $C\alpha$ et $C\beta$ sont de *63,2* et *34,9 ppm* respectivement. Cela signifie que les corrélations présentes sur le spectre sont celles d'une ou plusieurs *Ile* situées dans des régions de la protéine où la structure secondaire est en hélices α . Or il existe au niveau de la séquence deux *Ile* au niveau de la partie cytoplasmique N-terminale, trois *Ile* dans la boucle périplasmique, et trois *Ile* dans la partie cytoplasmique C-terminale hors l'hélice cytoplasmique, donc huit *Ile* présentes dans des régions en *RANDOM COIL*. Or très peu d'*Ile* en conformation *RANDOM COIL* sont identifiées sur les spectres (typiquement, les déplacements chimiques d'une *Ile* en conformation *RANDOM COIL* sont de *61,5 ppm* pour le $C\alpha$ et *38,5 ppm* pour le $C\beta$, d'après Berjanskii & Wishart, 2008 et Wishart & Nip, 1998). L'absence des corrélations de ces *Ile* pourrait provenir d'un phénomène lié aux conditions expérimentales qui font que dans les vésicules de *DOPC* à *248 K*, les résidus *Ile* des hélices transmembranaires sont dans une matrice rigide, et leur immobilisation accroît l'intensité des corrélations, tandis que les résidus situés dans les parties extramembranaires sont figés dans différentes configurations à la basse température. Dans ce cas, l'augmentation de la température expérimentale pour l'obtention des corrélations homonucléaires pourrait apporter plus d'homogénéité aux mouvements des boucles, aidant à obtenir plus d'informations.

La différence peut s'expliquer également par les mouvements des résidus ; en effet, les *Ile* situées dans la membrane peuvent être animés par des mouvements lents, de même pour les deux *Thr* situées dans l'hélice cytoplasmique. Il en résulte des corrélations intenses. La troisième *Thr* située dans la boucle périplasmique (*T60*) serait animée d'un mouvement rapide, donnant lieu à des corrélations intenses également. Les autres *Ile* seraient animées d'un mouvement intermédiaire, impliquant des corrélations moyennes non visibles sur le spectre. Dans ce cas, contrairement à l'hypothèse précédente, la diminution de la température pourrait ralentir les mouvements des boucles, faisant passer les mouvements des résidus du régime intermédiaire à un régime lent caractérisé par l'apparition de nouvelles corrélations (Helmus *et al.*, 2010).

Le tracé de l'évolution du signal de certaines corrélations ([Figure 2.4](#) et [Figure 2.5](#)) montre une relaxation plus lente que dans le cas où tous les carbones de la protéines sont marqués. En effet, la diminution commence entre 50 et 100 ms de temps de mélange dans le cas de carbones séparés par une seule liaison. Elle est décalée jusqu'à 100 à 200 ms pour les carbones séparés par deux ou trois liaisons. Cet effet est dû à la réduction du nombre de spins dans le système. Dans l'échantillon uniformément marqué, l'aimantation se dissipe facilement entre les spins proches. Le marquage sélectif réduit énormément le nombre de spins voisins, et augmente ainsi le temps de relaxation du système. La [figure 2.4](#) montre l'évolution des corrélations des carbones séparés par une liaison, d'une distance de 1,5 Å. L'intensité maximale est observée pour un temps de mélange de 50 ms. Pour des distances de 3 Å, l'intensité maximale est atteinte pour 100 ms de temps de mélange. Celle-ci s'élève à 200 ms pour des distances supérieures à 3 Å.

À partir des différents spectres 2D ^{13}C - ^{13}C obtenus, les déplacements chimiques identifiés ont été répertoriés. La [figure 2.7](#) montre les systèmes de spin des trois *Thr* de la protéine. Celles-ci ont été attribuées : *Ta* représenterait la *Thr 60* située au niveau de la boucle périplasmique, *Tb* la *Thr 123* de l'interface entre l'hélice cytoplasmique et la partie C-terminale en *RANDOM COIL* probablement et *Tc* représenterait la *Thr 116* située au milieu de l'hélice cytoplasmique. Pour les *Ile*, malgré l'identification de huit différents systèmes de spin, aucune information ne permet de les identifier ou les attribuer.

d Comparaison des résultats obtenus sur (Ile, Thr)-MscL par rapport à ceux obtenus sur U-MscL

La [figure 2.10](#) montrant deux spectres 2D obtenus sur les échantillon U-MscL et IT-MscL illustre bien la nécessité du marquage sélectif dans le cas de la protéine membranaire mécanosensible MscL (voir aussi les figures ci-dessous représentant certaines régions de la [figure 2.10](#)). En effet, le spectre obtenu sur l'échantillon uniformément marqué présente une bonne résolution, mais l'imbrication des corrélations empêche l'identification des systèmes de spin, étape primordiale dans l'attribution spectrale. En réduisant le nombre de spins via un marquage sélectif de deux ou plusieurs familles d'acides aminés, la résolution est améliorée, et ceci a permis d'identifier les systèmes de spin des trois Thr présentes au niveau de la protéine. La réduction du nombre de spin a permis également d'augmenter le temps de relaxation, ce qui permet d'obtenir des spectres avec des temps de mélange pouvant atteindre quelques centaines de *ms* afin d'obtenir des corrélations inter-résiduelles. L'article décrivant la mise en place de ce projet et les résultats obtenus sur ces deux échantillons est présenté en annexe.

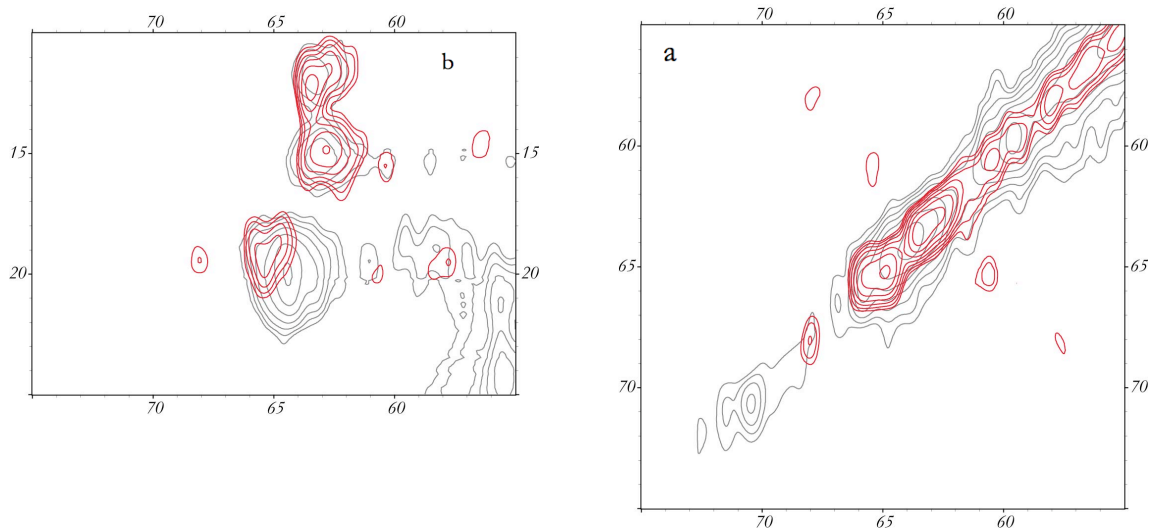


Figure 2.11 : Comparaison des spectres uniformément et sélectivement marqués. Agrandissement de la région des corrélations $C\alpha-C\beta$ (a) où les pics des Thr apparaissent lorsque la protéine est sélectivement marquée (en rouge) alors qu'elles sont absentes du spectre pour un échantillon uniformément marqué (temps d'acquisition dans les deux cas : 14 heures). Ceci est observé également pour les corrélations $C\alpha-C\gamma$ et $C\beta-C\gamma$ (b). Les corrélations des Thr apparaissent au niveau du spectre de l'échantillon uniformément marqué pour des temps d'acquisition de 56 heures.

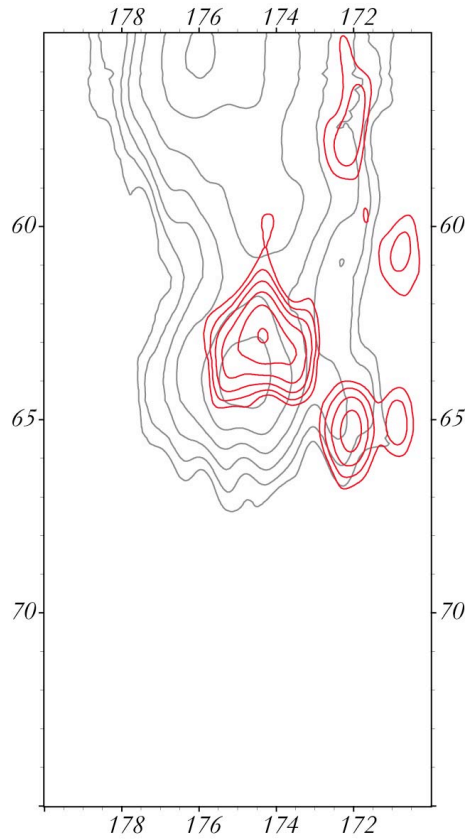


Figure 2.12 : Comparaison des spectres uniformément et sélectivement marqués. Agrandissement de la région des corrélations $C'-C\alpha$ où le marquage sélectif permet l'identification des corrélations.

III Analyse de l'échantillon sélectivement marqué *AKVY-Ec-MscL* - Apport des algorithmes de prédiction de déplacement chimiques

Après avoir appliqué la *SIV* comme technique de synthèse de la protéine membranaire *Ec-MscL* permettant son marquage sélectif, améliorant ainsi la lisibilité des spectres, il fallait adopter une stratégie de marquage permettant de couvrir tous les résidus de la protéine. Même si la structure de la protéine n'est pas résolue, les travaux entrepris sur ses homologues *Sa-MscL* et *Mt-MscL* ainsi que les analyses structurales et fonctionnelles obtenues *in vitro* par des stratégies de mutagenèse couplées à des analyses par *patch clamp* sur *Ec-MscL*, et les approches d'étude de la structure *in silico*, tout cela a permis de générer quelques modèles structuraux adaptés aux différentes observations effectuées (Corry & Martinac, 2008). En parallèle, différents laboratoires étudient depuis une vingtaine d'années, la relation étroite qui existe entre les déplacements chimiques observés des noyaux 1H , ^{13}C et ^{15}N et la structure tridimensionnelle de la molécule étudiée (Latek *et al.*, 2007 ; Wishart *et al.*, 1992 ; Wishart & Case, 2001 ; Zhao *et al.*, 2010). Ainsi, à partir de ces analyses, différents algorithmes ont été développés pour prédire une valeur de déplacement chimique à partir d'une structure tridimensionnelle. Parmi ces algorithmes, *SPARTA* présente une fiabilité testée sur toutes les protéines composant la banque de données, montrant un très faible écart entre les valeurs expérimentales et les valeurs prédites (Shen & Bax 2007). Les déplacements chimiques des acides aminés de plusieurs modèles structuraux de la protéine *Ec-MscL* ont été prédits, et séparés en fonction de la famille d'acides aminés. En analysant la répartition de ces déplacements chimiques prédits, quatre familles d'acides aminés ont été choisies pour tester cette stratégie. Il s'agit des résidus *Ala*, *Lys*, *Val* et *Tyr*. La synthèse de la protéine a été réalisée directement en liposomes.

a Prédiction des déplacements chimiques

Quinze modèles de *Ec-MscL* ont été utilisés pour la prédiction des déplacements

chimiques. Parmi ces modèles, trois représentent des états d'expansion lors de l'ouverture de la protéine (Sukharev *et al.*, 2001). Les douze autres modèles sont des *snapshots* de dynamique moléculaire à différents intervalles, à partir de quatre différents modèles issus de différents alignements de séquence (Valadié, 2003). Les déplacements chimiques ont été prédits sur chacun de ces quinze modèles. Les valeurs obtenues sur les carbones C' , $C\alpha$ et $C\beta$ ont été utilisées pour tracer les données 2D $C'-C\alpha$ et $C\alpha-C\beta$. La [figure 3.1](#) montre les corrélations $C\alpha-C\beta$ de 13 familles d'acides aminés. Les valeurs prédites des résidus *Ile* et *Thr* ne sont pas présentées dans la [figure 3.1](#). Cependant, les valeurs obtenues montrent une plus grande variance concernant les déplacements chimiques des $C\alpha$ des *Ile*, situés entre 57 et 65 *ppm*, et une très faible variance au niveau des déplacements chimiques des $C\beta$ (38,2 - 38,8 *ppm*). Les valeurs prédites des *Thr* montrent le même profil que celui obtenu expérimentalement, au niveau des corrélations $C\alpha-C\beta$, mais des valeurs différentes au niveau des déplacements chimiques des C' où un écart de 2 à 3 *ppm* est observé. La [figure 3.1](#) montre que les corrélations $C\alpha-C\beta$ des résidus *Ala*, *Lys*, *Val* et *Tyr* ne se recouvrent pas.

b Production de l'échantillon (*Ala*, *Lys*, *Val*, *Tyr*)-*MscL* en liposomes de *DOPC*

Pour tester cette stratégie de marquage, un nouveau protocole de synthèse de la protéine *Ec-MscL* par *SIV* a été appliqué. Il s'agit de la synthèse de la protéine en présence de liposomes pré-formés. Ce protocole a été mis en place pour la protéine *Ec-MscL* dans le laboratoire d'A. GHAZI. L'insertion dans la bicouche lipidique est spontanée et ne nécessite pas la présence de cofacteur ou de chaperon (Berrier *et al.*, 2011). La synthèse de l'échantillon (*Ala*, *Lys*, *Val*, *Tyr*)- ^{13}C , ^{15}N -*MscL* représente la première tentative de synthèse d'une protéine membranaire *in vitro*, directement en liposomes, avec une mise à l'échelle du protocole dans le but de l'étude par *RMN* en phase solide. À la fin de la synthèse, les protéoliposomes sont récupérés par ultracentrifugation. Le profil de migration de l'échantillon final montre la présence d'une bande intense représentant la protéine *Ec-MscL*, et des bandes de protéines insérées ou adhérentes aux liposomes ([Figure 3.2](#)).

c Analyse de (Ala, Lys, Val, Tyr)-MscL par RMN

Différentes expériences $2D\ ^{13}C\text{-}^{13}C\ DARR$ ont été effectuées sur l'échantillon (Ala, Lys, Val, Tyr)- $^{13}C,^{15}N$ -MscL avec des temps de mélange compris entre 10 et 500 ms. Les spectres présentent globalement une faible résolution, et à aucun temps de mélange des corrélations des différents carbones de la Tyr ne sont identifiables. Cependant, la résolution n'a pu être obtenue précisément en raison de l'absence de pics isolés. Le spectre obtenu avec un temps de mélange de 10 ms contient des corrélations $C'\text{-}C\alpha$, $C\alpha\text{-}C\beta$ et $C\beta\text{-}C\gamma$ très intenses des résidus Ala, Lys et Val, ainsi que des corrélations entre carbones séparés par deux ou plusieurs liaisons, mais à une moindre intensité (Figure 3.3). Les valeurs de déplacements chimiques des $C\alpha$ et des $C\beta$ correspondent à celles prédites par le programme S_{PARTA} . En augmentant le temps de mélange, les corrélations présentes pour 10 ms de temps de mélange sont plus intenses et de nouvelles corrélations apparaissent impliquant des noyaux éloignés au sein du même acide aminé, dans le cas des corrélations $C\alpha\text{-}C\epsilon$ des Lys, ou alors des noyaux de deux résidus différents, très probablement des corrélations entre une Ala et une Val juxtaposées (les corrélations encadrées de la figure 3.4). Cependant, l'analyse de l'évolution de l'intensité de ces signaux en fonction du temps de mélange n'apporte aucune information structurale car aucune identification n'est possible. L'évolution du signal montre un maximum d'intensité autour de 250 ms de temps de mélange (Figure 3.5). Ceci est supérieur au maximum d'intensité des corrélations des spectres de l'échantillon uniformément marqué, ce qui prouve l'utilité du marquage sélectif pour l'acquisition de spectres à des longs temps de mélange.

Différents tests ont montré que la protéine obtenue est active. Concernant les protéines présentes dans les liposomes, et visibles sur un gel $SDS\text{-}PAGE$, différents tests ont été effectués, où les protéoliposomes ont été traités à des concentrations variables d'urée pour dénaturer les protéines éventuellement adhérentes aux liposomes, et où les échantillons sont placés en haut, ou en bas d'un gradient de saccharose. Après ultracentrifugation, les différentes fractions ont été analysées sur gels d'électrophorèse, et ne montraient la présence d'aucune autre entité protéique détectable au bleu de $COOMASSIE$, en dehors de la bande contenant les protéoliposomes. De plus, les bandes vues sur le gel

SDS-PAGE de la [figure 3.2](#) ne sont pas des produits marqués. Un gradient de saccharose effectué à la fin de la *SIV* permettrait probablement de se débarrasser de ces protéines.

La stratégie adoptée montre des spectres où les corrélations ne se recouvrent pas lorsqu'il s'agit des familles de résidus marqués, mais au sein de la même famille, les résidus ne sont pas identifiables entre eux, ce qui empêche l'attribution. La stratégie de marquage basée sur la prédiction des déplacements chimiques à partir de quinze modèles de la protéine est efficace. Mais le choix des résidus fut mauvais, étant donné que les *Ala* et les *Val* sont parmi les résidus les plus présents dans la protéine (les deux représentent 20% de la protéine), et que de très nombreuses paires d'acides aminés marqués sont présentes au niveau de la séquence.

IV Analyse de l'échantillon sélectivement marqué *FIMPR-Ec-MscL* - Analyse de la séquence

La stratégie adoptée ici pour la *SIV* de la protéine membranaire *Ec-MscL* et son étude par *RMN* en phase solide consiste à identifier au sein de la séquence, les familles d'acides aminés qui présentent des résidus voisins formant des paires uniques identifiables par des corrélations spécifiques sur les spectres *2D* homo- ou hétéronucléaires. En effet, au niveau des échantillons étudiés précédemment, des corrélations inter-résiduelles ont été observées. Ces corrélations ont été utiles pour l'attribution de la seule *Thr T123*, mais ne pouvaient servir dans l'attribution dans le cas de l'échantillon (*Ala, Lys, Val, Tyr*)-*Ec-MscL* à cause de la présence de plusieurs paires *Ala-Val* dans la séquence. En combinant l'identification de paire unique à une approche consistant à éviter de marquer à la fois les acides aminés les plus fréquents, un premier ensemble a été adopté pour tester cette stratégie.

a Marquage basé sur l'analyse de la séquence

La [figure 4.1](#) montre la stratégie adoptée, ayant servi pour le marquage. Celui-ci montre la présence de huit paires uniques situées dans toutes les parties de la protéine.

b Production de l'échantillon (*Phe, Ile, Met, Pro, Arg*)-*MscL* en liposomes de *DOPC*

L'échantillon avec un marquage au niveau des résidus *Phe, Ile, Met, Pro* et *Arg* a été obtenu par *SIV* directement en liposomes de *DOPC*, avec dans un premier temps un ratio protéine/lipide de $1/4$. Ceci a permis d'obtenir un premier jeu de données de spectres de corrélations *2D* homonucléaires. La sensibilité était faible, l'échantillon de protéo-liposomes a été donc solubilisé en micelles de détergent puis reconstitué en liposomes avec un ratio protéine/lipide de $1/2$ afin d'améliorer la sensibilité.

c Analyse de (Phe, Ile, Met, Pro, Arg)-MscL par RMN

La [figure 4.2](#) montre un spectre obtenu sur l'échantillon reconstitué en liposomes avec un ratio protéine/lipide de $1/4$, en appliquant la séquence *DARR* avec un temps de mélange de 20 ms et à une température de 238 K . Le signal est faible et la résolution très basse. Les corrélations $C\alpha-C\beta$ des cinq types de résidus sont présentes. En augmentant le temps de mélange, peu de nouvelles informations apparaissent ([Figure 4.3](#)). D'autres expériences ont été effectuées à des températures plus élevées, mais les expériences effectuées à 238 K montraient des pics plus fins, malgré un rapport signal/bruit plus faible. Pour toutes températures confondues, le signal présentait une augmentation de l'intensité jusqu'à 100 ms de temps de mélange.

Pour obtenir un meilleur rapport signal/bruit, le ratio protéine/lipide a été augmenté à $1/2$. Les spectres de la [figure 4.4](#) montrent une amélioration de la résolution. Les pics sont plus fins, et deux corrélations inter-résiduelles apparaissent. Il s'agit d'une corrélation à $65/48,5\text{ ppm}$ probablement un pic croisé $IC\alpha-PC\delta$ présent à partir de 50 ms . Une autre corrélation à $65/59,5\text{ ppm}$ apparaît pour des temps de mélange supérieurs à 100 ms . Il peut s'agir d'une corrélation entre le $C\alpha$ d'une *Ile* et d'un $C\alpha$ d'une *Phe* ou d'une *Pro*. Pour la première corrélation, il peut s'agir de la paire unique ${}_{68}I-{}_{69}P$ mais l'absence de corrélations entre les $C\alpha$ de ces résidus rend improbable cette attribution. De plus, l'apparition de cette corrélation à un temps de mélange de 50 ms témoigne d'une distance très faible entre les deux noyaux, ce qui est possible avec deux résidus *Ile* et *Pro* situés l'un en face de l'autre et suffisamment proches dans l'espace. Il peut s'agir donc de deux résidus Ile_i et Pro_{i+4} dans une hélice α et séparés par un tour d'hélice. Dans ce cas, les résidus seraient suffisamment proches pour que les corrélations soient aussi intenses. Les résidus ${}_{41}I$ et ${}_{44}P$ situés au niveau de l'hélice transmembranaire *TM1* seraient les résidus identifiés par cette corrélation. En dehors de ces deux résidus, la seule partie non-structurée de la protéine qui pourrait contenir des résidus *Ile* et *Pro* proches est la partie située entre l'hélice transmembranaire *TM2* et l'hélice cytoplasmique où sont situés les résidus ${}_{95}I$, ${}_{99}I$, ${}_{109}P$, ${}_{113}P$ et ${}_{115}P$. Aucun de ces cas de figures ne peut être totalement confirmé ni infirmé. Des données complémentaires issues de corrélations hétéro-

nucléaires informeraient s'il s'agit de résidus voisins au niveau de la séquence ou non. Pour la deuxième corrélation, l'existence d'une paire ${}_{78}F\text{-}{}_{79}I$ et ${}_{92}I\text{-}{}_{93}F$, ainsi que des résidus *Ile* et *Phe* plus ou moins proches dans l'espace, empêchent également l'attribution.

Les spectres présentés dans la [figure 4.5](#) montrent l'accumulation, pour chaque temps de mélange, de quatre *FIDs* obtenant ainsi des spectres de 56 *b*. Sur ces spectres apparaissent dans la région 130-160 *ppm* les corrélations des noyaux du cycle phényle avec les $C\alpha$ et les $C\beta$ des *Phe*. Les corrélations entre le carbone $C\xi$ et les autres carbones des *Arg* apparaissent également.

Finalement, l'analyse des intensités des pics isolés en fonction du temps de mélange permet de montrer une diminution rapide pour la corrélation à 55-30 *ppm* ($C\alpha\text{-}C\beta$ de *Met*) et de celle à 59-15 *ppm* ($C\alpha\text{-}C\varepsilon$ de *Met* ou $C\alpha\text{-}C\delta$ de *Ile*) à partir de 100 *ms* de temps de mélange. La corrélation à 60-48 *ppm* montre une diminution moins abrupte, il s'agit d'une corrélation $PC\alpha\text{-}PC\delta$. Finalement la corrélation $RC\delta\text{-}RC\beta$ à 41,5-28,5 *ppm*, montre une intensité constante entre 100 et 500 *ms* de temps de mélange. Pour les corrélations $C\alpha\text{-}C\beta$, une diminution rapide a déjà été observée dans le cas des autres échantillons, et est probablement due à la présence d'autres noyaux ${}^{13}C$ à proximité des deux noyaux participant à la corrélation, ce qui provoque une relaxation plus rapide. Pour les corrélations $C\alpha\text{-}C\varepsilon$ de *Met* ou $C\alpha\text{-}C\delta$ de *Ile*, l'augmentation du temps de mélange devrait normalement augmenter l'intensité, si la distance entre les deux noyaux ne varie que peu pendant l'accumulation des données. La diminution de l'intensité pourrait provenir du fait que les noyaux ne font pas partie du même résidu, mais sont proches dans l'espace, avec des noyaux $C\varepsilon$ de la *Met* ou $C\delta$ de l'*Ile* très mobiles. Finalement, pour la corrélation $RC\delta\text{-}RC\beta$, la constance de l'intensité du signal pourrait être la conséquence de l'addition de plusieurs signaux provenant de plusieurs chaînes latérales d'*Arg*.

La comparaison des spectres avec deux différents ratios protéine/lipide montre une nette amélioration au niveau du rapport signal/bruit. La résolution est légèrement améliorée, et les déplacements chimiques ne sont pas modifiés. Ceci amène à penser que lors de la *SIV* directement en liposomes, l'échantillon de protéoliposomes est

suffisamment pur pour des études structurales par RMN en phase solide. Cette stratégie est prometteuse, car elle permet d'obtenir des informations structurales sans avoir à attribuer entièrement les résidus de la protéine. Cependant, le choix des résidus était probablement très ambitieux, au niveau du nombre de résidus marqués, mais aussi au niveau des corrélations inter-résiduelles attendues. Par la même stratégie, d'autres marquages sont possibles : la combinaison des résidus *Glu*, *Phe*, *Lys* et *Thr* donnerait accès aux paires ${}_{54}F\text{-}{}_{55}K$, ${}_{116}T\text{-}{}_{117}K$ et ${}_{123}T\text{-}{}_{124}E$. De plus, d'un côté les trois *Thr* sont attribuées, aidant ainsi à attribuer au moins deux résidus (${}_{117}K$ et ${}_{124}E$); d'un autre côté, d'autres paires sont présentes ${}_{5}K\text{-}{}_{6}E$, ${}_{6}E\text{-}{}_{7}F$, ${}_{9}E\text{-}{}_{10}F$, ${}_{106}K\text{-}{}_{107}E$, ${}_{117}K\text{-}{}_{118}E$ et ${}_{130}K\text{-}{}_{131}E$. Ainsi six *Glu* (sur huit) et cinq *Lys* (sur neuf) possèdent un résidu voisin marqué. Par ailleurs, cet échantillon ne contient aucun des cinq acides aminés les plus fréquents dans la protéine.

V Analyse de l'échantillon sélectivement marqué *DGLS-Ec-MscL* - Une combinaison des approches prédictives

Après avoir analysé les données issues de la prédiction des déplacements chimiques, de nouveaux modèles par homologie ont été créés et de nouveaux calculs de déplacements chimiques ont été faits. Ces résultats ainsi que l'analyse de la séquence à la recherche de paires d'acides aminés dans une région de la protéine pouvant subir des modifications, ont permis de choisir un nouvel ensemble.

a Stratégie de marquage

Pour la prédiction des déplacements chimiques, ce sont quarante modèles aléatoires issus de quatre différents alignements de séquences qui ont été utilisés. Une seule vérification est effectuée au niveau du serveur *WHATIF*, montrant des longueurs de liaisons standards. Quelques angles de torsion ont été trouvés à s'écarter légèrement des angles standard, ainsi que les angles dièdres des résidus *Pro*. Les atomes d'hydrogène ont été rajoutés aux modèles grâce au programme *Reduce* puis les modèles ont été revérifiés montrant les mêmes erreurs. Néanmoins, tous les modèles ont été utilisés pour la prédiction des déplacements chimiques afin de conserver la stochasticité de l'approche. Ainsi, le programme *SPARTA* donne 40 valeurs pour chacun des atomes suivants ^{15}N , ^1HN , $^1\text{H}\alpha$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ et $^{13}\text{C}'$ de chaque résidu. Les valeurs sont séparées par famille de résidus, et la distribution des corrélations $\text{C}\alpha\text{-C}\beta$ est analysée. Comme est mentionné lors de l'exposition des résultats, les résidus montrant une plus grande distribution sont *Asp*, *Met*, *Asn*, *Ser* et *Thr*. En prenant les résidus *Asp* et *Ser*, une paire unique $_{34}\text{S-35}\text{S}$ est disponible, mais en marquant conjointement les résidus *Leu*, une paire $_{18}\text{D-19}\text{L}$ est marquée à l'interface entre le cytoplasme et la membrane plasmique au niveau du début de l'hélice transmembranaire *TM1*, une paire $_{35}\text{S-36}\text{L}$ située après la paire $_{34}\text{S-35}\text{S}$ au sein de l'hélice *TM1*, et trois paires *L-L*, $_{47}\text{L-48}\text{L}$ située à l'interface entre la membrane plasmique et l'espace périplasmique à la fin de l'hélice *TM1*, $_{121}\text{L-122}\text{L}$ et $_{128}\text{L-129}\text{L}$ situées au niveau de l'hélice cytoplasmique de la protéine. Et finalement, le marquage des résidus *Gly*

permettrait de distinguer la paire $_{47}L-_{48}L$, et forme une paire unique avec un résidu *Asp* situé au niveau de la boucle périplasmique.

b Production de l'échantillon (*Asp, Gly, Leu, Ser*)-*Ec-MscL* en liposomes de *DOPC*

L'échantillon a été synthétisé après (*Phe, Ile, Met, Pro, Arg*)-*Ec-MscL* mais avant l'acquisition des premiers spectres sur ce dernier. Ainsi, en voyant la nette amélioration au niveau du rapport signal/bruit, il s'est avéré judicieux de resolubiliser l'échantillon préparé avec un ratio de protéine/lipide de $1/4$ afin de doubler le ratio. Les protéoliposomes ainsi obtenus sont conservés à -20°C jusqu'à leur analyse au spectromètre de RMN.

c Analyse de (*Asp, Gly, Leu, Ser*)-*Ec-MscL* par RMN

La [figure 5.4](#) montre quatre spectres $2D\ ^{13}\text{C}-^{13}\text{C}$ obtenus sur l'échantillon en appliquant la séquence *DARR* avec des temps de mélange de $20, 50, 100$ et $200\ \text{ms}$. La résolution est assez faible, et seules les corrélations des résidus *Leu* et *Gly* sont présentes. Des corrélations des résidus *Asp* sont confondues avec quelques unes des résidus *Leu*. L'analyse de l'intensité de quatre pics isolés de la [figure 5.5](#) montre un maximum d'intensité pour un temps de mélange de $100\ \text{ms}$ pour la corrélation $C\alpha-C\beta$ de *Leu* ou *Asp*. Les pics $175,5/20,7$ et $175,4/23,8$ impliquant des corrélations $LC'-C\delta$ montrent un maximum d'intensité vers $50\ \text{ms}$ de temps de mélange. Au sein du même résidu, une corrélation $LC'-C\delta$ correspond à une distance de $6\ \text{Å}$. Or l'évolution rapide du signal témoigne d'une proximité entre les deux carbonés. La distance est inférieure à $1,5\ \text{Å}$ étant donné que le signal croît plus vite que celui de la corrélation $C\alpha-C\beta$ de *Leu* ou *Asp*. Il s'agit donc de corrélations inter-résiduelles entre deux résidus très proches où la chaîne latérale de l'un est à moins de $1,5\ \text{Å}$ du C' de l'autre. En analysant la séquence de la protéine *Ec-MscL*, la région susceptible de comporter des résidus *Leu* proches de $1,5\ \text{Å}$ ou moins pourrait être l'hélice cytoplasmique où les paires $_{121}L-_{122}L$ et $_{128}L-_{129}L$ sont présentes. Dans ce cas la corrélation pourrait être entre une des deux leucines de la paire $_{121}L-_{122}L$, la leucine $_{128}L$ et une des deux leucines de la paire $_{128}L-_{129}L$. En effet, les hélices

cytoplasmiques de la protéine forment un faisceau rigide lorsque la protéine est fermée (Anishkin *et al.*, 2003). La structure de *Mt-MscL* obtenue en 1998 montrait des hélices dans un état non-naturel à cause de l'acidité du milieu (Chang *et al.*, 1998). Abaisser le *pH* lors de la reconstitution de la protéine en liposomes pourrait légèrement déstructurer les hélices cytoplasmiques, ce qui conserverait les interactions des *Leu* voisines au niveau de la séquence, et diminuerait les interactions entre résidus de deux monomères différents. Pour distinguer la paire $_{121}L-_{122}L$ de la paire $_{128}L-_{129}L$, l'obtention des corrélations des *Asp*, marquées dans cet échantillon, serait suffisante, vu que la paire $_{128}L-_{129}L$ est précédée par une *Asp* ($_{127}D$). Ce qui ramène à un autre problème rencontré pour cet échantillon, qui est l'absence des corrélations des *Asp* et des *Ser*; en effet, aucun des spectres obtenus ne présente des informations de ces résidus. Ce résultat est gênant, bien que prévisible; en effet, ces résidus ont été choisis car l'analyse de leurs déplacements chimiques prédits montraient une dispersion entre les résidus, ainsi que des fluctuations pour chaque résidu (Figures 5.3 et 5.6).

Les valeurs des écart-types ont été calculées pour chaque famille d'acides aminés. Celles-ci sont situées entre 0,2 *ppm* pour les *Pro* et 0,85 *ppm* pour les *Ser*. En dessous de la valeur de 0,4 *ppm* pour les *Gly*, se trouvent les valeurs d'écart-type des résidus *Pro*, *Met*, *Ala*, *Phe*, *Leu*, *Gln* et *Glu* (Figure 5.7). Les spectres obtenus sur les échantillons (*Ala*, *Lys*, *Val*, *Tyr*)-*MscL* ont montré des corrélations intenses pour les *Ala*. Les *Pro* de l'échantillon (*Phe*, *Ile*, *Met*, *Pro*, *Arg*)-*MscL*, et les *Leu* de l'échantillon (*Asp*, *Gly*, *Leu*, *Ser*)-*MscL* montrent également des corrélations intenses. Les résidus *Lys*, *Asn*, *Asp*, *Val*, *Arg*, *Thr*, *Ile* et *Ser* présentent des valeurs supérieures à celle des *Gly*. Il a été précédemment observé, pour certaines *Ile* de l'échantillon (*Ile*, *Thr*)-*MscL*, et pour des résidus *Arg* de l'échantillon (*Phe*, *Ile*, *Met*, *Pro*, *Arg*)-*MscL*, une absence de corrélations due justement à des mouvements dans le régime intermédiaire. Les *Thr* font partie du deuxième groupe, mais ont été néanmoins identifiées et attribuées. Les valeurs des prédictions des déplacements chimiques par cette méthode, ainsi que leur distribution, contribueraient ainsi à la compréhension des fluctuations observées au niveau des spectres.

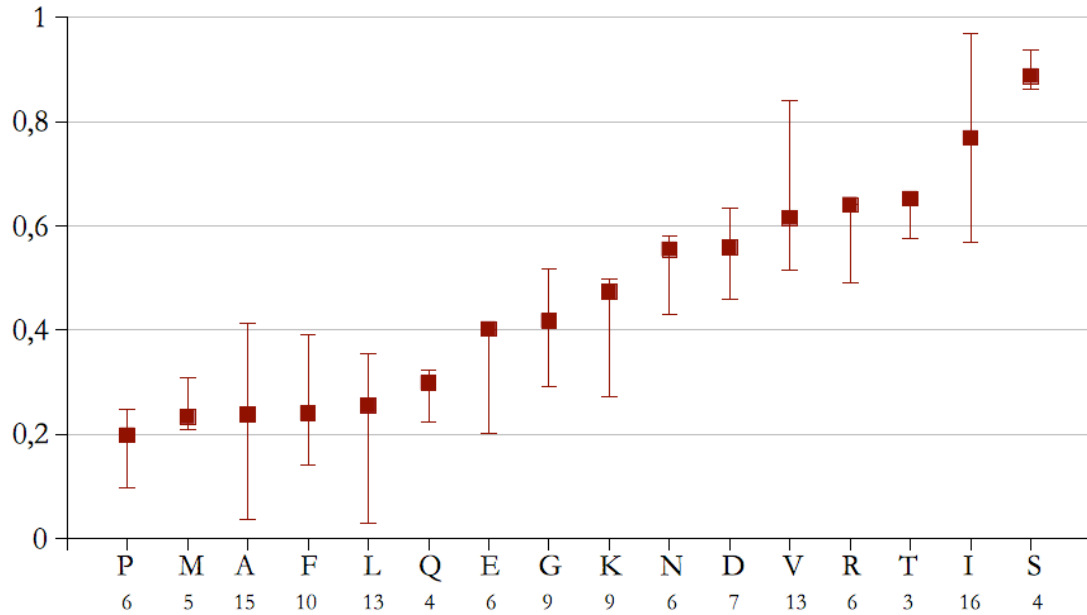


Figure 5.7 : Variation de l'écart-type à la moyenne des déplacements chimiques prédits en fonction de la famille d'acides aminés. Les résidus ont été classés en fonction de la valeur de l'écart-type. Pour chaque famille de résidus, le nombre d'acides aminés dans la protéine est signalé en dessous du code correspondant. Les barres représentent la distribution des acides aminés, entre hélices (barre du dessus), et les autres structures secondaires de la protéine (barre du bas).

VI Rôle de l'environnement lipidique - Analyse de l'échantillon *AGKVY-Ec-MscL*

Les travaux effectués sur la protéine *Ec-MscL* précédemment dans différents laboratoires ont montré le rôle primordial joué par la bicouche lipidique dans l'activation de la protéine (Elmore & Dougherty, 2003 ; Perozo *et al.*, 2002 ; Powl *et al.*, 2005) :

- Afin d'observer l'effet d'une rigidification de la membrane, un mélange *DPPC/DOPC* a été choisi. La température de transition de la *DOPC* est de -20°C ; celle de la *DPPC* est de 40°C . Les expériences étant réalisées à 258 K (-15°C), la membrane contenant la protéine est très probablement rigide, sauf dans le cas où une compartimentation a lieu au sein des vésicules.

- Le mélange de lipides extraits de la membrane interne d'*E. coli* contient surtout des lipides polaires à chaîne d'acide gras insaturée. Le ratio lipides/protéine utilisé est de 1/1 pour former un réseau cristallin bidimensionnel (Saint *et al.*, 1998).

- Le but de l'utilisation des lipides d'asolectine était d'observer la protéine en RMN comme elle est étudiée en *PATCH CLAMP*.

- L'ajout de lysophospholipides tendait à favoriser la conformation ouverte de la protéine. Malheureusement, ces deux dernières tentatives n'ont pas permis d'obtenir des spectres interprétables.

Outre la comparaison du comportement de la protéine dans différentes bicouches lipidiques, cet échantillon a présenté l'occasion d'utiliser une autre méthode d'extraction des molécules de détergent que la technique largement répandue des *bio beads*. L'avantage majeur du protocole de reconstitution des protéoliposomes à l'aide des molécules de β -cyclodextrine est le temps d'extraction. L'adsorption se fait beaucoup plus rapidement qu'avec les *bio beads*. Cependant, la purification des protéoliposomes sur gradient de saccharose nécessite un temps assez long. De plus, il s'agit d'une étape supplémentaire où

de nouvelles pertes d'échantillon sont à prévoir. Pour l'échantillon *AGKVY-Aso* beaucoup de pertes sont survenues à ce niveau-là, ce qui a empêché l'obtention de données spectroscopiques. Pour l'échantillon *AGKVY-DOL*, la reconstitution a été faite en *DOPC*, puis après le gradient de saccharose, les lysophospholipides ont été rajoutés, en resuspendant les protéoliposomes, puis en centrifugeant pour la récupération de l'échantillon final, ce qui a empêché également l'obtention d'une quantité suffisante pour la spectroscopie RMN. En ce qui concerne les échantillons *AGKVY-Eco* (ratio lipides/protéine de 1/1) et *AGKVY-DPDO* (ratio lipides/protéine de 2/1), les spectres obtenus montrent plus de signal dans le deuxième cas ([Figure 6.1](#)). Étant donné que dans le deuxième cas, il y a moins de protéine dans le rotor, ce résultat paraît contradictoire. L'analyse de ces spectres montre une différence particulièrement au niveau des corrélations $C\alpha C\beta$ des résidus *Ala* ([Figure 6.2](#)), ce qui a permis d'identifier cinq différents systèmes de spin (pour 15 résidus *Ala* présents dans la protéine *Ec-MscL*; voir [tableau 6.1](#)). En revanche, aucune corrélation $C\alpha C\beta$ des résidus *Lys* n'est visible sur le spectre de l'échantillon *AGKVY-Eco*. Le spectre obtenu sur l'échantillon *AGKVY-DPDO* est similaire à celui obtenu sur l'échantillon *AKVY-MscL* en *DOPC*. Les pics présents sur le spectre *AGKVY-DPDO* mais absents sur *AGKVY-Eco* représenteraient des corrélations figées dans le premier cas, à cause de la membrane, adoptant plusieurs conformations dans le cas d'une membrane plus fluide, ce qui provoque l'absence de signal. Il peut donc s'agir simplement d'un effet de la fluidité, comme lors d'un changement de température. Un criblage plus détaillé serait nécessaire pour déterminer le réel effet des lipides au niveau des spectres, en faisant varier la longueur de la chaîne des acides gras, ou la nature de la tête polaire, en procédant progressivement vers des mélanges complexes.

Une autre comparaison a été faite en remplaçant l'eau H_2O par de l'eau deutérée D_2O pour voir si cela modifierait la résolution en deutérant certains protons échangeables au niveau des groupement $-OH$ et $-NH$. Cela aurait pu aider à réduire la diffusion de spin pour les carbones proches de ces groupements. Cependant, aucune différence n'a été observée. Cela est probablement dû au fait que la majorité des résidus marqués dans cet échantillon se trouve dans les régions rigides de la protéine, au niveau des hélices transmembranaires. Les résidus situés dans les parties mobiles exposées au solvant ne

participent que très peu aux corrélations présentes sur le spectre. La deutération avant la reconstitution en liposomes pourrait être testée. Pour une deutération plus homogène des protons échangeables, l'utilisation d'un tampon deutéré dans le kit de *STV* pourrait être envisagée, mais des tests sont indispensables, pour voir si la deutération ne perturbe pas la machinerie de transcription/traduction *in vitro*.

VII Références

- Anishkin** A, Gendel V, Sharifi NA, et al. On the conformation of the COOH-terminal domain of the large mechanosensitive channel MscL. *J. Gen. Physiol.* **2003**;121(3):227-244.
- Berjanskii** MV, Wishart DS. Application of the random coil index to studying protein flexibility. *J. Biomol. NMR.* **2008**;40(1):31-48.
- Berrier** C, Guilvout I, Bayan N, et al. Coupled cell-free synthesis and lipid vesicle insertion of a functional oligomeric channel MscL. MscL does not need the insertase YidC for insertion in vitro. *Biochim. Biophys. Acta.* **2011**;1808(1):41-46.
- Berrier** C, Park K-H, Abes S, et al. Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry.* **2004**;43(39):12585-12591.
- Bertini** I, Bhaumik A, De Paëpe G, et al. High-resolution solid-state NMR structure of a 17.6 kDa protein. *J. Am. Chem. Soc.* **2010**;132(3):1032-1040.
- Cady** SD, Luo W, Hu F, Hong M. Structure and function of the influenza A M2 proton channel. *Biochemistry.* **2009**;48(31):7356-7364.
- Castellani** F, van Rossum B, Diehl A, et al. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature.* **2002**;420:98-102.
- Chang** G, Spencer RH, Lee AT, Barclay MT, Rees DC. Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science.* **1998**;282(5397):2220-2226.
- Corry** B, Martinac B. Bacterial mechanosensitive channels: experiment and theory. *Biochim. Biophys. Acta.* **2008**;1778:1859-1870.
- Duma** L, Abergel D, Ferrage F, et al. Broadband dipolar recoupling for magnetization transfer in solid-state NMR correlation spectroscopy. *Chemphyschem.* **2008**;9(8):1104-1106.
- Elmore** DE, Dougherty DA. Investigating lipid composition effects on the mechanosensitive channel of large conductance (MscL) using molecular dynamics simulations. *Biophys. J.* **2003**;85(3):1512-1524.
- Etzkorn** M, Martell S, Andronesi OC, et al. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2007**;46(3):459-462.
- Franks** WT, Wylie BJ, Schmidt HL, et al. Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105:4621-4626.
- Frericks** HL, Zhou DH, Yap LL, Gennis RB, Rienstra CM. Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: *E. coli* cytochrome bo₃ oxidase. *J. Biomol. NMR.* **2006**;36(1):55-71.
- Grommek** A, Meier BH, Ernst M. Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.* **2006**;427(4-6):404-409.
- Helmus** JJ, Surewicz K, Surewicz WK, Jaroniec CP. Conformational flexibility of Y145Stop human prion protein amyloid fibrils probed by solid-state nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **2010**;132(7):2393-2403.
- Lange** A, Luca S, Baldus M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **2002**;124(33):9704-9705.
- Latek** D, Ekonomiuk D, Kolinski A. Protein structure prediction: combining de novo modeling with sparse

- experimental data. *J Comput Chem.* **2007**;28(10):1668-1676.
- Li Y**, Berthold DA, Frericks HL, Gennis RB, Rienstra CM. Partial (13)C and (15)N chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. *Chembiochem.* **2007**;8(4):434-442.
- Loquet A**, Bardiaux B, Gardiennet C, et al. 3D structure determination of the Crh protein from highly ambiguous solid-state NMR restraints. *J. Am. Chem. Soc.* **2008**;130(11):3579-3589.
- Manolikas T**, Herrmann T, Meier BH. Protein structure determination from 13C spin-diffusion solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2008**;130(12):3959-3966.
- Miroux B**, Walker JE. Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **1996**;260(3):289-298.
- Perozo E**, Kloda A, Cortes DM, Martinac B. Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nat. Struct. Biol.* **2002**;9(9):696-703.
- Powl AM**, Carney J, Marius P, East JM, Lee AG. Lipid interactions with bacterial channels: fluorescence studies. *Biochem. Soc. Trans.* **2005**;33(Pt 5):905-909.
- Rogé J**, Betton J-M. Use of pIVEX plasmids for protein overproduction in Escherichia coli. *Microb. Cell Fact.* **2005**;4:18.
- Saint N**, Lacapère J-J, Gu LQ, et al. A hexameric transmembrane pore revealed by two-dimensional crystallization of the large mechanosensitive ion channel (MscL) of Escherichia coli. *J. Biol. Chem.* **1998**;273(24):14667-14670.
- Shi L**, Lake EMR, Ahmed MAM, Brown LS, Ladizhansky V. Solid-state NMR study of proteorhodopsin in the lipid environment: secondary structure and dynamics. *Biochim. Biophys. Acta.* **2009**;1788(12):2563-2574.
- Sperling LJ**, Berthold DA, Sasser TL, Jeisy-Scott V, Rienstra CM. Assignment strategies for large proteins by magic-angle spinning NMR: the 21-kDa disulfide-bond-forming enzyme DsbA. *J. Mol. Biol.* **2010**;399(2):268-282.
- Sukharev S**, Durell SR, Guy HR, Sukharev SI. Structural models of the MscL gating mechanism. *Biophys. J.* **2001**;81:917-936.
- Szeverenyi NM**, Sullivan MJ, Maciel GE. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson.* **1982**;47(3):462-475.
- Valadié H**, Lacapère J-J, Sanejouand YH, Etchebest C. Dynamical properties of the MscL of Escherichia coli: a normal mode analysis. *J. Mol. Biol.* **2003**;332:657-674.
- Varga K**, Aslimovska L, Watts A. Advances towards resonance assignments for uniformly-13C, 15N enriched bacteriorhodopsin at 18.8 T in purple membranes. *J. Biomol. NMR.* **2008**;41(1):1-4.
- Weingarth M**, Bodenhausen G, Tekely P. Broadband carbon-13 correlation spectra of microcrystalline proteins in very high magnetic fields. *J. Am. Chem. Soc.* **2009**;131(39):13937-13939.
- Wishart DS**, Case DA. Use of chemical shifts in macromolecular structure determination. *Meth. Enzymol.* **2001**;338:3-34.
- Wishart DS**, Nip AM. Protein chemical shift analysis: a practical guide. *Biochem. Cell Biol.* **1998**;76(2-3):153-163.
- Wishart DS**, Sykes BD, Richards FM. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry.* **1992**;31:1647-1651.
- Zhao Y**, Alipanahi B, Li SC, Li M. Protein secondary structure prediction using NMR chemical shift data. *J Bioinform Comput Biol.* **2010**;8(5):867-884.

Conclusions

E Conclusions

I La surexpression bactérienne de la protéine *Ec-MscL* pour son étude par *RMN* en phase solide

Le protocole de surexpression de la protéine *Ec-MscL* dans les bactéries *BL21* a permis d'obtenir après purification et reconstitution de la protéine des quantités avoisinantes le milligramme à partir d'un litre de culture en milieu minimum enrichi en glucose marqué au ^{13}C et du chlorure d'ammonium marqué au ^{15}N . L'étude de la protéine par *RMN* en phase solide a montré des spectres présentant une bonne résolution. Cependant les pics ne sont pas suffisamment séparés pour permettre une étude structurale. En effet, les systèmes de spin de quelques groupes de résidus sont visibles, mais l'équivalence résidu-système de spin n'est pas présente. L'intensité des corrélations identifiées augmente rapidement entre 10 et 50 ms de temps de mélange, pour devenir négligeable au-delà de 200 ms. Des données accumulées à des temps de mélange supérieurs ne montrent aucun signal à 500 ms. L'absence de systèmes de spin isolés ne permet pas d'étudier des interactions à courte distance, et celles à longue distance ne peuvent pas être observées en raison de la diminution rapide de l'intensité des corrélations. Il était donc nécessaire de réduire le nombre de résidus marqués au niveau de la protéine afin d'améliorer la lisibilité des spectres. Pour cela, la bibliographie regorge d'astuces métaboliques. L'application de ces méthodes rajoute un coût considérable, car d'un côté les produits marqués (glycérol, pyruvate, ... marqués au ^{13}C) sont plus chers que le glucose marqué au ^{13}C , d'un autre côté le rendement de production est plus faible. Le choix dans ce travail a été porté sur la synthèse *in vitro* pour marquer spécifiquement toute une famille de résidus et faciliter l'identification des systèmes de spin sur les spectres. Malgré les prix élevés des kits utilisés, les marquages spécifiques, par utilisation de la quantité suffisamment nécessaire pour synthétiser la protéine, présentent une réduction importante du coût de l'échantillon, surtout lorsque ce coût est ramené à la quantité de protéine obtenue.

II La synthèse *in vitro* de la protéine Ec-MscL pour son étude par RMN en phase solide

i. Apport de la SIV

La synthèse *in vitro* a été utilisée dans le but de réduire le nombre de systèmes de spin et procéder à un marquage sélectif de la protéine Ec-MscL afin de faciliter son étude par RMN en phase solide, et simplifier les étapes d'attribution. Par ailleurs, la SIV associée à un marquage sélectif de 3-5 résidus permet de diminuer le coût d'un échantillon (pour plus de détail, voir Abdine *et al.*, 2011), avec un rendement proche de 50% d'incorporation des acides aminés rajoutés dans le milieu (introduction d'acides aminés nécessaires pour la synthèse de 20 mg de protéine, et obtention de 9 mg de Ec-MscL). Pour un premier essai, les résidus choisis ont été les isoleucines et les thréonines. La protéine Ec-MscL contient 16 Ile et 3 Thr. Le marquage de ces résidus avait pour but de tester la capacité de la méthode en termes de résolution et de sensibilité. D'un côté, la résolution obtenue n'a permis d'identifier que la moitié des systèmes de spin des Ile. D'un autre côté, les trois Thr sont identifiées grâce à une sensibilité accrue au niveau des spectres. L'attribution de ces résidus a été possible. En parallèle, l'analyse de l'intensité des corrélations a montré un avantage majeur par rapport au marquage uniforme. En effet, les spectres obtenus à des temps de mélange supérieurs à 200 ms montrent un signal intense. Des corrélations inter-résiduelles de longue distance ont été identifiées, et ont contribué à l'attribution des 3 Thr. La comparaison des résultats obtenus sur l'échantillon sélectivement marqué et celui uniformément marqué montre clairement l'intérêt de la synthèse *in vitro* pour étudier par RMN en phase solide des protéines membranaires insérées en liposomes. La réduction du nombre de spins marqués a permis de simplifier les spectres 2D des corrélations ^{13}C - ^{13}C . Malgré la mauvaise résolution observée dans la dimension de l'azote ^{15}N , le marquage sélectif par synthèse *in vitro* montre une amélioration au niveau du spectre 2D des corrélations ^{15}N - ^{13}C . Il s'est donc avéré indispensable d'utiliser la synthèse *in vitro*, mais il fallait des astuces évitant de marquer les résidus un par un. Pour cela, différentes stratégies ont été élaborées, utilisant successivement les algorithmes de prédiction de déplacement chimique, l'analyse de la séquence et de la composition de la protéine, puis une

combinaison des approches prédictives et des informations structurales des données bibliographiques.

ii. Approche prédictive

À partir de modèles structuraux de la protéine *Ec-MscL*, les déplacements chimiques théoriques ont été prédits à l'aide du programme *SPARTA*, et les corrélations des C' avec les $C\alpha$ ainsi que celles des $C\alpha$ avec les $C\beta$ ont été tracées. Sur le graphe montrant ces dernières, quatre familles de résidus (*Ala*, *Lys*, *Val* et *Tyr*) ont été choisies car leurs corrélations ne se recouvraient pas. Ces résidus ont été marqués sélectivement par synthèse *in vitro*, et l'échantillon obtenu a été analysé par RMN en phase solide. Les spectres obtenus ne montraient aucun recouvrement entre les corrélations de résidus appartenant à deux différentes familles, mais la résolution n'était pas suffisante pour obtenir des informations structurales. Les spectres obtenus ne montrent que très peu de pics isolés. L'analyse de l'intensité en fonction du temps de mélange de ces pics isolés a permis de conforter l'idée proposée, selon laquelle la synthèse *in vitro* permet d'obtenir des spectres à des temps de mélange très élevés. Cependant, le signal diminue plus rapidement dans le cas de l'échantillon *AKVY-Ec-MscL* que dans le cas précédent de l'échantillon *IT-Ec-MscL*. Ceci est dû au nombre de spins étant donné que 173 ^{13}C sont présents dans le premier, et 108 dans le deuxième, mais également au nombre de spins voisins car dans le cas de *IT-Ec-MscL*, seuls 3 doublets *I-I* sont marqués, contre 7 doublets et 2 triplets pour *AKVY-Ec-MscL*, contribuant sans doute à l'élargissement des pics de corrélations, réduisant ainsi la dispersion au niveau du spectre.

iii. Approche par analyse de la séquence

L'analyse de la séquence a été effectuée dans le but d'identifier des paires d'acides aminés présentes une fois, tout en évitant de choisir plusieurs résidus très fréquents dans la protéine. Ainsi, en choisissant les résidus *Phe*, *Ile*, *Met*, *Pro* et *Arg*, ce sont huit paires uniques au niveau de la séquence qui sont marquées. L'analyse de l'échantillon obtenu par synthèse *in vitro* a permis d'identifier les systèmes de spin des résidus. Deux ratios

protéine/lipide ont été testés montrant un gain important en signal/bruit mais pas en résolution. L'analyse des pics isolés a permis d'identifier uniquement les corrélations inter-résiduelles $IC\alpha-F/PC\alpha$ et $IC\alpha-PC\delta$. Ces corrélations ne permettent pas d'attribuer la paire d'acides aminés dans les deux cas, du fait du manque d'informations supplémentaires. La paire $_{58}I-_{59}P$ est située dans la boucle périplasmique, alors que les paires $_{78}F-_{79}I$ et $_{92}I-_{93}F$ sont dans l'hélice $TM2$. Comme pour les autres échantillons, la variation de la fluidité de la membrane, en variant sa composition ou la température d'acquisition des spectres, pourrait aider à attribuer les corrélations identifiées.

iv. Approche combinatoire

En reprenant la stratégie adoptée pour le marquage de l'échantillon $AKVY-Ec-MscL$, des calculs de prédiction de déplacements chimiques ont été effectués en variant les modèles structuraux. Ceci avait pour rôle la création de modèles différents pour évaluer une marge d'erreur sur le déplacement chimique prédit. Cependant, l'utilisation d'un nombre insuffisant de *snapshots* pourrait fausser l'évaluation de cette marge d'erreur. Ici, quarante modèles sont utilisés. L'analyse des écart-types à la moyenne des valeurs prédites pour cinq, dix, vingt, trente et quarante modèles a montré une stabilisation entre trente et quarante, prouvant ainsi que quarante modèles est un nombre suffisant pour obtenir une évaluation correcte du déplacement chimique et de la marge d'erreur sur cette évaluation. Par exemple, pour la thréonine T_{60} , les écart-types sont de 0,590, 0,526, 0,462, 0,435 et 0,438 ppm pour 5, 10, 20, 30 et 40 modèles respectivement. Pour la thréonine T_{116} ces valeurs sont de 0,302, 0,281, 0,205, 0,232 et 0,229 ppm. À partir des prédictions de déplacements chimiques sur les quarante modèles, deux familles de résidus ont été choisies, les aspartates et les sérines car leurs tracés de corrélations théoriques $C\alpha-C\beta$ montraient le plus de variabilité. De plus, les six aspartates présentes dans la protéine sont réparties de manière homogène au niveau de la séquence. Sur les cinq sérines présentes, une ($_2S$) est située en N-terminal, et deux ($_{136}S$ et $_{141}S$) au niveau de la partie C-terminale de la protéine. L'absence des corrélations de ces résidus est prévisible, car les extrémités des protéines sont souvent sujettes à des fluctuations. Les deux autres sérines ($_{34}S$ - $_{35}S$) sont situées au niveau de l'hélice transmembranaire $TM1$. L'absence de leurs corrélations

est étonnante. Au niveau des aspartates, les corrélations auraient pu apporter des informations importantes sur la relation entre la structure secondaire et le déplacement chimique dans les conditions expérimentales étudiées. Les autres résidus choisis dans le marquage étaient les leucines et les glycines. La combinaison des quatre familles de résidus montrait une série de paires voisines au niveau de la séquence, et aussi au niveau de la structure prédite de la protéine. Les corrélations entre carbones des résidus *Leu* et *Gly* sont visibles sur les spectres *2D* obtenus. L'analyse de l'intensité des corrélations isolées en fonction du temps de mélange montre la présence d'une interaction intense entre la chaîne principale d'une leucine et la chaîne latérale d'une autre, même à des temps de mélange inférieurs à *50 ms*. Ceci témoigne d'une distance inférieure à 2 \AA . Il pourrait s'agir dans ce cas de deux leucines situées au niveau des hélices cytoplasmiques. La présence d'autres informations pourrait aider à identifier dans un premier temps les résidus impliqués dans ces corrélations, et déterminer la distance les séparant. Ceci pourrait avoir une importance majeure dans la détermination d'informations structurales concernant l'état de la protéine dans le système membranaire utilisé dans les conditions expérimentales définies.

v. Rôle des lipides

L'échantillon *AGKVY-Ec-MscL* a été synthétisé *in vitro* en détergent dans le but d'étudier le rôle de l'environnement lipidique. Le choix des lipides est basé sur les données expérimentales existant sur la protéine. Pour les deux échantillons testés, les spectres obtenus ont montré une différence au niveau du signal, mais aussi au niveau de la résolution de certains pics, permettant d'identifier de nouveaux systèmes de spin des résidus alanines. Dans le cas d'une protéine membranaire comme *Ec-MscL*, voir un effet de la composition lipidique n'est nullement surprenant. En effet, la protéine interagit très fortement avec son environnement hydrophobe. Une bicouche rigide favoriserait les corrélations des résidus rigides qui se trouveraient figés en une conformation stable augmentant ainsi l'intensité des interactions. Une bicouche plus fluide empêcherait cela, car les segments transmembranaires pourraient dans ce cas adopter différentes conformations, condition intéressante pour étudier des résidus situés dans les zones

extramembranaires de la protéine. Des résultats supplémentaires sont nécessaires pour corrélérer la nature des lipides utilisés aux informations spectroscopiques. Un autre test a été effectué en substituant l'eau par de l'eau lourde, mais aucune différence majeure n'est signalée.

Par ailleurs, une étude biophysique détaillée de la distribution de la protéine reconstituée en liposomes (à un ratio lipides/protéine de 10/1 poids/poids) a montré récemment le regroupement des protéines par cluster au sein de la bicouche (Grage *et al.*, 2011). Cela pourrait expliquer la faible dispersion observée au niveau des spectres RMN et l'élargissement des raies. La reconstitution de protéines uniformément ou sélectivement marquées, en présence de protéines non marquées pourrait réduire cet effet par la formation de clusters mixtes, mais provoquerait une diminution du signal liée à celle du pourcentage de protéines marquées dans le rotor.

Les différents échantillons synthétisés ont permis de montrer le rôle que peut jouer la synthèse *in vitro* dans l'allègement des corrélations au niveau des spectres bidimensionnels obtenus sur la protéine membranaire *Ec-MscL* pour son étude par RMN en phase solide. Malgré une résolution du même ordre que celles ayant servi à déterminer la structure de protéines microcristallines ou fibrillaires, les données spectrales ne permettent pas d'obtenir les informations nécessaires pour une protéine membranaire *in situ*. Différentes stratégies ont été proposées au niveau de la biosynthèse, mais les progrès apportés à la spectroscopie de RMN en phase solide constituent un axe essentiel d'amélioration de la qualité des données. La résolution peut être améliorée par l'augmentation de la vitesse de rotation (Bertini *et al.*, 2010 ; Laage *et al.*, 2008) ou en diminuant la température lors de l'acquisition des spectres avec les cryosondes (Rossi *et al.*, 2010) ; l'utilisation de la technique de polarisation dynamique nucléaire ou DNP (pour DYNAMIC NUCLEAR POLARIZATION ; Maly *et al.* 2008) pourrait augmenter le rapport signal/bruit. Cette technique utilise l'effet OVERHAUSER (Overhauser, 1953) par le transfert de polarisation du spin des électrons au noyau étudié (pour des récentes applications de la DNP à des systèmes biologiques, voir Akbey *et al.*, 2010 ; Bajaj *et al.*, 2009 ; Bechniger *et al.*, 2011).

vi. Coût de la *SIV*

Malgré le prix élevé des kits de *SIV* utilisés, cette technique permet de réduire le coût d'un échantillon de protéine membranaire marqué spécifiquement aux ^{13}C et ^{15}N pour un nombre d'acides aminés marqués inférieur à six. Lorsque la protéine est surexprimée en bactéries le rendement est faible, car les substrats marqués aux ^{13}C et ^{15}N du milieu sont utilisés par la bactérie pour tout son métabolisme, dont la surexpression de la protéine. Un marquage sélectif *in vivo* est possible en exploitant les différentes voies de biosynthèse (glycérol, pyruvate, succinate, etc.), mais ce type de marquage augmenterait le coût de l'échantillon, car d'un côté les produits utilisés sont plus chers, et d'un autre côté le rendement de surexpression est plus faible. De même, le marquage spécifique par incorporation d'acides aminés marqués augmente le coût, car les quantités d'acides aminés utilisés est beaucoup plus importante que lors de l'incorporation des acides aminés en *SIV*. En utilisant la *SIV*, même si le prix des acides aminés est élevé, ceux-ci sont utilisés à de faibles quantités, et le rendement d'incorporation avoisine les 50%.

vii. Limites de la *SIV*

Les extraits bactériens utilisés pour effectuer la *SIV* contiennent différentes enzymes du métabolisme des acides aminés. Le marquage de certains résidus peut subir un brouillage métabolique. Ceci peut être observé dans le cas des résidus *Asp*, *Asn*, *Glu*, *Gln* et *Ser*. Il en convient donc d'utiliser une grande quantité de ces acides aminés pour un marquage efficace. Les résidus *Arg*, *Cys*, *Glu*, *Trp* et *Met* peuvent subir des dégradations lors d'une incubation prolongée. De plus, le tampon utilisé dans la *SIV* est à base de glutamate. Il est donc impossible de marquer, avec les kits *RTS*, les résidus *Glu* et *Gln*.

III Conclusion générale

La protéine mécanosensible à large conductance *MscL* est une protéine membranaire impliquée dans l'osmorégulation bactérienne. Lors d'un choc hypo-osmotique, l'entrée massive de l'eau entraîne la courbure de la membrane, ce qui provoque l'ouverture du canal *MscL* pour prévenir l'éclatement de la bactérie. Outre la tension membranaire, l'ouverture de la protéine peut être catalysée par différentes molécules jouant sur la fluidité ou la courbure de la bicouche lipidique. Les études antérieurement effectuées sur la protéine ont prouvé la nécessité de garder la protéine dans son environnement natif pour observer une fonction. Pour une étude structurale de la protéine mécanosensible à large conductance *MscL* dans la bicouche lipidique, la résonance magnétique nucléaire en phase solide paraît indispensable pour l'obtention de données structurales à haute résolution. La voie traditionnelle impliquant une surexpression bactérienne de la protéine avec un marquage uniforme a permis d'obtenir un échantillon pur et fonctionnel. Les données expérimentales obtenues sur cet échantillon montrent une bonne résolution, mais la complexité des spectres est un obstacle majeur pour l'attribution des systèmes de spin et l'obtention de données structurales. L'introduction du système de synthèse acellulaire pour les études structurales a servi dans le cas de la protéine *MscL* à obtenir des échantillons sélectivement marqués, réduisant le nombre de corrélations et améliorant ainsi la lisibilité des spectres. Le premier échantillon synthétisé et marqué uniquement au niveau de deux familles de résidus, les isoleucines et les thréonines, a permis d'obtenir un jeu de données montrant clairement l'intérêt de la stratégie. En effet, malgré la faible amélioration de la résolution, la diminution du nombre de spins observés, grâce au marquage sélectif, a permis d'obtenir des informations structurales, et d'attribuer les trois thréonines présentes dans la protéine.

Différentes stratégies ont ensuite été appliquées pour trouver la bonne combinaison de résidus marqués. La première consiste à marquer des résidus ayant des corrélations théoriques qui ne se recouvrent pas. Pour cela, la stratégie nécessite un modèle structural théorique de la protéine, et un algorithme de prédiction des

déplacements chimiques. Cette stratégie a servi pour la synthèse de l'échantillon marqué au niveau des alanines, lysines, valines et tyrosine. Malgré une nette séparation entre les corrélations des résidus de nature différente, les spectres obtenus n'ont pas permis l'obtention d'informations structurales à cause d'une faible dispersion des déplacements chimiques au sein des acides aminés de même nature. La deuxième stratégie adoptée consistait à rechercher au sein de la séquence, les résidus, qui, par un marquage sélectif, entraînaient la création de paires uniques d'acides aminés marqués, afin d'associer à cette paire une éventuelle corrélation unique au niveau des spectres, permettant donc d'identifier les résidus. Cinq familles de résidus ont été choisies, et un échantillon marqué au niveau des résidus phénylalanines, isoleucines, méthionines, prolines et arginines a été synthétisé. Peu de corrélations inter-résiduelles sont identifiées empêchant l'exploitation de la stratégie. La dernière stratégie de marquage employée consistait en une combinaison des données prédictives et des informations structurales connues de la protéine. Ainsi, les deux familles de résidus montrant le plus de flexibilité au niveau des corrélations théoriques, les aspartates et les sérines, et deux autres familles de résidus présentes dans des zones stratégiques de la protéine, les glycines et les leucines, ont été choisies pour synthétiser un échantillon. Les spectres ont permis d'obtenir les corrélations des glycines et des leucines, et aucune information n'était présente pour les aspartates ou les sérines. Des corrélations inter-résiduelles sont observées, mais leur exploitation nécessite des informations complémentaires. Les données hétéronucléaires auraient été d'une aide cruciale pour tous les échantillons, mais la résolution au niveau de l'azote a empêché l'obtention de ces informations. D'autres tests ont été effectués, montrant un effet dû au changement de la composition de la membrane. L'effet de la température a été également étudié montrant un comportement complexe, entre l'intensification des corrélations des résidus hydrophobes à basses températures, et l'apparition de nouvelles corrélations à hautes températures.

IV Références

- Abdine** A, Verhoeven MV, Warschawski DE. Cell-free expression and labeling strategies for a new decade in solid-state NMR. *N. Biotechnol.* **2011**;28(3):272-276.
- Akbey** Ü, Franks WT, Linden A, et al. Dynamic nuclear polarization of deuterated proteins. *Angew. Chem. Int. Ed. Engl.* **2010**;49(42):7803-7806.
- Bajaj** VS, Mak-Jurkauskas ML, Belenky M, Herzfeld J, Griffin RG. Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization-enhanced solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2009**;106(23):9244-9249.
- Bechinger** B, Resende JM, Aisenbrey C. The structural and topological analysis of membrane-associated polypeptides by oriented solid-state NMR spectroscopy: established concepts and novel developments. *Biophys. Chem.* **2011**;153(2-3):115-125.
- Bertini** I, Emsley L, Lelli M, et al. Ultrafast MAS solid-state NMR permits extensive ¹³C and ¹H detection in paramagnetic metalloproteins. *J. Am. Chem. Soc.* **2010**;132(16):5558-5559.
- Grage** SL, Keleshian AM, Turdzeladze T, et al. Bilayer-mediated clustering and functional interaction of MsCL channels. *Biophys. J.* **2011**;100(5):1252-1260.
- Laage** S, Marchetti A, Sein J, et al. Band-selective ¹H-¹³C cross-polarization in fast magic angle spinning solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2008**;130(51):17216-17217.
- Maly** T, Debelouchina G, Bajaj VS, et al. Dynamic nuclear polarization at high magnetic fields. *J Chem Phys.* **2008**;128(5):052211.
- Overhauser** AW. Polarization of Nuclei in Metals. *Phys. Rev.* **1953**;92(2):411.
- Renault** M, Bos MP, Tommassen J, Baldus M. Solid-State NMR on a Large Multidomain Integral Membrane Protein: The Outer Membrane Protein Assembly Factor BamA. *J. Am. Chem. Soc.* **2011**;133(12):4175-4177.
- Rossi** P, Swapna GVT, Huang YJ, et al. A microscale protein NMR sample screening pipeline. *J. Biomol. NMR.* **2010**;46(1):11-22.

Perspectives

F Perspectives

Le travail présenté dans ce manuscrit traite d'abord de l'application de la synthèse *in vitro* pour l'étude structurale de la protéine membranaire *MscL* par résonance magnétique nucléaire en phase solide, à travers le développement de plusieurs stratégies de marquage sélectif, afin de faciliter l'utilisation de la RMN pour des protéines membranaires *in situ*.

Au niveau de la synthèse *in vitro*, des tests en spectrométrie de masse sont encore en cours, pour quantifier l'incorporation des acides aminés marqués et détecter l'éventuelle présence de protéines issues de synthèse partielle. Pour la synthèse *in vitro* en présence de micelles de détergent, le *triton-x-100* a montré son efficacité pour solubiliser la protéine membranaire, sans perturber la transcription et la traduction dans le milieu acellulaire. Or, ce détergent présente un inconvénient : bien que les billes de polystyrène permettent de l'extraire, l'extraction ne peut se faire complètement (Holloway, 1973). Des molécules de détergent sont probablement présentes dans les protéoliposomes obtenues. Ces impuretés modifient le comportement de la bicouche lipidique en la fluidifiant (Park & Opella, 2010). Deux solutions se présentent face à ce problème : changer de détergent ou changer de méthode d'extraction. Le changement de détergent est faisable, et parmi les détergents testés en synthèse *in vitro* le *n*-dodecyl- β -D-maltoside (*DDM*) présentant le meilleur profil sur les aspects de rendement de synthèse, la pureté du détergent et sa faible concentration micellaire critique pourrait être testé (Berrier *et al.*, 2004). Cependant, l'utilisation du *DDM* rajouterait un coût supplémentaire et non négligeable au protocole de synthèse. L'autre possibilité est de changer de méthode d'extraction ; les méthodes traditionnelles de dialyse ou de dilution présentent les mêmes inconvénients que les *bio beads*. La β -cyclodextrine a été testée, mais aucune donnée n'est disponible concernant son efficacité vis-à-vis du *triton-x-100* (Degrip *et al.*, 1998). Des études récentes ont montré que le cholate de sodium présent dans la solution pendant l'extraction entraîne la disparition du pic correspondant au *triton-x-100* sur un chromatogramme *HPLC* en phase inverse (Varhac *et al.*, 2009).

Après le choix du détergent et la méthode de son extraction, une autre question

revient souvent quant à la nature du lipide utilisé pour la reconstitution, la *DOPC*. Le choix d'étudier la protéine membranaire *MscL* en liposomes de *DOPC* provient des études antérieures (Perozo *et al.*, 2002). La longueur des chaînes carbonées de la molécule permet une insertion efficace de la protéine, et lui permet également d'avoir un comportement variable. En effet, la *DOPC* favorise la conformation fermée de la protéine, et l'addition de molécules de lysophospholipides ou l'application d'une tension membranaire permet l'ouverture du canal. D'autres lipides ont été testés et ont montré un changement au niveau des spectres *RMN*. Cette voie est probablement intéressante à exploiter dans le but d'améliorer la résolution des informations spectroscopiques. La reconstitution de la protéine dans des liposomes de *DPPC* pourrait figer la protéine dans sa conformation fermée, en augmentant probablement la résolution des corrélations issues des résidus intra-membranaires. La *DOPE* pourrait être également testée, étant donné que des récents résultats ont montré que l'insertion de la protéine serait plus efficace en liposomes de *DOPC/DOPE* (30/70) qu'en *DOPC* seule (Grage *et al.*, 2011), comme en témoigne les observations par microscopie confocale. Cet article traite également de l'effet des lipides sur l'organisation de la protéine *MscL* en clusters, ce qui suggère que les interactions entre les protéines pourrait jouer un rôle non négligeable sur l'élargissement des raies au niveau des spectres *RMN*. La reconstitution en liposomes de protéines marquées et non marquées pourrait être envisagée afin de réduire les interactions entre deux pentamères voisins. D'autres environnements membranaires ou mimant la membrane sont également envisageables: les nanodisques (Borch & Hamann, 2009) et les amphipols (Popot *et al.*, 2011). Récemment, le système de nanodisques a été utilisé pour la reconstitution d'une protéine possédant un domaine membranaire et son étude par *RMN* en phase solide (Kijac *et al.*, 2007). Ceci pourrait être appliqué dans le cas de *Ec-MscL*. Idéalement, l'organisation des nanodisques rendrait le système plus ordonné, améliorant ainsi la résolution au niveau des spectres. Les amphipols peuvent être également testés car les complexes formés par ces molécules avec les protéines membranaires supportent la déshydratation et pourraient donc être étudiés à l'état solide.

Ensuite, une fois la protéine reconstituée en liposomes, pour l'acquisition des données spectroscopiques, de nouvelles séquences d'impulsions pourraient être testées en

rotation à l'angle magique, en utilisant par exemple un transfert *INEPT* (pour *INSENSITIVE NUCLEI ENHANCED BY POLARIZATION TRANSFER*) au lieu de la polarisation croisée (Fyfe *et al.*, 1995 ; Riek *et al.*, 1999) ou *DREPT* (pour *DIPOLAR ENHANCED POLARIZATION TRANSFER*) capables de donner une meilleure résolution au niveau des parties flexibles de la protéine (Xu *et al.*, 2010), mais à plus haute température. Une autre technique de RMN en phase solide peut également être utilisée, la rotation à l'angle magique d'un échantillon orienté (Glaubitx & Watts, 1998) ou *MAOSS* (pour *MAGIC ANGLE ORIENTED SAMPLE SPINNING*). Pour cela, les membranes contenant la protéine sont déposées sur des plaques de verre, et les plaques sont empilées à l'intérieur du rotor. Cet arrangement permettrait une amélioration de la résolution. Par ailleurs, malgré le double marquage aux ^{13}C et ^{15}N , les expériences bidimensionnelles utilisant la séquence *specific-CP* n'ont pas abouti à des données exploitables. L'amélioration de la résolution au niveau de la dimension de l'azote apporterait des informations indispensables pour une attribution complète de la protéine (Schuetz *et al.*, 2010). L'utilisation d'une autre séquence d'impulsions pourrait être abordée également.

Quant aux méthodes de marquage par synthèse *in vitro*, différents nouveaux échantillons pourraient être élaborés. La première stratégie employée, et qui a servi à produire l'échantillon (*Ala, Lys, Val, Tyr*)-*Ec-MscL*, a montré des lacunes vis-à-vis du manque d'informations vis-à-vis de la dispersion des signaux au sein de chaque famille d'acides aminés. La stratégie employée pour synthétiser l'échantillon (*Phe, Ile, Met, Pro, Arg*)-*Ec-MscL* pourrait être le point de départ pour une attribution de la protéine. Le premier choix effectué était probablement très ambitieux, car les cinq familles d'acides aminés choisies représentent près du tiers de la protéine. Cependant, par la même stratégie, d'autres combinaisons d'acides aminés présentent des paires uniques à des endroits stratégiques de la protéine. La combinaison des résidus *Glu, Phe, Lys* et *Thr* donnerait accès aux paires $_{54}\text{F}-_{55}\text{K}$, $_{116}\text{T}-_{117}\text{K}$ et $_{123}\text{T}-_{124}\text{E}$. De plus, d'un côté les trois *Thr* sont attribuées, aidant ainsi à attribuer au moins deux résidus ; d'un autre côté, d'autres paires sont présentes $_{5}\text{K}-_{6}\text{E}$, $_{6}\text{E}-_{7}\text{F}$, $_{9}\text{E}-_{10}\text{F}$, $_{106}\text{K}-_{107}\text{E}$, $_{117}\text{K}-_{118}\text{E}$ et $_{130}\text{K}-_{131}\text{E}$. Ainsi six *Glu* (sur huit) et cinq *Lys* (sur neuf) possèdent un résidu voisin marqué. Par ailleurs, cet échantillon ne contient aucun des cinq acides aminés les plus fréquents dans la protéine. Toutes ces

informations séquentielles serviront énormément pour l'attribution des résidus. La dernière stratégie adoptée, combinant les résultats de prédiction de déplacement chimique à l'analyse de la séquence pourrait s'adapter à des études fonctionnelles, en combinant un marquage à une variation de la composition lipidique ou du pH . Le marquage des résidus aux interfaces pourrait servir de témoin pour déterminer l'effet des phospholipides (longueur de la chaîne d'acides gras, nature de la tête polaire, complexité de la bicouche, etc.). Parmi ces résidus, les plus intéressants seraient les *Gly*, *Pro* et *Met* (situés aux interfaces entre la membrane et l'extérieur ou l'intérieur, sans compter les résidus hydrophobes *Ala*, *Val* et *Leu* situés aux interfaces également, mais fortement présents dans la séquence). Les résidus *Arg* et *Lys* pourraient servir de témoin pH (Kloda *et al.*, 2006). Par ailleurs, le dernier chapitre a dévoilé des différences au niveau des spectres de protéine reconstituée dans différents environnements lipidiques. Cette piste pourrait être également exploitée, en la combinant aux marquages présentés ci-dessus. La diminution du rapport lipide/protéine est envisageable, permettant d'un côté d'accroître la quantité de protéine dans le rotor, et favorisant d'un autre côté une organisation cristalline bidimensionnelle qui peut augmenter la résolution.

À moyen terme, l'application des différentes stratégies de marquage *in vivo* pourrait être abordée. Ainsi, l'utilisation du glycérol marqué au niveau des carbones ^{13}C et ^{15}C ou au niveau du carbone ^{13}C pourrait être utilisée (Castellani *et al.*, 2002). Les spectres obtenus sur les différents échantillons obtenus montrent des signaux intenses au niveau de la région des méthyles. Ces groupements très riches en protons contribuent à un effet de diffusion de spin, d'une part à cause de leur proximité des chaînes d'acides gras des phospholipides, également riches en protons, d'autre part à cause de la séquence d'impulsion utilisée, où l'impulsion principale est au niveau des protons. La relaxation est donc plus lente au niveau des protons des groupements méthyles (Horwitz *et al.*, 1972). L'extinction de ces corrélations pourrait donc améliorer la résolution des spectres. Le pyruvate est le précurseur métabolique des résidus *Ala*, *Val*, *Leu* et *Ile* (Ishima *et al.*, 1999), et le groupement méthyle du pyruvate correspond aux groupements méthyles de ces résidus. Le pyruvate deutéré au niveau du groupement méthyle pourrait être utilisé. Cependant, des tests de criblages de conditions de culture seraient nécessaires, car le

pyruvate devrait être utilisé avec une autre source de carbone pour assurer le marquage des autres résidus. Des marquages inverses pourraient être réalisés également par surexpression bactérienne, avec comme source de carbone du glucose marqué au ^{13}C , en présence d'acides aminés non marqués (Krishnarjuna *et al.*, 2011 ; Vuister *et al.*, 1994). Finalement, une autre idée pouvant améliorer la résolution au niveau des spectres serait de co-reconstituer des monomères marqués et d'autres non marqués de la protéine *Ec-MscL*. Cela pourrait réduire les interactions entre résidus voisins mais appartenant à deux monomères différents (Grage *et al.*, 2011).

À long terme, une fois l'attribution effectuée, le calcul des distances à partir d'interactions intra- et inter-résiduelles contribuera à la résolution de la structure de la protéine. Des méthodes pourraient être développées ensuite pour favoriser la conformation ouverte de la protéine, pour essayer de contribuer par des informations à haute résolution *in situ* de comprendre le mécanisme d'ouverture de la protéine. Pour cela, l'utilisation de lysophospholipides pourrait être intéressante (Perozo *et al.*, 2002). L'étude de certains mutants pourrait également apporter des informations sur des conformations intermédiaires de la protéine, car la mutation de certains résidus a montré une stabilisation de conformations intermédiaires, comme dans le cas de l'étude structurale par microscopie électronique réalisée sur le mutant G22N (Yoshimura *et al.*, 2008).

L'utilisation de peptides amphipathiques s'insérant partiellement dans la membrane pourrait être envisagée également, comme par exemple, les peptides antimicrobiens (Su *et al.*, 2011), les peptides participant aux mécanismes d'endocytose (Cui *et al.*, 2011) mais aussi des peptides viraux (Shchelokovskyy *et al.*, 2011). La déshydratation des liposomes pourrait agir également sur les protéines et provoquer leur ouverture (Mallikarjunaiah *et al.*, 2011).

Toutes les stratégies développées au cours de ce travail, et exposées dans ce manuscrit pourraient servir de point de départ pour l'étude d'autres protéines membranaires dont l'activité est régie essentiellement par la membrane lipidique.

Références

- Berrier** C, Park K-H, Abes S, et al. Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry*. **2004**;43:12585–12591.
- Borch** J, Hamann T. The nanodisc: a novel tool for membrane protein studies. *Biol. Chem.* **2009**;390(8):805-814.
- Castellani** F, van Rossum B, Diehl A, et al. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*. **2002**;420:98–102.
- Cui** H, Lyman E, Voth GA. Mechanism of membrane curvature sensing by amphipathic helix containing proteins. *Biophys. J.* **2011**;100(5):1271-1279.
- Degrip** WJ, Vanoostrum J, Bovee-Geurts PH. Selective detergent-extraction from mixed detergent/lipid/protein micelles, using cyclodextrin inclusion compounds: a novel generic approach for the preparation of proteoliposomes. *Biochem J.* **1998**;330(Pt 2):667-674.
- Fyfe** CA, Wongmoon KC, Huang Y, Grondey H. INEPT experiments in solid-state NMR. *J. Am. Chem. Soc.* **1995**;117(41):10397-10398.
- Glaubit**z C, Watts A. Magic angle-oriented sample spinning (MAOSS): A new approach toward biomembrane studies. *J. Magn. Reson.* **1998**;130(2):305-316.
- Grage** SL, Keleshian AM, Turdzeladze T, et al. Bilayer-mediated clustering and functional interaction of MscL channels. *Biophys. J.* **2011**;100(5):1252-1260.
- Holloway** PW. A simple procedure for removal of triton X-100 from protein samples. *Anal. Biochem.* **1973**;53(1):304-308.
- Horwitz** AF, Horsley WJ, Klein MP. Magnetic Resonance Studies on Membrane and Model Membrane Systems: Proton Magnetic Relaxation Rates in Sonicated Lecithin Dispersions. *Proc. Natl. Acad. Sci. U.S.A.* **1972**;69(3):590-593.
- Ishima** R, Louis JM, Torchia DA. Transverse C-13 relaxation of CHD2 methyl isotopomers to detect slow conformational changes of protein side chains. *J. Am. Chem. Soc.* **1999**;121(49):11589-11590.
- Kijac** AZ, Li Y, Sligar SG, Rienstra CM. Magic-angle spinning solid-state NMR spectroscopy of nanodisc-embedded human CYP3A4. *Biochemistry*. **2007**;46(48):13696-13703.
- Kloda** A, Ghazi A, Martinac B. C-terminal charged cluster of MscL, RKKEE, functions as a pH sensor. *Biophys. J.* **2006**;90:1992–1998.
- Krishnarjuna** B, Jaipuria G, Thakur A, D'Silva P, Atreya HS. Amino acid selective unlabeled for sequence specific resonance assignments in proteins. *J. Biomol. NMR.* **2011**;49(1):39-51.
- Mallikarjunaiah** KJ, Leftin A, Kinnun JJ, et al. Solid-State H-2 NMR Shows Equivalence of Dehydration and Osmotic Pressures in Lipid Membrane Deformation. *Biophys. J.* **2011**;100(1):98-107.
- Park** SH, Opella SJ. Triton X-100 as the « short-chain lipid » improves the magnetic alignment and stability of membrane proteins in phosphatidylcholine bilayers for oriented-sample solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2010**;132(36):12552-12553.
- Perozo** E, Cortes DM, Somporapisut P, Kloda A, Martinac B. Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature*. **2002**;418(6901):942-948.
- Popot** J-L, Althoff T, Bagnard D, et al. Amphipols from A to Z. *Annu Rev Biophys.* **2011**;40:379-408.

- Riek R**, Wider G, Pervushin K, Wüthrich K. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proc. Natl. Acad. Sci. U.S.A.* **1999**;96(9):4918-4923.
- Schuetz A**, Wasmer C, Habenstein B, et al. Protocols for the sequential solid-state NMR spectroscopic assignment of a uniformly labeled 25 kDa protein: HET-s(1-227). *Chembiotech.* **2010**;11(11):1543-1551.
- Shchelokovskyy P**, Tristram-Nagle S, Dimova R. Effect of the HIV-1 fusion peptide on the mechanical properties and leaflet coupling of lipid bilayers. *New J. Phys.* **2011**;13(2):025004.
- Su Y**, Waring AJ, Ruchala P, Hong M. Structures of β -Hairpin Antimicrobial Protegrin Peptides in Lipopolysaccharide Membranes: Mechanism of Gram Selectivity Obtained from Solid-State Nuclear Magnetic Resonance. *Biochemistry.* **2011**;50(12):2072-2083.
- Varhac R**, Robinson NC, Musatov A. Removal of bound Triton X-100 from purified bovine heart cytochrome bc1. *Anal. Biochem.* **2009**;395(2):268-270.
- Vuister GW**, Kim SJ, Wu C, Bax A. 2D and 3D NMR study of phenylalanine residues in proteins by reverse isotopic labeling. *J. Am. Chem. Soc.* **1994**;116(20):9206-9210.
- Xu J**, Soong R, Im S-C, Waskell L, Ramamoorthy A. INEPT-based separated-local-field NMR spectroscopy: a unique approach to elucidate side-chain dynamics of membrane-associated proteins. *J. Am. Chem. Soc.* **2010**;132(29):9944-9947.
- Yoshimura K**, Usukura J, Sokabe M. Gating-associated conformational changes in the mechanosensitive channel MscL. *Proc. Natl. Acad. Sci. U.S.A.* **2008**;105:4033-4038.

Annexes

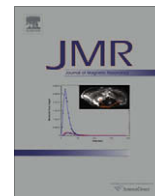
I Structural study of the membrane protein MscL using cell-free expression and solid-state NMR

Alaa Abdine^a, Michiel A. Verhoeven^{a, b}, Kyu-Ho Park^{c, 1}, Alexandre Ghazi^c, Eric Guittet^b, Catherine Berrier^c, Carine Van Heijenoort^b and Dror E. Warschawski^a

^a UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France

^b UPR 2301, CNRS, ICSN, F-91190 Gif-sur-Yvette, France

^c UMR 8619, CNRS and Université Paris-Sud 11, IBBMC, F-91405 Orsay, France



Communication

Structural study of the membrane protein MscL using cell-free expression and solid-state NMR

Alaa Abdine^a, Michiel A. Verhoeven^{a,b}, Kyu-Ho Park^{c,1}, Alexandre Ghazi^c, Eric Guittet^b, Catherine Berrier^c, Carine Van Heijenoort^b, Dror E. Warschawski^{a,*}^a UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France^b UPR 2301, CNRS, ICSN, F-91190 Gif-sur-Yvette, France^c UMR 8619, CNRS and Université Paris-Sud 11, IBBMC, F-91405 Orsay, France

ARTICLE INFO

Article history:

Received 25 December 2009

Revised 1 February 2010

Available online 11 February 2010

Keywords:

Solid-state NMR spectroscopy

Magic-angle spinning

Membrane proteins

Isotope labeling

Structural biology

ABSTRACT

High-resolution structures of membrane proteins have so far been obtained mostly by X-ray crystallography, on samples where the protein is surrounded by detergent. Recent developments of solid-state NMR have opened the way to a new approach for the study of integral membrane proteins *inside* a membrane. At the same time, the extension of cell-free expression to the production of membrane proteins allows for the production of proteins tailor made for NMR. We present here an *in situ* solid-state NMR study of a membrane protein selectively labeled through the use of cell-free expression. The sample consists of MscL (mechano-sensitive channel of large conductance), a 75 kDa pentameric α -helical ion channel from *Escherichia coli*, reconstituted in a hydrated lipid bilayer. Compared to a uniformly labeled protein sample, the spectral crowding is greatly reduced in the cell-free expressed protein sample. This approach may be a decisive step required for spectral assignment and structure determination of membrane proteins by solid-state NMR.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

In recent years, solid-state NMR has shown its efficiency on studying microcrystalline soluble proteins in the solid-state [1–3], and it is now applied to molecules that must be studied by solid-state NMR in their native state, like fibrils [4–6], macromolecular assemblies [7] and membrane proteins in a lipid environment [8–15]. In the latter case, spectral resolution remains a crucial issue, first because the local order of proteins inserted in lipid bilayers is often not as high as that of micro-crystals and second, because membrane proteins are often composed of repetitive hydrophobic amino acids in α -helices. Spectral congestion can be overcome by removing a number of resonances from the NMR spectra, by isotopically labeling only a subset of the protein nuclei, through different biochemical methods [1,4,13,16,17]. Another major challenge is the production of large quantity of functional proteins. The cytotoxicity of overexpression is aggravated, in the case of membrane proteins, by the limited membrane surface available. The recently developed cell-free expression (CFE) approach for protein production is particularly well adapted for

selective labeling of amino acids, and it has proven its efficiency in solution-state NMR studies [18–21]. It is applied here to a solid-state NMR study, where it will be even more useful due to crucial spectral crowding issues.

In this communication, we also present the first NMR study of the mechano-sensitive channel of large conductance from *E. coli* (MscL), a 75 kDa homopentameric membrane protein that functions as a pressure valve in the bacterial inner membrane. While the structure of a closed form of the analogous Tb-MscL pentamer from *Mycobacterium tuberculosis* has been determined by X-ray crystallography [22], the opening/closing of the channel can only occur when the protein is embedded in a lipid bilayer [23]. Furthermore, the recent structure determination of a tetrameric analog from *Staphylococcus aureus* [24] shows an unexpected structural variety of these ion channels, and has revived the interest for the structure determination of membrane proteins in their native environment.

2. Results

Small scale cell-free expression of functional MscL in the presence of detergents has been optimized before [25]. By scaling up this protocol, we have synthesized several milligrams of MscL ¹³C, ¹⁵N-labeled on the 16 isoleucine and on the 3 threonine residues. The high repetition of isoleucines makes them a good probe

* Corresponding author. Fax: +33 1 58 41 50 24.

E-mail address: Dror.Warschawski@ibpc.fr (D.E. Warschawski).¹ Present address: URA 3015, CNRS and Institut Pasteur, Département de Virologie, 25 rue du Dr. Roux, F-75015 Paris, France.

for resolution, whereas the low amount of threonines is a good test for sensitivity. For comparison, we have also prepared a uniformly ^{13}C , ^{15}N -labeled MscL sample by expression in *E. coli*. After purification, reconstitution of MscL is feasible in a large number of lipid mixtures and we have obtained NMR spectra in DOPC, DOPC/DPPC and soybean asolectin, with several lipid-to-protein ratios. Since the resolution was the same (data not shown), the data presented here corresponds to MscL reconstituted in a DOPC membrane, in which it is as active as in a natural membrane [26].

The state of our proteoliposomes was checked by several approaches: (i) crosslinking experiments with formaldehyde and disuccinyl suberate showed the integrity of the pentamer on SDS-PAGE gel; (ii) sucrose flotation on a discontinuous sucrose gradient of 50/30/20/0% (before and after washing them with urea, to exclude proteins that would merely sit on the membrane surface), resulted in a single band at the interface of 20–30% corroborating protein charged liposomes; (iii) ^{31}P NMR at room temperature resulted in a typical hydrated fluid bilayer spectrum; (iv) electron microscopy resulted in a typical proteoliposome picture with no signs of protein precipitation; (v) the functionality of the proteins was assayed using standard patch-clamp methods [25]; (vi) finally, ^1H and ^{13}C NMR spectra were acquired prior to and after each NMR set of experiments and showed no sign of dehydration or lipid degradation.

With a sample typically containing 3 mg of proteins, 12 mg of lipids and 50% (w/w) of water, a 1D ^{13}C solid-state NMR spectrum can be obtained in about an hour (Fig. 1). Spectral crowding in the C' (160–180 ppm) and $\text{C}\alpha$ (40–60 ppm) regions is important (Fig. 1a). The 1D spectrum of MscL specifically labeled through CFE, recorded and processed with identical parameters is shown in Fig. 1b. The sensitivity of this spectrum is satisfactory since threonine residues are visible in the $\text{C}\alpha\text{C}\beta$ region (55–70 ppm). In addition, crowding reduction clearly improves the spectral reading, as predicted by the simulations (see Fig. S1 in the Supplementary Data, SD).

Protein structural studies by solid-state NMR rely on 2D ^{13}C – ^{13}C spectra where the first limiting step consists in assigning resonances coming from each residue. Fig. 2 shows the comparison of 2D ^{13}C – ^{13}C solid-state NMR spectra, obtained through the DARR experiment [27], on both samples described previously: MscL either uniformly ^{13}C and ^{15}N labeled produced in *E. coli* (Fig. 2a) or selectively ^{13}C and ^{15}N labeled by CFE on all isoleucine and threonine residues (Fig. 2b), both reconstituted in a hydrated lipid bilayer. The current spectral resolution does not allow the assignment of over 700 different carbons, as seen in Fig. 2a. As expected, since the specifically labeled sample contains only 108 ^{13}C spins, the corresponding spectra appear spectacularly less congested. With the same acquisition parameters, the $\text{C}\alpha\text{C}\beta$ threonine resonances are visible in Fig. 2b, while they are buried in the noise in Fig. 2a (they can be distinguished with longer acquisition times, see Fig. S2 in the SD).

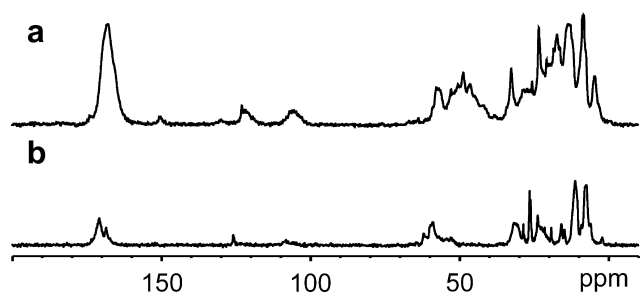


Fig. 1. 176 MHz 1D ^{13}C CP MAS NMR spectra of labeled MscL in unlabeled DOPC liposomes, spinning at 11,000 Hz and 258 K: (a) ^{13}C , ^{15}N labeled and (b) (Ile, Thr)- ^{13}C , ^{15}N labeled.

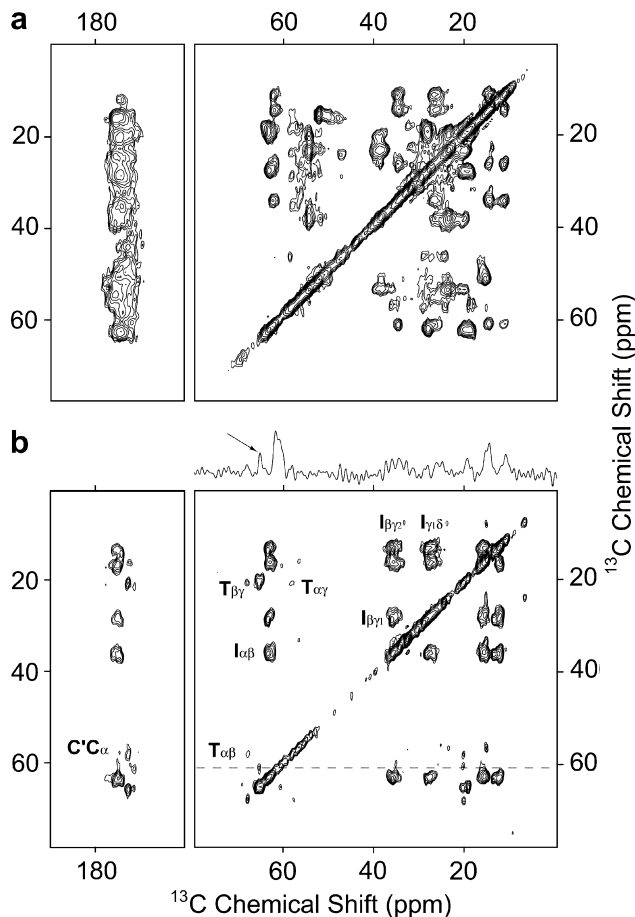


Fig. 2. 2D ^{13}C – ^{13}C DARR NMR (50 ms mixing time, 14 h of acquisition) spectra of MscL/DOPC: (a) ^{13}C , ^{15}N labeled protein and (b) (Ile, Thr)- ^{13}C , ^{15}N labeled protein, with a cross section through the threonine $\alpha\beta$ region shown on top. The arrow indicates a threonine cross signal which is 140 Hz wide (0.80 ppm). Some unlabeled lipid signals appear on the diagonal, mostly around 30 ppm.

The DARR experiment was performed at several mixing times between 5 and 500 ms to monitor the build-up of each correlation. At 500 ms, long range cross signals between isoleucines and threonines could be observed in the selectively labeled sample (see Fig. 3). Since there is no adjacent isoleucine–threonine pair in MscL, this opens the way for future long distance measurements and structure determination.

3. Discussion

In vitro synthesis offers a good alternative to biosynthesis. With a good yield and a very simple purification step, our membrane protein can be cell-free expressed in 2 days, as opposed to 8 days in *E. coli*. Moreover, for a moderate number of labeled amino acids (up to 5), the cost of a CFE sample is comparable to that of a uniformly labeled sample produced *in vivo*, and much less than that of a specifically labeled *in vivo* sample. Finally, CFE bypasses cytotoxicity and membrane targeting issues, as well as greatly minimizing amino acid scrambling.

In all cases, our proteins are well folded, as shown by various biochemical and biophysical tests described above in the Section 2. Our sample is also homogeneous, as confirmed by the presence of narrow individual resonances (see for example the cross sections at 60.7 ppm on Fig. 2b and Fig. S2 in the SD), comparable to recently published spectra [13]. There are very few such individual resonances because the hydrophobic residues are all in a similar

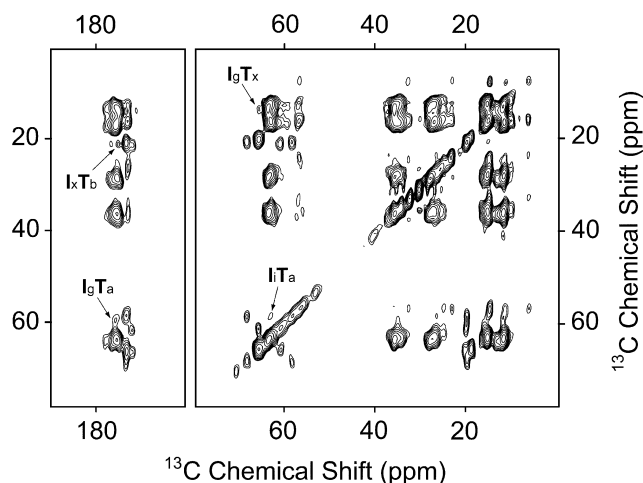


Fig. 3. 2D ^{13}C - ^{13}C DARR NMR (500 ms mixing time, total experiment time: 56 h) spectra of (Ile, Thr)- ^{13}C , ^{15}N labeled MscL/DOPC. Some long range cross signals are indicated on the spectrum (Ix and Tx indicate that the specific isoleucine or threonine could not be identified).

chemical environment. The resolution is sufficient to fully characterize the 3 threonine sets of spin systems. In the case of the 16 isoleucine residues of MscL, up to 10 different sets of spin systems can be distinguished (see Figs. S3 and S4 and Tables S1 and S2 in the SD).

Crowding reduction by CFE results from the reduced number of spins and spin couplings. For instance, the specifically labeled sample contains less ^{13}C - ^{13}C pairs of spins from neighboring residues, which modifies the interactions, spin diffusion and relaxation experienced by the remaining spins. Long-range contacts in the specifically labeled sample will thereby suffer less from relayed polarization transfers in correlation experiments, hence resulting in potentially more accurate deduced distances.

MscL is a good candidate for testing various assignment strategies in solid-state NMR because it contains a high percentage (40%) of hydrophobic residues. Clearly, sequential information is required to assign the resonances. While isotopic labeling of a small number of amino acids reduces the probability of adjacent amino acid pairs that are required for sequential assignment, larger well-chosen sets of labeled amino acids will allow access to complete assignment, using ^{13}C and ^{15}N NMR. Several ways can be followed, inspired from solution-state NMR studies: the combinatorial approach suggested by the Otting group [18] or the “dual combinatorial” approach of the Dotsch group [20]. Both approaches will have to be adapted to the particular case of solid-state NMR where carbons and nitrogens are observed, rather than protons, and where the dipolar interaction dominates the scalar one. Another labeling strategy could focus solely on amino acids that are important for the channel function [19,28], and CFE will become a necessity for the study of potential open mutants that would be lethal for bacterial growth. Finally, knowing the structure of a model [29] or an analog of our protein, an approximate NMR spectrum can be simulated, using for example the SPARTA package [30]. This simulated spectrum can, in turn, help us choose an efficient labeling strategy that would minimize the spectral overlap [31]. Regardless, the required specific labeling is most efficiently achieved by CFE and we demonstrate the feasibility of this strategy in combination with solid-state NMR.

4. Conclusions

Combined with new technological developments, such as Dynamic Nuclear Polarization, cryo-NMR or very high fields, solid-state NMR is currently adapting to the study of larger molecules,

where both high-resolution and high sensitivity are still challenges to attain. We have shown here that our CFE approach is efficient to simplify an overcrowded NMR spectrum and that it provides a useful approach for the first steps of assigning membrane proteins by solid-state NMR. It is competitive and complementary to other labeling schemes that have been suggested to alleviate assignment obstacles. We also expect our strategy to be very helpful to extract secondary structure and dynamics information, as well as distance constraints and thus for structure determination by solid-state NMR. Solid-state NMR, together with novel biochemistry techniques for protein isotopic labeling, is an efficient technique for studying the ion channel MscL inside a hydrated lipid bilayer. Such an approach is very general and should therefore be applicable to other channels or membrane protein structural studies.

5. Experimental

5.1. Membrane protein expression

The *mscL* gene with a C-terminal His₆-Tag, cloned under the T7 promoter into pVEX2.3 plasmid was used [25]. The recombinant plasmid was introduced in the *E. coli* BL21(λDE3) strain (Stratagene) by electroporation. Precultures were grown overnight in LB medium, which was removed by centrifugation (5000g, 4 °C, 20 min) before inoculation of selective M9 medium (3 L with 100 μg/mL ampicillin) containing $^{13}\text{C}_6$ -glucose (3 g/L) and ^{15}N -ammonium chloride (1 g/L, Spectra Stable Isotopes). Cultures were grown in shaking bottles at 37 °C. When the optical density at 600 nm reached 0.6–0.8, protein expression was induced for 3 h by 1 mM IPTG (Sigma–Aldrich). Cells were harvested by centrifugation and stored overnight at –20 °C. The bacterial pellet was resuspended in buffer (50 mM NaH₂PO₄, 10 mM NaCl, 2 mM MgSO₄, 5% sucrose, 1 mM phenylmethylsulphonyl fluoride, 10 μg/mL DNase, 2 mM mercapto ethanol, 1 mM dithiothreitol), cells were broken using a French press at 10⁴ psi (twice). Cell debris was removed by centrifugation (5000g, 4 °C, 20 min) and membranes were isolated by ultra-centrifugation (300,000g, 4 °C, 30 min). The membrane pellet was resuspended for 30 min at room temperature in buffer A (50 mM Na₂HPO₄, pH 7, 200 mM NaCl) containing 2% Triton X-100 and 10 mM imidazole. Unsolubilized material was discarded by a second ultra-centrifugation (300,000g, 4 °C, 30 min) and the supernatant was immediately purified on an affinity column.

In vitro expression of MscL was performed with the RTS9000 kit (Roche) in combination with the Amino Acid Sampler (Roche) as follows: 150 μg of plasmid were incubated for 23 h at 30 °C in the presence of the lysate containing 31 and 35 mg of uniformly ^{13}C , ^{15}N -labeled Thr and Ile respectively (Spectra Stable Isotopes), while the remaining 18 unlabeled amino acids were added from the Amino Acid Sampler. Although a large variety of non ionic detergents could be used [29], Triton X-100 was added in a final concentration of 0.2% in a final reaction volume of 10 mL. At the end of the run, the reaction mixture was diluted 5 times in buffer A containing 1% Triton and centrifuged (5000g, 4 °C, 15 min) to remove precipitates before FPLC purification. Recent progress in CFE efficiency suggests that the amount of amino acids added could be cut by a factor of 6 (A. Pedersen, private communication), and the lysate could be home made, reducing the cost by an additional significant factor.

5.2. Protein purification and membrane reconstitution

FPLC chromatography of the His₆-tagged MscL was performed using a 1 mL Ni²⁺ charged HiTrap chelating HP column (GE). After removal of unbound material, Triton X-100 concentration was adjusted to 0.2% if necessary. Unspecifically bound material was

washed out using buffer A supplemented with 100 mM imidazole, and MscL was eluted at a concentration of 500 mM imidazole. In order to reduce the imidazole concentration, the fractions containing the protein were pooled and dialysed twice against 1 L of buffer B (10 mM HEPES KOH, pH 7.4, 100 mM KCl, 0.2% Triton X-100), using a dialysis cassette (Pierce) with a cut-off value of 10 kD. The amount of produced proteins was quantified with the Micro-BCA kit (Pierce) and indicated a yield of ~ 1 mg/L for the sample produced *in vivo* and 800 $\mu\text{g}/\text{mL}$ for the sample produced *in vitro*.

The buffer containing purified MscL was combined with buffer B containing DOPC/Triton mixed micelles to achieve a final lipid/detergent ratio of 1/3 and a lipid/protein ratio of 4/1 (w/w). The detergent (~ 12 mg) was then removed by addition of twice 120 mg of SM2-BioBeads as described before [32]. The proteoliposomes were pelleted (300,000 g, 4 °C, 30 min) and transferred into 50 μL 4 mm NMR rotors (Bruker) and stored at -20 °C until used for data acquisition. Several other sample conditions were tested, using DOPC/DPPC:1/1 (Avanti) or soybean asolectin (Fluka), and lipid/protein ratios from 4/1 to 1/1 (w/w). The phospholipid content of the sample was checked by the method of Rouser [33]. Water content was estimated by weighing samples produced in parallel prior to and after 16 h in vacuum.

5.3. NMR

Solid-state NMR experiments were run on a Bruker 700 MHz spectrometer using a triple resonance 4 mm MAS probe. ^{13}C chemical shifts were referenced relative to adamantane as a secondary reference [34]. Reducing the temperature may improve spectral resolution and sensitivity by increasing Boltzmann polarization, freezing molecular motion and improving the efficiency of sample-spinning, proton decoupling and cross polarization. On the other hand, dipolar couplings and T_1 relaxation times are increased, and multiple conformations are frozen out, resulting in potentially broader lines [8,10,14]. Samples were maintained at 258 K for the duration of the measurements at a sample-spinning rate of 11,000 Hz (± 3 Hz). The temperature inside the spinning rotor was checked by measuring the transition temperature of DOPC in the same conditions. Several experiments were performed at higher temperatures (288 K), and did not give a better spectral resolution (data not shown). At comparable resolution, experiments were therefore performed at a lower temperature to extend the lifetime of the protein.

1D ^{13}C MAS NMR experiments were performed with ramped CP (1.5 ms contact time and rf power of 50 kHz), TPPM ^1H decoupling (rf power of 100 kHz), 1024 scans and 4 s repetition time. 2D ^{13}C - ^{13}C correlation experiments were performed using the analogous DARR or PDSO experiments [27,35,36] with similar ramped CP and TPPM ^1H decoupling during indirect evolution and acquisition periods. ^1H and ^{13}C rf powers were 100 and 50 kHz respectively, except for the ^1H DARR mixing pulse set at 11 kHz (equal to the spinning rate), and varied between 5 and 500 ms. Sampling parameters were 10 μs for 2048 points (direct dimension) by 10 μs for 256 rows (indirect dimension), with 64 scans per row and 3 s repetition time. Total time was approximately 14 h. A few experiments were conducted with 256 scans per row, amounting to 56 h acquisition time. Data were processed with nmrPipe [37], employing zero filling and 90° shifted sine bell apodization in the indirect dimension, 10 Hz exponential line broadening in the direct dimension and automatic baseline correction in both dimensions. Only the positive contours are displayed.

Acknowledgments

This work was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (to A.A.), grants and

post-doctoral fellowships (to M.V. and K.-H.P.) by the CNRS (UMR 7099, UPR 2301, UMR 8619), the ANR (ANR-06-JCJC0014), the Université Paris Diderot and the Université Paris-Sud. We thank Daniel Lévy for the electron microscopy, Cathy Etchebest for the molecular dynamics simulations, Philippe Devaux and Francesca Zito for stimulating scientific discussions.

Appendix A. Supplementary data

1D ^{13}C NMR simulations of MscL (Fig. S1), long 2D ^{13}C - ^{13}C PDSO NMR spectrum of U- ^{13}C labeled MscL (Fig. S2), zooms in the 2D ^{13}C - ^{13}C DARR NMR spectra of (Ile, Thr)- ^{13}C labeled MscL (Figs. S3 and S4), and partial chemical shift characterization of (Ile, Thr)- ^{13}C labeled MscL (Tables S1 and S2).

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmr.2010.02.003.

References

- [1] F. Castellani, B. Van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy, *Nature* 420 (2002) 98–102.
- [2] A. Loquet, B. Bardiaux, C. Gardiennet, C. Blanchet, M. Baldus, M. Nilges, T. Malliavin, A. Bockmann, 3D structure determination of the Crh protein from highly ambiguous solid-state NMR restraints, *J. Am. Chem. Soc.* 130 (2008) 3579–3589.
- [3] L.J. Sperling, A.J. Nieuwkoop, A.S. Lipton, D.A. Berthold, C.M. Rienstra, High resolution NMR spectroscopy of nanocrystalline proteins at ultra-high magnetic field, *J. Biomol. NMR* 46 (2010) 149–155.
- [4] A. Goldbourn, L.A. Day, A.E. McDermott, Assignment of congested NMR spectra: carbonyl backbone enrichment via the Entner–Doudoroff pathway, *J. Magn. Reson.* 189 (2007) 157–165.
- [5] H. Heise, Solid-state NMR spectroscopy of amyloid proteins, *ChemBioChem* 9 (2008) 179–189.
- [6] A. Lange, Z. Gattin, H. Van Melckebeke, C. Wasmer, A. Soragni, W.F. van Gunsteren, B.H. Meier, A combined solid-state NMR and MD characterization of the stability and dynamics of the HET-s(218–289) prion in its amyloid conformation, *ChemBioChem* 10 (2009) 1657–1665.
- [7] S. Sun, A. Siglin, J.C. Williams, T. Polenova, Solid-state and solution NMR studies of the CAP-Gly domain of mammalian dynactin and its interaction with microtubules, *J. Am. Chem. Soc.* 131 (2009) 10113–10126.
- [8] M. Hiller, L. Krabben, K.R. Vinothkumar, F. Castellani, B.J. Van Rossum, W. Kuhlbrandt, H. Oschkinat, Solid-state magic-angle spinning NMR of outer-membrane protein G from *Escherichia coli*, *ChemBioChem* 6 (2005) 1679–1684.
- [9] A. Lange, K. Giller, S. Hornig, M.F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR, *Nature* 440 (2006) 959–962.
- [10] H.L. Frericks, D.H. Zhou, L.L. Yap, R.B. Gennis, C.M. Rienstra, Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: *E. coli* cytochrome bo3 oxidase, *J. Biomol. NMR* 36 (2006) 55–71.
- [11] Y. Li, D.A. Berthold, H.L. Frericks, R.B. Gennis, C.M. Rienstra, Partial (^{13}C) and (^{15}N) chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy, *ChemBioChem* 8 (2007) 434–442.
- [12] S. Ganapathy, A.J. van Gammeren, F.B. Hulsbergen, H.J.M. de Groot, Probing secondary, tertiary, and quaternary structure along with protein-cofactor interactions for a helical transmembrane protein complex through ^1H spin diffusion with MAS NMR spectroscopy, *J. Am. Chem. Soc.* 129 (2007) 1504–1505.
- [13] M. Eitzkorn, S. Martell, O.C. Andronesi, K. Seidel, M. Engelhard, M. Baldus, Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy, *Angew. Chem. Int. Ed. Engl.* 46 (2007) 459–462.
- [14] S.D. Cady, T.V. Mishanina, M. Hong, Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding, *J. Mol. Biol.* 385 (2009) 1127–1141.
- [15] L. Shi, M.A. Ahmed, W. Zhang, G. Whited, L.S. Brown, V. Ladizhansky, Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump—structural insights, *J. Mol. Biol.* 386 (2009) 1078–1093.
- [16] M. Hong, Determination of multiple-torsion angles in proteins by selective and extensive ^{13}C labeling and two-dimensional solid-state NMR, *J. Magn. Reson.* 139 (1999) 389–401.
- [17] M. Hiller, V.A. Higman, S. Jehle, B.J. van Rossum, W. Kuhlbrandt, H. Oschkinat, [^{13}C]-labeling of aromatic residues—getting a head start in the magic-angle-spinning NMR assignment of membrane proteins, *J. Am. Chem. Soc.* 130 (2008) 408–409.

- [18] K. Ozawa, P.S. Wu, N.E. Dixon, G. Otting, N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes, *FEBS J.* 273 (2006) 4154–4159.
- [19] I. Lehner, D. Basting, B. Meyer, W. Haase, T. Manolikas, C. Kaiser, M. Karas, C. Glaubitz, The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric, *J. Biol. Chem.* 283 (2008) 3281–3288.
- [20] S. Sobhanifar, S. Reckel, F. Junge, D. Schwarz, L. Kai, M. Karbyshev, F. Lohr, F. Bernhard, V. Dotsch, Cell-free expression and stable isotope labelling strategies for membrane proteins, *J. Biomol. NMR* 46 (2010) 33–43.
- [21] H.J. Kim, S.C. Howell, W.D. Van Horn, Y.H. Jeon, C.R. Sanders, Recent advances in the application of solution NMR spectroscopy to multi-span integral membrane proteins, *Prog. Nucl. Magn. Reson. Spectrosc.* 55 (2009) 335–360.
- [22] G. Chang, R.H. Spencer, A.T. Lee, M.T. Barclay, D.C. Rees, Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel, *Science* 282 (1998) 2220–2226.
- [23] A.M. Powl, J.M. East, A.G. Lee, Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL, *Biochemistry* 47 (2008) 12175–12184.
- [24] Z. Liu, C.S. Gandhi, D.C. Rees, Structure of a tetrameric MscL in an expanded intermediate state, *Nature* 461 (2009) 120–124.
- [25] C. Berrier, K.H. Park, S. Abes, A. Bibonne, J.M. Betton, A. Ghazi, Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent, *Biochemistry* 43 (2004) 12585–12591.
- [26] E. Perozo, D.M. Cortes, P. Sompornpisut, A. Kloda, B. Martinac, Open channel structure of MscL and the gating mechanism of mechanosensitive channels, *Nature* 418 (2002) 942–948.
- [27] K. Takegoshi, S. Nakamura, T. Terao, ^{13}C - ^1H dipolar-assisted rotational resonance in magic-angle spinning NMR, *Chem. Phys. Lett.* 344 (2001) 631–637.
- [28] Y. Li, R. Wray, C. Eaton, P. Blount, An open-pore structure of the mechanosensitive channel MscL derived by determining transmembrane domain interactions upon gating, *FEBS J.* 23 (2009) 2197–2204.
- [29] H. Valadie, J.J. Lacapere, Y.H. Sanejouand, C. Etchebest, Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis, *J. Mol. Biol.* 332 (2003) 657–674.
- [30] Y. Shen, A. Bax, Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology, *J. Biomol. NMR* 38 (2007) 289–302.
- [31] M.J. Sweredoski, K.J. Donovan, B.D. Nguyen, A.J. Shaka, P. Baldi, Minimizing the overlap problem in protein NMR: a computational framework for precision amino acid labeling, *Bioinformatics* 23 (2007) 2829–2835.
- [32] D. Levy, A. Bluzat, M. Seigneuret, J.L. Rigaud, A systematic study of liposome and proteoliposome reconstitution involving Bio-Bead-mediated Triton X-100 removal, *Biochim. Biophys. Acta* 1025 (1990) 179–190.
- [33] G. Rouser, S. Fkeischer, A. Yamamoto, Two dimensional thin layer chromatographic separation of polar lipids and determination of phospholipids by phosphorus analysis of spots, *Lipids* 5 (1970) 494–496.
- [34] C.R. Morcombe, K.W. Zilm, Chemical shift referencing in MAS solid state NMR, *J. Magn. Reson.* 162 (2003) 479–486.
- [35] N.M. Szeverenyi, M.J. Sullivan, G.E. Maciel, Observation of spin exchange by two-dimensional Fourier transform ^{13}C cross polarization-magic-angle spinning, *J. Magn. Reson.* 47 (1982) 462–475.
- [36] A. Grommek, B.H. Meier, M. Ernst, Distance information from proton-driven spin diffusion under MAS, *Chem. Phys. Lett.* 427 (2006) 404–409.
- [37] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6 (1995) 277–293.

Supplementary data

Structural study of the membrane protein MscL using cell-free expression and solid-state NMR

Alaa Abdine, Michiel A. Verhoeven, Kyu-Ho Park, Alexandre Ghazi, Eric Guittet,
Catherine Berrier, Carine Van Heijenoort and Dror E. Warschawski*

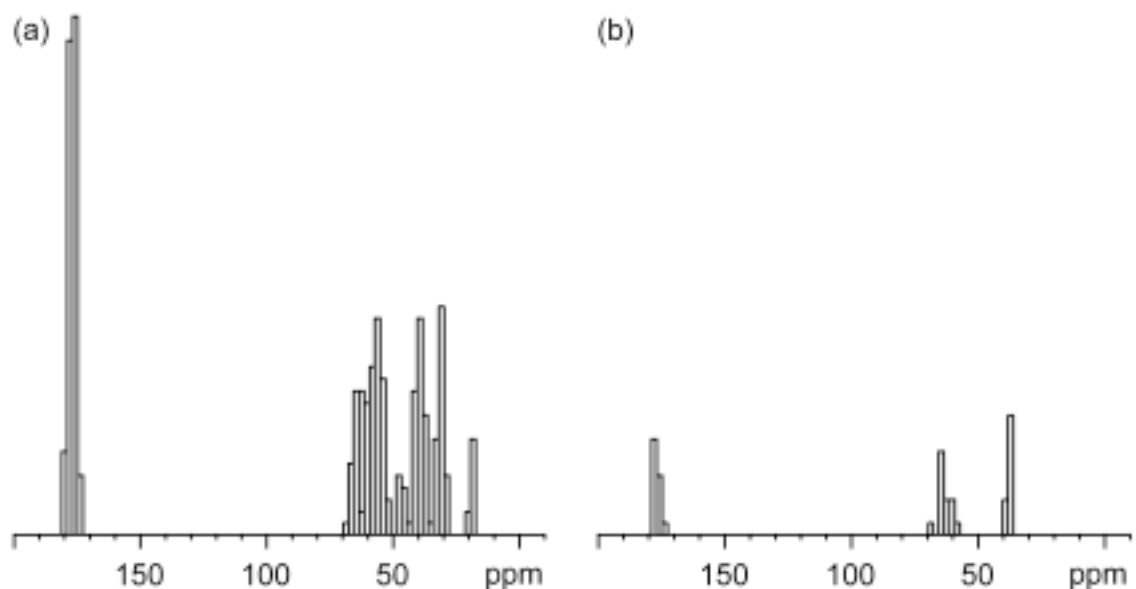


Figure S1. ^{13}C C' , $\text{C}\alpha$ and $\text{C}\beta$ chemical shift simulations using SPARTA of a truncated *E. coli* MscL homology model based on the *M. tuberculosis* MscL X-ray structure and obtained after MD simulation: (a) $\text{U-}^{13}\text{C}, ^{15}\text{N}$ labeled and (b) (Ile, Thr)- $^{13}\text{C}, ^{15}\text{N}$ labeled. The chemical shift values are a weighted average of 6 independent pentameric MscL molecules, hence 30 monomers. Resonances from DOPC and spinning side bands appear in the spectra, not in the simulations.

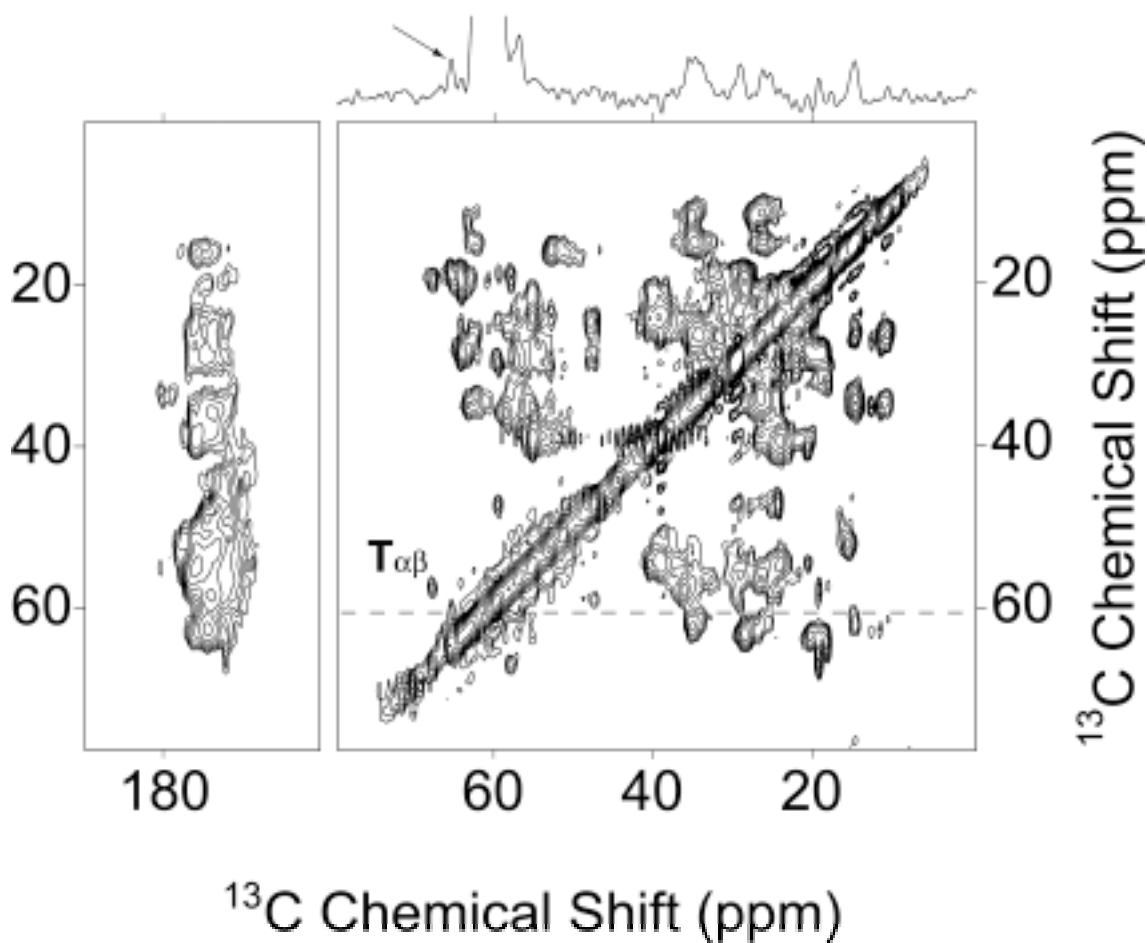


Figure S2. 2D ^{13}C - ^{13}C PDS NMR (20 ms mixing time, total experiment time: 56 hours) spectra of U- ^{13}C , ^{15}N labeled MscL/DOPC. In the case of the uniformly labeled protein, 256 scans per row are necessary to be able to observe the individual threonine cross signals, which are narrow, as can be seen on the cross section shown on top. The arrow indicates a threonine cross signal which is 120 Hz wide (0.68 ppm).

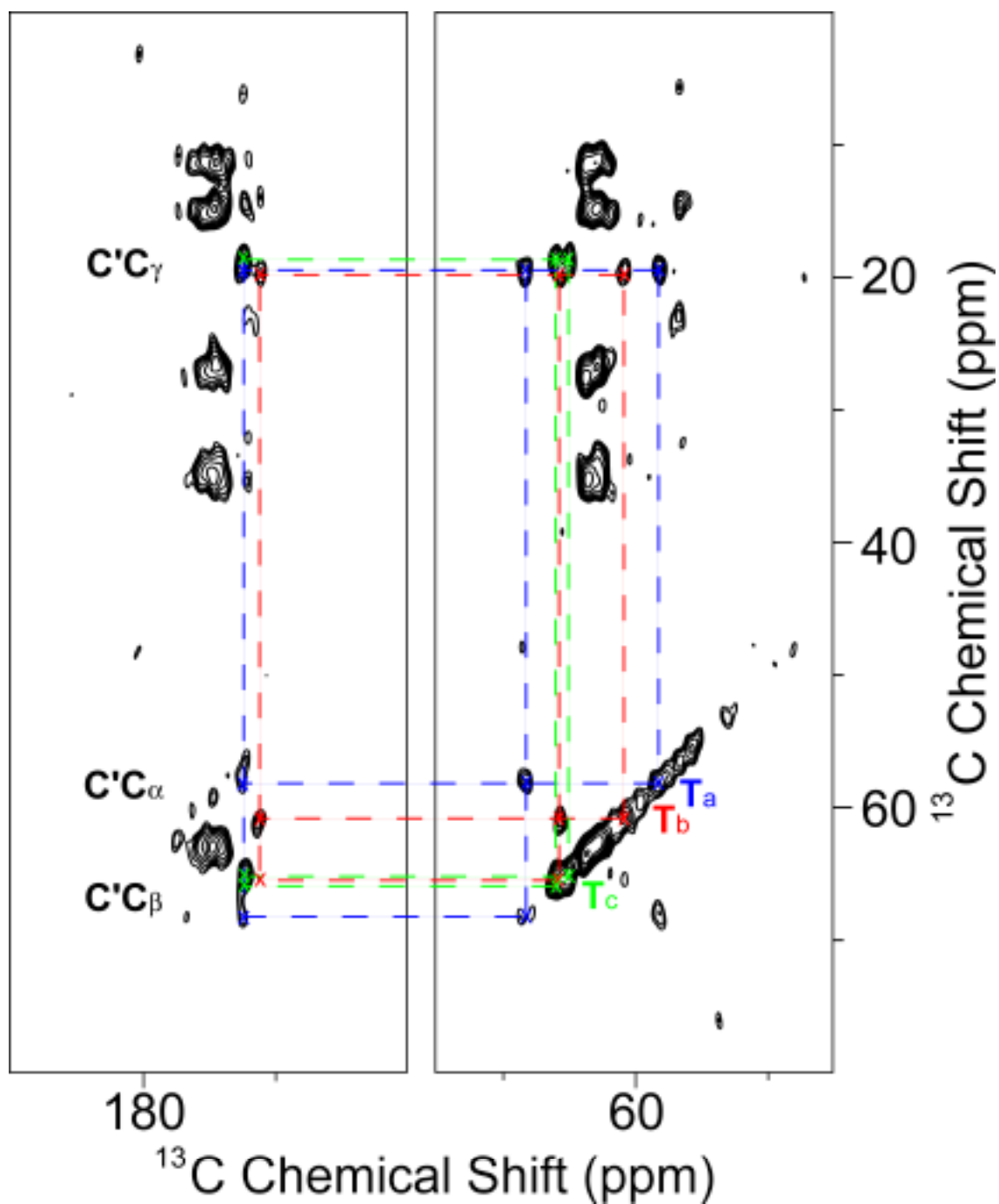


Figure S3. Magnifications of the 2D ^{13}C - ^{13}C DARR NMR (100 ms mixing time, 56 hours of acquisition) spectrum of (Ile, Thr)- ^{13}C , ^{15}N labeled MscL/DOPC relevant for the characterization of the 3 Thr residue spin systems in MscL.

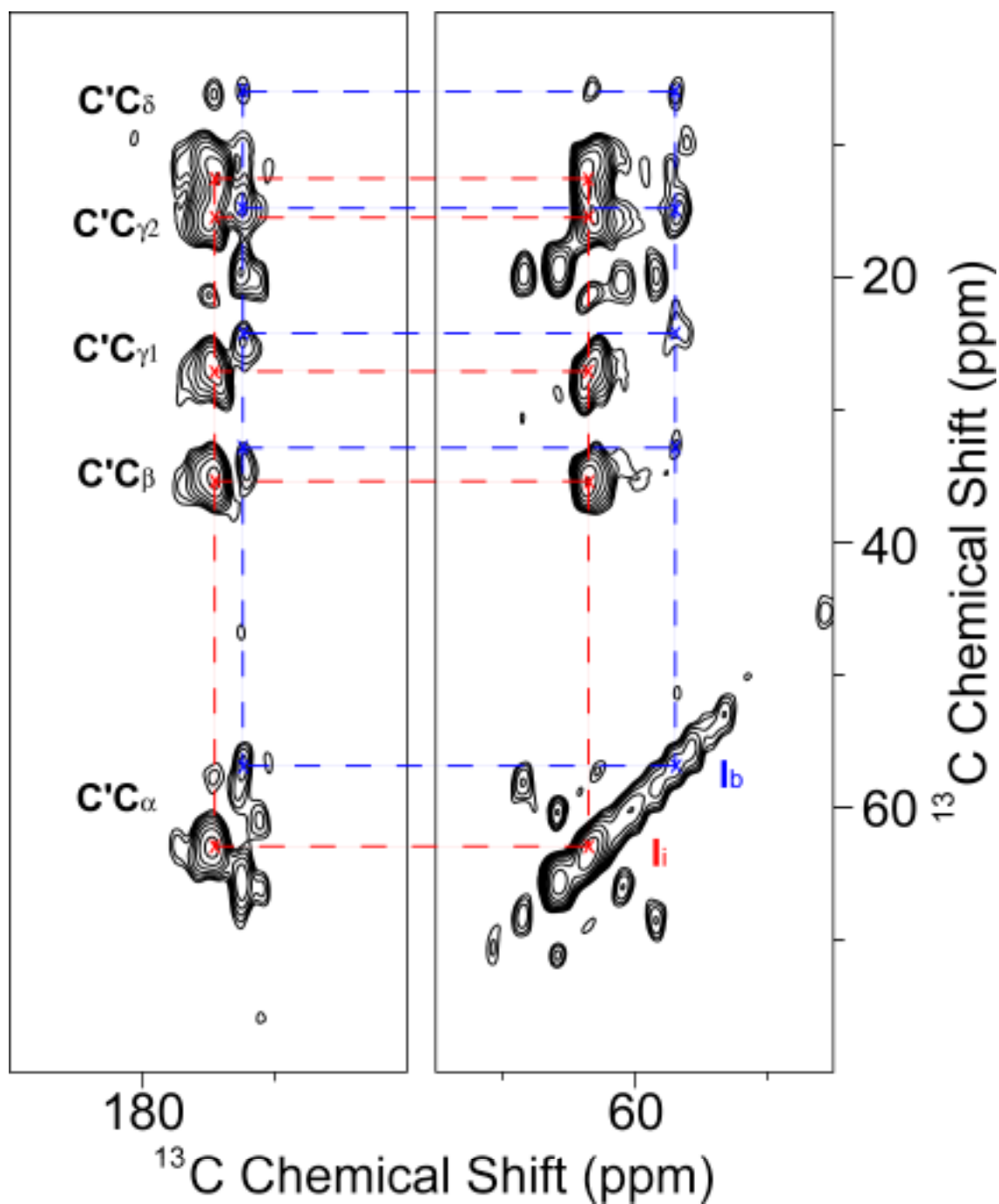


Figure S4. Magnifications of the 2D ^{13}C - ^{13}C DARR NMR (200 ms mixing time, 14 hours of acquisition) spectrum of (Ile, Thr)- ^{13}C , ^{15}N labeled MscL/DOPC relevant for the characterization of the Ile residue spin systems in MscL. Only the spin system for Ib and Ii are shown here.

Table S1. Partial chemical shift characterization of threonine residues in (Ile, Thr)-¹³C/¹⁵N labeled MscL. Peaks were picked from a series of spectra with several DARR mixing times (50, 100, 200, 500 ms) and the error is indicated as well. Since no sequential assignment was made, the amino acids were given an arbitrary alphabetic label. # indicates the number of cross peaks observed and used for the characterization.

amino acid	C'	Cα	Cβ	Cγ	
Ta	shift (ppm)	172.3	58.2	68.3	19.6
	error (ppm)	0.2	0.3	0.2	0.2
	#	19	21	22	20
Tb	shift (ppm)	170.9	60.7	65.7	19.9
	error (ppm)	0.2	0.2	0.2	0.2
	#	15	22	27	20
Tc	shift (ppm)	171.7	64.9	65.8	18.5
	error (ppm)	0.2	0.1	0.1	0.1
	#	5	7	7	9

Table S2. Partial chemical shift characterization of isoleucine residues in (Ile, Thr)-¹³C/¹⁵N labeled MscL. Peaks were picked from a series of spectra with several DARR mixing times (50, 100, 200, 500 ms) and the error is indicated as well. Since no sequential assignment was made, the amino acids were given an arbitrary alphabetic label. # indicates the number of cross peaks observed and used for the characterization.

amino acid	C'	Cα	Cβ	Cγ1	Cγ2	Cδ	
Ia	shift (ppm)	172.2	55.9	35.0	24.1	14.3	9.9
	error (ppm)	0.2	0.1	-	0.1	-	0.1
	#	5	5	1	2	1	2
Ib		171.8	56.7	32.5	23.0	14.7	6.0
		0.3	0.1	0.2	0.2	0.3	0.3
		6	13	13	13	12	5
Ic		171.7	57.2	36.3	23.1	14.7	11.2
		-	0.1	-	0.2	-	-
		1	5	1	3	1	1
Id		172.5	58.5	36.2	27.1	15.6	11.7
		-	0.1	-	0.3	0.2	0.1
		1	3	1	2	3	2
Ie		174.3	59.1	36.0	26.4	15.6	12.0
		0.1	0.1	0.1	0.1	0.1	0.2
		3	6	2	2	3	2
If		-	60.2	34.9	-	15.5	-
		-	0.2	0.1	-	0.2	-
		-	4	3	-	3	-
Ig		175.1	60.8	33.7	26.6	15.4	10.8
		0.1	0.1	-	0.2	0.2	0.2
		2	4	1	3	3	3
Ih		174.7	62.5	33.9	26.4	14.6	10.9
		0.1	0.1	0.2	0.1	0.2	0.2
		6	10	9	9	9	5
Ii		175.2	63.5	35.3	27.7	14.9	12.0
		0.3	0.1	0.1	0.1	0.1	0.1
		5	7	7	7	7	7
Ij		171.7	65.4	36.0	28.3	-	12.3
		0.2	0.1	0.2	0.2	-	0.1
		3	3	2	3	-	3

II Cell-free expression and labeling strategies for a new decade in solid-state NMR

Alaa Abdine^a, Michiel A. Verhoeven^a and Dror E. Warschawski^a

^a UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France





Cell-free expression and labeling strategies for a new decade in solid-state NMR

Alaa Abdine, Michiel A. Verhoeven and Dror E. Warschawski

UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France

Although solid-state NMR and cell-free expression have recently become standard methods in biology, the combination of the two is still at a very early stage of development. In this article, we will explore several approaches by which cell-free expression could help solid-state NMR in its quest for biomolecular structure and mechanism elucidation. Far from being just another structure determination technique, this quest is motivated by the unique possibility of using solid-state NMR to determine the high resolution structure of a membrane protein *within its native environment*, the lipid membrane. We will examine the specific sample preparation requirements that such a goal imposes and how cell-free expression can play a key role in such a protocol.

Introduction

After 50 years of development, solid-state nuclear magnetic resonance (SSNMR) has finally entered, in 2002, the small club of techniques providing the RCSB Protein Structure Database with high resolution atomic structures of biomolecules. About 20 SSNMR protein structures have been deposited in the past decade, 12 of them in the past two years (<http://www.drorlist.com/nmr/SPNMR.html>). Following the lesson learned from solution-state NMR, SSNMR will benefit immensely from developments in biochemistry, protein overproduction, sample preparation and isotopic labeling strategies. At the same time, Cell-Free Expression (CFE) is becoming an affordable technique for a protein biochemistry laboratory. Although SSNMR and CFE are entering the standard toolbox of the structural biologist, both techniques are still confronted with specific bottlenecks when tackling *membrane proteins*. We believe that CFE and SSNMR, together, could join forces to try and increase the throughput of membrane protein structure determination. In this article, we will consider several approaches by which CFE can help to design proteins that are isotopically labeled, specifically for the various steps of molecular structure determination by SSNMR: resonance assignments, distance measurements and structure–function relationship determination.

Corresponding author: Warschawski, D.E. (Dror.Warschawski@ibpc.fr)

Solid-state NMR

Nuclear magnetic resonance comes in different flavors and can be applied to a variety of samples, from individual molecules to the entire human body. As an example, soluble proteins under 30 kDa tumble in water with a correlation time of under 20 ns, which is fast enough to obtain high resolution ^1H NMR spectra in a high magnetic field and allow for the determination of a three-dimensional molecular structure. With such an approach, solution-state NMR has provided the structures of some 6000 soluble proteins in the past 20 years [1].

When the molecules do not tumble fast enough, spectral resolution is lost. The field of SSNMR applies not only to solid molecules, but also when the molecules are too large, the solution is too viscous or the temperature is too low. In that case, other approaches have to be followed, to compensate for nature's fate and to regain high spectral resolution. One such approach is magic-angle spinning (MAS), where the solid sample is spun very fast, almost mimicking the fast molecular tumbling in solution. The spinning rate is limited by current technology to around 70 kHz, or even 30 kHz on standard equipment, which is hardly enough to compete with nature in ^1H NMR spectroscopy, but largely enough for most ^{13}C or ^{15}N biomolecular NMR experiments. SSNMR is a fast growing technique that benefits immensely from constant technological developments such as high magnetic fields or fast spinning NMR probes [2].

Many intriguing biological molecules are 'solid' in the NMR sense, or can become solid (dried, frozen or crystallized) for the purpose of the experiment. When a crystal is large enough, the molecular structure can be determined by X-ray diffraction. When it is too small (also known as micro-crystalline), it can still be studied by SSNMR [3]. Some biological macromolecules are solid by nature. Amyloid peptides are soluble, but they polymerize to form solid fibers that are unsuitable for solution-state NMR or X-ray crystallography, but not for SSNMR [4]. Last, but not least, membrane proteins embedded in their hydrated lipid bilayers are not strictly solid, but they are solid in the NMR sense: they tumble too slowly to be studied by solution-state NMR (or by X-ray crystallography, for that matter). SSNMR has the immense advantage of being suited to the study of membrane proteins *within* their native environment [5]. All these approaches, combined with major technological developments, have allowed for the elucidation of about 20 protein structures by solid-state NMR since 2002 (<http://www.drорlist.com/nmr/SPNMR.html>).

NMR, whether in solution or in the solid state, is not a very sensitive technique. A typical sample contains 1 mM of protein, which requires the production of mg quantities and the development of protein overexpression systems. In addition, NMR often requires isotopic labels in the protein (mostly ^{13}C , ^{15}N and sometimes ^2H), incorporated uniformly or at specific places. As mentioned above, standard solution-state NMR is limited to relatively small molecules (under 100 kDa), otherwise the spectral resolution is lost. Because SSNMR does not rely on molecular tumbling, resolution (the inverse of the linewidth of a spectral line) is independent of molecular size, but strongly depends on sample homogeneity. For a given resolution, nuclei with similar magnetic environments will resonate at similar frequencies and may not be distinguishable one from another (spectral crowding). This limit can be overcome by designing a molecule (whatever the size) where only a *subset* of the atoms is 'NMR-visible', while the remaining atoms are invisible. Although a significant technological breakthrough was achieved in the past decade, that has allowed for structure determination of biomolecules by SSNMR, the next breakthrough will come from biotechnological developments that will help produce, in large quantities, proteins that will be engineered to incorporate isotopic labels at specific places. In this context, cell-free synthesis could well play a major role in the next decade.

Cell-free expression for solid-state NMR

Cell-free expression is a major development in structural biology, for both NMR and X-ray crystallography, because it allows for fast and easy overproduction of proteins, both wild-type and mutant, with or without non-natural amino acids or additional cofactors. For NMR, CFE offers the additional possibility of incorporating a selected subset of isotopically labeled amino acids, with very little metabolic scrambling. In addition, the rapid development of cell-free lysates now allows for a wide range of post-translational modifications [6].

For SSNMR, CFE offers several methods of preparing a membrane protein sample [7]. Membrane proteins are notoriously difficult to overproduce in cells, due to the availability of limited membrane surface. In addition, membrane proteins often require a complex targeting mechanism to transport them inside their host

membrane. In many cases, they end up as inclusion bodies in the cell, largely unfolded, sometimes irreversibly [8]. By contrast, CFE can produce a membrane protein sample in two days, directly into lipid liposomes or nanodiscs, without the addition of detergents and with a purification method that does not require a tag. Should this method fail, several other options exist, such as using CFE in the presence of detergents or other surfactants. In that case, a tag is required for purification and the protein needs to be reconstituted in a lipid membrane for another two days and a slightly lower yield [9,10]. With CFE, the choice of lipid or lipid mixture, concentration, pH and ionic strength in the final sample is very flexible.

Labeling strategies

Cellular versus cell-free expression

Creative biochemists have come up with a variety of methods to produce molecules tailor made for NMR, by incorporating or removing isotopes at specific places in the molecule. The most popular method is uniform labeling in *E. coli*, using $^{15}\text{NH}_4\text{Cl}$ as a nitrogen source, ^{13}C -labeled glucose as a carbon source and/or $^2\text{H}_2\text{O}$ as a deuterium source [11]. When uniform labeling results in an overcrowded spectrum, one can use other sources of carbon, such as $^{13}\text{C}_2$ -glycerol or $^{13}\text{C}_{1,3}$ -glycerol, to produce partially labeled proteins [12]. Another way is to feed the bacteria directly with amino acids that are expected to end up in the protein, either with labeled amino acids (specific labeling), or with labeled nutrients together with 'cold' amino acids (reversed labeling [13]). The resulting proteins are never isotopically pure, because of metabolic scrambling, but they are of great help in the first steps of a protein NMR study. All these approaches have their advantages and disadvantages, including a cost that depends on the desired labeling scheme (see below).

CFE has opened up new ways to produce those samples, and new types of samples are available. Isotopic labels are introduced into the protein via the labeled amino acids in the CFE system. Amino acids are available with various patterns, uniformly or specifically ^{13}C , ^{15}N and/or ^2H labeled [14,15]. If a uniformly ^{13}C , ^{15}N , ^2H -labeled protein is required, all 20 uniformly labeled amino acids have to be incorporated, which can be prohibitively expensive, although affordable commercial mixtures are already available [16]. By contrast, for specific labeling, CFE is a very competitive method that is soon to become routine in an NMR laboratory. The efficiency of this approach in SSNMR has been shown in the context of heavily crowded membrane protein spectra [17,18], as can be seen in Fig. 1 with the 2D ^{13}C NMR spectra of the mechano-sensitive ion channel MsL, a 75-kDa pentameric α -helical protein reconstituted in a hydrated DOPC bilayer. Spectrum 1a shows a uniformly labeled protein, while spectrum 1b shows a specifically labeled protein on isoleucines and threonines only. Crowding in the specifically labeled spectrum is reduced, allowing the characterization of 10 out of 16 labeled isoleucines, and of all three threonine spin systems [18].

After obtaining a well-resolved set of NMR spectra, and before measuring distances between atoms that will contribute to the protein structure determination, one needs to assign each resonance to an atom on a specific amino acid. A simplistic approach would consist of making 20 different samples, each containing *one* labeled amino acid type, easily assigned by NMR. The cost of each

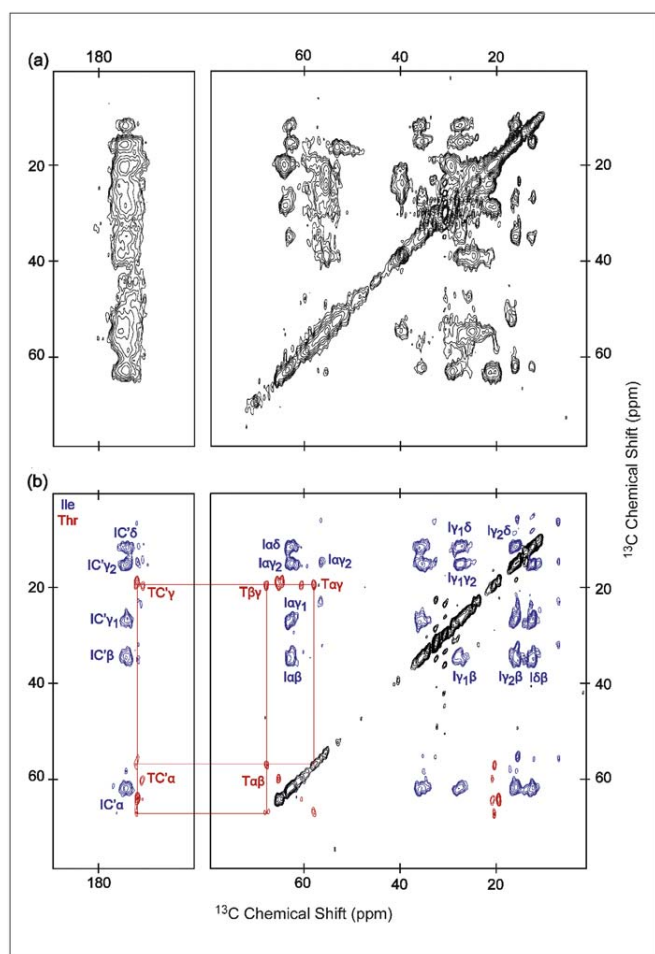


FIGURE 1
2D ^{13}C - ^{13}C DARR NMR (100 ms mixing time) spectra of MscL/DOPC: **(a)** U- ^{13}C , ^{15}N -labeled protein produced in *E. coli* and **(b)** (Ile₁₆, Thr₃)- ^{13}C , ^{15}N -labeled protein produced *in vitro*. Identified isoleucine cross-peaks are colored in blue, while threonine cross-peaks are in red. The assignment pattern for one of the three threonines is indicated on the spectrum.

sample and the amount of NMR data that would be required point toward searching for better approaches. The main purpose of this article is to explore the various CFE strategies that one could adopt when studying a biomolecule by SSNMR.

Combinatorial approaches

For the initial and crucial assignment step, several combinatorial approaches have been suggested, tested and refined for solution-state NMR, based on the sole knowledge of the protein amino acid sequence. Rather than labeling all amino acids at once, which would often result in an overcrowded NMR spectrum, the idea is to make a reasonable number of samples, each with a different labeling scheme and a reasonable number of labels, covering all amino acids with all samples. The Otting group [19] suggests making five different samples, each comprising seven labeled amino acids, including only *one* of the five most abundant residues. The less abundant amino acids appear in several samples and can be identified by comparison between the five sets of NMR spectra. The 'dual combinatorial' approach of the Dötsch group [20] aims at identifying unique *pairs* of neighboring residue types,

labeling these types and assigning those residues by observing unique cross-peaks. While the first approach allows the assignment of residue *types* (all glycines, all valines, among others), the second approach is less general, but allows *specific* residue assignment of glycine 22 and valine 23 (among others), as long as the (glycine 22–valine 23) pair is the only glycine–valine neighboring pair in the protein sequence. A preliminary example of this method is shown in Fig. 2 where a sample of MscL was ^{13}C -labeled on all phenylalanines, isoleucines, methionines, prolines and arginines using CFE, resulting in 9 unique pairs of labeled neighboring residue types along the protein sequence. Although the unique cross-peaks have not yet been localized on this spectrum with a short mixing time, many intra-residue cross-peaks have been identified.

Complete assignment of a protein is a difficult task, and sometimes 45% assignment is enough to continue and determine a protein structure [5]. Nevertheless, in some cases, a particularly important residue needs to be assigned unambiguously, despite spectral overlap. Site-directed mutagenesis offers the possibility of growing mutant proteins, both in cells and *in vitro*, that would replace all but one residue of a certain type by another type. The remaining residue could then be isotopically labeled and identified with certainty by NMR. In such a case, CFE can be preferred over biosynthesis, as exemplified by Lehner *et al.* [17] when they managed to identify the key residue for transport in the EmrE membrane protein by SSNMR.

These approaches have been designed for solution-state NMR, where ^1H - ^{15}N spectra are acquired, amide protons play an essential role, prolines are excluded, and spectral overlap is worse in α -helices than in β -sheets. In SSNMR, ^{13}C - ^{13}C and ^{13}C - ^{15}N spectra

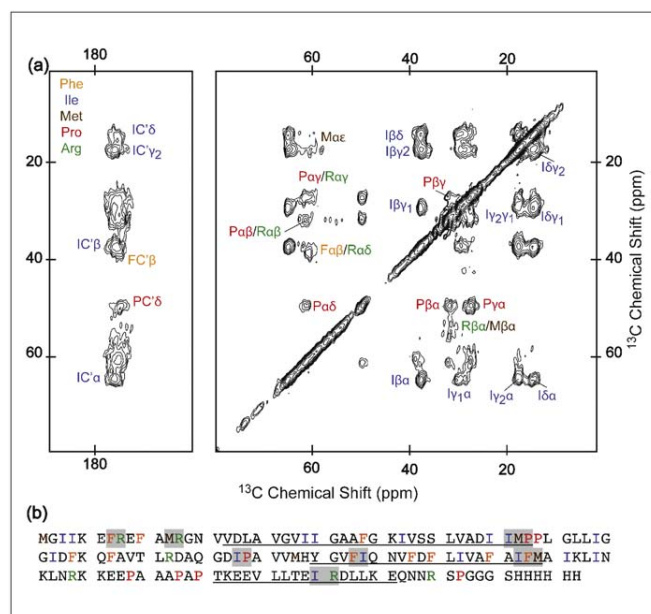


FIGURE 2
(a) 2D ^{13}C - ^{13}C DARR NMR (50 ms mixing time) spectrum of (Phe₁₀, Ile₁₆, Met₅, Pro₇, Arg₆)- ^{13}C , ^{15}N -labeled MscL/DOPC produced *in vitro*. Preliminary identified intra-residue cross-peaks are indicated on the spectrum. **(b)** MscL amino acid sequence with the corresponding color coding. The nine unique pairs of labeled neighboring residue types are highlighted in gray. The three α -helices, of which the first two are transmembrane, are underlined.

are acquired, prolines are welcome, and spectral overlap is comparable within each secondary structure. Resolution is generally worse in SSNMR than in its solution-state counterpart, because magnetic interactions are stronger and because magnetic and conformational inhomogeneities are not entirely averaged out by molecular motion or packing. Another difference stems from the nature of polarization transfers. In solution-state NMR, polarization is transferred from one nucleus to its neighbor through chemical bonds, via the scalar coupling. In SSNMR, transfer is accomplished through space, via the dipolar coupling. Although dipolar transfer can be voluntarily limited to very short distances so as to resemble scalar transfers, it can also be extended over longer distances, and probe the environment of a nucleus beyond its neighbor. In other words, while this approach has been designed, in solution-state NMR, mostly for helping in the crucial residue assignment step, it can also be used in SSNMR and extended for long distance measurements and structure determination.

Predictive approaches

One major goal of specific labeling is to reduce the spectral crowding on NMR spectra. Although combinatorial approaches only require prior knowledge of the protein amino acid sequence, they do not take into account the regions of the spectra where those amino acids will appear and, potentially, overlap. Because a structural model, or the structure of an analog protein, is often available, the expected NMR spectra can be predicted and help choose the labeling scheme that will minimize spectral overlap. A variety of (usually free) software packages have been designed for this purpose, such as SPARTA [21], and this approach has been refined by Sweredoski *et al.* [22], again in the case of solution-state NMR, ^1H and ^{15}N detection.

The limit of such an approach is that of the NMR spectral resolution. The position and crowding of cross-peaks can be predicted, but not their linewidth. In SSNMR, ^{13}C linewidth can go from around 0.5 ppm in a quasi-crystalline membrane protein [23] to 0.8 ppm in a membrane with a high lipid and water content [18]. As an example, using a computer model for the membrane channel MscL [24], we have predicted that alanine, lysine, valine and tyrosine ^{13}C signals would not overlap. Through CFE, we have synthesized a protein with those four amino acid types ^{13}C -labeled. Indeed, as can be seen in Fig. 3, the various amino acid types do not overlap. Unfortunately, due to low spectral resolution of this membrane preparation, all amino acids of the *same* type overlap with one another and could not be assigned. For example, 13 valine carbon lines (of 0.8 ppm width each) are overlapped within a 3 ppm wide square. In addition, with this particular labeling scheme, no unique pair of neighboring amino acids was found either. By contrast, although most observed cross-peaks are intra-residues, a couple of alanine–valine *inter-residue* cross-signals are observed in Fig. 3, confirming that distances could be extracted with this approach, for future structure determination.

Structure–function approaches

Structure determination is most useful for understanding biological functions and molecular mechanisms. In this context, protein function can be understood by knowing a rough overall structure (secondary or low resolution structure) and a precise structure of

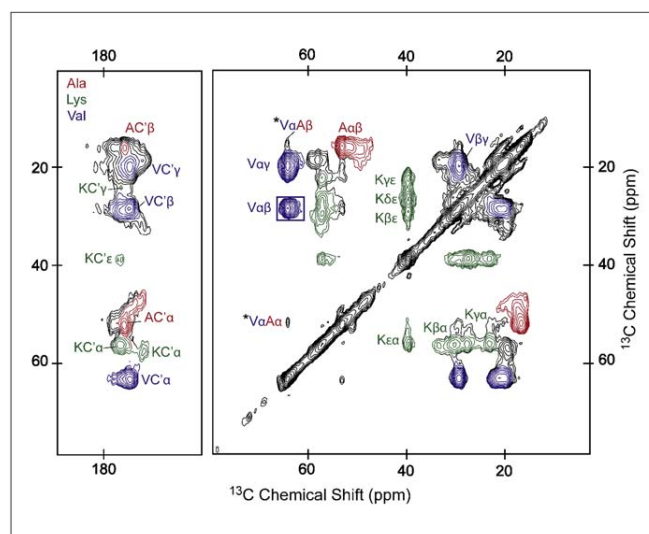


FIGURE 3

2D ^{13}C – ^{13}C DARR NMR (100 ms mixing time) spectrum of (Ala₁₅, Lys₉, Val₁₃, Tyr₁)- ^{13}C , ^{15}N -labeled MscL/DOPC produced *in vitro*. Each identified set of intra-residue cross-peaks is colored differently. The blue square contains all 13 valine C α to C β cross-peaks. Cross-peaks indicated with a star correspond to inter-residue peaks between valine C α and alanine C α and C β .

only the crucial parts of the protein (active site, dimerization site, among others). Regardless of prior assignment or structural determination, another labeling strategy could therefore focus solely on regions that are expected to be important for the protein function, or whose structure needs to be refined. Once those crucial amino acids are identified, they can be isotopically labeled, observed by NMR and distances can be measured. Upon ligand binding, channel opening or conformational change, for example, these distances can be used as probes of molecular motion, and help visualizing the protein mechanism.

In the context of a structure–function study, the possibility of studying the structure of a mutant with a different functional phenotype is of great importance. Unfortunately, such mutants are sometimes lethal for bacterial growth. For example, a membrane channel that would be permanently open would cause a lethal cell leakage. In such a case, CFE is a very attractive alternative production method for structural biology, because the produced mutant can consequently be studied by electron microscopy [25] or by solid-state NMR.

Cost

Although costs are highly variable, the cost difference between expression methods is an argument that is often put forward. Besides the labor cost, which is lower for CFE because the protocol is faster, it is therefore interesting to compare the approximate protein production cost for each approach, as of today. The price of 1 mg of uniformly ^{13}C , ^{15}N -labeled protein expressed in *E. coli* is mostly the price of 3 g of uniformly ^{13}C -labeled glucose, in the order of 600 €. The price of 1 mg of a specifically labeled protein expressed *in vitro* through the Roche system is a combination of the price of the Roche RTS kit (300 €) and that of the ^{13}C , ^{15}N -labeled amino acids (10 mg, around 50 €, of each). The price of 1 mg of a specifically labeled protein expressed in *E. coli* would be

that of the labeled amino acids, which would have to be added in larger quantities (100 mg approximately). For six ^{13}C , ^{15}N -labeled residues, CFE would produce 1 mg of protein for 600 €, whereas biosynthesis would cost 3000 € per mg. Using a commercial kit such as RTS is advantageous for its convenience and reliability, for saving the time (and manpower) to make the lysate and also for managing the stocks. However, there are many protocols available for making lysates [26], which would reduce the cost of CFE by an additional significant factor.

Conclusion

In this article, we have provided an overview of several approaches that could be followed for protein studies by solid-state NMR. In addition to the methods developed for solution-state NMR, a similar effort has to be devoted to adapt them for their solid-state counterpart. Among those, we have described several labeling schemes developed in our laboratory on the membrane channel MscL, using cell-free expression or expression in *E. coli*, that will hopefully lead to a protein structure determination by SSNMR.

The rate of protein structure deposition at the RCSB Protein Structure Database, and determined by solid-state NMR, is not going to slow down soon. We hope that membrane protein structures will be a significant number of those in the next decade. While NMR still benefits from technological development (dynamic nuclear polarization, cryo-NMR, among others), progress in biomolecular NMR will also originate from the wet laboratory, by optimizing sample preparation, for which we believed cell-free expression of membrane proteins may well play a major role.

Acknowledgements

This work was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (to AA), by the CNRS (UMR 7099), the ANR (ANR-06-JCJC0014) and the Université Paris Diderot. We thank Carole Gardienet and Cathy Etchebest for help with the simulations, Carine Van Heijenoort for the NMR, Anders Pedersen, Catherine Berrier, Kyu-Ho Park and Francesca Zito for various tricks during sample preparations, and for helpful discussions.

References

- 1 Billeter, M. *et al.* (2008) Solution NMR structure determination of proteins revisited. *J. Biomol. NMR* 42, 155–158
- 2 Lesage, A. (2009) Recent advances in solid-state NMR spectroscopy of spin $I = 1/2$ nuclei. *Phys. Chem. Chem. Phys.* 11, 6876–6891
- 3 Bertini, I. *et al.* (2010) High-resolution solid-state NMR structure of a 17.6 kDa protein. *J. Am. Chem. Soc.* 132, 1032–1040
- 4 Wasmer, C. *et al.* (2008) Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319, 1523–1526
- 5 Cady, S.D. *et al.* (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463, 689–692
- 6 He, M. (2008) Cell-free protein synthesis: applications in proteomics and biotechnology. *Nat. Biotechnol.* 25, 126–132
- 7 Abdine, A. *et al.* (2011) Cell-free membrane protein expression for solid-state NMR. In *Protein NMR Techniques*, 3rd Edition (Shekhtman A. and Burz D.S. eds), chapter 5, Methods in Molecular Biology series, Humana Press.
- 8 Wagner, S. *et al.* (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol.* 24, 364–371
- 9 Park, K.H. *et al.* (2007) Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem. J.* 403, 183–187
- 10 Park, K.H. *et al.* In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *Nat. Biotechnol.* (this issue)
- 11 Goto, N.K. and Kay, L.E. (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* 10, 585–592
- 12 Castellani, F. *et al.* (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420, 98–102
- 13 Heise, H. *et al.* (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15871–15876
- 14 Kainosho, M. *et al.* (2006) Optimal isotope labelling for NMR protein structure determinations. *Nature* 440, 52–57
- 15 Hiller, M. *et al.* (2008) [2,3-(^{13}C)-labeling of aromatic residues-getting a head start in the magic-angle-spinning NMR assignment of membrane proteins. *J. Am. Chem. Soc.* 130, 408–409
- 16 Etezady-Esfarjani, T. *et al.* (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. *J. Biomol. NMR* 39, 229–238
- 17 Lehner, I. *et al.* (2008) The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J. Biol. Chem.* 283, 3281–3288
- 18 Abdine, A. *et al.* (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* 204, 155–159
- 19 Ozawa, K. *et al.* (2006) ^{15}N -labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *FEBS J.* 273, 4154–4159
- 20 Sobhanifar, S. *et al.* (2010) Cell-free expression and stable isotope labelling strategies for membrane proteins. *J. Biomol. NMR* 46, 33–43
- 21 Shen, Y. and Bax, A. (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* 38, 289–302
- 22 Sweredoski, M.J. *et al.* (2007) Minimizing the overlap problem in protein NMR: a computational framework for precision amino acid labeling. *Bioinformatics* 23, 2829–2835
- 23 Shi, L. *et al.* (2009) Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump – structural insights. *J. Mol. Biol.* 386, 1078–1093
- 24 Valadié, H. *et al.* (2003) Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis. *J. Mol. Biol.* 332, 657–674
- 25 Yoshimura, K. *et al.* (2008) Gating-associated conformational changes in the mechanosensitive channel MscL. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4033–4038
- 26 Schneider, B. *et al.* (2010) Membrane protein expression in cell-free systems. *Methods Mol. Biol.* 601, 165–186

III Cell-free membrane protein expression for solid-state NMR

Alaa Abdine, Kyu-Ho Park and Dror E. Warschawski

Corresponding author: D. E. Warschawski, UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France ; e-mail: Dror.Warschawski@ibpc.fr

Methods in Molecular Biology

Protein NMR Techniques, Third Edition, 2011

Edited by Alex Shekhtman and David Burz

Running head: Membrane proteins for solid-state NMR

Cell-free membrane protein expression for solid-state NMR

Alaa Abdine, Kyu-Ho Park and Dror E. Warschawski

Corresponding author: D. E. Warschawski, UMR 7099, CNRS and Université Paris Diderot, IBPC, 13 rue Pierre et Marie Curie, F-75005 Paris, France ; e-mail: Dror.Warschawski@ibpc.fr

Abstract

Although cell-free expression is a relative newcomer to the biochemical toolbox, it has already been reviewed extensively, even in the more specialized cases such as membrane protein expression, nanolipoprotein particles and applications to crystallography and nuclear magnetic resonance (NMR). Solid-state NMR is also a newcomer in the structural biology toolbox, with its own specificities in terms of sample preparation. Cell-free expression and solid-state NMR are a promising combination that has already proven useful for the structural study of membrane proteins in their native environment, the hydrated lipid bilayer. We describe below several protocols for preparing MscL, a mechanosensitive membrane channel, using cell-free expression destined for a solid-state NMR study. These protocols are flexible and can easily be applied to other membrane proteins, with minor adjustments.

Key words: in vitro synthesis, integral membrane proteins, membrane protein reconstitution, liposomes, solid-state NMR

1. Introduction

Cell-free expression is one of the major new developments in structural biology, both for nuclear magnetic resonance (NMR) and X-ray crystallography, because it allows for overproduction of proteins, both wild-type and mutant, that are often produced with prohibitively low yields using classical biosynthetic methods. Proteins that are expressed cell-free end up in the reaction container, hence in water where membrane proteins precipitate, often irreversibly. Cell-free expression was first thought to be impractical for membrane protein expression because the system did not include a biological apparatus for targeting the protein to the membrane (1). Providing the medium with detergent was also considered risky since detergent could perturb protein expression by interfering with the ribosome, polymerase or any other essential component of the cell-free expression system. These obstacles were overcome in 2004, when several groups, using optimized cell lysates, developed protocols for membrane protein expression *in vitro*, in presence of detergents (2-4).

For X-ray crystallography and solution-state NMR, the major techniques used in structural biology today, membrane protein structure determination is still a challenge. First, because the aforementioned necessity of obtaining large quantity of functional proteins is aggravated, *in vivo*, by the limited membrane surface available (5). Second, because membrane proteins are hydrophobic and have to be manipulated in detergents at all times when they are extracted from their native environment, the lipid bilayer (6). Detergents or other surfactants are used for membrane solubilization and during protein purification (which often requires an affinity tag, such as a poly-histidine stretch). They can interfere with protein crystallization or make aggregates that are too large for solution-state NMR. Last but not least, detergents can interfere with protein folding or function, and the protein structure or dynamics determined in a detergent environment is not necessarily representative of the native structure or dynamics (7, 8). Nevertheless, since 1985, over two hundred membrane protein structures have been determined by using X-ray crystallography (9) and, since 1997, about thirty structures by using solution-state NMR (10). In this context, membrane protein cell-free expression has been developed and reviewed, especially for isotope labeling strategies, another issue for NMR, where *in vitro* synthesis offers a very efficient and versatile alternative that greatly minimizes amino-acid scrambling (11, 12).

Solid-state NMR is an alternative for membrane protein structure determination *inside* the hydrated lipid bilayer, where the protein is correctly folded and stable for a long time.

Sample preparation for solid-state NMR is therefore different than for solution-state NMR. If the protein is expressed cell-free in the presence of detergent, it needs to be purified, renatured and reconstituted in a membrane bilayer. Compared to a protein expressed in a cell, the solubilization step is avoided and the purification step is greatly simplified and accelerated, reducing the chances of spurious proteolysis. Reconstitution of the protein in a lipid membrane is the price to pay to obtain a sample where the membrane protein is almost certainly in its native state, and where the protein function can be checked and monitored (**13, 14**). All these steps are quite flexible, and they will be described below for the expression of the mechanosensitive channel MscL. This protocol is general for membrane proteins, although details such as the nature or the concentration of detergent can vary from one protein to the next.

Cell-free expression in presence of liposomes or nanodiscs is an alternative sample preparation approach that has been developed more recently, and that has yet to be proven general for membrane proteins (**15-17**). When possible, it presents many advantages for solid-state NMR studies. First and foremost, the protein is already in its final state, saving time and avoiding many biochemical steps where the protein may be partially lost or inactivated. Second, the protein is purified simply by centrifugation, alleviating the necessity of adding an affinity tag. In addition, no detergent is used that can interfere and that needs to be removed. Importantly, the protein is in its native environment at all times, where it is well folded, functional and stable (**18**). Finally, this approach is quite flexible, allowing for a variety of labeling strategies (**19**). In our hands, cell-free expression of MscL in the presence of liposomes has proven efficient (**3, 14, 18, 19**) and we will describe it below, as it is the best method so far for providing a solid-state NMR sample.

The protocols presented here make use of commercial kits for cell-free expression, which we found advantageous for its convenience, reliability, for saving the time and manpower to make the lysate, and also for managing the stocks. Cell lysates can also be prepared, following published protocols from various cell types such as bacteria, wheatgerm, or others (**20, 11**). We describe below an optimized procedure for expressing the membrane protein MscL, using the commercial Roche/5Prime continuous exchange vessel and kits, with an *Escherichia coli* extract. Since, for each sample preparation protocol, it is necessary to check for protein integrity, we also describe several tests that should be performed to assess the quality of the sample.

2. Materials

2.1 Cell-free expression of the mechanosensitive channel MscL in detergent micelles.

All solutions are prepared with autoclaved nuclease-free Milli-Q water (*see Note 1*).

1. Thermoregulated shaker for the cell-free reaction vessel.
2. Nuclease-free 50 mL tubes.
3. Nuclease-free 1.5 mL tubes.
4. Nuclease-free glass pipettes.
5. Nuclease-free pipette tips (0 - 10 μ L, 10 - 200 μ L, 200 - 1000 μ L, and 1 - 5 mL), autoclaved at 121 °C for 20 min.
6. pIVEX-2.3-mscL plasmid (**3**). Encodes the *E. coli* MscL C-terminally fused to a His₆ tag under the control of the T7 promoter. Dissolve in pure nuclease-free water and store aliquoted at -20 °C, at 0.5 μ g/ μ L.
7. RTS 9000 cell-free expression kit (Roche/5Prime). Contains lyophilized *E. coli* lysate, reaction mix, feeding mix, reconstitution buffer, continuous exchange reaction vessel and a syringe (*see Fig. 1, Notes 2, 3, 4 and 5*). Store at -80 °C.
8. Unlabeled amino-acids (powder), store at -20 °C.
9. ¹³C/¹⁵N labeled amino-acids (powder), store at -20 °C.
10. Dithiothreitol (DTT): Prepare 40 mM stock, store at -20 °C.
11. 20% (w/v) Triton X-100: Prepare 200 mL, stir until homogeneous, store at 4 °C for up to six months.
12. Unlabeled amino-acid solutions: Prepare 168 mM solutions of each amino-acid except for leucine (140 mM) and the labeled amino-acids, Ile and Thr (*see step 13*), with the appropriate solution (*see Note 6*). The amount of amino-acid to be incorporated is calculated in Section 3.1 and is indicated in **Table 1**. The appropriate volume of each unlabeled amino-acid is transferred to a 1.5 mL vial, and sonicated until the solution looks clear. The individual solutions are transferred into a 50 mL tube. Tyr and Leu are added last, to avoid precipitation, prepare fresh before using, keep on ice.
13. Labeled amino-acid solutions: In our case, Ile and Thr. Dissolve 23 mg of Ile and 3.9 mg

- of Thr in 1.0 and 0.20 mL of reconstitution buffer respectively (around 168 mM), into 1.5 mL tubes, prepare fresh before using, keep on ice (*see* Section 3.1, **Table 1** and **Note 7**).
14. FPLC purification system.
 15. 1 M NaOH: To adjust pH, store at room temperature.
 16. 0.1 M NiSO₄: Store at room temperature.
 17. 20% Ethanol: Store at room temperature for up to a month.
 18. AEBSF : 4-(2-Aminoethyl) benzene sulfonyl fluoride hydrochloride (or equivalent serine protease inhibitors), prepare a 100 mM stock solution in water, store at 4 °C for up to six months.
 19. Chelating column (5 mL): Stored in 20% ethanol at 4 °C.
 20. FPLC buffer A1: 50 mM Na₂HPO₄-NaOH, pH 8 (7.1 g/L), 300 mM NaCl (17.5 g/L), 10 mM imidazole (680 mg/L), 4.0 % Triton X-100 (40 g/L), store at 4 °C.
 21. FPLC buffer A2: 50 mM Na₂HPO₄-NaOH, pH 8 (7.1 g/L), 300 mM NaCl (17.5 g/L), 10 mM imidazole (680 mg/L), 0.2 % Triton X-100 (2 g/L), store at 4 °C.
 22. FPLC buffer B: 50 mM Na₂HPO₄-NaOH, pH 8 (7.1 g/L), 300 mM NaCl (17.5 g/L), 500 mM imidazole (34 g/L), 0.2 % Triton X-100 (2 g/L), store at 4 °C.
 23. FPLC sample loading buffer: 50 mM Na₂HPO₄-NaOH, pH 8 (7.1 g/L), 300 mM NaCl (17.5 g/L), 10 mM imidazole (680 mg/L), 1% Triton X-100 (10 g/L), store at 4 °C.
 24. 10 x dialysis buffer: 0.1 M HEPES-KOH, pH 7.5, 1 M KCl and 2% (w/v) Triton X-100, store at 4 °C for up to six months.
 25. 1 x dialysis buffer: Prepare 2 L by diluting 200 mL of 10 x buffer in 1750 mL of water, adjusting the pH to 7.5 using KOH, and adding water to 2 L final volume. Store at 4 °C for up to six months.
 26. 2 dialysis cassettes with a 10 kDa MW cut-off.
 27. Methanol.
 28. DOPC: 1,2-dioleoyl-sn-glycero-3-phosphocholine powder, store at -20 °C.
 29. 4 M KOH: To adjust pH, store at room temperature.
 30. 0.8% Triton X-100: Prepare 10 mL, dissolve and stir until homogeneous, store at 4 °C for up to six months.
 31. Wet polystyrene beads: Wash 5 g of 300-1200 μm polystyrene beads with 25 mL of pure methanol, then wash four times with 25 mL of Milli-Q water. Beads are kept in water at 4 °C and are stable for months, provided a weekly renewal of Milli-Q water.

32. HEPES-KCl buffer: 10 mM HEPES-KOH, pH 7.5, 100 mM KCl, store at 4 °C for up to six months.

2.2 Cell-free expression of the mechanosensitive channel MscL in liposomes

All solutions are prepared with autoclaved nuclease-free Milli-Q water (*see Note 1*).

1. Thermoregulated shaker for the cell-free reaction vessel.
2. Mini-Extruder with two 1 mL gastight microsyringes and 0.1 µm cut-off polycarbonate membranes.
3. Nuclease-free 50 mL tubes.
4. Nuclease-free 1.5 mL tubes.
5. Nuclease-free glass pipettes.
6. Autoclaved nuclease-free pipette tips (0 - 10 µL, 10 - 200 µL, 200 - 1000 µL, and 1 - 5 mL), at 121 °C for 20 min.
7. pIVEX-2.3-mscL plasmid: For protein expression in liposomes, an affinity tag such as a poly-histidine stretch is preferable but not mandatory.
8. RTS 9000 cell-free expression kit (Roche/5Prime). Contains lyophilized *E. coli* lysate, reaction mix, feeding mix, reconstitution buffer, a continuous exchange reaction vessel and a syringe, (*see Fig. 1, Notes 2, 3, 4 and 5*), store at -80 °C.
9. Unlabeled amino-acids (powder), store at -20 °C.
10. ¹³C/¹⁵N labeled amino-acids (powder), store at -20 °C.
11. DTT (dithiothreitol): Prepare 40 mM stock, store at -20 °C.
12. 4 M KOH: To adjust pH, store at room temperature.
13. AEBSF: 100 mM stock in water, store at 4 °C for up to six months.
14. DOPC.
15. Chloroform.
16. HEPES-KCl buffer (*see Section 2.1*).
17. DOPC liposome preparation: Prepare a chloroform solution of 10 mg/mL of DOPC (100 mg of DOPC in 10 mL of dry chloroform). Dry the lipids and then re-suspend with RTS reconstitution buffer to obtain a 20 mg/mL aqueous solution. Sonicate 2.5 mL of the solution for 5 minutes, and then extrude it 13 times with a 0.1 µm filter, to obtain a liposome solution containing 40 to 50 mg of lipids, store at 4 °C. This is the amount of

lipids that is expected in the final sample. DOPC can be replaced by other lipids or lipid mixtures, at a higher or lower concentration.

18. Prepare HEPES solutions for solubilizing Tyr, Trp and Phe amino-acids (*see* **Notes 6** and **7**): HEPES 60 mM at pH 13 for Tyr, at pH 1 for Trp and at pH 7.5 for Phe. Adjust the pH using KOH or HCl. Keep on ice.
19. Unlabeled amino-acid solutions: Prepare 168 mM solutions of each amino-acid except for leucine (140 mM) and the labeled amino-acids (*see* step 20), with the appropriate solution (*see* **Note 6**). The amount of amino-acid to be incorporated is calculated in Section 3.1 and is indicated in **Table 1**. The appropriate volume of each unlabeled amino-acid is transferred to a 1.5 mL vial, and sonicated until the solution looks clear. The vials are transferred into a 50 mL tube. Tyr and Leu are added last, to avoid precipitation, prepare fresh before using, keep on ice.
20. Labeled amino-acids solutions (*see* Section 3.1, **Table 1** and **Note 7**): In our case Arg, Ile, Pro, Met and Phe. Weigh 11 mg of Arg, 23 mg of Ile and 8.9 mg of Pro. Dissolve the amino-acids in 0.39 mL, 1.0 mL and 0.46 mL of reconstitution buffer respectively (around 168 mM). 8.2 mg of Met are weighted and dissolved in 0.33 mL of reconstitution buffer with 50 μ L of the DTT solution. 18 mg of Phe are weighted, dissolved in 0.65 mL of the appropriate HEPES solution (*see* step 18), and sonicated for one minute. Prepare fresh before using, keep on ice.

2.3 Sample characterization

All solutions described should be prepared using Milli-Q water and stored at room temperature unless stated otherwise.

1. FA diluted solution: 37% formaldehyde in water, store at 4 °C.
2. Dimethyl sulfoxide (DMSO).
3. 1 M NaOH: To adjust pH.
4. 1 M HCl: To adjust pH.
5. 1 M Tris-HCl, pH 8.
6. SDS running buffer: 100 mM Tris, 100 mM HEPES, 0.1% SDS, adjust to pH 8 using NaOH or HCl.
7. Staining buffer: 0.5% Coomassie Brilliant Blue R-250 dye, 50% ethanol and 10% acetic

acid.

8. Destaining buffer: 20% ethanol and 10% acetic acid.
9. 2 x SDS loading buffer: 8% SDS, 40% glycerol, 0.4% bromophenol blue, 0.4 M DTT, 0.2 M Tris, adjust to pH 8 using NaOH or HCl.
10. 1 x SDS loading buffer: dilute 1 mL of 2 x SDS loading buffer in 1 mL of water.
11. DSS solution: 50 mM disuccinimidylsuberate in DMSO.
12. HEPES-KCl buffer (*see* Section 2.1)
13. Polyvinylidene fluoride (PVDF) membranes.
14. TBS-Tween buffer 1: 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05 % Tween-20.
15. Monoclonal anti-histidine-peroxidase conjugate antibodies solution (for the dot blot):
Dilute 2,000 x from the stock solution with TBS-Tween buffer 1, prepare fresh.
16. Diaminobenzidine solution: Dissolve 1 tablet of diaminobenzidine and 1 tablet of urea hydrogen peroxide in 5 mL of water, prepare fresh.
17. Western-blotting detection kit.
18. Non-fat dry milk (powder).
19. Chemiluminescence reagents (H₂O₂ and luminol), store at 4 °C.
20. Towbin buffer: 25 mM Tris, 192 mM glycine, 20% isopropanol and 0.1% SDS. Store without isopropanol at 4 °C, add isopropanol just before use.
21. Ponceau S solution: 0.1% Ponceau S, 5% acetic acid.
22. TBS-Tween buffer 2: TBS-Tween buffer 1 with 5% (w/v) non-fat dry milk, freshly made.
23. Mouse anti-histidine antibodies solution (for the western-blot): Dilute 5,000 x from the stock solution with TBS-Tween buffer 1, prepare fresh.
24. Anti-mouse-peroxidase coupled antibodies solution (for the western-blot): Dilute 10,000 x from the stock solution with TBS-tween buffer 1, prepare fresh.
25. Micro BCA Protein Assay kit (Pierce).
26. Nuclease-free 1.5 mL tubes.
27. 1 mg/mL Bovine serum albumine (BSA) stock solution.
28. Solubilization buffer: 0.2% Triton X-100, solubilized in HEPES-KCl buffer, store at 4 °C for up to six months.
29. Acetone, stored at -20 °C.
30. Nitric acid solution: 1%, store at room temperature in a safety cabinet. Handle under the hood.

31. Perchloric acid solution: 70%, store at room temperature in a safety cabinet. Handle under the hood.
32. Ammonium molybdate solution: Dissolve 2.5 g in 100 mL of water, store at 4 °C.
33. Ascorbic acid solution: Dissolve 10 g in 100 mL of water, prepare fresh, do not keep more than one week at 4 °C.
34. Phosphate stock solution: Dissolve 10 mg KH_2PO_4 in 100 mL of water, store at 4 °C.
35. Nuclease-free 3 mL ultracentrifugation tubes.
36. Sucrose.
37. Sucrose solutions: 15%, 30% and 45% (w/v) sucrose, solubilized in HEPES-KCl buffer, store at 4 °C.

2.4 NMR sample preparation

4 mm diameter rotors for high-resolution magic-angle spinning solid-state NMR (*see Fig. 2*).

3. Methods

The methods outlined below describe: (1) calculation of the amount of incorporated amino-acids, (2) cell-free expression of the mechanosensitive channel MscL in detergent micelles, purification of protein/detergent micelles, protein reconstitution into liposomes, (3) cell-free expression directly into liposomes, (4) sample characterization and (5) NMR sample preparation.

3.1 Calculation of the amount of amino-acids to be incorporated during protein expression.

The concentration of amino-acids of each type to be incorporated in the protein cell-free expression system ($V = 110$ mL, *see Fig. 1*) is usually around 1 mM. For a given protein, each amino-acid concentration should depend on its occurrence, n , in the protein sequence. In our case, we have found good results with final concentrations of amino-acids to be $n/10$, expressed in mM. The corresponding weight of each amino-acid, of molecular weight MW_i , is therefore $(n \times V \times MW_i / 10)$, expressed in mg. The corresponding volume of each amino-acid solution of concentration C_i is $(n \times V / 10) / C_i$, expressed in mL. The weights and final volumes calculated for this protocol are

summarized in **Table 1**.

3.2 Cell-free expression of the mechanosensitive channel MscL in detergent micelles

The method below describes the production of an MscL sample for solid-state NMR, using a cell-free protein expression system in detergent micelles. Although any labeling scheme could be performed, the sample described below was labeled on isoleucines and threonines, using $^{13}\text{C}/^{15}\text{N}$ -labeled amino-acids, and was subsequently used in solid-state NMR experiments (**14**).

3.2.1 Day one: Cell-free protein expression in detergent micelles

1. Thaw the DTT, RTS reconstitution buffer and MscL plasmid at room temperature.
2. Thaw the other components of the RTS kit (*E. coli* lysate, reaction mix and feeding mix) on ice.
3. Prepare the unlabeled amino-acid solutions (*see* Section 2.1).
4. Prepare the labeled amino-acid solutions (*see* Section 2.1).
5. Add the labeled amino-acid solution to the unlabeled amino-acid solution.
6. Add 3 mL of 40 mM DTT, sonicate (80 W for 5 minutes or until clear) and store on ice.
7. Reconstitute the lyophilized *E. coli* lysate in 5.2 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
8. Reconstitute the lyophilized reaction mix in 2.2 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
9. Reconstitute the lyophilized feeding mix in 80 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
10. Prepare the feeding solution by adding 26 mL of the reconstituted amino-acid solution (mixture of unlabeled and labeled amino-acid solutions) and 3 mL of 40 mM DTT to the feeding mix (80 mL).
11. Prepare the reaction solution by adding the reconstituted reaction mix (2.2 mL), 2.7 mL of the reconstituted amino-acid solution and 0.3 mL of 40 mM DTT to the reconstituted *E. coli* lysate (5.2 mL). Take a 20 μL aliquot of this solution for later comparison on gel electrophoresis.
12. Add 300 μL of the MscL plasmid ($\sim 150 \mu\text{g}$) to the reaction solution.
13. Add 200 μL of the 20% Triton X-100 solution to the reaction solution.
14. Add 2.2 mL of the 20% Triton X-100 solution to the feeding solution.

15. Open both screws of the reaction compartment and fill the reaction compartment with the reaction solution using a nuclease-free pipette (*see Fig. 1*). Remove air bubbles by slightly tapping the vessel.
16. Open both screws of the feeding compartment and fill the feeding compartment with the feeding solution, using the provided syringe (*see Fig. 1*). Remove air bubbles by slightly tapping the vessel.
17. Insert the reaction vessel into a thermoregulated shaker.
18. Set the shaking speed to 800 rpm.
19. Set the temperature to 30 °C.
20. Incubate the reaction for 22 hours.

3.2.2 Day two: Purification of detergent / protein micelles

1. After 22 hours, stop the cell-free protein expression. Take a 20 µL aliquot of the reaction run-on solution for gel electrophoresis.
2. Remove the reaction run-on mix using a pipette, add 100 µL of the AEBSF solution to prevent protease digestion, and stir at room temperature for 15 minutes (*see Note 9*).
3. Dilute the sample to a final volume of 50 mL using FPLC buffer A1, and centrifuge at 10,000 g for 15 minutes at 4 °C.
4. Connect the column to the FPLC system (*see Note 10*). Pass 4-5 column volumes (CV) of distilled water through the column, to wash away the ethanol solution in which it has been kept.
5. Degas the FPLC buffers for 10 minutes in an ultrasonic bath. Wash the pump inlets with the degassed solutions. Connect the chelating column into the FPLC system. Wash the system with FPLC buffer A1 until the UV and conductivity baselines are stable.
6. At the end of the sample centrifugation (*see step 3*), dilute the supernatant (reaction run-on mix) with FPLC sample loading buffer to a final volume of 100 mL (final dilution ~10x).
7. Filter this solution through a 0.45 µm filter.
8. Resuspend the pellet with 5 mL of FPLC buffer A1 and remove a 20 µL aliquot for gel electrophoresis.
9. Introduce half (50 mL) of the diluted reaction run-on mix into the column (*see Note 11*). Collect the flow through in the first 50 mL tube.
10. Wash off the non-specifically bound material by raising the concentration of imidazole to

100 mM, by mixing 80% of FPLC buffer A1 with 20% of FPLC buffer B (6 CV total).

The concentration of detergent remains high (around 3%). Collect the flow through in the second 50 mL tube.

11. To elute the protein, switch FPLC buffer A1 to buffer A2, to reduce the detergent concentration to 0.2%. Collect the flow through in the second 50 mL tube.
12. Increase the imidazole concentration to 200 mM by mixing 60% of solution A2 with 40% of solution B (2 CV total), and collect eluant fractions of 2 mL each (*see Note 12*).
13. Increase the imidazole concentration to 500 mM by switching to 100% FPLC buffer B (6 CV total), and collect eluant fractions of 2 mL each.
14. Wash the column with FPLC buffer A1 (10 CV or until the UV and conductivity baselines are stable), then repeat steps 9 to 13 with the remaining 50 mL (second half of the diluted reaction run-on mix, *see Note 11*).
15. At the end of the FPLC protein purification run (*see Note 10*), the potentially interesting fractions are separated into approximately 40 tubes, whereas the flow-through is contained in two 50 mL tubes. Remove a 20 μ L aliquot from each of these 42 tubes, for gel electrophoresis.
16. Determine which fractions contain the histidine-tagged protein by performing a dot blot (*see Section 3.4.4*).
17. Wash the two dialysis cassettes in 1 x dialysis buffer.
18. For each half of the solution, pool the fractions containing the protein into a dialysis cassette.
19. Dialyze against 1 L of 1 x dialysis buffer, overnight at 4 °C.

3.2.3 Day three: Sample characterization and reconstitution into liposomes.

1. Dialyze a second time against 1 L of 1 x dialysis buffer for two hours at 4 °C. Remove a 20 μ L aliquot for gel electrophoresis.
2. Assess cell-free expression level, FPLC purification efficiency, and oligomeric state by SDS-PAGE (*see Sections 3.4.1 and 3.4.3*).
3. After SDS-PAGE, wash the gel in water and perform a western-blot transfer (*see Section 3.4.5*).
4. Estimate the protein concentration. Since MscL does not contain any tryptophan it has to be quantified by using the BCA method (*see Section 3.4.6 and Note 13*).

5. BCA quantification of our MscL preparation indicated that each half of the solution contains approximately 4.8 mg of protein in a 10 mL volume that will be reconstituted in 19.2 mg of DOPC liposomes (protein/lipid ratio of 1/4). DOPC can be replaced by other lipids or lipid mixtures, at a higher or lower concentration (**21**). Lipids can also be replaced by other mixtures to reconstitute the protein into bicelles or nanodiscs (**22, 23**).
6. Split the solutions into four fractions of 5 mL, each containing 2.4 mg of protein and 10 mg of Triton X-100. To each fraction, add 9.6 mg of DOPC solubilized in 2.5 mL of the 0.8% Triton X-100 solution, such that the final lipid/detergent ratio is 1/3. Stir slowly for 15 minutes at room temperature.
7. To each fraction of 7.5 mL, add 300 mg of wet polystyrene beads (detergent/bead ratio is 1/10) and put the four tubes on a rotator, for removing the detergent by adsorption onto the beads. Incubate overnight at 4 °C. For alternative methods, *see Note 14*.

3.2.4 Day four: Preparation of the sample for solid-state NMR

1. Add another 300 mg of wet polystyrene beads to each tube. Incubate for another two hours, at 4 °C.
2. After 2 hours, set the tubes vertically for 5 minutes and discard the pellet and the beads.
3. Divide the supernatant into 4 tubes, and centrifuge at 100,000 g for 30 minutes at 4 °C. Remove a 20 µL aliquot of one supernatant for gel electrophoresis, discard the rest and resuspend the pellets with 1 mL of the HEPES-KCl buffer for each pellet. Split the new resuspended pellets into 2 tubes, and centrifuge at 100,000 g for 30 minutes at 4 °C. Repeat one more time so that the entire sample fits into a single tube. Remove a final 20 µL aliquot of the supernatant for gel electrophoresis. Dry the final pellet under argon and store it at -20 °C.
4. The reconstitution is assessed by SDS-PAGE (*see* Section 3.4.1), and by proteoliposome density characterization on a sucrose gradient (*see* Section 3.4.9). Lipid and water contents are also estimated (*see* Section 3.4.8 and 3.4.10, and **Note 15**).
5. The average quantity of protein obtained with the RTS 9000 kit is 10 mg in 10 mL of reaction mix, which is generally sufficient to prepare two or three NMR samples. Transfer ~30 mg of pelleted sample to the NMR rotor (*see* Section 3.5).
6. Store the remaining pellet at -20 °C.

3.3 Cell-free expression of the mechanosensitive channel MscL in liposomes.

The method below describes the production of an MscL sample for solid-state NMR, using a cell-free protein expression system directly into lipid vesicles. Although any labeling scheme could be performed, the sample described below was labeled on arginines, isoleucines, methionines, phenylalanines and prolines, using $^{13}\text{C}/^{15}\text{N}$ -labeled amino-acids, and was subsequently used in solid-state NMR experiments (**19**). Lipids used were synthetic DOPC, but other lipids or complex mixtures can also be used, such as asolectin (**18**) or nanodiscs (**17**).

3.3.1. Day one: Cell-free protein expression

1. Thaw the DTT solution, RTS reconstitution buffer and MscL plasmid at room temperature.
2. Thaw the other components of the RTS kit (*E. coli* lysate, reaction mix and feeding mix) on ice.
3. Prepare the DOPC liposomes (*see* Section 2.2).
4. Prepare the HEPES solutions (*see* Section 2.2).
5. Prepare the unlabeled amino-acid solutions (*see* Section 2.2).
6. Prepare the labeled amino-acid solutions (*see* Section 2.2).
7. Add the labeled amino-acid solution to the unlabeled amino-acid solution.
8. Add 3 mL of 40 mM DTT, sonicate (80 W for 5 minutes or until clear) and store on ice.
9. Reconstitute the lyophilized *E. coli* lysate in 2.7 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
10. Add the liposome solution (2.5 mL) to the *E. coli* lysate once the latter is completely dissolved.
11. Reconstitute the lyophilized reaction mix in 2.2 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
12. Reconstitute the lyophilized feeding mix in 80 mL of reconstitution buffer. Shake the bottle gently (*see* **Note 8**).
13. Prepare the feeding solution by adding 26 mL of the reconstituted amino-acid solution (mixture of unlabeled and labeled amino-acid solutions) and 3 mL of 40 mM DTT to the feeding mix solution (80 mL).
14. Prepare the reaction solution by adding the reconstituted reaction mix (2.2 mL), 2.7 mL of

the reconstituted amino-acid solution and 0.3 mL of 40 mM DTT to the reconstituted *E. coli* lysate with the liposomes (5.2 mL). Take a 20 μ L aliquot of this reaction solution for later comparison on gel electrophoresis.

15. Add 300 μ L of the MscL plasmid (~150 μ g) to the reaction solution.
16. Open both screws of the reaction compartment and fill the reaction compartment with the reaction solution using a nuclease-free pipette (*see Fig. 1*). Remove any air bubbles by slightly tapping the vessel.
17. Open both screws of the feeding compartment and fill the feeding compartment with the feeding solution, using the provided syringe (*see Fig. 1*). Remove any air bubbles by slightly tapping the vessel.
18. Insert the reaction vessel into a thermoregulated shaker.
19. Set the shaking speed to 800 rpm.
20. Set the temperature to 30 °C.
21. Incubate the reaction for 22 hours.

3.3.2 Day two: Preparation of the sample for solid-state NMR

1. After 22 hours, stop the cell-free protein expression. Take a 20 μ L aliquot of the reaction run-on solution for gel electrophoresis.
2. Remove the reaction run-on mix using a pipette, add 100 μ L of the AEBSF solution to prevent protease digestion, and stir at room temperature for 15 minutes (*see Note 9*).
3. Split the sample into 6 tubes, and centrifuge at 100,000 *g* for 30 minutes at 4 °C. Remove a 20 μ L aliquot of one supernatant for gel electrophoresis, discard the rest and resuspend the pellets with 1 mL of the HEPES-KCl buffer for each pellet. Split the new resuspended pellets into 3 tubes, and centrifuge at 100,000 *g* for 30 minutes at 4 °C. Repeat one more time so that the entire sample fits into a single tube. Remove a final 20 μ L aliquot of the supernatant for gel electrophoresis. Dry the final pellet under argon and store it at -20 °C.
4. Assess cell-free expression efficiency and oligomeric state by SDS-PAGE (*see Sections 3.4.2 and 3.4.3*). At this point, the expressed protein is generally almost pure.
5. If the protein was expressed with a poly-histidine tag, after SDS-PAGE, wash the gel in water and perform a western-blot transfer (*see Section 3.4.5*).
6. Estimate the protein and lipid concentration using the BCA and Rouser methods respectively (*see Sections 3.4.7 and 3.4.8, and Note 13*).

7. Characterize the proteoliposome density on a sucrose gradient (*see* Section 3.4.9). Water content is also estimated (*see* Section 3.4.10 and **Note 15**).
8. The average quantity of protein obtained with the RTS 9000 kit is 10 mg in 10 mL of reaction mix, which is generally sufficient to prepare two or three NMR samples. About 30 mg of pelleted sample is transferred into the NMR rotor (*see* Section 3.5).
9. Store the remaining pellet at -20 °C.

3.4 Sample characterization

3.4.1 Electrophoresis of protein expressed in detergent micelles

Perform SDS-PAGE on the following sample aliquots:

- (1) aliquots from the reaction mix before and after the cell-free expression reaction, to check out the expression of the targeted protein; (2) aliquots from the centrifugation pellet and supernatant, to check whether the protein was totally solubilized in detergent micelles or if there are some aggregates; (3) check the FPLC results by comparing aliquots from the purification flow through and the purified fractions; (4) check also the integrity of the protein after dialysis; (5) after the reconstitution, analyze the incorporation of the protein by comparing aliquots of the protein-detergent micelles with the proteoliposomes; (6) check also aliquots of the supernatant of all the wash steps.

3.4.2 Electrophoresis of protein expressed in liposomes

Perform SDS-PAGE on the following sample aliquots:

- (1) aliquots from the reaction mix before and after the cell-free expression reaction, to check out the expression of the targeted protein; (2) check also the centrifugation supernatants.

3.4.3 Electrophoresis for oligomeric state characterization

If the protein is an oligomer (like MscL), assess its integrity by cross-linking the protein with different free primary amine targeting reagents, like formaldehyde (FA) or disuccinimidylsuberate (DSS) (**3, 25**). *N.b.* the protein buffer must not contain primary amines (such as in Tris).

1. Mix 5 µg of purified MscL, either in detergent or liposomes, in 10 µL of HEPES-KCl buffer and 0.54 µL of the FA diluted solution.
2. Mix 5 µg of purified MscL, either in detergent or liposomes, in 10 µL of HEPES-KCl

buffer and 0.2 μL of 50 mM DSS.

3. Incubate both samples, without shaking, for 30 minutes at room temperature.
4. Stop the cross-linking reactions by adding 2 μL of 1 M Tris-HCl, pH 8, to each sample. Gently mix and incubate for 10 minutes at room temperature.
5. Add 3 μL of 2 x SDS loading buffer to each sample and mix.
6. Heat the FA sample to 60 °C and the DSS sample to 90 °C, for 5 minutes in a water bath.
7. Cool the samples and load on a gel.
8. After migration, wash the gel in water and either Coomassie blue-stain (*see Fig. 3*) or perform a western-blot transfer (*see Section 3.4.5*).

3.4.4. Dot blot

1. Deposit 2 μL of each fraction directly onto a PVDF membrane.
2. Allow the protein solution to air dry.
3. Prepare the TBS-Tween buffers 1 and 2.
4. Prepare the monoclonal antibodies solution.
5. When the protein spots are completely dry, incubate the membrane in the TBS-Tween buffer 2 for 30 minutes.
6. Wash the membrane twice for 5 minutes with the TBS-Tween buffer 1.
7. Incubate the membrane with the monoclonal antibodies solution for 1 hour.
8. Wash the membrane twice for 5 minutes with the TBS-Tween buffer 1.
9. Prepare the diaminobenzidine solution.
10. Dry the membrane on a tissue.
11. Incubate the membrane with the diaminobenzidine solution for a few seconds, until the spots are colored.
10. Dry the membrane on a tissue. The color remains on the spots containing the histidine-tagged protein, while it vanishes on the other spots.

3.4.5 Semi-dry western-blot

1. Prepare the Towbin buffer.
2. Prepare the transfer sandwich by superimposing two filter papers, the gel, the nitrocellulose membrane and another two filter papers in the Towbin buffer.
3. Place the transfer sandwich on the anode plate, and clip the cathode plate on top.

4. Clamp the potential at 25 V and blot for 20 minutes.
5. Prepare TBS-Tween buffers 1 and 2, mouse anti-histidine antibodies and anti-mouse-peroxidase coupled antibodies solutions.
6. Assess sample transfer and relative concentration by staining the membrane with Ponceau S solution. This is a reversible stain that can be removed by washing in TBS-Tween buffer 1.
7. After protein transfer and destaining, saturate the nitrocellulose membrane by incubating in TBS-Tween buffer 2 for 30 minutes.
8. Rinse the membrane quickly with TBS-Tween buffer 1.
9. Incubate for 45 minutes at room temperature with the mouse anti-histidine antibodies solution.
10. Wash the membrane for 4 x 5 minutes with TBS-Tween buffer 1.
11. Incubate for 30 minutes at room temperature with the anti-mouse-peroxidase coupled antibodies solution.
12. Wash the membrane for 5 x 5 minutes with TBS-Tween buffer 1.
13. Incubate the membrane for one minute with 1 mL of each of the chemiluminescence reagents (H₂O₂ and luminol).
14. In a light tight box, expose X-ray films to the membrane for varying times, usually between 30 seconds and 2 minutes, and develop the films in the photo developer.

3.4.6 Protein quantification in detergents

Protein concentration is measured using the BCA method (26, 27), using the test tube procedure of the Micro BCA Protein Assay Kit:

1. Prepare eight BSA standards by diluting the stock solution with the solubilization buffer (the final BSA concentration should fall between 0.5 and 20 µg/mL). Make three replicates of each dilution. Introduce 500 µL of each into 1.5 mL tubes (24 tubes in total), store at room temperature.
2. Dilute the unknown protein sample with the solubilization buffer to three different concentrations that are expected to be between 1 and 20 µg/mL. Make three replicates of each dilution. Transfer 500 µL of each sample into 1.5 mL tubes (9 tubes in total).
3. Add 500 µL of Micro BCA Working Reagent to each of the 1.5 mL tubes (24 BSA standards and 9 unknown proteins) and incubate for 60 minutes at 60 °C. Cool to room

temperature.

4. Measure the absorbance at 562 nm.
5. Plot a standard curve based on the absorbance of the BSA samples.
6. Deduce the protein concentration of each unknown protein sample.

3.4.7 Protein quantification in proteoliposomes

Lipids need to be removed from the sample by precipitating the proteins using cold acetone, followed by a centrifugation to separate the protein pellet from the supernatant containing the lipids:

1. Prepare the 24 BSA standard tubes (*see* Section 3.4.6, step 1).
2. Dilute the unknown protein sample with the solubilization buffer to three different concentrations that are expected to be between 1 and 20 $\mu\text{g}/\text{mL}$. Make three replicates of each dilution. Transfer 500 μL of each sample into 1.5 mL tubes (9 tubes in total).
3. Add 1 mL of cold acetone to each tube of unknown protein. Vortex the tubes and incubate for 60 minutes at $-20\text{ }^{\circ}\text{C}$. Centrifuge at 10,000 g for 10 minutes at room temperature. Discard the supernatant. Incubate the tubes for 30 minutes at room temperature to allow for acetone evaporation. Add 500 μL of the solubilization buffer, and vortex the tubes again.
4. Add 500 μL of Micro BCA Working Reagent to each of the 1.5 mL tubes (24 BSA standards and 9 unknown proteins) and incubate for 60 minutes at $60\text{ }^{\circ}\text{C}$. Cool to room temperature.
5. Measure the absorbance at 562 nm.
6. Plot a standard curve based on the absorbance of the BSA samples.
7. Deduce the protein concentration of each unknown protein sample.

3.4.8 Lipid quantification in proteoliposomes

The phospholipid content in the proteoliposome sample is assessed by the Rouser method, which measures the phosphate concentration in the sample (**28**). Full eye, face and skin protection is required for this method.

1. Set a heating block at $180\text{ }^{\circ}\text{C}$.
2. Wash the glass tubes in nitric acid solution before use, and dry in oven.
3. Prepare 5 phosphate standard samples by diluting the stock solution into five samples

containing 1 to 5 μg of KH_2PO_4 per tube. Classically, 5 μg of KH_2PO_4 give an absorbance of 0.9 at 800 nm. Store at 4 °C.

4. Prepare the ascorbic acid solution (*see* Section 2.3).
5. Collect three proteoliposomes volumes containing 1 to 5 μg of lipids approximately. Transfer the samples into clean glass tubes.
6. Add 0.65 mL of the perchloric acid solution and place the tubes in the heated block at 180 °C for 30 minutes of digestion, until the yellow color has disappeared.
7. Add 0.65 mL of the perchloric acid solution per phosphate standard tube (digestion is not necessary).
8. Put all the tubes on ice, and set the heated block at 100 °C (alternatively, a boiling water bath can be used)
9. Once cooled, add to the tubes: 3.3 mL water, 0.5 mL of the ammonium molybdate solution and 0.5 mL of the ascorbic acid solution. Agitate on a vortex after each addition.
10. Put the tubes in the heated block at 100 °C for 5 minutes.
11. Put the tubes on ice for 5 minutes.
12. Read the absorbance of the cooled samples at 800 nm.
13. Plot a standard curve based on the absorbance of the phosphate standard samples. The exact concentration in the phospholipid samples is calculated based on the standard curve.

3.4.9 Proteoliposomes density characterization.

A discontinuous sucrose flotation gradient analysis (**29, 21**) can be performed to check that the membrane protein is correctly reconstituted into proteoliposomes. The proteoliposomes can be layered *at the bottom* (see below) or *at the top* (see **Note 16**) of the sucrose layers.

1. After MscL reconstitution, add 120 mg of sucrose to 0.2 mL of the proteoliposome suspension (containing approximately 96 μg of protein) and mix gently by pipetting, until the sucrose has completely dissolved. The final volume is approximately 0.265 mL and the final sucrose concentration is 45%. Adjust the final volume to 0.4 mL with the buffered 45% sucrose solution.
2. Transfer the resulting suspension to a 3 mL ultracentrifuge tube, and keep it vertical in an appropriate rack for the following steps.
3. Carefully deposit 0.7 mL of the buffered 30% sucrose solution, with minimal flow along the tube wall to avoid layers mixing.

4. Repeat with 0.7 mL of the buffered 15% sucrose solution.
5. Ultracentrifuge at 100,000 g for 1.5 h in a swinging rotor at 18 °C.
6. At the end of the centrifugation, carefully remove and transfer the tube into the rack, keeping it vertical. Carefully collect 6 fractions of 0.3 mL, from top to bottom, into 1.5 mL tubes, and adjust the final volume to 1 mL with water.
7. Wash the bottom of the 3 mL tube with 100 μ L of 1x SDS loading buffer, and remove a 15 μ L aliquot for gel electrophoresis.
8. Centrifuge the 1.5 mL tubes at 16,000 g for 15 minutes at room temperature, to pellet the proteoliposomes and remove the sucrose.
9. Resuspended each pellet in 100 μ L of 1x SDS loading buffer, and take a 15 μ L aliquot of each for gel electrophoresis.
10. Analyze the aliquots by gel electrophoresis. The proteoliposomes usually appear at the interface between 30% and 45% sucrose. Protein-free liposomes lay at the top of the sucrose gradient, while lipid-free protein aggregates lay at the bottom of the sucrose gradient (*see Note 17*).

3.4.10 Water quantification in the sample

Water content is estimated by weighing 10 mg of sample prior to and after 16 h in vacuum.

3.5 NMR sample preparation

1. Weigh the empty rotor, cap, insert bottom, top and screw (*see Fig. 2*).
2. Introduce the insert bottom into the rotor.
3. Introduce a couple of mg of the pelleted sample into the rotor on the tip of a spatula. Place the rotor into a 1.5 mL tube and centrifuge it at 10,000 g for one minute at room temperature. Repeat this until approximately 30 mg of sample has been introduced into the rotor.
4. Place the insert top into the rotor, without the top screw. Clean the upper part the rotor with a precision wiper before capping tightly. At this point, the rotor containing sample should be stored at -20 °C until the NMR experiment is performed.
5. Introduce the rotor into the magic-angle spinning NMR probe and into the magnet of the NMR spectrometer. Spin the rotor at ~10 kHz for 15 minutes.
6. Extract the rotor and open it. Clean the upper part of the rotor again, in case a drop of

water has come out. Place the top screw, tighten it and then cap the rotor tightly. Weigh the full rotor to deduce the final sample mass. A typical sample consists of 3 mg proteins, 12 mg lipids and 15 mg water.

4. Notes

1. All solutions require Milli-Q water, but only solutions used inside the cell-free expression vessel need to be autoclaved nuclease-free, to make sure no ribonucleases are present that would damage the RNAs in the lysate.
2. Continuous-exchange cell-free protein expression with an RTS commercial kit from Roche/5Prime is advantageous for its convenience, reliability, for saving time and manpower in preparing the lysate, and for managing the stocks. The thermoregulated shaker, on the other hand, is not necessarily from Roche/5Prime.
3. Cell-free reactions should be performed on a small scale (RTS 100 or analogue kits) for optimization studies, including detergents or lipids. Protein yield should be evaluated in a cell-free kit equipped with a continuous-exchange system (e.g. RTS 500), where the yield is usually higher. Once the protocol is optimized, the proteins can then be expressed on larger scales (e.g. RTS 9000).
4. Cell-free kits were stored at -80 °C rather than -20 °C, which increases the lifespan of the kits to over a year. Freeze-thaw cycles should be avoided.
5. Circular DNA was used for cell-free expression of MscL. Linear PCR templates can also be used, but they require larger amounts of DNA. DNA quality is important for cell-free expression. A plasmid purification procedure with a good yield (typically 100 µg of plasmid DNA per standard MIDI-prep), including an anion-exchange step, was found to be necessary. DNA should be dissolved in either nuclease-free water (our case) or Tris solution (typically 10 mM, at a pH between 8 and 8.5), but not in a buffer containing EDTA, as it would change the free magnesium ion concentration and reduce the protein expression yield.
6. Some amino-acids are difficult to solubilize. Sonication or heating at 60 °C can improve solubility. Met (M) and Cys (C) have to be supplemented with DTT (4 mM and 8 mM respectively). Tyr (Y), Trp (W) and Phe (F) have to be solubilized in 60 mM HEPES in nuclease-free water, adjusting the pH to 13, 1 or 7.5 respectively, using KOH or HCl.

However, Trp (W), Asp (D), Asn (N), Cys (C) or Tyr (Y) will not dissolve completely and have to be used as suspensions. Alternatively, the Roche/5Prime RTS Amino-Acid Sampler provides appropriate stock solution of each individual *unlabeled* amino-acid.

7. Commercial mixtures of labeled amino-acids are also available (30). All amino-acids, except for Gln (Q), can be specifically incorporated into the expressed proteins. Some scrambling is observed with Ser (S), Asp (D) and Asn (N). Some metabolic degradation can occur with Arg (R), Cys (C), Trp (W), Met (M), Asp (D) and Glu (E) during prolonged incubation. In such a case, an increased amount of amino-acids can enhance the protein production.
8. Too vigorous mixing should be avoided while preparing the different lysates, as it may denature proteins or ribosomes in the extracts.
9. AEBSF is added after the expression is complete, to prevent protease degradation of the target product. Proteases inhibitors can also be added to the reaction chamber, but they should be tested on a small scale beforehand.
10. If the chelating column is not pre-charged with Ni ions, before connecting it to the FPLC system, open it at both ends and wash it with 2 CV of distilled water, 1 CV of a 100 mM EDTA, pH 8 solution, 3 CV of distilled water, 2 CV of 0.1 M NiSO₄ solution drop by drop and 5 CV of distilled water. Once the column is charged, an additional washing step with 5 CV of 20 % ethanol is necessary before storage at 4°C. At the end of the protein purification run, wash the column with water (5 CV), 0.5 M NaOH (5 CV), water (5 CV), then store in 20% ethanol.
11. The total reaction run-on is expected to contain approximately 10 mg of protein but also detergent and impurities which may bind to the column as well. Since the binding saturation of the column surface is on the order of 10 mg/mL, the reaction run-on is split in two so as to always remain below the saturation limit.
12. For MscL, the fractions containing the protein are eluted between 200 and 500 mM of imidazole, but this is highly protein dependent. Depending on the histidine-tag accessibility, the protein may be eluted at lower or higher imidazole concentration.
13. Protein concentration is usually estimated using UV fluorescence of aromatic amino-acids. If the protein is not sufficiently fluorescent, quantification can be performed using the commercial bicinchonic acid protein assay (Micro BCA kit) by Pierce. Care should be taken if lipids are present in the sample.

14. Instead of using polystyrene beads, detergent extraction can be performed using cyclodextrin inclusion compounds, as described by DeGrip *et al.* (24), or by dialysis (21).
15. Proteoliposomes can be dehydrated and rehydrated, but care should be taken, as the process may affect protein activity.
16. The discontinuous sucrose flotation gradient analysis to check that the membrane protein is correctly reconstituted can be performed with the proteoliposomes layered either *at the top* or *at the bottom* of the sucrose layers, or both (29, 21). If at the top, the proteoliposomes must be mixed with sucrose for a final sucrose concentration of 15%. The rest of the protocol is the same: After layering 0.7 mL of 45% sucrose, then 0.7 mL of 30% sucrose and then 0.7 mL of the proteoliposomes in 15% sucrose, the proteoliposomes usually appear at the interface between 30% and 45% sucrose.
17. If the amount of lipid-free protein aggregates is not negligible, the discontinuous sucrose flotation gradient can also be used as a purification technique, by layering the entire reconstituted proteoliposome suspension.

Acknowledgements

This work was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (to A.A.), by the CNRS (UMR 7099 and 8619), the ANR (ANR-06-JCJC0014), the Université Paris Diderot and the Université Paris-Sud 11. We thank Alexandre Ghazi, Catherine Berrier, Emmanuelle Billon-Denis and Michiel A. Verhoeven for helping us optimize the protocols presented here.

References

1. Lyford, L. K., and Rosenberg, R. L. (1999) Cell-free expression and functional reconstitution of Homo-oligomeric $\alpha 7$ Nicotinic Acetylcholine Receptors into Planar Lipid Bilayers. *J. Biol. Chem.* **274**, 25675-25681
2. Elbaz, Y., Steiner-Mordoch, S., Danieli, T., and Schuldiner, S. (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1519-1524
3. Berrier, C., Park, K. H., Abes, S., Bibonne, A., Betton, J. M., and Ghazi, A. (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585-12591
4. Klammt, C., Lohr, F., Schafer, B., Haase, W., Dötsch, V., Ruterjans, H., Glaubitz, C., and Bernhard, F. (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur. J. Biochem.* **271**, 568-580
5. Wagner, S., Bader, M. L., Drew, D., and de Gier, J. W. (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol.* **24**, 364-371
6. Eshaghi, S. (2009) High-throughput expression and detergent screening of integral membrane proteins. *Methods Mol. Biol.* **498**, 265-271
7. Tate, C. G. (2010) Practical considerations of membrane protein instability during purification and crystallisation. *Methods Mol. Biol.* **601**, 187-203
8. Breyton, C., Pucci, B., and Popot, J.-L. (2010) Amphipols and fluorinated surfactants: Two alternatives to detergents for studying membrane proteins in vitro. *Methods Mol. Biol.* **601**, 219-245
9. White, S. H. (2010) Membrane Proteins of Known Structure. University of California at Irvine. http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html Accessed 21 December 2010
10. Warschawski, D. E. (2010) Membrane Proteins of Known Structure Determined by NMR. Drorlist. <http://www.drorlist.com/nmr/MPNMR.html> Accessed 21 December 2010
11. Schneider, B., Junge, F., Shirokov, V. A., Durst, F., Schwarz, D., Dötsch, V., and Bernhard, F. (2010) Membrane protein expression in cell-free systems. *Methods Mol.*

12. Sobhanifar, S., Reckel, S., Junge, F., Schwarz, D., Kai, L., Karbyshev, M., Löhr, F., Bernhard, F., and Dötsch, V. (2010) Cell-free expression and stable isotope labelling strategies for membrane proteins. *J. Biomol. NMR* **46**, 33-43
13. Lehner, I., Basting, D., Meyer, B., Haase, W., Manolikas, T., Kaiser, C., Karas, M., and Glaubitz, C. (2008) The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J. Biol. Chem.* **283**, 3281-3288
14. Abdine, A., Verhoeven, M. A., Park, K.-H., Ghazi, A., Guittet, E., Berrier, C., Van Heijenoort, C., and Warschawski, D. E. (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* **204**, 155-159
15. Kalmbach, R., Chizhov, I., Schumacher, M. C., Friedrich, T., Bamberg, E., and Engelhard, M. (2007) Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. *J. Mol. Biol.* **371**, 639-648
16. Marques, B., Liguori, L., Paclat, M. H., Villegas-Mendéz, A., Rothe, R., Morel, F., Lenormand, J.-L. (2007) Liposome-mediated cellular delivery of active gp91(phox). *PLoS One* **2**, e856
17. Katzen, F., Fletcher, J. E., Yang, J. P., Kang, D., Peterson, T. C., Cappuccio, J. A., Blanchette, C. D., Sulchek, T., Chromy, B. A., Hoeprich, P. D., Coleman, M. A., and Kudlicki, W. (2008) Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J. Proteome Res.* **7**, 3535-3542
18. Berrier, C., Guilvout, I., Bayan, N., Park, K.-H., Mesneau, A., Chami, M., Pugsley, A. P., and Ghazi, A. (2011) Coupled cell-free synthesis and lipid vesicle insertion of a functional oligomeric channel MscL - MscL does not need the insertase YidC for insertion in vitro. *Biochim. Biophys. Acta* **1808**, 41-46
19. Abdine, A., Verhoeven, M. A., and Warschawski, D. E. (2011) Cell-free expression and labeling strategies for a new decade in solid-state NMR. *N. Biotechnol.* **in press**, doi:10.1016/j.nbt.2010.07.014
20. He, M. (2008) Cell-free protein synthesis: applications in proteomics and biotechnology. *N. Biotechnol.* **25**, 126-132

21. Rigaud, J.-L., and Lévy, D. (2003) Reconstitution of Membrane Proteins into Liposomes. *Methods Enzymol.* **372**, 65-86
22. Triba, M. N., Zoonens, M., Popot, J.-L., Devaux, P. F., and Warschawski, D. E. (2006) Reconstitution and alignment by a magnetic field of a β -barrel membrane protein in bicelles. *Eur. Biophys. J.* **35**, 268-275
23. Leitz, A. J., Bayburt, T. H., Barnakov, A. N., Springer, B. A., and Sligar, S. G. (2006) Functional reconstitution of β 2-adrenergic receptors utilizing self-assembling Nanodisc technology. *BioTechniques* **40**, 601-612
24. De Grip, W. J., Van Oostrum, J., and Bovee-Geurts, P. H. M. (1998) Selective detergent-extraction from mixed detergent/lipid/protein micelles, using cyclodextrin inclusion compounds: a novel generic approach for the preparation of proteoliposomes. *Biochem. J.* **330**, 667-674
25. Sukharev, S. I., Schroeder, M. J., and McCaslin, D. R. (1999) Stoichiometry of the Large Conductance Bacterial Mechanosensitive Channel of E. coli. A Biochemical Study. *J. Membrane Biol.* **171**, 183-193
26. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76-85
27. Wiechelmann, K. J., Braun, R. D., and Fitzpatrick, J. D. (1988) Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation. *Anal. Biochem.* **175**, 231-237
28. Rouser, G., Fkeischer, S., and Yamamoto, A. (1970) Two dimensional thin layer chromatographic separation of polar lipids and determination of phospholipids by phosphorus analysis of spots. *Lipids* **5**, 494-496
29. Laird, D. M., Eble, K. S., and Cunningham, C. C. (1986) Reconstitution of mitochondrial F0.F1-ATPase with phosphatidylcholine using the nonionic detergent, octylglucoside. *J. Biol. Chem.* **261**, 14844-14850
30. Etezady-Esfarjani, T., Hiller, S., Villalba, C., and Wüthrich, K. (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. *J. Biomol. NMR* **39**, 229-238

Figure Captions

Fig. 1. Roche/5Prime RTS 9000 reaction vessel. The reaction compartment (10 mL) and the feeding compartment (100 mL) are each accessible through two screws. The reaction compartment contains the cell lysate, plasmid DNA, detergent or pre-formed liposomes and is where the coupled transcription/translation reaction takes place. The separate feeding compartment provides additional ions, energy substrates, nucleotides and amino-acids, through a semipermeable membrane (MW cut off 10 kD). Simultaneously, by-products that may inhibit the reaction are diluted through the same membrane into the feeding compartment. This continuous exchange allows cell-free expression to last for up to 24 hours.

Fig. 2. 4 mm diameter rotors for high-resolution magic-angle spinning solid-state NMR. The sample (approximately 50 μ l) is contained in the center of the zirconium rotor, thanks to the Teflon insert bottom and top, and is kept tightly in place by the top screw and the Kel-F cap.

Fig. 3. Coomassie blue stained SDS-PAGE of MscL: (1) purified in detergent ; (2) cross-linked with formaldehyde ; (3) cross-linked with disuccinimidylsuberate. Formaldehyde generates five protein bands on the gel, corresponding to various oligomeric forms, from monomers to pentamers, while disuccinimidylsuberate generates mostly pentamers, confirming the pentameric nature of *E. coli* MscL (3, 25).

Table 1

Preparation of amino-acid solutions for MscL cell-free expression in a volume $V = 110$ mL (see Section 3.1 and Fig. 1). n is the number of each amino-acid type, of molecular weight MW_i , in the protein sequence. The weight of each amino-acid is $(n \times MW_i \times 1.1 \times 10^{-2})$, expressed in mg. Since each amino-acid is solubilized at 168 mM, except for leucine, at 140 mM, the corresponding volume of each amino-acid solution is $(n \times 11)/168$, except for leucine where it is $(n \times 11)/140$, expressed in mL.

Tables

Table 1:

Amino-acid	<i>MWi</i> (g/mol)	<i>n</i>	Amount (mg)	Volume (mL)
Ala (A)	89	15	15	0.98
Arg (R)	174	6	11	0.39
Asn (N)	132	6	8.7	0.39
Asp (D)	133	7	10	0.46
Cys (C)	121	0	0	0.10
Gln (Q)	146	4	6.4	0.26
Glu (E)	147	8	13	0.52
Gly (G)	75	13	11	0.85
His (H)	155	7	12	0.46
Ile (I)	131	16	23	1.0
Leu (L)	131	13	19	1.0
Lys (K)	146	9	14	0.59
Met (M)	149	5	8.2	0.33
Phe (F)	165	10	18	0.65
Pro (P)	115	7	8.9	0.46
Ser (S)	105	4	4.6	0.26
Thr (T)	119	3	3.9	0.20
Trp (W)	204	0	0	0.10
Tyr (Y)	181	1	2.0	0.10
Val (V)	117	13	17	0.85

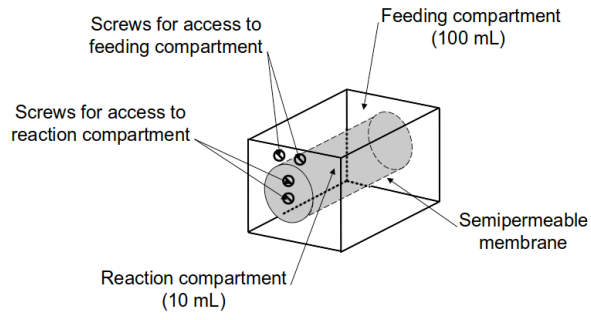


Fig. 1. Roche/5Prime RTS 9000 reaction vessel

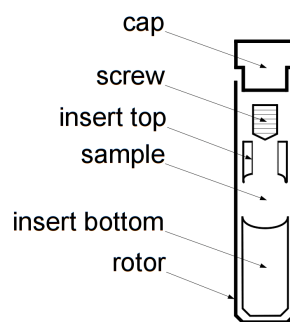


Fig. 2. 2.4 mm diameter rotors for high-resolution magic-angle spinning solid-state NMR

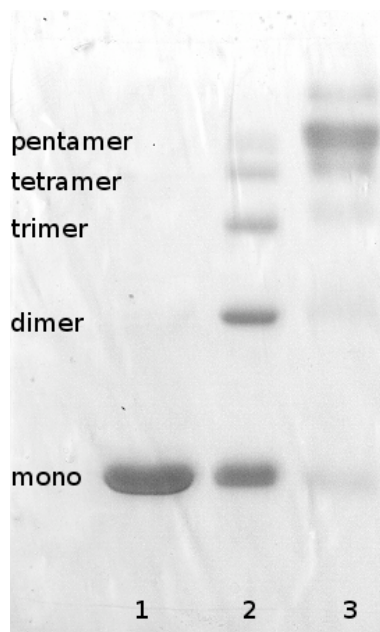


Fig. 3. Coomassie blue stained SDS-PAGE of MscL

Développement et applications de protocoles de synthèse *in vitro* du canal mécanosensible bactérien à large conductance, pour son étude structurale par résonance magnétique nucléaire en phase solide

Le génome de différents organismes contient près de 30% de séquences codantes pour des protéines membranaires. Celles-ci sont impliquées dans des processus biologiques tels que la signalisation cellulaire, la transduction énergétique ou le transport des métabolites. Cependant, leur étude structurale se heurte souvent aux problèmes de surexpression, ainsi qu'aux différents inconvénients liés à leur extraction de leur environnement natif.

Le canal mécanosensible à large conductance *MscL* d'*Escherichia coli* est une protéine intrinsèque de la membrane interne de la bactérie. L'activité de cette protéine est fortement dépendante de la membrane ; en effet, le canal s'ouvre lorsque la pression au niveau de la membrane augmente, lors d'un choc hypoosmotique, relarguant de l'eau et différents substrats afin que la bactérie retrouve un état osmotique adéquat à sa survie. Il est donc essentiel de recueillir des données structurales de la protéine dans son environnement natif. Pour cela, nous proposons une étude structurale de la protéine par résonance magnétique nucléaire en phase solide.

La surexpression de la protéine et son marquage aux isotopes ^{13}C et ^{15}N sont deux étapes clés d'un tel projet. Une surexpression bactérienne et un marquage uniforme de la protéine ont été effectués. Les données spectrales obtenues montrent une bonne résolution, mais l'imbrication des corrélations ne permet pas d'en extraire des données structurales. Afin de réduire le nombre de résidus marqués, un marquage spécifique a été appliqué par la voie de la synthèse *in vitro*. Le premier test a montré que l'approche permettait de réduire le temps d'acquisition, tout en diminuant la quantité d'informations. Différentes méthodes ont ensuite été appliquées pour trouver les meilleures combinaisons d'acides aminés à marquer spécifiquement en synthétisant la protéine *in vitro*. Ces approches utilisent les programmes de prédiction de déplacements chimiques ou l'analyse de la séquence à la recherche de paires d'acides aminés uniques. Une approche combinatoire a été également testée. Les différents échantillons synthétisés ont montré une bonne résolution, et ont permis l'identification de plusieurs corrélations.

Les méthodes développées au cours de ce travail pourront aider progressivement à déterminer la structure de la protéine *MscL*. Elles pourront également servir à d'autres protéines membranaires pour leur étude par résonance magnétique nucléaire.

