HAL
open science

Interaction dystrophine-membrane : structure 3D de fragments de la dystrophine en présence de phospholipides

Raphael dos Santos Morais

To cite this version:

Raphael dos Santos Morais. Interaction dystrophine-membrane : structure 3D de fragments de la dystrophine en présence de phospholipides. Médecine humaine et pathologie. Université de Rennes, 2017. Français. NNT : 2017REN1B062 . tel-01831209

HAL Id: tel-01831209

https://theses.hal.science/tel-01831209

Submitted on 5 Jul 2018

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

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

ANNÉE 2017

THÈSE / UNIVERSITÉ DE RENNES 1

sous le sceau de l'Université Bretagne Loire

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Biologie et Science de la Santé

Ecole doctorale Biologie-Santé

présentée par

Raphael Dos Santos Morais

Préparée à l'unité de recherche UMR CNRS 6290 IGDR
Institut de Génétique et de Développement de Rennes
UFR Sciences de la Vie et de l'Environnement (SVE),
au Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS et
sur la ligne SWING du Synchrotron SOLEIL

Interaction dystrophine/ membrane : structure 3D de fragments de la dystrophine en présence de phospholipides

Thèse soutenue à Rennes le 27 octobre 2017

devant le jury composé de :

Christine EBEL
Directrice de recherche CNRS - IBS / *rapportrice*

Erick DUFOURC
Directeur de recherche CNRS - CBMN / *rapporteur*

Véronique RECEVEUR-BRECHOT
Directrice de recherche CNRS - iSCB / e*xaminatrice (présidente du jury)*

Alexandros KOUTSIOUMPAS
Scientifique de ligne - JCNS at MLZ / *examinateur*

Javier PEREZ
Responsable de ligne de lumière- Synchrotron SOLEIL / *co-encadrant*

Olivier DELALANDE
Maître de conférences - Université de Rennes 1 / *co-encadrant*

Sophie COMBET
Chargée de recherche CNRS - LLB / *co-directrice de thèse*

Jean-François HUBERT
Professeur - Université de Rennes 1 / *directeur de thèse*

A toi Marine

REMERCIEMENTS

Je tiens à remercier le Conseil Régional de Bretagne, le Laboratoire Léon Brillouin et le Synchrotron SOLEIL de m'avoir octroyé un financement pour ces trois années de recherche. Je remercie également l'Association Française contre les Myopathies (AFM) dont le financement à grandement contribué à la réalisation de ces travaux de recherche.

Ces travaux ont été réalisés au sein (i) de l'équipe Structure et Interactions Moléculaires (SIM) de l'Institut de Génétique & Développement de Rennes (IGDR, UMR CNRS 6290), (ii) du Laboratoire Léon Brillouin (LLB, UMR12 CEA-CNRS) et (iii) de la ligne de lumière SWING du Synchrotron SOLEIL.

Je tiens à remercier Christine Ebel et Erick Dufourc d'avoir accepté de rapporter ce manuscrit, ainsi que Véronique Receveur-Brechot et Alexandros Koutsioumpas pour leur rôle d'examinateurs.

Je remercie mon directeur de thèse Jean-François Hubert, ma co-directrice de thèse Sophie Combet ainsi que mes co-encadrants Olivier Delalande et Javier Pérez. Merci à vous quatre, j'ai beaucoup appris à vos côtés. Si je devais résumer ces trois années en un seul mot, je choisirais le terme confiance. Merci d'avoir fait rimer ce mot avec autonomie, rigueur et épanouissement.

Je remercie le synchrotron SOLEIL, le Laboratoire Léon Brillouin, le centre Heinz Maier-Leibnitz et l'Institut Laue-Langevin pour l'attribution de temps faisceau, sur les instruments SWING, PACE, PA20, KWS-1, KWS-2, REFSANS, FIGARO et D22, indispensable pour mener à bien ce projet.

Je remercie toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce projet.

Je vous souhaite à toutes et tous le meilleur.

SOMMAIRE

ABREVIATIONS v
AVANT-PROPOS vii
PARTIE I : INTRODUCTION BIBLIOGRAPHIQUE 1
Chapitre 1: Des biomembranes aux modèles membranaires 1
1. Les biomembranes 1
1.1. Généralités 1
1.2. Les lipides membranaires 2
1.3. Dynamique de la membrane 5
2. Les modèles membranaires 6
2.1. Organisation des lipides 6
2.2. Les liposomes 9
2.3. Les micelles de détergent 10
2.4. Les amphipols 10
2.5. Les bicelles 11
2.6. Les nanodisques 11
2.7. Les monocouches de Langmuir 12
2.8. Les bicouches supportées 12
3. Le modèle membranaire « bicelle » 14
3.1. Généralités 14
3.2. L'importance de la concentration bicellaire critique 18
3.3. Le modèle idéal de bicelle 20
Chapitre 2: Mode d'interaction et structure des protéines membranaires 23
1. Les interactions protéine/lipide 23
1.1. Les protéines membranaires intégrales 23
1.2. Les protéines membranaires périphériques 24
2. Etude structurale des protéines membranaires 27
2.1. Généralités 27
2.2. Structure des protéines membranaires : approche par SAS 32
2.3. Structure des protéines membranaires : le modèle « bicelle » 34
Chapitre 3 : Dystrophine et dystrophinopathies 37
1. Dystrophinopathies et thérapies 37

1.1. Le muscle strié squelettique 37
1.2. Les dystrophinopathies 38
1.3. Exemples de thérapies 39
1.3.1. Thérapie génique 40
1.3.2. Thérapie cellulaire 42
2. Le gène *DMD* 42
3. La dystrophine 43
3.1. Organisation 43
3.2. Structure 45
3.3. Structure des protéines issues de délétions d'exons 47
3.4. Partenaires protéiques majeurs 49
3.5. Interaction avec les lipides membranaires 51
PARTIE II : MATERIEL ET METHODES 55
1. Production, extraction et purification des protéines recombinantes 55
1.1. Généralités 55
1.2. Culture et lyse bactérienne 56
1.3. Purification 56
1.3.1. Chromatographie d'affinité 56
1.3.2. Chromatographie d'exclusion stérique (SEC) 57
1.3.3. Quantification et concentration 57
1.3.4. Analyse SDS-PAGE 57
1.3.5. Immuno-empreinte (Western-blotting) 57
2. Préparation des modèles membranaires 58
2.1. Phospholipides 58
2.2. Les liposomes de DMPC 58
2.3. Les micelles de DHPC 59
2.4. Les bicelles 59
3. Caractérisation biophysique des objets d'étude 59
3.1. Estimation du R_h et de la MW apparente par SEC 59
3.2. Diffusion statique de la lumière (SLS) multi-angles (MALS) 60
3.3. Diffusion dynamique de la lumière (DLS) 61
3.4. Spectroscopie RMN ^{1}H et ^{31}P 62
4. Caractérisation des interactions protéine/lipide 63

4.1. Spectrofluorimétrie : fluorescence intrinsèque des tryptophanes 63
4.2. Dichroïsme circulaire (CD) 63
4.3. Thermophorèse micro-échelle (MST) 64
4.4. Marquage aux billes d'or 65
4.5. Coloration négative et microscopie électronique en transmission (TEM) 65
4.6. Cryo-micro électronique en transmission (Cryo-TEM) 66
4.7. Chimie-click et spectrométrie de masse 66
5. Caractérisation structurale des objets d'étude 67
5.1. Diffusion aux petits angles (small-angle scattering, SAS) 67
5.2. Analyse des données SAS 70
6. Réflectivité de neutrons 75
7. Modélisation moléculaire 80
PARTIE III : RESULTATS ET DISCUSSION 82
Contexte de travail et objectif de la thèse 82
Chapitre 1 : Caractérisation du modèle « bicelle » 84
1. Introduction 85
2. Publication 85
3. Résultats complémentaires 101
3.1. Micelles de DHPC et liposomes de DMPC 101
3.2. Analyse complémentaire des bicelles par SANS 102
3.3. Analyse complémentaire des bicelles par SAXS 104
4. Conclusion 107
Chapitre 2 : Analyse des interactions dystrophine/bicelle 109
1. Introduction 110
2. Publication 110
3. Résultats complémentaires 147
3.1. Stœchiométrie du complexe R1-3/bicelle anionique 147
3.2. R1-3/bicelles anionique : approche par cryo-EM 148
3.3. R11-15/bicelle : analyse par CG-MD 149
3.4. Digestion ménagée 150
3.5. Complexe protéine/bicelle étudié par SAXS 151
3.6. Thermostabilité de la dystrophine en présence de phospholipides 153
4. Conclusion 156

Chapitre 3 : Degré d'enfouissement de la dystrophine dans le sarcolemme 158
1. Introduction 158
2. Mesures NR en monocouches de Langmuir 158
3. Mesures NR en bicouches supportées 160
4. Conclusion 162
PARTIE IV : AUTRES CONTRIBUTIONS SCIENTIFIQUES 164
Chapitre 1 : Structure tridimensionnelle du domaine central de la dystrophine 164
1. Introduction 164
2. Publication 164
3. Résultats complémentaires 211
4. Conclusion 213
Chapitre 2 : Caractérisation des interactions dystrophine/actine 215
1. Introduction 215
2. Dystrophine/actine-F et thermophorèse 215
3. Cartographie des interactions dystrophine/actine 216
4. Conclusion 217
DISCUSSION GENERALE ET PERSPECTIVES 219
REFERENCES BIBLIOGRAPHIQUES 226
Résumé 241
Abstract 241

ABREVIATIONS

AAV	:	virus adeno-associé
ABD	:	domaine de liaison à l'actine
BMD	:	dystrophie musculaire de Becker
CBC	:	concentration bicellaire critique
CD	:	dichroïsme circulaire
CMC	:	concentration micellaire critique
DCG	:	complexe dystrophine et glycoprotéines associées
DDM	:	n-dodecyl β-D-maltoside
DHPC	:	dihexanoylphosphatidylcholine
DLS	:	diffusion dynamique de la lumière
DMD	:	dystrophie musculaire de Duchenne
DMPC	:	dimyristoylphosphatidylcholine
DMPS	:	dimyristoylphosphatidylsérine
DPC	:	dodécylphosphocholine
DPPC	:	dipalmitoylphosphatidylcholine
EDTA	:	acide éthylène diamine tétra-acétique
GST	:	glutathion-S-transférase
GUV	:	vésicule unilamellaire géante
IMP	:	protéine membranaire intégrale
LBD	:	domaine de liaison aux lipides
LUV	:	grande vésicule unilamellaire
MALS	:	diffusion de la lumière multi-angle
MLV	:	vésicules multilamellaires
MSP	:	« membrane scaffold protein »
MST	:	thermophorèse micro-échelle
MT	:	microtubule
MW	:	masse moléculaire
nNOS	:	oxyde nitrique synthase neuronale
pacFA PC	:	1-palmitoyl-2- (9- (3-pent-4-ynyl-3-H-diazirine-3-yl) -nonanoyl) -sn-glycéro- 3-phosphocholine
PDC	:	complexe protéine/détergent
PMP	:	protéine membranaire périphérique
R1	:	répétition numéro 1 du domaine central de la dystrophine
Rg	:	rayon de giration
Rh	:	rayon hydrodynamique
RI	:	indice de réfraction
RMN	:	résonnance magnétique nucléaire
SANS	:	diffusion des neutrons aux petits angles (small angle neutron scattering)
SAXS	:	diffusion des rayons-X aux petits angles (small angle X-ray scattering)
SDS-PAGE	:	« sodium dodecylsulfate-polyacrimide gel electrophoresis »
SEC	:	chromatographie d'exclusion stérique

SL	:	longueur de diffusion
SLD	:	densité de longueur de diffusion
SLS	:	diffusion statique de la lumière
SUV	:	petite vésicule unilamellaire
XRC	:	cristallographie aux rayons X
β-DG	:	β-D-dystroglycane

Ces travaux de recherche s'inscrivent dans la problématique générale des relations structure/fonction des protéines. Plus particulièrement, ils concernent une protéine essentielle au soutien de la membrane des cellules musculaires, la dystrophine. Des mutations génétiques sont à l'origine de l'absence ou de la production sous forme tronquée de cette protéine, et conduisant respectivement aux dystrophies musculaires de Duchenne et de Becker.

La dystrophine est une protéine essentielle des cellules musculaires au sein desquelles elle assure un rôle de charpente et de soutien du sarcolemme. Ce rôle est assuré grâce aux diverses interactions avec ses partenaires protéiques et également avec les lipides membranaires. Certaines mutations dans le gène *DMD* engendrent un déficit total en dystrophine, ce qui conduit à la dystrophie musculaire de Duchenne. Cette maladie très invalidante est associée à des phénotypes très sévères. Lorsque les mutations génétiques conduisent à la production de dystrophine sous forme tronquée plus ou moins fonctionnelle, la pathologie associée est dénommée dystrophie musculaire de Becker. Cette dernière est caractérisée par des phénotypes variés, allant de très sévères jusqu'à quasi asymptomatiques.

Une compréhension complète du rôle biologique de la dystrophine implique une caractérisation fine et précise du mode de liaison impliqué dans les interactions qu'elle assure avec le sarcolemme. L'objectif à long terme est de concevoir des thérapies de façon rationnelle ayant pour but le traitement des dystrophinopathies, notamment par le design de mini-dystrophines minimales fonctionnelles.

Ce manuscrit de thèse comporte trois parties principales et une quatrième partie, plus courte, qualifiée d'annexe par rapport au sujet principal de la thèse.

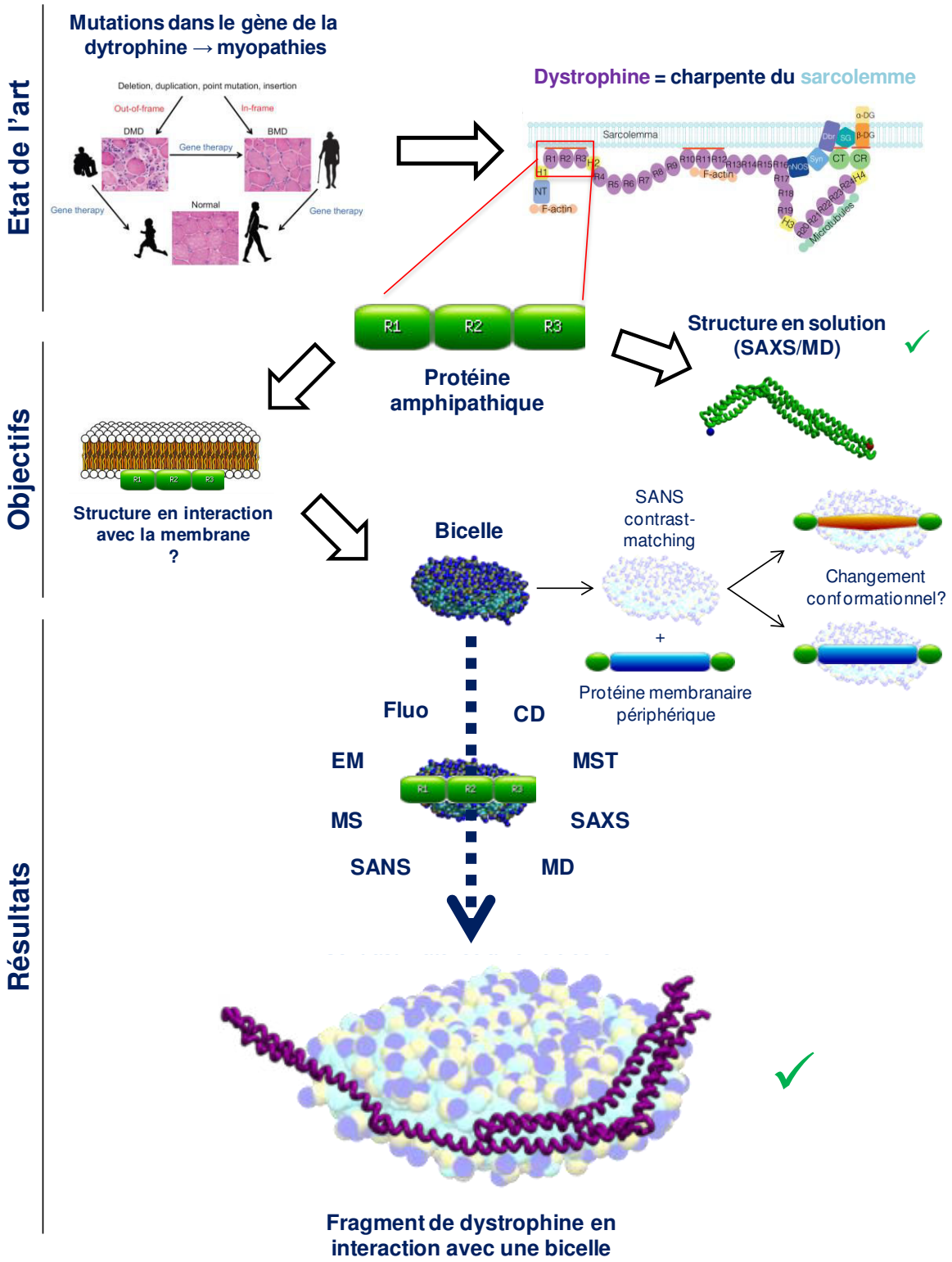
La partie I est une revue de la littérature comportant trois chapitres : le premier traite des biomembranes et de l'utilisation de modèles membranaires pour la réalisation d'études *in vitro*. Le deuxième chapitre concerne les interactions protéines/lipides. Enfin, le troisième est dédié à la dystrophine.

La partie II concerne les techniques et la méthodologie utilisées lors de ces travaux de thèse.

La partie III concerne l'ensemble des résultats obtenus et les discussions associées, tout d'abord la caractérisation du modèle membranaire « bicelle » puis son utilisation pour caractériser les interactions dystrophine/membrane. Le dernier chapitre de cette partie traite des interactions dystrophine/membrane aux interfaces.

La partie IV concerne les autres contributions scientifiques que j'ai apportées au cours de ma thèse, qui ne sont pas en lien direct avec la thématique principale.

Résumé graphique



PARTIE I :
INTRODUCTION BIBLIOGRAPHIQUE

Chapitre 1: Des biomembranes aux modèles membranaires

1. Les biomembranes

1.1. Généralités

Les cellules sont les « briques » de construction de base, comparables à des pièces de Lego®, de tout organisme vivant. Un corps humain contiendrait ~ 10^{14} cellules, de quoi faire rougir les meilleurs modèles de la célèbre firme danoise. Ces cellules sont majoritairement constituées d'eau et de macromolécules comme les acides nucléiques, les protéines et les lipides (Tableau 1).

Tableau 1 : Composition chimique approximative d'une cellule bactérienne, d'une levure et d'une cellule de mammifère. Adapté de [1].

		E. coli	*S. cerevisiae*	Cellule de mammifère
% masse totale	**Eau**	70	80	70
% masse sèche	**ADN**	3	0.1–0.6	1
	ARN	20	6–12	4
	Protéines	50–55	35–60	60
	Lipides	7–9	4–10	13

Les cellules sont séparées de leur environnement par des membranes biologiques, ou biomembranes, qui sont présentes chez tous les organismes avec un degré de complexité varié. Elles sont trouvées chez les virus enveloppés, les bactéries, les archées et les cellules eucaryotes. Ces biomembranes sont l'un des principaux constituants des cellules et ont pour fonction d'assurer une compartimentation. Elles jouent également un rôle de barrières sélectives pour les échanges intra- et inter-cellulaires. Dans les cellules eucaryotes, les membranes

permettent de séparer les différentes organelles composant la cellule (noyau, mitochondrie, réticulum endoplasmique, appareil de Golgi, …) de la partie cytosolique. Les membranes biologiques sont constituées d'une bicouche lipidique, ayant une épaisseur de ~ 5 nm, dans laquelle s'intègrent des protéines et possèdent un haut degré d'organisation et de complexité (Figure 1). De plus, elles sont organisées en bicouches asymétriques, ce qui signifie que les compositions des feuillets internes et externes sont différentes.

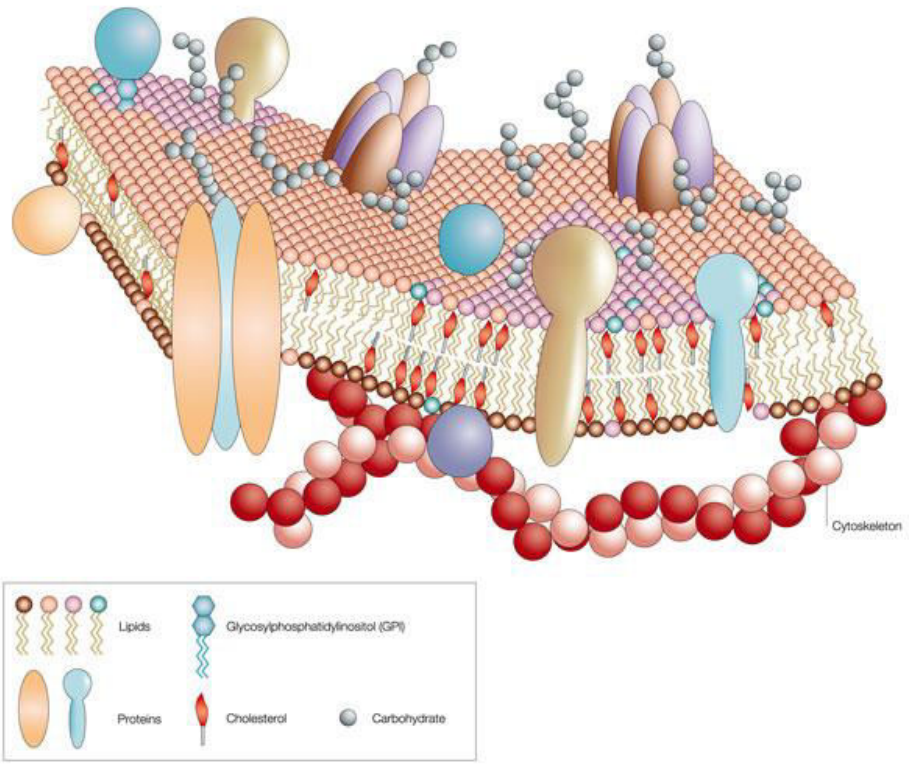


Figure 1 : Représentation schématique de la membrane plasmique composée de lipides et de protéines [2].

1.2.Les lipides membranaires

Les membranes cellulaires sont en grande partie composées de lipides. Ces lipides sont des molécules majoritairement amphiphiles. En effet, ils possèdent une tête polaire (hydrophile) et des chaînes aliphatiques apolaires (hydrophobes). Le nombre de carbones composant les chaînes grasses peut varier, tout comme le nombre d'insaturations, c'est-à-dire le nombre de double(s) liaison(s) C=C sur les chaînes aliphatiques. Parmi ces lipides, les plus représentés sont les diacylphosphoglycérides ou phospholipides (Figure 2). Ces phospholipides proviennent de l'estérification des trois fonctions alcool du glycérol par deux acides gras et par l'acide phosphorique. Le proton du groupement phosphate peut être substitué par une fonction

alcool (choline, éthanolamine, glycérol ou inositol) ou par un acide aminé de type sérine (Figure 2).

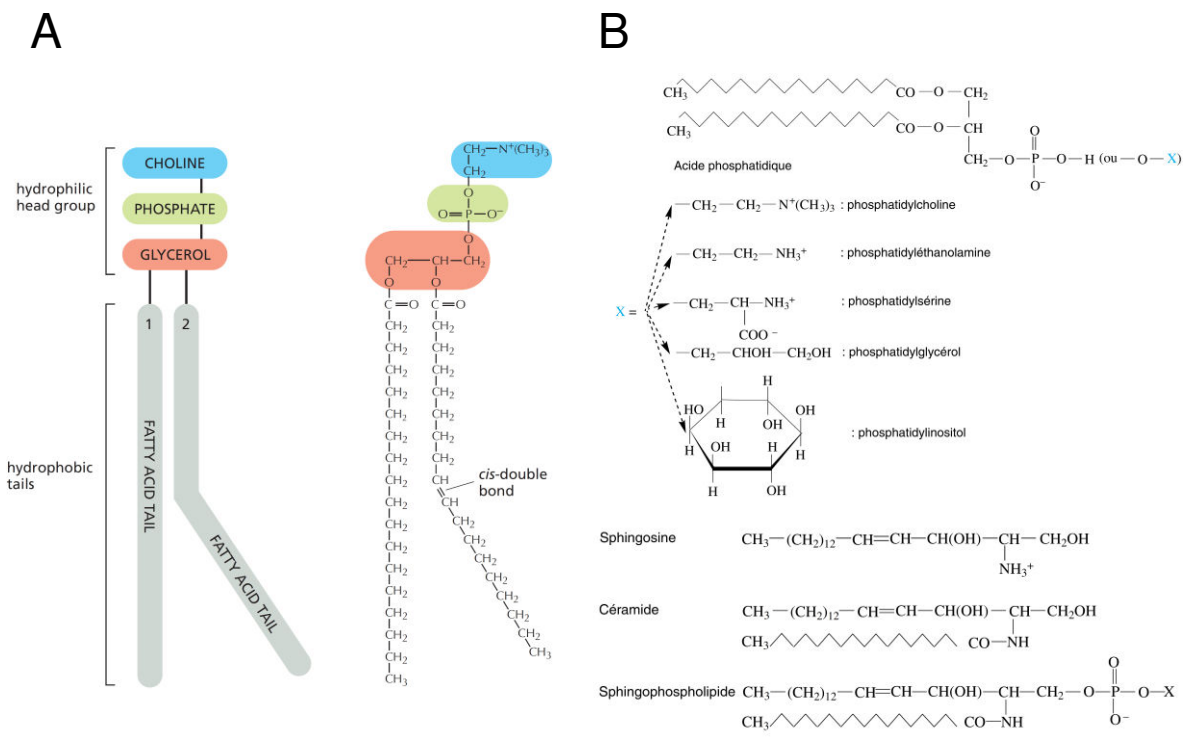


Figure 2 : (A) Détails de la structure hydrophile-hydrophobe des phospholipides : tête choline (bleu), groupement phosphate (vert), glycérol (orange) et chaînes carbonées (gris). (B) Formules chimiques de lipides composant les membranes biologiques. PC : phosphatidylcholine, PE : phosphatidyléthanolamine, PS : phosphatidylsérine, PG : phosphatidylglycérol, PI phosphatidylinositol. Adapté de [3,4].

Outre les phospholipides, les biomembranes sont également constituées de sphingophospholipides. Le nom est issu d'un de leurs composants, la sphingosine, qui est une molécule de type alcool aminé à longue chaîne carbonée (Figure 2). La formation d'une liaison amide entre la sphingosine et un acide gras donne un autre type de molécule appelé céramide (Figure 2). Finalement, lorsqu'un groupement phosphocholine (ou éthanolamine) substitue la fonction alcool du céramide, cette molécule est dénommée sphingophospholipide (Figure 2). Ce dernier type de lipides est notamment retrouvé dans les cellules neuronales où ils forment les gaines de myéline. Un autre type de lipides composant les membranes biologiques est celui des glycolipides. Cette catégorie est dénommée glycérolipide ou sphingoglycolipide (Figure 2), où la troisième fonction alcool du glycérol ou de la sphingosine forme une liaison O-osidique avec un carbohydrate (galactose ou acide sialique). Le dernier constituant majoritaire des membranes des cellules animales est le cholestérol, qui est un composé lipidique tétracyclique

jouant un rôle dans la fluidité des biomembranes. Tous ces lipides sont trouvés en proportion différente selon les types cellulaires, ainsi qu'en fonction des organelles et du feuillet de la membrane [4,5] (Figure 3).

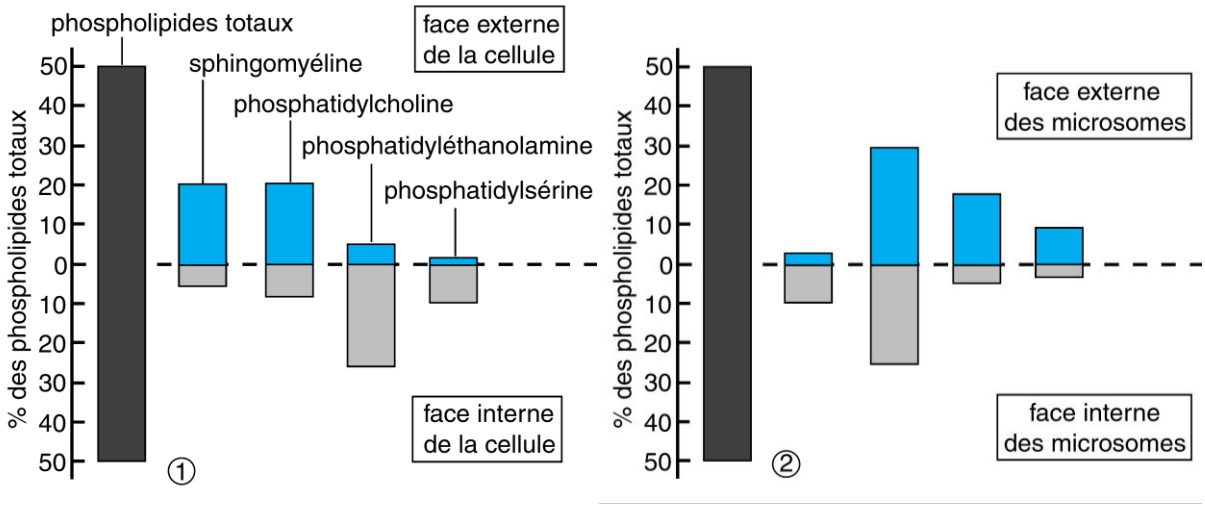


Figure 3 : (1) Distribution des lipides au sein des deux feuillets constituant la membrane plasmique d'érythrocytes et (2) de la membrane de microsomes de foie de rat. La répartition des phospholipides varie selon le type cellulaire, le type de membranes considéré, ainsi qu'entre les deux feuillets. Adapté de [4].

Un exemple de la composition en lipides d'une cellule, pertinent pour introduire ces travaux de thèse, est celui des cellules musculaires (Tableau 2).

Tableau 2 : Composition lipidique des membranes des cellules musculaires. Les proportions sont données en pourcentage des lipides totaux et l'étude a été réalisée sur du muscle strié squelettique de rat. Adapté de [6].

	Sarcolemme	Réticulum endoplasmique	Mitochondrie
Phosphatidylcholine (PC)	46	58	49
Phosphatidyléthanolamine (PE)	22	29	39
Phosphatidylsérine (PS)	18	9	9
Sphingomyéline (SM)	14	3	1

La membrane plasmique des cellules musculaires est appelée sarcolemme. Les phospholipides de type PC y sont représentés à hauteur d'environ 46%, les PE à 22% et ceux de types PS à 18%, ce dernier type de phospholipides étant connu pour être bien plus représenté dans le feuillet interne de la membrane plasmique [7]. Le sarcolemme comporte également de la sphingomyéline et tous les lipides cités ci-avant sont présents en proportion différente selon le

type de membrane des cellules musculaires (membrane plasmique, réticulum endoplasmique et membrane mitochondriale).

1.3.Dynamique de la membrane

Le modèle de « mosaïque fluide » décrivant les membranes biologiques a été proposé en 1972 par Singer et Nicolson [8]. Plus de 40 ans plus tard, ce modèle reste approprié pour comprendre et décrire la structure, la fonction ainsi que la dynamique des membranes biologiques [9]. Cette notion de « mosaïque » provient de l'hétérogénéité de la membrane au niveau de sa composition et le terme « fluide » a été proposé car les molécules peuvent diffuser au sein de la membrane. Ainsi, les lipides sont en mouvement continu dans la membrane (Figure 4).

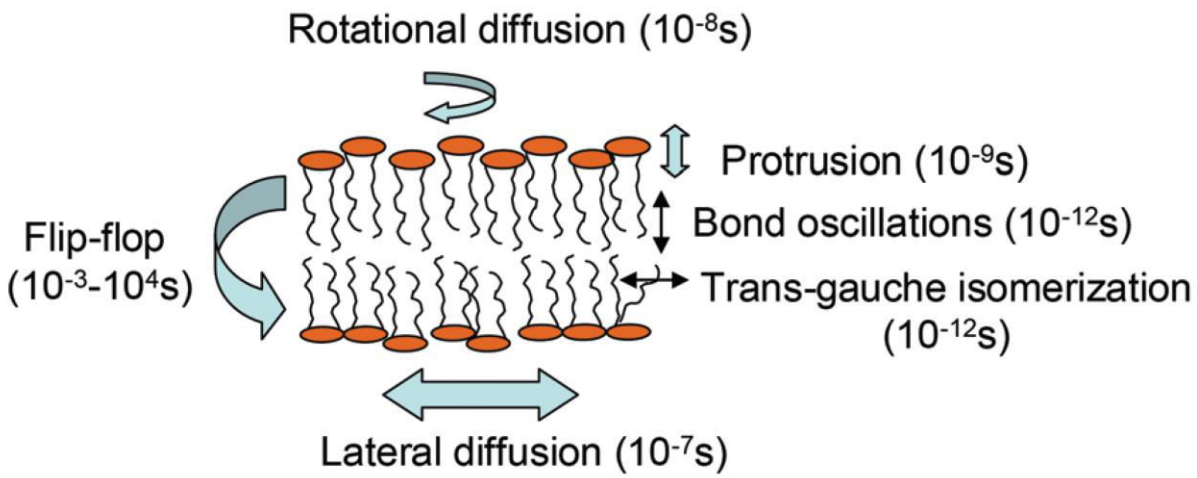


Figure 4 : Ordre de grandeur des temps de corrélation des divers mouvements effectués par les lipides au sein des membranes biologiques [10].

Les lipides peuvent diffuser de manière rotationnelle, translationnelle et également changer de feuillet (Figure 4). Ce dernier phénomène est appelé « flip-flop » et peut être catalysé par des enzymes dénommées flippases ou floppases. L'exposition au solvant de certaines queues hydrophobes (protrusion), tout comme l'oscillation des liaisons C-C, sont également des phénomènes retrouvés au sein de la membrane. Lors de l'augmentation de la température, les liaisons « trans » entre deux carbones des queues hydrophobes peuvent s'isomériser en liaisons de type « gauche » (rotamère) (Figure 4), ce qui conduit à une diminution de l'épaisseur de la bicouche.

Bien que toujours pertinent pour décrire les membranes biologiques, ce modèle de « mosaïque fluide » a été enrichi de quelques nouveautés, selon Félix M. Goñi [11] :

1. Une forte densité de protéines transmembranaires peut être retrouvée dans la membrane.
2. Certaines protéines membranaires périphériques ne sont pas nécessairement localisées systématiquement à la membrane mais peuvent s'y trouver de manière transitoire.
3. La phase fluide des lipides n'est plus considérée comme la seule ayant un intérêt biologique.
4. Certaines parties de la membrane, dans des cas très précis, peuvent ne pas être organisées en bicouches.
5. Les membranes ne sont pas parfaitement planes mais sont courbées en certains endroits du fait de leur composition en lipides et/ou en protéines.
6. Des domaines enrichis d'un ou plusieurs type(s) de lipides et/ou de protéines impliquent une hétérogénéité latérale de la membrane et assurent des fonctions spécifiques.
7. Dans certains cas précis, les lipides peuvent être sujets à un échange rapide de feuillet (« flip-flop »).

2. Les modèles membranaires

2.1. Organisation des lipides

La principale phase adoptée par des lipides membranaires dans une solution aqueuse est la phase lamellaire. La partie hydrophile des lipides est exposée vers le solvant aqueux et les queues hydrophobes se regroupent entre elles. Selon la température, les lipides peuvent adopter principalement deux phases : (*i*) la phase gel-ordonnée (L_β) ou solide (S_0) qui est retrouvée généralement à basse température ; (*ii*) la phase fluide (L_α) ou liquide-désordonnée (L_d), qui est retrouvée à plus haute température (Figure 5A). La température à laquelle s'effectue la transition gel-fluide des lipides est dénommée *Tm*. Ce paramètre varie selon la longueur des chaînes carbonées, le nombre d'insaturations et peut également varier selon la nature de la tête hydrophile (Figure 5A, graphe). Typiquement, plus la chaîne carbonée est longue, plus la température de transition est élevée. *A contrario*, pour des longueurs de chaînes identiques, la présence d'une ou plusieurs insaturation(s) fait baisser *Tm*. Enfin, en ce qui concerne la nature de la tête hydrophile, à longueur de chaînes égale, les phospholipides de types PC ont un *Tm* plus bas que ceux de type PS et PE (Figure 5A, graphe).

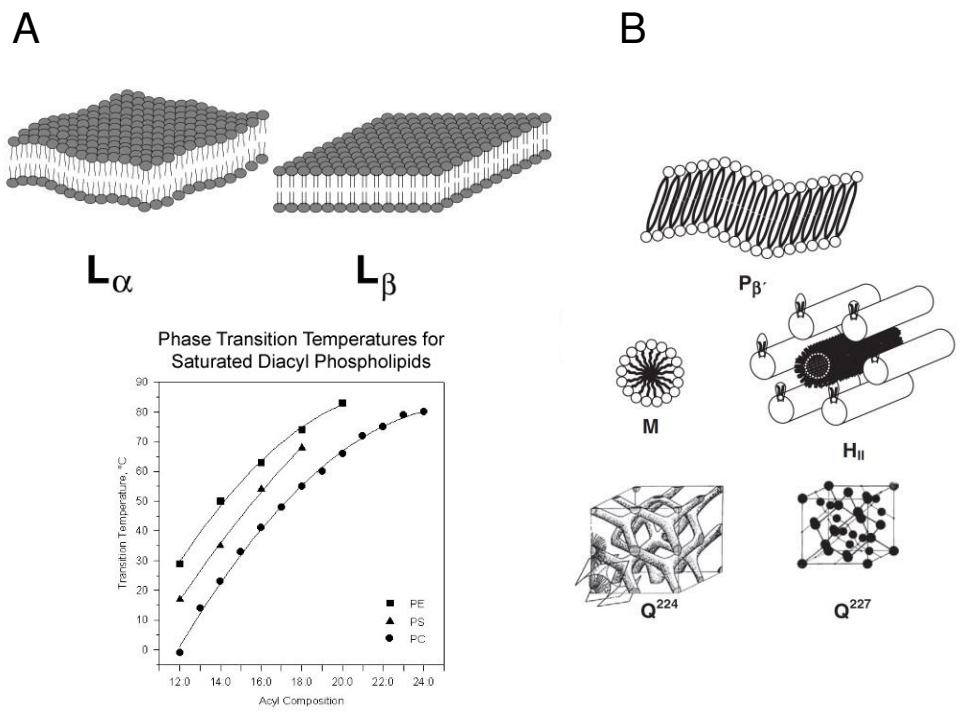


Figure 5 : (A) Représentation de l'arrangement des phospholipides en bicouche en phase fluide (L_α) ou en phase gel (L_β) [12] et température de transition de phase (*Tm*) de certains phospholipides selon la longueur des chaînes carbonées et la nature de la tête hydrophile (graphe [13]). (B) Illustration d'autres types de morphologies que les lipides peuvent adopter. (P_β'), ondulée lamellaire (« ripple phase »). (M), micellaire. (H_{II}), hexagonale inversée. (Q^{224}), phase cubique inversée continue. (Q^{227}), phase cubique inversée discontinue. Adapté de [11].

En plus de la phase lamellaire, les lipides peuvent adopter d'autres morphologies *in vitro*. Des travaux pionniers ayant établi des diagrammes de phase de l'organisation des lipides dans l'eau ont été réalisés avec des lécithines extraites de jaune d'œuf [14]. Ainsi, *in vitro*, les lipides peuvent s'organiser en micelles, micelles inversées, ou encore former une architecture plus élaborée comme la phase cubique (Figure 5B). Le polymorphisme des assemblages de lipides dépend d'un équilibre entre des forces attractives et répulsives (Figure 6) : la tension interfaciale lipide/eau. La force attractive $(F_{L/W})$ (négative) est due à la tension superficielle $(\gamma_{L/W})$ des groupes acyles hydrophobes avec l'eau, ce qui tend à minimiser la surface de l'interface. D'autres interactions comprennent les dipôles des têtes hydrophiles, les liaisons hydrogène et les forces de van der Waals au sein des chaînes hydrophobes de la bicouche. La force répulsive opposée (positive) est due à encombrements stériques, provenant à la fois des têtes (F_{head}) et des queues (F_{chain}). D'une manière globale, la pression répulsive des têtes et des queues contrebalance la tension attractive interfaciale lipide/eau. L'équilibre des forces régit le profil

de pression latérale résultant des lipides et peut affecter leur organisation au sein des membranes.

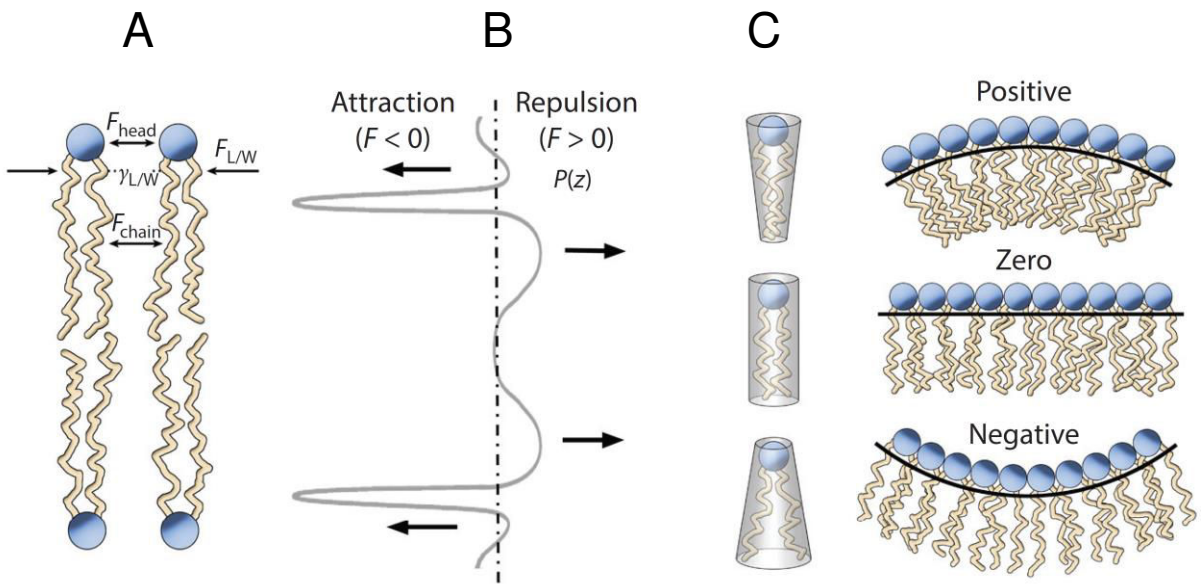


Figure 6 : (A). Forces attractives (pression, $F_{L/W}$) agissant à l'interface lipide/eau en raison de l'effet hydrophobe, en équilibre avec les forces répulsives des têtes polaires (F_{head}) et des queues hydrophobes (F_{chain}). (B) Profil correspondant de la pression latérale en fonction de la profondeur de la bicouche. (C) Selon l'aire moléculaire de la tête hydrophile, le volume moléculaire des queues hydrophobes et la longueur des chaînes carbonées, les phospholipides adoptent une certaine géométrie. La courbure spontanée pour une monocouche lipidique est due à un déséquilibre des forces latérales illustrées en (A). Cette courbure est soit positive, nulle ou négative [15,16].

Dans les années 1970, J. Israelachvili a montré que la morphologie des auto-assemblages de molécules amphiphiles, et donc de lipides, dépend surtout de la géométrie de la molécule, selon l'équation ci-après [17,18] :

$$x = \frac{V_c}{a_0 . l_c} \qquad \text{Eq 1}$$

où V_c est le volume de la partie hydrophobe, a_0 l'aire moléculaire de la tête hydrophile et l_c la longueur des chaînes carbonées. En fonction de la valeur de x, les lipides s'auto-assemblent selon diverses morphologies :

- $x \leq 1/3$, formation de micelles globulaires
- $1/3 \leq x \leq 1/2$, formation de micelles allongées
- $1/2 \leq x \leq 1$, formation de bicouches

- $x \geq 1$, formation de micelles inversées

L'aire moléculaire des têtes hydrophiles des phospholipides dépend de leur nature, zwitterionique (PC, PE) ou chargée (PS, PI, PG). Le volume de la partie hydrophobe dépend, quant à lui, de la longueur des chaînes carbonées. La longueur des chaînes carbonées l_c ainsi que le volume de la partie hydrophobe V_c montrent une dépendance linéaire selon le nombre d'atomes de carbone *(N)* composant la queue hydrophobe. Ces deux paramètres peuvent être estimés par la règle empirique de Tanford [19] selon :

$$l_c\,(en\ \text{Å}) = 1{,}265 * (N - 1) + 1{,}5 \qquad \text{Eq 2}$$

$$V_c\,(en\ \text{Å}^3) = 2 * (27{,}4 + 26{,}9 * N) \qquad \text{Eq 3}$$

Ainsi, au fil des dernières décennies, divers modèles membranaires ont vu le jour avec pour objectif de mieux comprendre l'organisation au sein des biomembranes et également de caractériser les phénomènes s'y déroulant.

2.2.Les liposomes

Les liposomes sont des vésicules composées de phospholipides organisés en bicouche enfermant une cavité aqueuse (Figure 7). Ils sont obtenus facilement car les phospholipides ont tendance à former spontanément des vésicules multilamellaires (MLV) quand ils se trouvent dans un environnement aqueux. En apportant de l'énergie au système, soit par extrusion, soit par sonification, il est possible de former des vésicules unilamellaires de différentes tailles. On distingue les « small unilamellar vesicles » (SUV, diamètre de 20-100 nm) et les « large unilamellar vesicles » (LUV, diamètre > 100 nm) selon le degré de courbure (Figure 7). Ils sont fréquemment utilisés pour la caractérisation des interactions protéine/lipide et beaucoup de travaux de recherche s'intéressent également à leur utilisation dans l'administration de médicaments [20]. Un dernier type de vésicule unilamellaire est celui des « giant unilamelar vesicles » (GUV, diamètre > 300 nm) (Figure 7). Ce dernier modèle membranaire est plutôt dédié à la compréhension de mécanismes à l'échelle cellulaire [21].

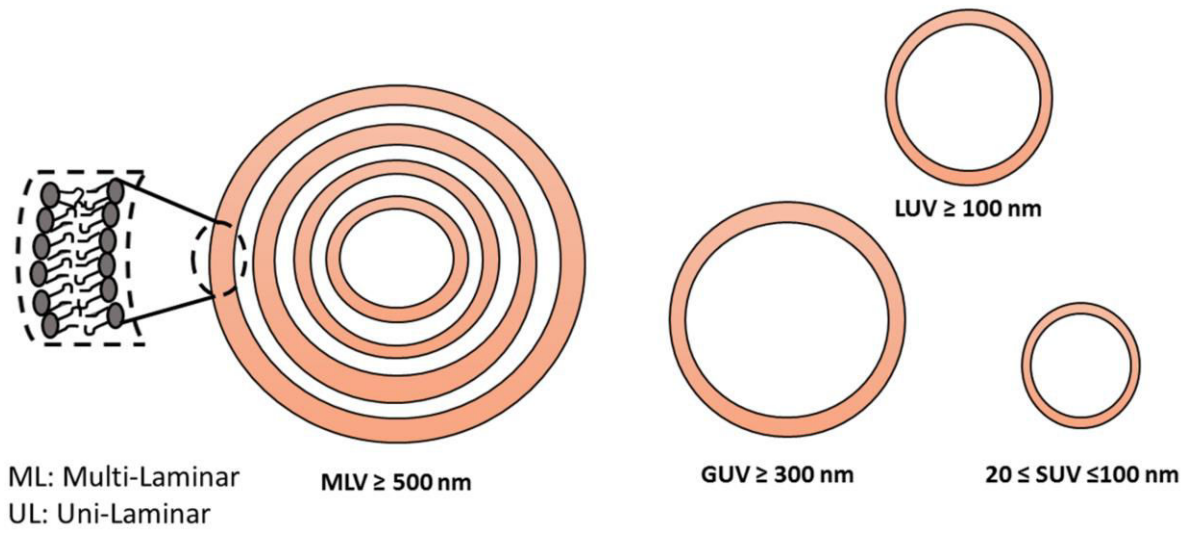


Figure 7 : Représentation schématique des différents types de liposomes selon leur taille [22].

2.3.Les micelles de détergent

Les micelles sont composées de petites molécules amphiphiles, dont les têtes polaires s'orientent vers le solvant aqueux alors que les queues s'organisent pour former un cœur hydrophobe (Figure 8a). Cet arrangement se produit au-dessus de leur concentration micellaire critique (CMC). Les micelles de détergent forment de relativement petits agrégats dont la taille, le nombre d'agrégation et la forme varient selon le type de détergent utilisé [23]. Les micelles sont utilisées pour la caractérisation des interactions protéine/lipide, mais aussi pour la détermination de structure de protéines membranaires solubilisées dans une couronne de détergent, que ce soit par résonance magnétique nucléaire (RMN) ou par cristallographie des rayons X (XRC) [24]. Cependant, ce modèle ne présente pas de bicouche et certains auteurs ont proposé que la perte d'activité observée pour certaines protéines membranaires pourrait être inhérente au manque de dynamique interne dans les micelles de détergent par rapport aux modèles membranaires possédant une bicouche [25].

2.4.Les amphipols

Les amphipols [26,27] sont des polymères amphipathiques donc par définition possédant une face hydrophile et une face hydrophobe. Ils sont utilisés pour stabiliser les protéines membranaires intégrales à la place des détergents (Figure 8b). Le complexe protéine/amphipol se veut plus stable que celui obtenu avec des détergents puisque ces polymères ont été conçus afin qu'ils aient une très forte affinité avec les protéines membranaires. Ainsi, la concentration

critique est extrêmement faible et permet de travailler avec de faibles concentrations de protéines membranaires et sans micelles de détergent libres en solution [28].

2.5.Les bicelles

Le modèle bicelle (Figure 8c) est décrit de façon plus détaillée que les autres modèles membranaires (voir partie **3.** ci-après) car nous l'avons essentiellement utilisé pour répondre à la problématique de cette thèse.

2.6.Les nanodisques

Les nanodisques sont des modèles membranaires composés d'une bicouche lipidique entourée par des protéines dites « MSP » (« membrane scaffold protein ») [29] ou par des copolymères d'acide styrènique et maléique (SMA) [30] (Figure 8, d et e). Dans le cas des nanodisques composés de MSP, l'ajustement de leur taille nécessite la purification de protéines plus ou moins grandes et l'utilisation d'un nombre de lipides précis pour former des disques stables [31,32].

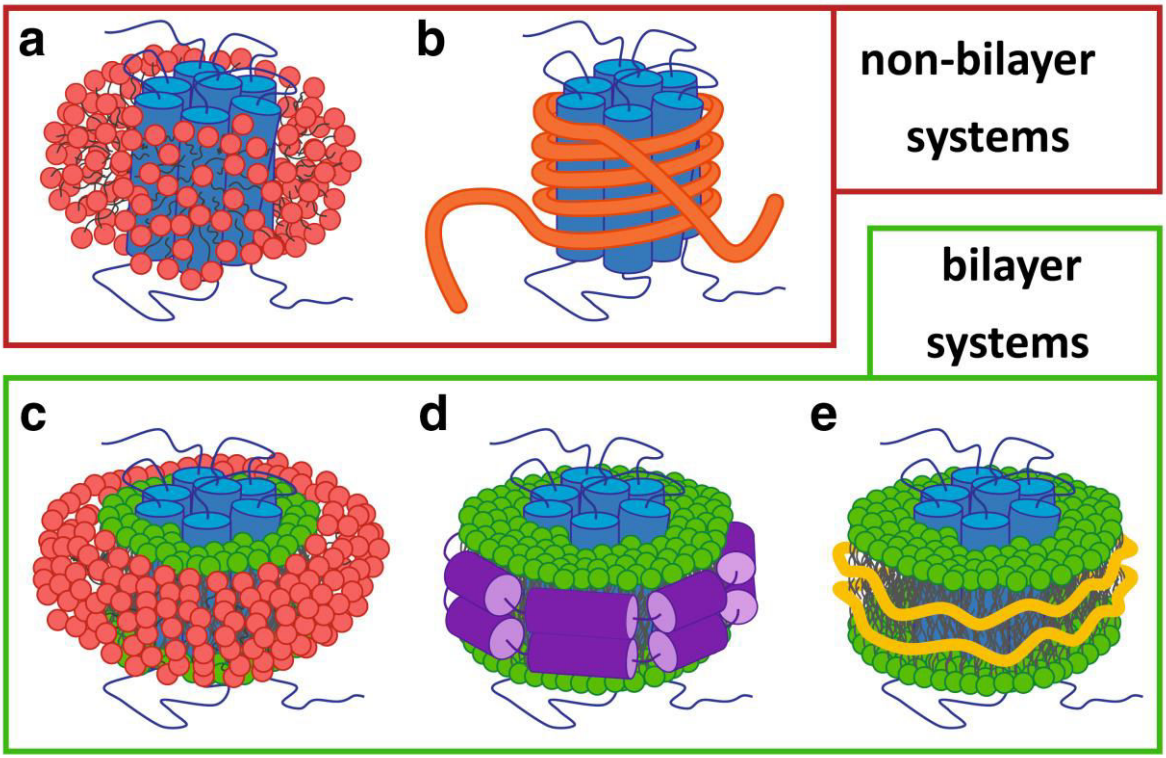


Figure 8 : Illustration d'une protéine membranaire intégrale incorporée dans (a) une couronne de détergent, (b) un amphipol, (c) une bicelle ou (d, e) dans un nanodisque composé de MSP (d) ou de SMA (e). Les modèles a et b ne possèdent pas de bicouche lipidiques contrairement aux modèles c, d et e [33].

2.7.Les monocouches de Langmuir

Ce type de modèle membranaire nécessite l'utilisation d'une cuve de Langmuir (Figure 9A). Ce dispositif est composé d'une cuve et de deux barrières en Teflon®, remplie d'une solution aqueuse. Une monocouche de molécules amphiphiles est formée à l'interface air/eau par simple dépôt. Dans le cas des phospholipides, les têtes hydrophiles sont orientées vers la solution aqueuse et les queues hydrophobes vers l'air. La tension de surface peut être mesurée avec une plaque de Wilhelmy et ajustée par compression de la monocouche grâce aux barrières. Dans ce dispositif, les phospholipides peuvent adopter différentes phases en fonction de la tension de surface [34], qui est ajustée par l'augmentation de la compression réalisée grâce aux barrières. La phase gazeuse (G) correspond à un état où peu de molécules sont présentes par unité de surface. La phase liquide expansée (LE) est l'équivalent de la phase fluide en phase lamellaire. Les têtes hydrophiles y sont proches mais les queues hydrophobes restent mobiles. La phase liquide condensée (LC), quant à elle, est le pendant de la phase gel en phase lamellaire. Finalement, la phase solide (S), qui n'est pas accessible par les phospholipides, mais peut l'être par certains acides gras, se traduit par l'alignement des chaînes carbonées. En cas de compression supplémentaire de l'état solide, la monocouche collapse et forme des empilements et des agrégats afin de minimiser l'énergie du système.

2.8.Les bicouches supportées

Ce modèle membranaire, comme son nom l'indique, consiste en une bicouche de phospholipides absorbée sur un bloc (« wafer ») de silicium (Figure 9B) ou sur du mica. La bicouche peut être obtenue simplement par fusion de liposomes [35] ou plus exactement par fusion de SUV ayant un rayon de courbure suffisamment prononcé pour pouvoir fusionner spontanément au contact du support.

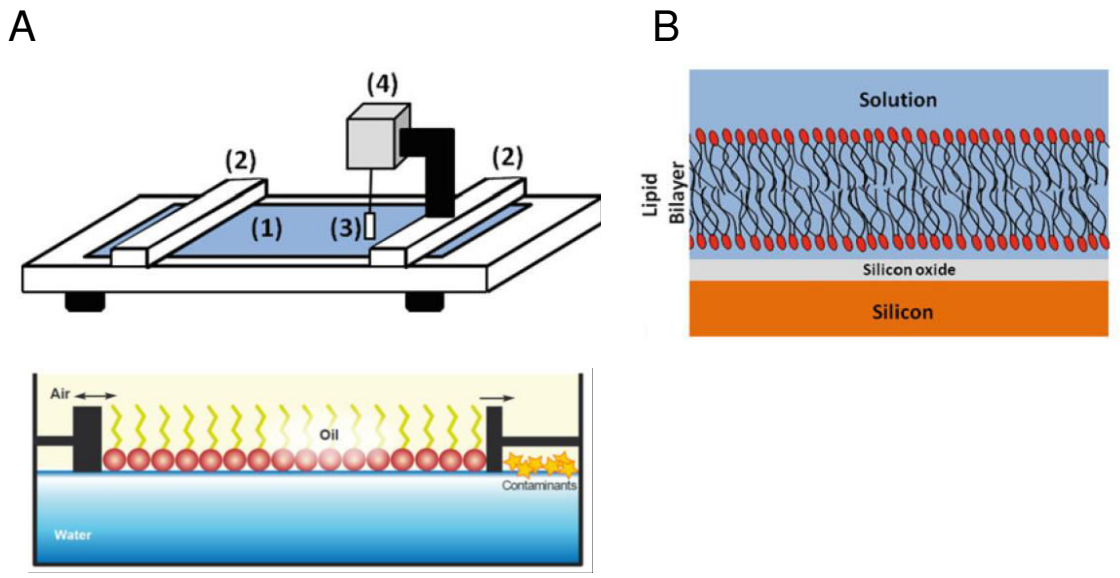


Figure 9 : (A) Illustration d'une cuve de Langmuir remplie d'une solution aqueuse (1). Une barrière mobile (2) permet la compression de la monocouche à la surface. Une plaque de Wilhelmy (3) attachée à un capteur de tension de surface (4) permet de suivre la compression et la stabilité de la monocouche, ainsi que la variation de la pression de surface pendant l'étude des interactions protéine/lipide [36]. Les molécules de phospholipides se propagent spontanément sur l'interface air/eau jusqu'à former une monocouche [37]. Les queues hydrophobes s'orientent vers l'air, alors que les têtes polaires s'orientent vers la partie aqueuse. Ce dispositif permet de mimer l'un des feuillets d'une bicouche lipidique. (B) Représentation schématique d'une bicouche de phospholipides absorbée sur un « wafer » de silicium. Adapté de [36,37].

3. Le modèle membranaire « bicelle »

3.1. Généralités

Les bicelles sont des auto-assemblages de lipides la plupart du temps constituées d'un mélange de phospholipides à chaînes longues et de phospholipides à chaînes courtes ou de détergents. Ces mélanges lipidiques peuvent adopter différentes morphologies selon plusieurs paramètres (Figure 10) [38–40] dont le ratio molaire des constituants q et la température.

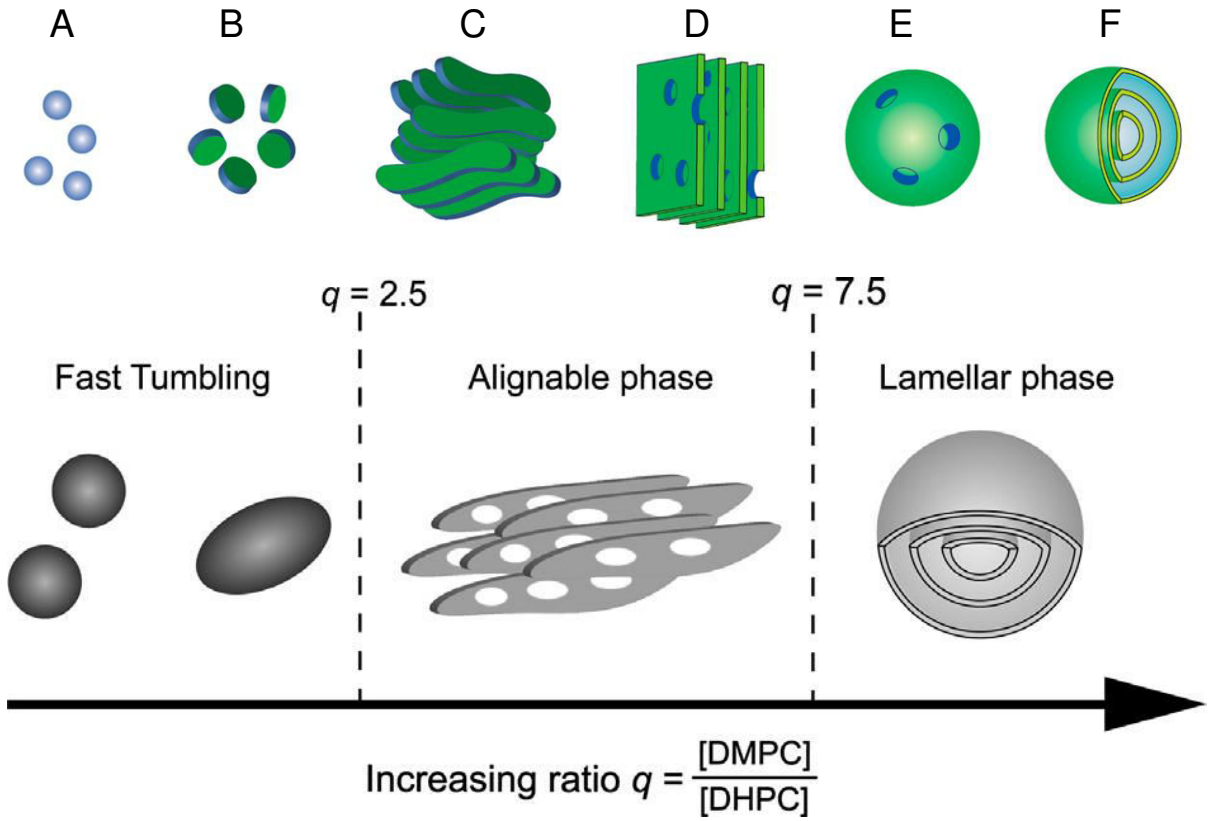


Figure 10 : (Haut) Représentations schématiques du polymorphisme des bicelles selon le ratio molaire de leurs constituants. Micelles mixtes (A), disques (B), rubans vermiculaires (C), lamelles perforées (D), vésicules multilamellaires avec ou sans pores toroïdaux (E et F). (Bas) (Bas) Exemples de morphologies adoptées par des bicelles, ainsi que leur capacité à s'aligner dans un champ magnétique, selon q. Adapté de [41,42].

La morphologie la plus courante admise pour un système bicellaire est celle du disque (Figure 10E), où les phospholipides à chaînes longues forment une bicouche entourée par des phospholipides à chaînes courtes ou par des détergents formant un tore. Le terme « bicelle » (qui vient de « bilayered micelle ») a été proposé en 1995 par Sanders et Landis [43]. Cependant, des études pionnières décrivant des objets discoïdaux ont été effectuées bien avant sur des

mélanges de lécithines et de sels biliaires ou des analogues [44–46] (Figure 11) et sur des mélanges de phospholipides [47,48].

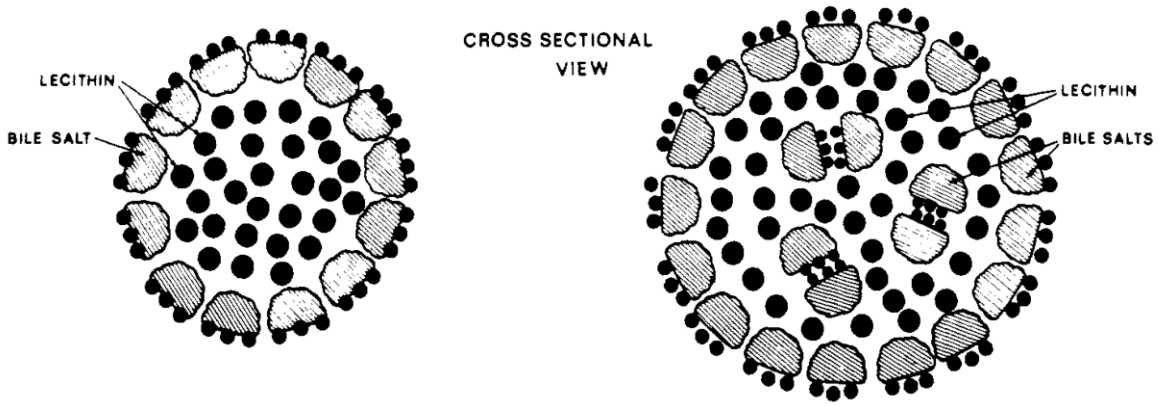


Figure 11 : Modèles schématiques, montrés en coupe transversale, de la structure adoptée par un mélange de lécithines et de sels biliaires [45].

Un des paramètres les plus importants régissant la taille des bicelles est le ratio molaire de leur constituants, c'est-à-dire le ratio lipides à chaînes longues/lipides à chaînes courtes ou détergents. Ce ratio est noté q par convention et peut être facilement ajusté en fonction de la taille finale des bicelles souhaitée pour une application donnée. Ce terme q ne doit pas être confondu avec Q, le vecteur de diffusion (voir ci-après partie II). Plus q a une valeur élevée, plus le diamètre des bicelles est grand. La valeur de ce ratio molaire permet, par convention, de distinguer deux types de bicelles :

- Les plus grandes ($q \geq 2{,}5$), lorsqu'elles sont composées de phospholipides, peuvent s'aligner dans un champ magnétique dans certaines conditions, dont une température supérieure à *Tm*. Les bicelles s'orientent naturellement de façon perpendiculaire au champ magnétique [49], alors que l'ajout d'ions lanthanides permet un alignement parallèle [50]. Le développement de bicelles composées de phospholipides portant un groupement biphényle permet de maintenir une orientation parallèle en se passant d'ions lanthanides [51,52] (Figure 12). Ces bicelles sont couramment utilisées en RMN du solide [41,42] et seraient organisées en lamelles perforées plutôt qu'en disques dans ces conditions [53]. Elles restent néanmoins des assemblages de phospholipides. Outre la RMN du solide, ce type de bicelles ($q \geq 2{,}5$) est de plus en plus utilisés en XRC [54–56] pour déterminer la structure des protéines membranaires mais, dans ce cas, la DHPC est la plupart du temps remplacé par du CHAPSO.

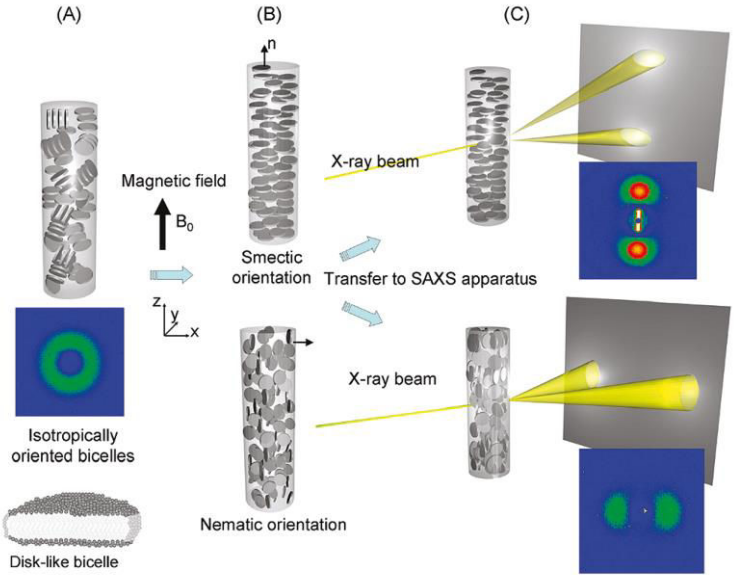


Figure 12 : Schémas et images du détecteur de diffusion aux petits angles (SAXS) montrant l'orientation des bicelles DMPC et TBBPC (1-myristoyl-2-[4-(4-biphenyl)butanoyl]-sn-glycero-3-phosphocholine), avant orientation (A) et après orientation (B) dans un champ magnétique, puis après leur transfert dans le dispositif SAXS (C) [52].

- Les bicelles qui ne s'alignent pas dans un champ magnétique sont dites isotropes (q < 2,5) et sont couramment utilisées en RMN du liquide [41,42] pour l'étude de l'objet bicelle en soi et également pour la caractérisation des interactions protéine/lipide et dans le cadre de la détermination de la structure de protéines membranaires. La morphologie discoïdale admise [57,58] des bicelles pour des valeurs de q < 0,5 reste largement débattue. Les bicelles seraient vraisemblablement davantage des micelles mixtes lorsque q est inférieur à 1 [59].

D'une façon générale, les bicelles ont été caractérisées par microscopie électronique [39,53,57], diffusion dynamique de la lumière (DLS) [39,57] et modélisation moléculaire [60–62]. Cependant la RMN reste la technique la plus utilisée lors des études dédiées aux bicelles [42]. Seulement quelques études ont été réalisées par SAXS [51,52,63,64]. Dans la littérature, à notre connaissance, la seule étude SAXS effectuée sur des bicelles prenant en compte les valeurs à petits Q et pas uniquement celle des valeurs intermédiaires a été réalisée par Yang *et al.* [63] sur des bicelles DPPC/DH_7PC (Figure 13). Quelques études ont été effectuées par diffusion des neutrons aux petits angles (SANS) [40,65–69] dont celle de Luchette *et al.* [65] qui est, à notre connaissance, la seule consacrée à la caractérisation par SANS des bicelles isotropes DMPC/DHPC hydrogénées.

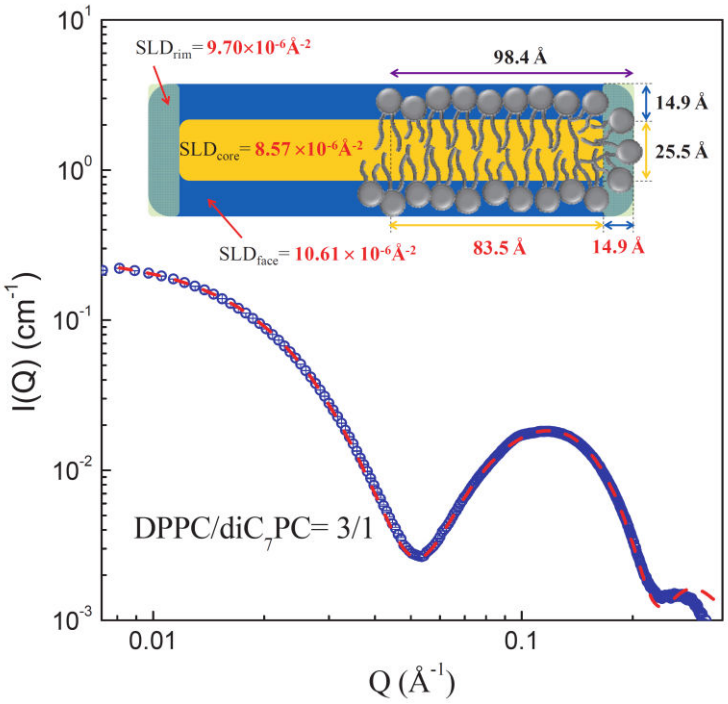


Figure 13 : Courbe SAXS d'un mélange bicellaire DPPC/DH$_7$PC pour q = 3 à 25° C, ajustée avec le modèle « core-shell bicelle » [63].

Pour mieux mimer la bicouche membranaire, des phospholipides insaturés, du cholestérol [70], des phospholipides anioniques [71], ou de la sphingomyéline [72] ont été incorporés avec succès dans la partie bicouche. Des bicelles ont même été formées à partir de lipides d'*E. coli* [73]. Récemment, Beaugrand et *al.* [74] proposent d'utiliser des bicelles composées d'un mélange phospholipides/monoalkylphosphocholines (MAPCHO), les MAPCHO étant déjà utilisés pour purifier certaines protéines membranaires. Malgré la variété de compositions retrouvée dans la littérature, le système actuellement le mieux décrit est constitué de dimyristoylphosphatidylcholine/dihexanoylphosphatidylcholine (DMPC 14:0 PC et DHPC 6:0 PC) depuis sa première caractérisation par RMN du solide [49]. Le mélange lipidique DMPC/CHAPSO a également fait l'objet de plusieurs études [58,68]. Une des plus notables est celle réalisée par Li *et al.* [68] dans laquelle un diagramme de phase du mélange DMPC/CHAPSO (q = 3) ainsi qu'une mise à jour de celui du mélange DMPC/DHPC ont été rapportés (Figure 14).

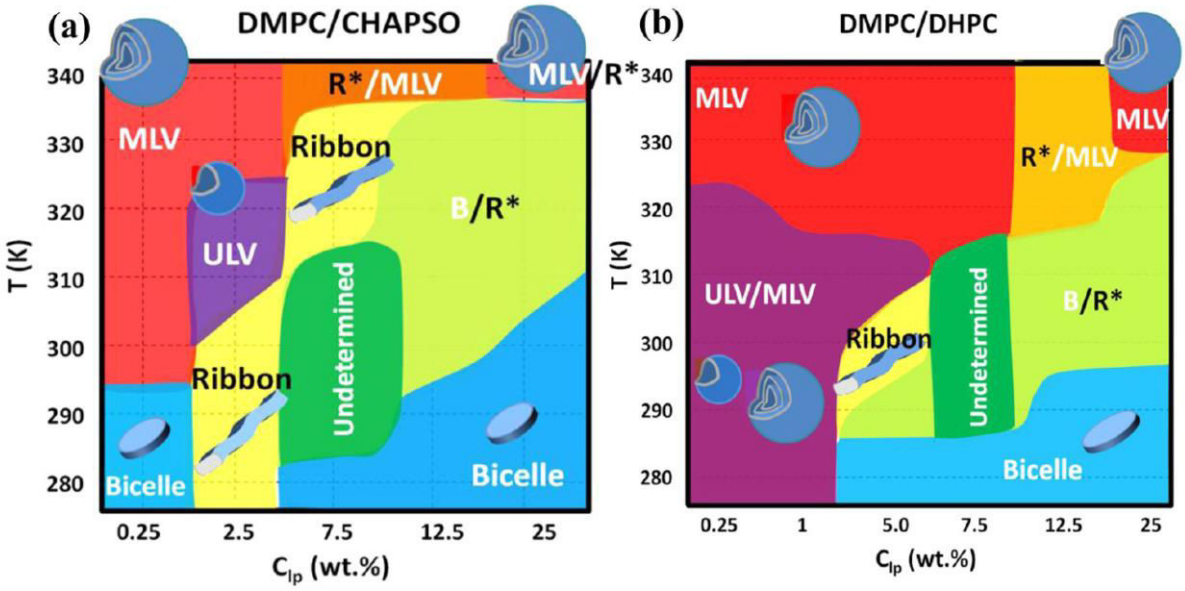


Figure 14. Morphologies adoptées spontanément par les mélanges lipidiques en fonction de la température et de la concentration en lipides, pour un ratio molaire q = 3. (a) DMPC/CHAPSO, (b) DMPC/DHPC. R* représente un éventuel mélange de rubans et de lamelles perforées. Les régions avec une coexistence possible de différentes morphologies sont représentées en utilisant des couleurs combinées (par exemple, l'orange représente un mélange de R* (jaune) et de MLV (rouge), le vert clair représente un mélange de bicelles (bleu) et de R* (jaune), et le violet foncé représente un mélange de ULV (violet) et MLV (rouge)) [68].

Malgré leur grande polyvalence, la morphologie et la taille des bicelles sont très sensibles au ratio q, à la température ainsi qu'à la dilution [39,75]. Ainsi, plusieurs précautions sont à prendre lors de leur utilisation. Dans ces travaux de thèse, nous avons utilisé uniquement des bicelles isotropes composées de DMPC, DMPS et DHPC, principalement pour $q \approx 1$. Uniquement les points importants inhérents à l'utilisation de ce type de bicelle précisément sont abordés par la suite.

3.2. L'importance de la concentration bicellaire critique

Le rapport molaire des deux phospholipides, noté q, est le paramètre principal qui régit la taille des bicelles. Pour un système bicellaire DMPC/DHPC, on a :

$$q = \frac{[\mathrm{DMPC}]}{[\mathrm{DHPC}]} \qquad \text{Eq 4}$$

Un point à prendre en compte dans les systèmes bicellaires est qu'une proportion de DHPC est en équilibre rapide entre la bicelle et la forme libre en solution. Par analogie avec la CMC, Beaugrand *et al.* [59] ont introduit la notion de concentration bicellaire critique (CBC). Cette

notion de $[DHPC]_{libre}$ avait déjà était considérée, par d'autres auteurs [39,75]. Ainsi, le ratio molaire effectif est désigné q_{eff} lorsque la proportion de DHPC libre est prise en considération et est défini par :

$$q_{eff} = \frac{[\mathrm{DMPC}]}{[\mathrm{DHPC}]_{total} - [\mathrm{DHPC}]_{libre}} \qquad \text{Eq 5}$$

Beaugrand *et al.* [59] ont obtenu une valeur de $[DHPC]_{libre}$ de 6 mM, que nous utilisons dans ces travaux de thèse (Figure 15). La CBC a une valeur bien inférieure à la CMC des micelles de DHPC (14-15 mM) [76] . Une explication simple est que les molécules de DHPC sont plus stables dans une morphologie bicellaire. Ceci pourrait expliquer la décroissance de la CBC de $q = 0$ à $q = 0{,}5$ avant d'atteindre une constante jusqu'à $q = 2$ (Figure 15).

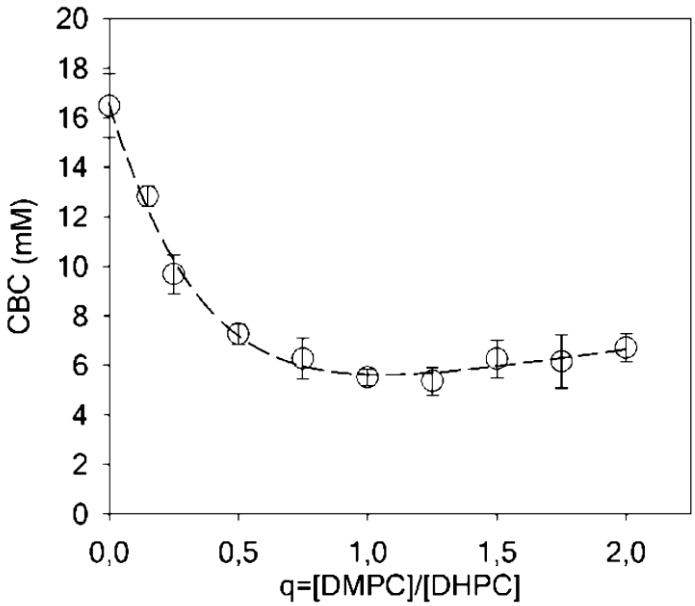


Figure 15. Evolution de la concentration bicellaire critique (CBC) en fonction de q. Notons que pour $q = 0$, CBC = CMC (DHPC seule en solution) [59].

En cas de très fortes concentrations en lipides, la CBC peut être considérée comme négligeable car $[DHPC]_{libre}$ ne représente qu'une faible proportion de la concentration totale en DHPC. *A contrario*, dans des régimes plus dilués, ce paramètre doit absolument être pris en compte afin de ne pas induire une forte augmentation de q_{eff} pouvant générer des biais d'interprétation. Ainsi, le polymorphisme des bicelles présenté dans les diagrammes de phase précédents semblerait plutôt être dû à une variation de q_{eff} et non pas à la concentration de lipides en soi. D'ailleurs, Beaugrand *et al.* [59] proposent des limites de concentrations dans

lesquelles le ratio molaire q_{eff} peut être considéré comme constant. Plus récemment, Mineev *et al.* [58] proposent même de déterminer d'abord la CBC puis de diluer les bicelles dans du solvant contenant cette même concentration afin de ne pas déplacer l'équilibre et ainsi maintenir q_{eff} constant. Cette considération a été implicitement suggérée par Draney *et al.* [77] lors de la caractérisation de bicelles DMPC/DPC (dodécylphosphocholine) par chromatographie d'exclusion stérique basse pression. En effet, les colonnes ont été préalablement équilibrées avec une concentration en DPC correspondant à sa CMC (1.5 mM). Notons également que la proportion de DMPC libre est de l'ordre de 6 nM [78], ce qui est cependant négligeable dans la plupart des conditions expérimentales.

3.3.Le modèle idéal de bicelle

Dans le cas où les bicelles adoptent une morphologie discoïdale, plusieurs modèles de bicelles ont été proposés dans la littérature pour corréler q_{eff} au rayon R de la bicelle. Ici R concerne uniquement la bicouche des modèles. Ces modèles sont basés, soit sur l'aire de la tête hydrophile [75,79], soit sur le volume des phospholipides [80] (Figure 16).

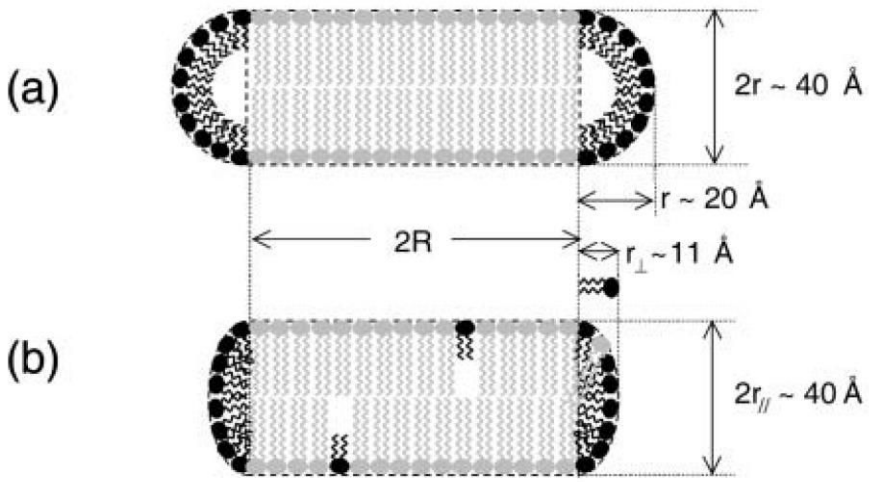


Figure 16 : Représentation schématique d'une bicelle avec des bords circulaires (a) ou ellipsoïdaux (b) [80].

- (Figure 16a) modèle de Vold et Prosser [79] basé sur l'aire des têtes hydrophiles avec l'hypothèse que les volumes moléculaires sont identiques et en utilisant des bords circulaires.

$$R = \frac{1}{2} rq \left[\pi + \left(\pi^2 + \frac{8}{q} \right)^{1/2} \right] \qquad \text{Eq 6}$$

où *r* correspond à l'épaisseur de la partie toroïdale.

- (Figure 16a) Glover *et al.* [75] : modèle amélioré de Vold et Prosser, basé sur l'aire des têtes hydrophiles, en considérant des volumes moléculaires différents pour la bicouche et le tore et en utilisant des bords circulaires.

$$R = \frac{1}{2} krq \left[\pi + \left(\pi^2 + \frac{8k}{q} \right)^{1/2} \right] \qquad \text{Eq 7}$$

où *k* est le ratio des aires moléculaires des deux constituants (DHPC et DMPC).

- (Figure 16b) Triba *et al.* [80] : modèle basé sur les volumes moléculaires des phospholipides et en utilisant des bords ellipsoïdaux.

$$R = \frac{1}{4} r_{\perp} q_v \left[\pi + \sqrt{\pi^2 + \frac{32}{3q_v}} \right] \qquad \text{Eq 8}$$

où q_v est le ratio des volumes moléculaires des deux constituants (DHPC et DMPC) multiplié par *q* dans le cas d'une ségrégation parfaite.

- (Figure 16b) Mineev *et al.* [58] à partir de Triba *et al.* [80] :

$$R_t = r_{\perp} + \frac{r_{\perp} q_{eff}}{4\Lambda} \left[\pi + \left(\pi^2 + \frac{32\Lambda}{3q_{eff}} \right)^{1/2} \right] \qquad \text{Eq 9}$$

où Λ est le ratio des volumes moléculaires des deux constituants. Ici, R_t fait référence au rayon total de la bicelle et non plus uniquement à celui de la bicouche lipidique.

Notons que l'équation proposée par Triba *et al.* [80] provient de la combinaison des volumes du tore et de la bicouche.

Le volume du tore est déterminé par :

$$V_{tore} = \pi\, r_{\perp} r_{/\!/} \left(\pi R + \frac{4}{3} r_{\perp}\right) \qquad \text{Eq 10}$$

Le volume de la partie en bicouche est déterminé par :

$$V_{bilayer} = 2\pi r_{/\!/} R^2 \qquad \text{Eq 11}$$

Lors de ces travaux de thèse, nous avons utilisé le dernier modèle (Eq. 9) qui est basé sur les volumes moléculaires des phospholipides. A notre avis, et selon celui d'autres auteurs [81,82], ce modèle est plus fiable que celui utilisant les aires moléculaires des lipides. En effet, les volumes moléculaire sont très similaires en-dessous et au-dessus de la température de transition gel-fluide [83,84] contrairement à l'aire moléculaire occupée par la tête hydrophile [84]. De plus, l'équation ne dépend pas de l'épaisseur de la bicouche lipidique. Nous faisons également l'approximation d'une ségrégation parfaite entre les phospholipides à chaînes longues et à chaînes courtes. Grâce à ce modèle, il possible d'estimer la taille, le nombre d'agrégation et donc la MW des bicelles (Figure 17).

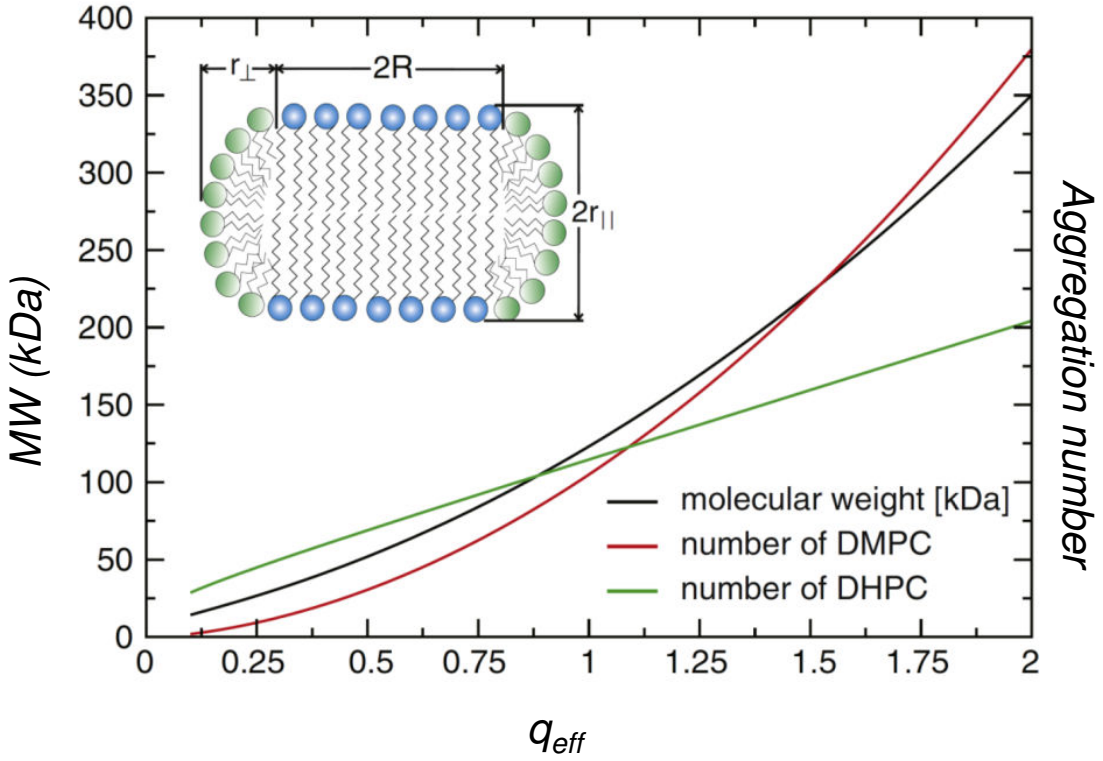


Figure 17 : Nombre d'agrégation, MW et taille des bicelles en fonction de q_{eff}. Adapté de [81].

Chapitre 2: Mode d'interaction et structure des protéines membranaires

1. Les interactions protéine/lipide

En-dehors des lipides, les biomembranes sont également composées de protéines faisant partie intégrante de la structure de la membrane. Elles assurent une multitude de fonctions au sein de la cellule (récepteur, transporteur, canaux, enzyme, …). Les protéines membranaires sont divisées en deux catégories, les protéines dites « intégrales » ou « intrinsèques » et celles dites « périphériques » ou « extrinsèques » (Figure 17).

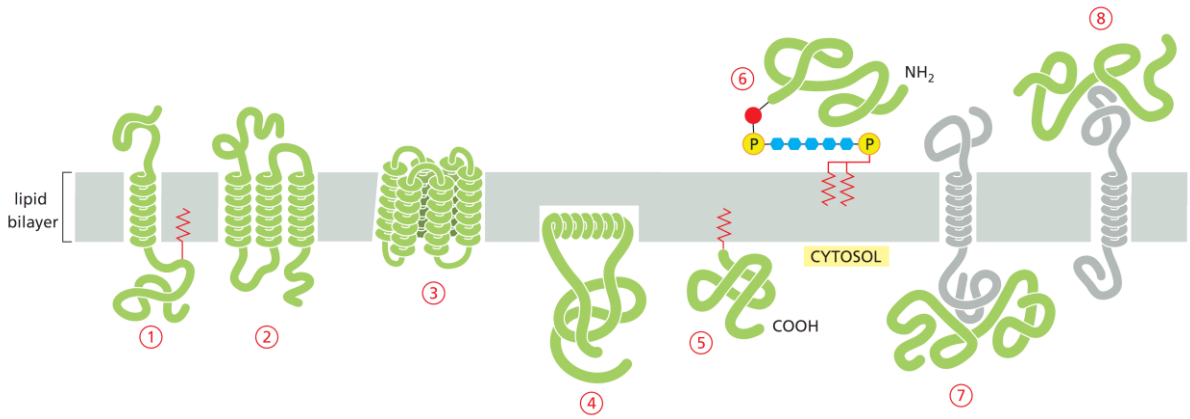


Figure 17 : Représentation schématique des différentes classes de protéines membranaires (en vert) selon leur mode d'interaction avec la membrane. (1 à 3) sont des protéines intrinsèques ou intégrales et (4 à 8) sont des protéines extrinsèques ou périphériques [3].

1.1. Les protéines membranaires intégrales

Les protéines membranaires intégrales (IMP) possèdent des domaines transmembranaires enchevêtrés dans la partie hydrophobe de la bicouche lipidiques (Figure 17, 1-3). Ces parties hydrophobes peuvent être structurées en hélices α à un ou plusieurs domaines transmembranaires, en tonneaux β ou une mélange des deux comme dans le cas des tonneaux $(\beta/\alpha)_8$. Une notion importante des interactions IMP/lipide est celle de la correspondance hydrophobe ou « hydrophobic matching » (Figure 18). Cela implique que la surface de la partie hydrophobe d'une protéine intégrale concorde avec l'épaisseur de la partie hydrophobe de la bicouche lipidique dans laquelle elle se trouve. En effet, il a été montré que certaines protéines voient leur activité diminuer ou augmenter en fonction de la longueur des chaînes carbonées

des phospholipides. Un exemple est celui de la pompe à calcium Ca^{2+}-ATPase [85,86]. Dans le cas de phospholipides de type PC, cette protéine possède une activité dépendante de la longueur des chaînes carbonées avec une distribution quasi-gaussienne centrée sur des chaînes hydrophobes longues de 18 atomes de carbone. La même observation a été faite pour diacylglycérol kinase d'*E. coli* [85,87].

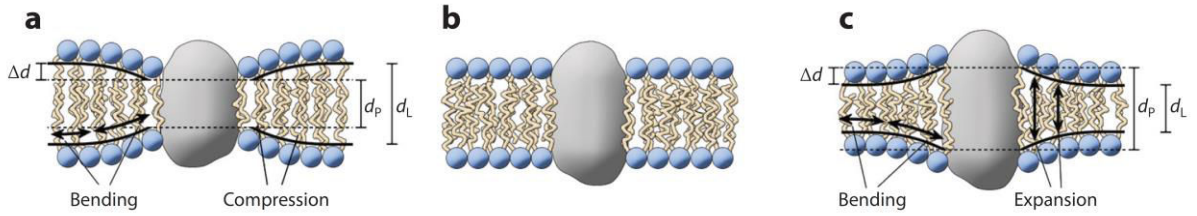


Figure 18 : Les protéines membranaires intégrales sont incorporées dans la bicouche lipidique et sont en interaction avec les chaînes hydrocarbonées des queues hydrophobes des phospholipides. Selon la correspondance des parties hydrophobes, l'épaisseur de l'interface protéine/lipide (d_P) est (a) inférieure, (b) égale à ou (c) supérieure à l'épaisseur moyenne de la bicouche moyenne (d_L). La déformation de la bicouche implique la courbure, la compression et/ou l'étirement des chaînes hydrocarbonées [15,16].

La correspondance hydrophobe apparaît comme le paramètre le plus important pour qu'une protéine intégrale puisse conserver son activité, puisque la majorité de la protéine est au contact de la partie hydrophobe des lipides. Cependant, des études montrent qu'une différence d'activité catalytique existe chez certaines protéines selon la nature des têtes polaires des lipides [88,89] et attestent donc de leur importance.

1.2.Les protéines membranaires périphériques

Les protéines périphériques (PMP), quant à elles, possèdent à la fois une large partie soluble dans le cytosol et peuvent interagir avec les lipides membranaires (Figures 17, 4-8 et 22). Ces interactions peuvent se faire *via* une hélice α amphipathique (Figure 17, 4), une boucle hydrophobe ou encore par interaction électrostatique directement *via* les têtes polaires des lipides ou par l'intermédiaire de ponts calciques. Un autre type de protéines périphériques inclut celles qui sont ancrées dans la bicouche par l'intermédiaire d'une liaison covalente avec les lipides membranaires (myristoylation, palmitoylation, …) (Figures 17, 5) ou *via* une ancre GPI (glycosylphosphatidylinositol) également appelée glypiation (Figures 17, 6). Un dernier type de protéines membranaires périphériques concerne celles localisées à la membrane *via* des interactions avec des protéines membranaires intégrales (Figure 17, 7 et 8).

- ***Ancrage à la membrane***

Certaines protéines membranaires périphériques possèdent une modification post-traductionnelle de type myristoylation, palmitoylation ou encore glypiation dans le cas d'une ancre GPI. Ces modifications post-traductionnelles leur permettent de littéralement s'ancrer à la membrane *via* l'insertion de la partie apolaire dans la bicouche lipidique. Les palmitoylations se font la plupart du temps sur un résidu cystéine de la protéine *via* la formation d'une liaison thioester avec l'acide gras. Les exemples les plus connus de ce type de modification post-traductionnelle sont l'hémagglutinine et la pompe à proton M2, deux protéines du virus de la grippe [90]. Dans le cas d'une glypiation, le groupement PE d'un phospholipide forme une liaison covalente avec la fonction C-terminale de la protéine d'intérêt [91].

- ***Les têtes polaires des lipides***

Un exemple de protéine interagissant avec les têtes polaires des lipides est celui de l'annexine V [92]. Ce mode d'interaction protéine/lipide est particulier car il s'effectue par l'intermédiaire de ponts calciques avec les têtes PS des lipides. Cette interaction spécifique avec les têtes PS est exploitée en laboratoire pour détecter des cellules en apoptose [93].

- ***La courbure membranaire***

Les domaines BAR doivent leur nom à trois protéines dans lesquels ces motifs ont été retrouvés (« Bin, Amphiphysin and Rvs »). Ce sont des structures en forme d'arc d'une longueur d'environ 20 nm formées par la dimérisation d'un faisceau d'hélices α. Les domaines BAR dits classiques interagissent avec des parties de la membrane ayant une forte courbure positive, alors que les domaines I-BAR reconnaissent des zones présentant une courbure négative et les domaines F-BAR des zones faiblement courbées [94] (Figure 19A). Contrairement aux domaines BAR, les motifs ALPS (« amphipathic lipid packing sensor ») ne semblent pas reconnaître la courbure membranaire mais plutôt le défaut de « packing » des lipides qui en découle (Figure 19B). Ces motifs ont la particularité d'être non structurés seuls en solution et de se structurer en hélice α lors de leur interaction avec la membrane. Ils sont composés de 20 à 40 acides aminés et ont initialement été découverts au sein de la protéine ArfGAP1 (« ADP-ribosylation factor GTPase-activating protein 1 ») [95]. Il a été démontré que cette protéine possède une meilleure affinité pour des liposomes composés de lipides cylindriques dans lesquels des lipides coniques ont été insérés [95]. La même tendance a été observée avec des liposomes contenant des lipides avec des mono-insaturations. Dans ces deux cas, les liposomes présentent un défaut de

« packing ». L'interaction de ce motif ALPS avec les lipides membranaires se fait *via* l'insertion de résidus hydrophobes dans les zones présentant un défaut de « packing » [96].

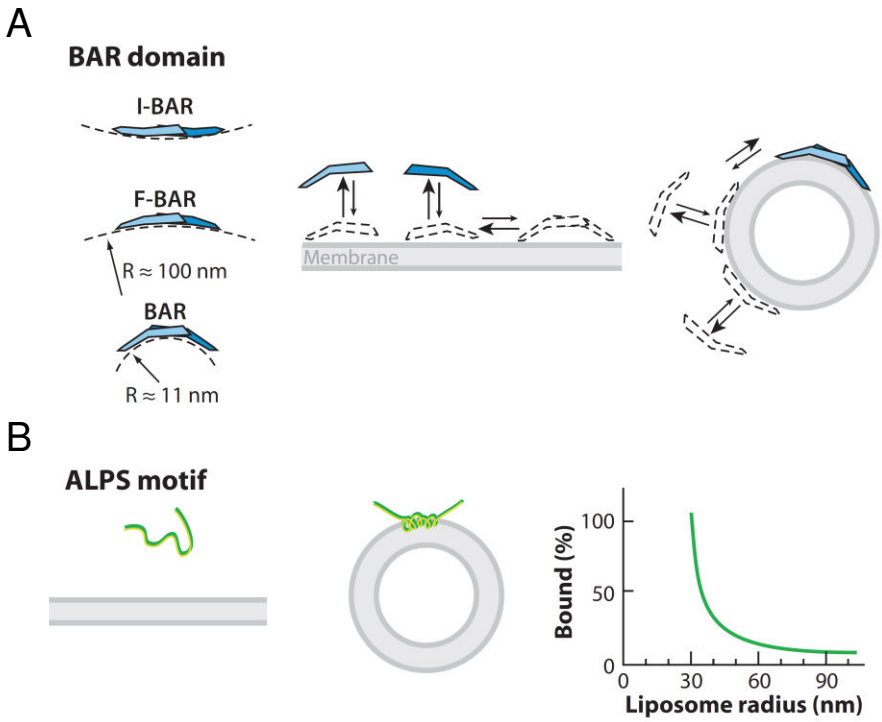


Figure 19. Mécanismes de détection de la courbure de la membrane. (A) les domaines protéiques BAR sont des dimères étendus adaptés aux membranes possédant différents degrés de courbures. La faible affinité de dimérisation pourrait augmenter la sensibilité à la courbure membranaire. Par des effets d'orientation et de concentration, une membrane incurvée pourrait faciliter la dimérisation des domaines BAR et inversement. (B) Les motifs ALPS senseurs de courbure membranaire sont des séquences peptidiques intrinsèquement dépliées. Ils forment des hélices α amphipathiques à la surface des membranes ayant une forte courbure. Leur reploiement et leur adsorption à la membrane sont sensibles aux changements du rayon de courbure membranaire (B, graphe). Adapté de [94].

Un autre exemple est celui de l'α-synucléine. Cette protéine est essentiellement retrouvée au niveau du tissu cérébral et est associée à l'étiologie de la maladie de Parkinson [97]. Similairement aux motifs ALPS, l'α-synucléine possède un domaine N-terminal non structuré en solution qui se reploie en hélices α amphipathiques au contact de lipides membranaires (Figure 20). Néanmoins, sa composition en acides aminés est différente de celle des motifs ALPS. Lors de l'interaction avec des micelles de SDS, l'α-synucléine adopte une conformation en hélices α composée de deux hélices α antiparallèles courbées et connectées par un « linker » étendu (Figure 20D) [98]. *A contrario*, d'autres auteurs proposent que cette protéine adopte une

conformation allongée lors de l'interaction avec des vésicules phospholipidiques (Figure 20D) [99], soulignant l'importance du modèle membranaire choisi.

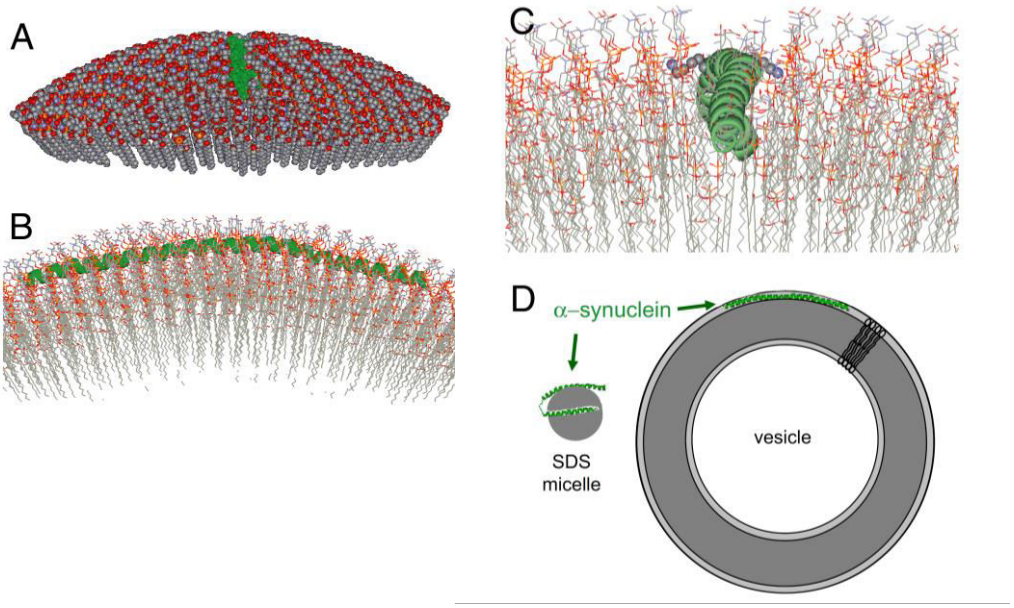


Figure 20 : (A, B et C) modèle de l'α-synucléine en interaction avec une partie d'une vésicule phospholipidique adoptant une structure étendue. (D) Comparaison de la conformation adoptée en présence de vésicules et de micelles de SDS. Adapté de [99].

2. Etude structurale des protéines membranaires

2.1. Généralités

Les protéines sont composées d'un enchaînement d'acides aminés, appelé séquence ou structure primaire. Ces acides aminés peuvent s'organiser localement en structure de type hélice α, feuillet β, boucle ou encore coude. Cet agencement est appelé structure secondaire. Le positionnement de ces structures les unes par rapport aux autres dans l'espace définit la structure tertiaire des protéines. Finalement, si la protéine contient plusieurs chaînes polypeptidiques, comme dans le cas de l'hémoglobine qui est organisée en homotétramère, cette organisation est qualifiée de structure quaternaire (Figure 21).

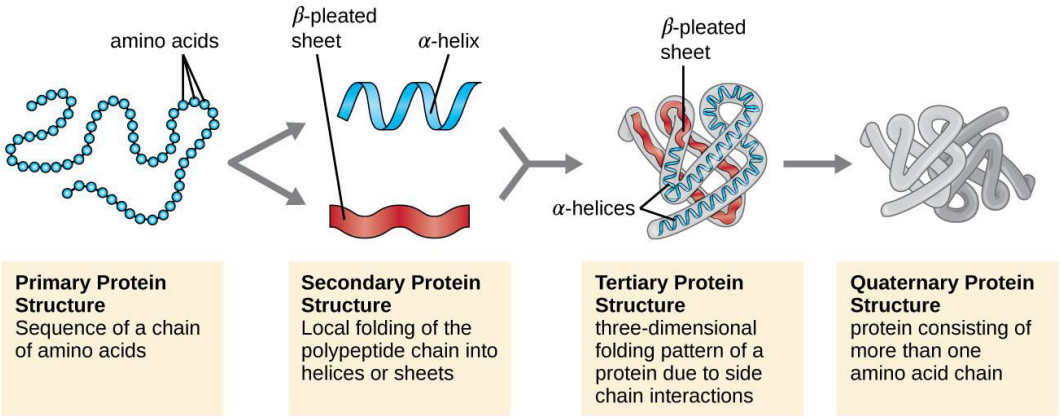


Figure 21 : Représentation des quatre niveaux d'organisation des protéines [100].

Les protéines peuvent être classées en quatre groupes principaux selon leurs caractéristiques structurales : (*i*) les protéines globulaires ; (*ii*) les protéines fibreuses qui possèdent une structure étendue ; (*iii*) les protéines intrinsèquement désordonnées ; et (*iv*) les protéines membranaires. Les protéines membranaires peuvent être structurées en hélices α, avec un ou plusieurs domaine(s) transmembranaire(s), en tonneau β, ou encore en un mélange d'hélices α et de feuillets β. Quelques exemples de structure sont présentés Figure 22.

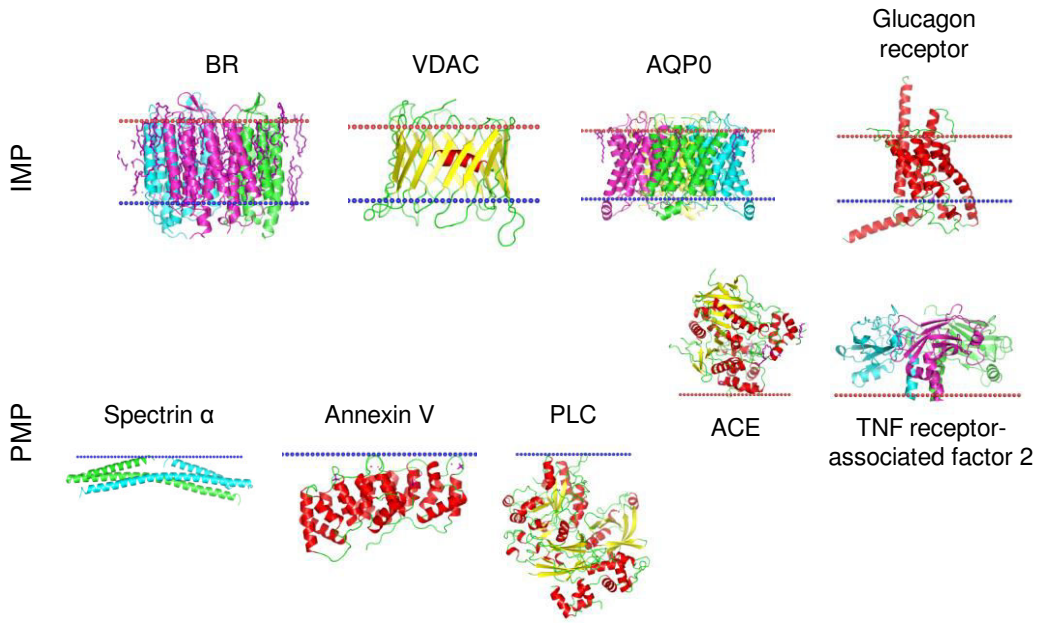


Figure 22 : Exemples de structure de protéines membranaires intégrales (IMP) et périphériques (PMP) avec leur orientation par rapport à la membrane. Feuillet interne (bleu), feuillet externe (roue). AQP0: aquaporine 0, VDAC : « voltage-dependent anion channel », BR : bactériorhodopsine, PLC: phospholipase C, ACE: acétylcholinestérase, TNF: « tumor necrosis factor ». Les images proviennent de la base de données OPM [101].

Environ 26% du protéome humain a été prédit comme étant composé de protéines membranaires [102]. Dans un raisonnement ou un gène code une protéine, ce qui est largement sous-estimé, cela signifie que parmi les ~ 20000 protéines prédites [103] à partir de l'annotation du génome, ~ 5000 sont des protéines membranaires. Cependant, tout organisme confondu, seulement ~ 700 structures 3D de protéines membranaires sont disponibles à ce jour, dont ~ 110 sont des protéines humaines [104]. Ces chiffres mettent en évidence la difficulté à manipuler les protéines membranaires du point de vue de leur solubilisation et/ou de leur purification [24] (Figure 23).

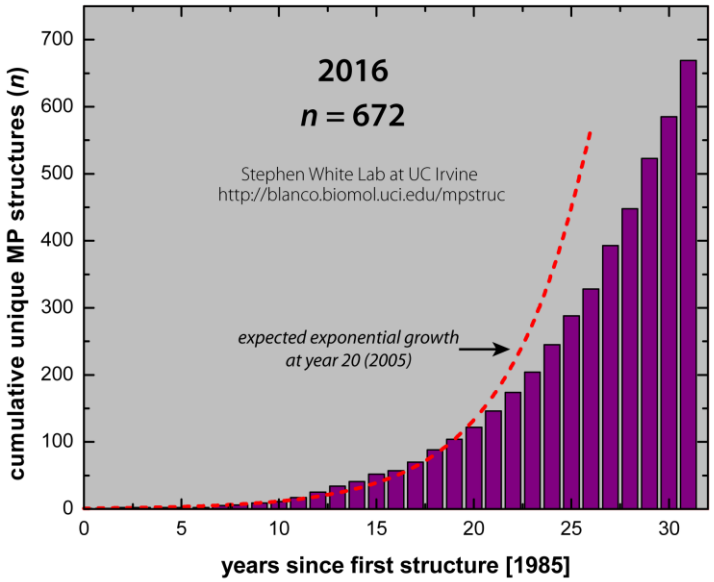


Figure 23 : Progrès dans la détermination des structures des protéines membranaires. Seules les structures uniques sont considérées. Cela signifie que les protéines du même type provenant de différentes espèces sont incluses, mais les structures provenant de différentes mutagenèses sont exclues, tout comme celles qui ne diffèrent que par rapport à l'état physiologique ou au type de substrat lié [104,105].

Obtenir des informations structurales sur les protéines membranaires est un vrai challenge depuis plus de 40 ans [106]. Cependant, des progrès considérables ont été réalisés depuis la première détermination de la structure de la bactériorhodopsine par diffraction des électrons [107] (1975) et depuis l'obtention de la première structure haute résolution d'une protéine membranaire intégrale, le centre réactionnel photosynthétique (1985), par XRC [108]. Classiquement, la RMN et la XRC sont les techniques de référence pour obtenir des informations sur la structure 3D des protéines. Nous ne pouvons pas exclure la montée en puissance de la microscopie électronique (EM) [109] avec l'amélioration des caméras, désormais à détection directe d'électrons (Figure 24). Sur ~ 124000 structures protéiques que compte la

PDB, 90% ont été résolues par XRC, 8.5% par RMN et moins d'1% par cryo-EM. Hormis ces techniques, d'autres approches expérimentales sont également disponibles, comme la cristallographie aux électrons (EC) en utilisant des cristaux 2D de protéines et même, plus récemment, l'utilisation de cristaux 3D. Cette dernière technique est désormais connue sous le nom de microED [110].

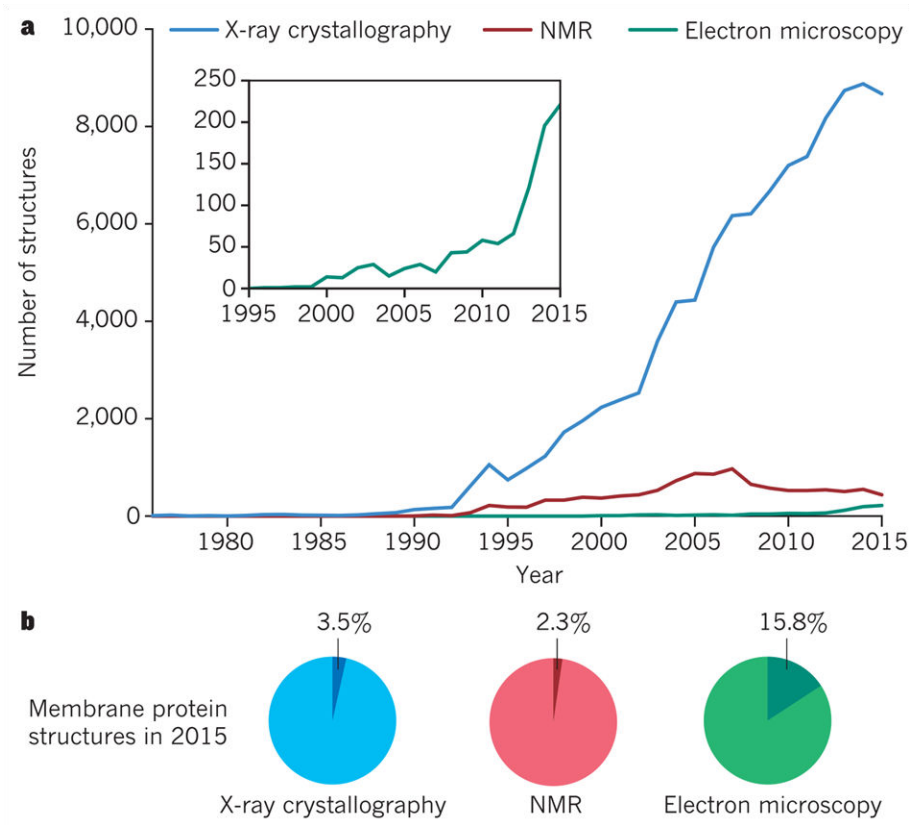


Figure 24 : Evolution des approches de biologie structurale au cours des 40 dernières années. (a) Nombre de structures enregistrées dans la PDB, déterminées par XRC (bleu), RMN (rouge) et EM (vert) entre 1975 et 2015. Insert : nombre de structures déterminées par EM entre 1995 et 2015, mettant en évidence la croissance récente de l'utilisation de cette méthode. (b) Pourcentage de structures de protéines membranaires résolues en 2015 en utilisant chacune des trois techniques [109].

Ainsi, la diversité de techniques disponibles pour un biologiste structural s'avère relativement riche. Cependant, chaque technique possède ses avantages et inconvénients (Tableau 3) et les problèmes inhérents à chaque technique sont souvent similaires pour une protéine soluble ou membranaire. Les principales difficultés se situent souvent en amont, lors des étapes de purification/solubilisation des protéines membranaires [111]. Globalement, toutes les techniques de diffraction nécessitent *de facto* d'obtenir des cristaux. Ainsi, la quantité de protéines nécessaire pour un criblage des conditions de cristallogenèse peut s'avérer

conséquente. Dans le cas où des cristaux sont obtenus, leur taille relativement petite peut constituer un nouveau frein à la résolution de structure. Cependant la microED [112] et l'utilisation de laser à électrons libres [113] permettent d'utiliser des cristaux de plus petites tailles. Un autre point de discussion toujours inhérent à ces techniques est : la structure cristalline est-elle la même que celle adoptée par la protéine en solution ? Il est reconnu que certaines structures cristallines ne sont pas le reflet exact de la structure en solution, tout comme l'état d'oligomérisation qui peut être artéfactuel et causé par les conditions de cristallogenèse. Ces techniques de diffraction, notamment la XRC, restent néanmoins la référence pour obtenir des structure à haute résolution et sans limite de taille. En ce qui concerne la RMN, cette méthode possède le grand avantage d'obtenir des structures de molécules en solution. Néanmoins, pour résoudre la structure d'une protéine, une marquage isotopique ^{13}C et/ou ^{15}N est parfois nécessaire. De plus, la taille limite de la protéine n'excède souvent pas 30 kDa, bien que la structure de protéines de plus hautes masses moléculaires ait été résolue par RMN [114]. Un vrai point fort de la cryo-EM est que la quantité d'échantillon nécessaire est vraiment très faible (~ 5 µL à 10 mg/L) [115], largement moins que pour toutes les autres techniques. Une des faiblesses de la cryo-EM réside cependant dans la taille des macromolécules caractérisables (~ 200 kDa), bien que Merk *et al.* [116] aient repoussé, en juin 2016, à la fois la limite inférieure de taille (< 100 kDa) et celle de la résolution de structure atteignable par cette méthode (1.8 Å). Cependant, certaines protéines ne sont pas compatibles avec ces techniques du fait de leur taille, leur forme, leur flexibilité ou encore leur organisation structurale. Les techniques de diffusion aux petits angles que sont le SANS et le SAXS sont alors des alternatives et/ou compléments de choix (voir ci-après).

Tableau 3 : Comparaison des avantages et inconvénients des différentes techniques utilisées en biologie structurale. §La quantité d'échantillon utilisée varie selon le criblage des conditions de cristallogenèse.

	Quantité d'échantillon	Résolution	Structure	Taille limite de l'échantillon	Marquage
XRC	variable§	haute	cristalline	pas de limite	non
EC	variable	haute	cristalline	pas de limite	non
Micro ED	variable	haute	cristalline	pas de limite	non
RMN	élevée	haute	solution	< 30-50 kDa	^{13}C et/ou ^{15}N
Cryo-EM	faible	moyenne à haute	glace vitreuse	≥ ~ 200 kDa	non
SAS	moyenne	basse	solution	pas de limite	deutération/D_2O (SANS)

Les approches *in silico* font également partie de la panoplie des approches disponibles pour obtenir des informations structurales de protéines membranaires [117]. Dans le cas des IMP, de nombreuses études biophysiques indiquent que les interactions protéine/lipide ou protéine/détergent sont essentielles au repliement ainsi qu'à la stabilité des protéines membranaires intégrales. Cependant, peu de données structurales sont disponibles à propos de ces interactions. La base de données MemProtMD [118] contient ~ 3000 structures d'IMP, identifiées dans la PDB, insérées dans des bicouches lipidiques et simulées à l'aide de dynamique moléculaire gros-grain [119]. Cette approche permet d'apporter des informations concernant les interactions IMP/membrane. La base de données OPM [101] permet quant à elle d'obtenir des informations concernant le positionnement des protéines membranaires par rapport à la membrane [120,121], que celles-ci soient intégrales ou périphériques.

2.2.Structure des protéines membranaires : approche par SAS

Certaines protéines membranaires restent incompatibles avec les méthodes citées ci-avant. D'un point de vue général, le SAXS et le SANS sont de puissants alternatives et/ou compléments pour obtenir des informations structurales d'une protéine, non seulement seule en solution, mais aussi en interaction avec un et ou des partenaire(s) [122,123]. Plus précisément, l'utilisation de la variation de contraste en SANS est la seule technique permettant d'obtenir des informations sur un partenaire impliqué dans un complexe macromoléculaire. Une des études les plus probantes réalisées par SAXS/SANS a contribué à la résolution de la structure

du ribosome bactérien [124]. Désormais, la base de données SASBDB (« Small Angle Scattering Biological Data Bank ») [125] inventorie les modèles obtenus en tout ou partie à partir de données SAS.

- ***SAXS***

Le SAXS permet d'obtenir des informations structurales sur des protéines membranaires, ou plus exactement des complexes protéine/détergents (PDC). Des travaux pionniers réalisés par Berthaud *et al.* [126] ont permis de proposer en 2012 la structure de la couronne de détergent (DDM) entourant une aquaporine, en combinant SAXS et modélisation moléculaire gros-grain. En couplant le SAXS et la chromatographie d'exclusion stérique, les auteurs ont pu obtenir spécifiquement le signal du PDC en s'affranchissant de la contribution du signal des micelles de détergents libres. Le signal SAXS a été comparé à des modèles contenant la structure cristallographique connue de la protéine entourée par des modèles gros-grain de DDM. Le logiciel Memprot [127], développé par J. Pérez et A. Koutsioumpas, est disponible depuis 2015 et peut être utilisé pour la caractérisation d'autres PDC. Skar-Gislinge *et al.* ont, quant à eux, démontré la possibilité d'utiliser le SAXS pour obtenir la forme et la localisation d'une protéine membranaire, le cytochrome P450, intégrée dans des nanodisques [128].

- ***SANS***

La technique de variation de contraste en SANS est la seule technique permettant d'obtenir spécifiquement le signal d'un composant dans un complexe macromoléculaire, en masquant le signal provenant des autres composants [129]. Ceci est possible grâce à l'ajustement du niveau de deutération des composants dans un mélange H_2O/D_2O approprié. En effet, contrairement aux rayons X, les neutrons interagissent avec le noyau des atomes et sont donc sensibles au changement isotopique. Ainsi, les noyaux des atomes d'hydrogène et de deutérium n'ont pas la même capacité à diffuser les neutrons et c'est cette propriété qui est exploitée dans la technique de variation de contraste (se référer à la partie II pour plus d'informations). Cette technique a été appliquée avec succès pour des complexes protéine/protéine et protéine/ADN [129] et également pour des protéines membranaires dans des environnements membranomimétiques [130]. Un premier exemple notable concerne une étude pionnière ayant recours à la technique de variation de contraste. En 1997, Hunt *et al.* [131] ont déterminé l'état d'oligomérisation de la bactériorhodopsine dans des liposomes de DMPC. Une preuve de principe qui déjà laissait présager la possibilité d'utiliser le SANS pour la caractérisation structurale de protéines

membranaires. Un deuxième exemple plus récent concerne une étude menée par Gabel *et al.* [132]sur la conformation du transporteur FhaC, structuré en tonneau β, et entouré par une couronne de détergent de type β-octyl glucoside (β-OG). En couplant SANS et modélisation moléculaire, une structure de la protéine similaire à celle résolue par XRC a été obtenue. Un dernier exemple concerne l'étude de la protéine de phage pb5 lors de la sa liaison à son récepteur membranaire, la protéine FhuA d'*E. coli* [133]. En combinant l'utilisation de protéines hydrogénées et deutérées, et en masquant le signal du surfactant entourant FhuA, Breyton *et al.* [133] ont obtenu des structures à base résolution (i) des protéines seules, (ii) des protéines lors de leur interaction et (iii) du complexe dans son entier. Ainsi, les auteurs ont démonté que les deux protéines maintenaient leur conformation lors de l'interaction ligand/récepteur.

Une grande partie des études citées ci-dessus ont eu recours à l'utilisation de micelles de détergent. Bien que celles-ci soient considérées comme appropriées au maintien de la structure tridimensionnelle des protéines membranaires, d'autres modèles peuvent être utilisés et gagnent en intérêt afin de fournir un environnement lipidique considéré comme plus proche de conditions natives [33,62,134]. Parmi eux, on trouve les modèles décrits dans le chapitre précédent, à savoir les nanodisques et les bicelles, qui se sont avérés des alternatives efficaces, bien que tout deux présentent avantages et inconvénients [33,62,134]. Dans le cas des nanodisques, Maric *et al.* [135] ont proposé une méthode intéressante pour utiliser ce modèle membranaire lors d'expérience SANS en vue d'analyser des protéines membranaires intégrales incorporées dans un environnement lipidique de type natif. Cependant, dans le cas des protéines membranaires périphériques, il est essentiel de maintenir une tête PC sur tout le modèle membranaire pour éviter des interactions non spécifiques qui pourraient se produire entre la MSP et la protéine d'intérêt dans les nanodisques. De plus, l'ajustement de la taille des nanodisques requiert l'utilisation de plusieurs formes tronquées de MSP [136,137], ce qui nécessite un travail supplémentaire de purification de protéines. Finalement, l'utilisation de MSP sous forme deutérée peut venir compliquer la purification surtout au niveau des rendements.

2.3.Structure des protéines membranaires : le modèle « bicelle »

Le modèle membranaire de bicelle est un modèle de choix afin d'obtenir des informations structurales sur des protéines membranaires. Les bicelles ont l'avantage de présenter une bicouche lipidique contrairement aux micelles de détergents dont la structure et les propriétés physico-chimiques peuvent être très différentes de celles des phospholipides membranaires. Dans les cas de bicelles capables de s'aligner dans un champ magnétique, cette faculté permet

d'obtenir des spectres de meilleure qualité et donc facilite les études structurales des protéines membranaires réalisées en RMN du solide. Pour les bicelles isotropes, ou « fast-tumbling bicelles », leur rotation rapide permet également d'obtenir des spectres de haute qualité lors d'expériences de RMN du liquide. Un exemple d'utilisation des bicelles s'orientant dans un champ magnétique est celui de l'obtention de la structure du cytochrome b5 pleine taille [138] par RMN du solide. Dans le cas de l'utilisation des bicelles isotropes pour la caractérisation de la structure de protéines membranaires, un exemple des plus notables est celui de l'obtention de la structure atomique d'une protéine de l'enveloppe du virus HIV [139] par RMN du liquide (Figure 25). Cette protéine est impliquée dans l'initiation de l'infection des cellules.

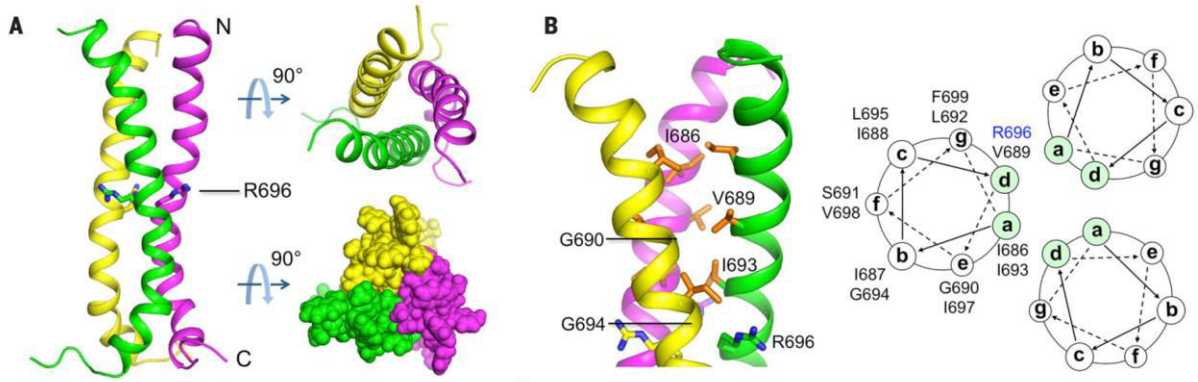


Figure 25 : Structure RMN de la protéine gp41HIV1D (677-716) sous forme de trimère et incorporée dans des bicelles. (A) Représentation en rubans de la structure de plus basse énergie de l'ensemble calculé. (B) Partie N-terminale de la structure qui possède des résidus hydrophobes (orange) structurés en « coiled-coil ». Adapté de Dev *et al.* [139].

En plus de leur utilisation en RMN, les bicelles sont de plus en plus utilisées en XRC. Cette nouvelle méthode de cristallisation peut être un complément, voire une alternative, à la cristallisation en phase cubique [140]. La première publication rapportant l'utilisation de bicelles pour cristalliser une protéine membranaire date de 2002. Ces travaux ont été réalisés par Faham et Bowie [141] sur la bactériorhodopsine en utilisant des bicelles DMPC/CHAPSO (3:1). En 2015, 19 structures de protéines membranaires ont été obtenues par XRC en utilisant les bicelles [54]. Rasmussen *et al.* ont publié la structure d'un récepteur couplé aux protéines G (GPCR), l'adrénorécepteur β_2. Plus récemment, Hosaka *et al.* [142] ont obtenu une structure à haute résolution du canal VDAC1 (voltage-dependent anion channel 1) et proposé un nouvel état d'oligomérisation.

Tableau 4 : Tableau récapitulatif des protéines membranaires cristallisées dans des bicelles. 7 TM : récepteur à 7 domaines transmembranaires. Adapté de Poulos *et al.* [54].

Nom de la protéine	Code PDB	Composition des bicelles	Résolution (Å)
Serpentines (7 TM récepteurs)			
Bacteriorhodopsin (BR)	1KME	DMPC/CHAPSO (3:1)	2.0
BR (température ambiante)	1XJI	DTPC/CHAPSO (3:1)	2.2
β2 Adrenergic GPCR	2R4R	DMPC/CHAPSO (3:1)	3.4
Xanthorhodopsin	3DDL	DMPC/NM (0.6:1)	1.9
BR	4HYX, 4HWL		2.0
Proteorhodopsin	4JQ6		2.3
Enzymes			
Protéase rhomboide	2XTV	DMPC/CHAPSO (2.6:1)	1.7
Transglycosylase	3VMT	DMPC/CHAPS (3:1)	2.3
Cellulose synthase	4P02	DMPC/CHAPS (2.34:1)	2.65
Transporteurs			
LeuT-Leu	3USG	DMPC/CHAPSO (2.8:1)	2.5
LeuT-SeMet	3USL	DMPC/CHAPSO (2.52:1)	2.7
c10 ring ATP synthase	3U2F	DMPC/CHAPSO (2:1)	2.0
Maltose transporter	4JBW	DMPC/CHAPSO (2.7:1)	3.9
Canaux ioniques			
Voltage-gated Na channel	3RVY	DMPC/CHAPSO	2.7
Voltage-gated Ca channel	4MS2	DMPC/CHAPSO	2.75
Tonneaux β			
VDAC	3EMN	DMPC/CHAPSO (2.8:1)	2.3
TamA	4C00	DMPC/CHAPSO (2.43:1)	2.3
HdBamAΔ3	4K3C	DMPC/CHAPSO (2.8:1)	2.9
NgBamA	4K3B	DMPC/CHAPSO (2.8:1)	3.2

Une des avancées majeures concernant la résolution par XRC de la structure des protéines membranaires avec le modèle bicelle est sans doute le couplage de laser à électrons libres avec des bicelles constituées de détergents contenant des atomes d'iode, mis en place par Nakane *et al.* [143] (PDB : 5B34). Cette méthode a été appliquée à la bactériorhodopsine et permet de résoudre le problème de la phase par diffusion anomale et remplacement isomorphe. Ces résultats ouvrent la voie à la détermination de la structure *de novo* de protéines membranaires. Ainsi, les bicelles sont des membranomimétiques polyvalents pouvant être utilisés en RMN et XRC en vue de la caractérisation structurale de protéines membranaires. Etonnamment, aucune publication à notre connaissance ne fait référence à leur utilisation en microscopie électronique dans le cas d'une caractérisation structurale de protéines membranaires. Les nanodisques leur sont préférés à ce jour [144].

Chapitre 3 : Dystrophine et dystrophinopathies

1. Dystrophinopathies et thérapies

1.1.Le muscle strié squelettique

Les vertébrés possèdent trois types de muscle : cardiaque, lisse et strié squelettique. Ce dernier type est le plus commun et le corps humain en compte 590 différents. Les muscles squelettiques sont attachés aux os par les tendons et ils produisent tous les mouvements des parties du corps, l'une par rapport à l'autre. Contrairement au muscle lisse et au muscle cardiaque, le muscle squelettique est sous contrôle volontaire [145]. Le muscle strié squelettique est composé de nombreuses fibres de forme cylindrique, de 0,02 à 0,08 mm de diamètre chacune. Chaque fibre est délimitée par une membrane plasmique appelée sarcolemme dans le cas des cellules musculaires. Il existe un haut degré d'organisation dans les fibres musculaires, avec la présence d'une série de bandes alternativement sombres (bande A) et légères (bande I). Chaque bande s'étend perpendiculairement à la longueur de la fibre [146] (Figure 26).

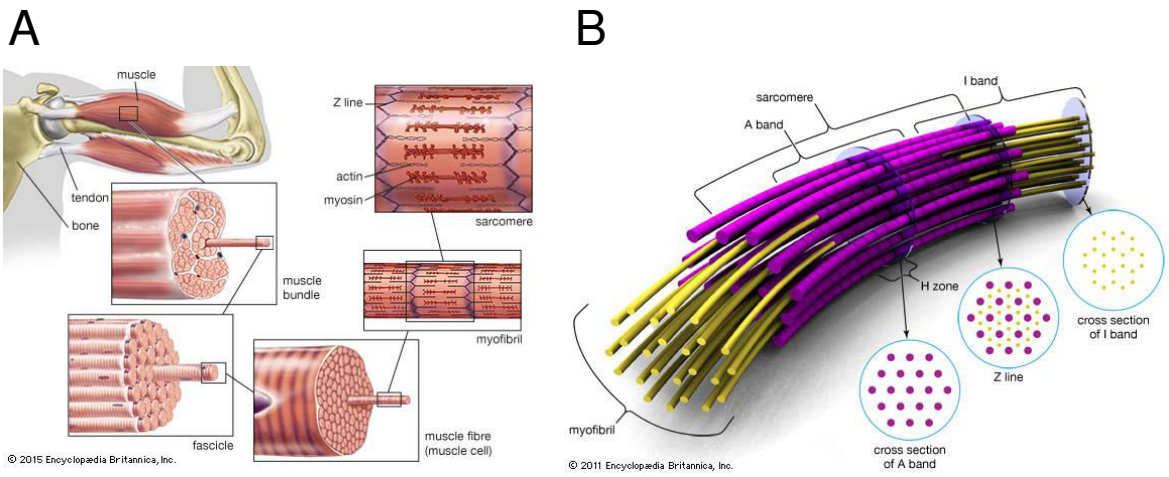


Figure 26 : (A) Représentation et organisation du muscle strié squelettique à différentes échelles [145]. (B) Représentation de l'intérieur d'une fibre musculaire, sans le sarcolemme [146].

La dystrophine une protéine essentielle de soutien de la membrane des cellules musculaires et est retrouvée dans les trois catégories de muscles en diverses proportions [147]. Dans les fibres musculaires elle ne représente qu'environ 0.002% des protéines des cellules du muscle strié squelettique [148].

1.2.Les dystrophinopathies

Comme leur nom l'indique, les dystrophinopathies sont des maladies liées à la dystrophine. Ce sont des maladies génétiques récessives liées au chromosome X causées par des mutations dans le gène *DMD*. Ces mutations peuvent être des délétions, des insertions et également des mutations ponctuelles et touchent la quasi-totalité des exons avec une occurrence différente [149]. Deux grands types de dystrophies musculaires peuvent être distingués. La dystrophie musculaire de Duchenne (DMD) est une forme très sévère de myopathie touchant environ 1 sur 5000 naissances masculines [150]. Son nom provient de celui du Dr. Duchenne de Boulogne qui a grandement contribué à la description de cette pathologie. Dans le cas d'une DMD, les mutations dans le gène conduisent à un décalage du cadre de lecture accompagné d'un déficit total en protéine. L'absence de dystrophine conduit à une rupture du sarcolemme suivie de cycles de dégénérescence/régénération (Figure 27), de la nécrose, de l'inflammation et finalement un remplacement progressif du muscle par du tissu graisseux. Les patients atteints par cette pathologie décèdent fréquemment avant l'âge de 30 ans. La cause majeure de décès est liée à des complications cardio-respiratoires.

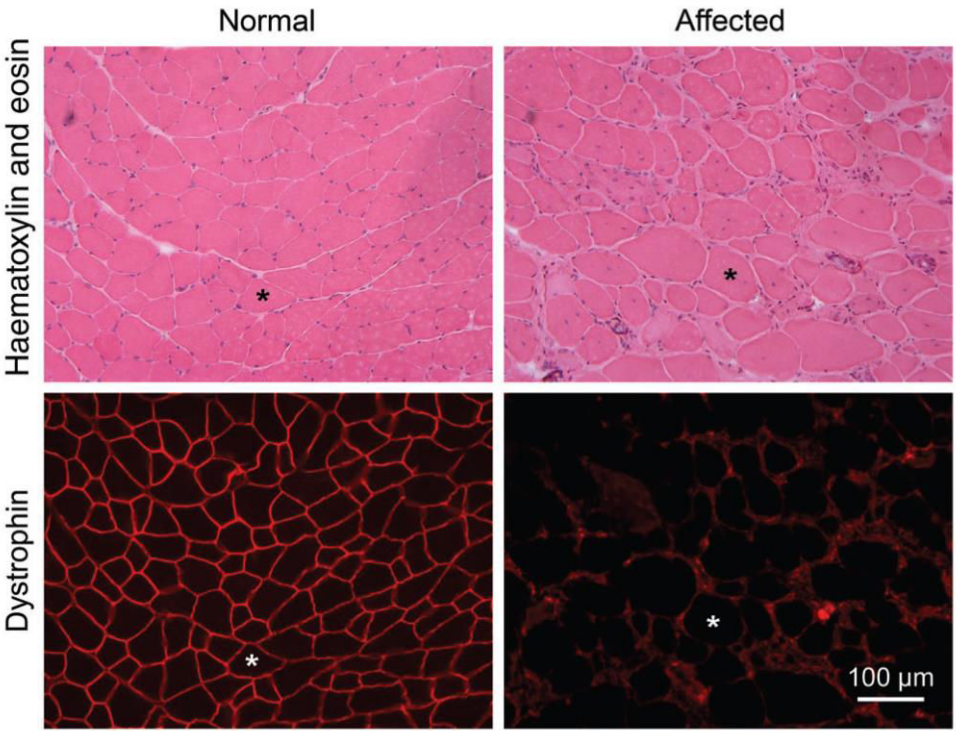


Figure 27 : Comparaison entre des fibres musculaires d'une souris normale (gauche) et d'une souris atteinte d'une dystrophinopathie (droite). (Haut) Coupe transversales de tissus musculaires colorés à l'hématoxyline (noyaux en bleu) et à l'éosine (cytoplasme en rose). (Bas) Immuno-localisation de la dystrophine par immunofluorescence. Dans le muscle normal, la

dystrophine est localisée au sarcolemme. La présence de noyaux au centre des fibres musculaires (bleu) indique une régénération récente de ces fibres. Les astérisques servent de point de repère et se trouvent au milieu d'une fibre musculaire [151].

Dans le cas de la dystrophie musculaire de Becker [152] (nom venant du Dr. E. Becker, BMD), les mutations dans le gène ne décalent pas le cadre de lecture et engendrent la production de dystrophine sous une forme tronquée plus ou moins fonctionnelle. Les phénotypes des patients BMD varient très fortement, allant de profils comparables à ceux de patients DMD à des formes quasi asymptomatiques. L'occurrence des BMD est très difficile à chiffrer à cause de l'hétérogénéité phénotypique.

1.3. Exemples de thérapies

Plus de 30 ans après la découverte du gène impliqué dans ces myopathies, le gène *DMD*, les dystrophies musculaires de Duchenne et de Becker restent actuellement incurables [153]. Plusieurs thérapies actuellement au stade d'essai clinique ont pour but d'améliorer la fonction musculaire, de diminuer la progression de la maladie et même de transformer certains patients DMD en patients BMD de phénotype le moins sévère possible (Figure 28).

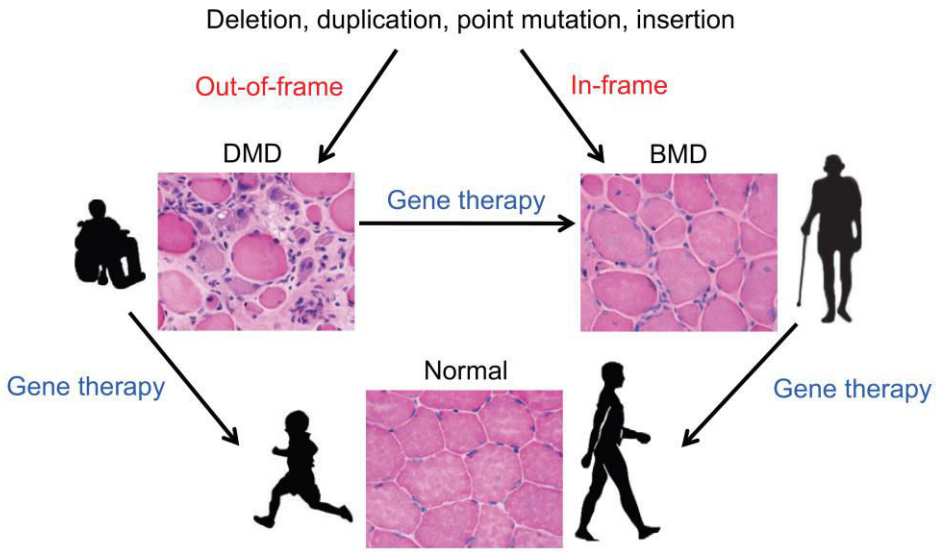


Figure 28 : Représentation des objectifs de thérapies géniques envisagées. Les mutations décalant le cadre de lecture entraînent une myopathie de Duchenne (DMD), sévère, alors que celles ne le décalant pas conduisent à une dystrophie musculaire de Becker (BMD), plus ou moins sévère. Une stratégie thérapeutique vise à changer le phénotype DMD en un phénotype moins sévère retrouvé chez certains patients BMD. Adapté de [154].

1.3.1. Thérapie génique

➢ ***Saut d'exon***

Le saut d'exon (ou « exon skipping ») consiste en l'utilisation d'oligonucléotides anti-sens (AON) qui s'hybrident sur l'exon porteur de la mutation sur le pré-ARNm. Cette hybridation conduit à un saut d'exon permettant de restaurer un cadre de lecture et ainsi de produire une dystrophine raccourcie mais supposée fonctionnelle. Un exemple de ces AON est l'Eteplirsen® qui s'hybride sur l'exon 51 (Figure 29) qui fait partie du « hot spot » d'occurrence de mutations. Cette molécule est développée par la société Sarepta Therapeutics et des essais cliniques sont en cours.

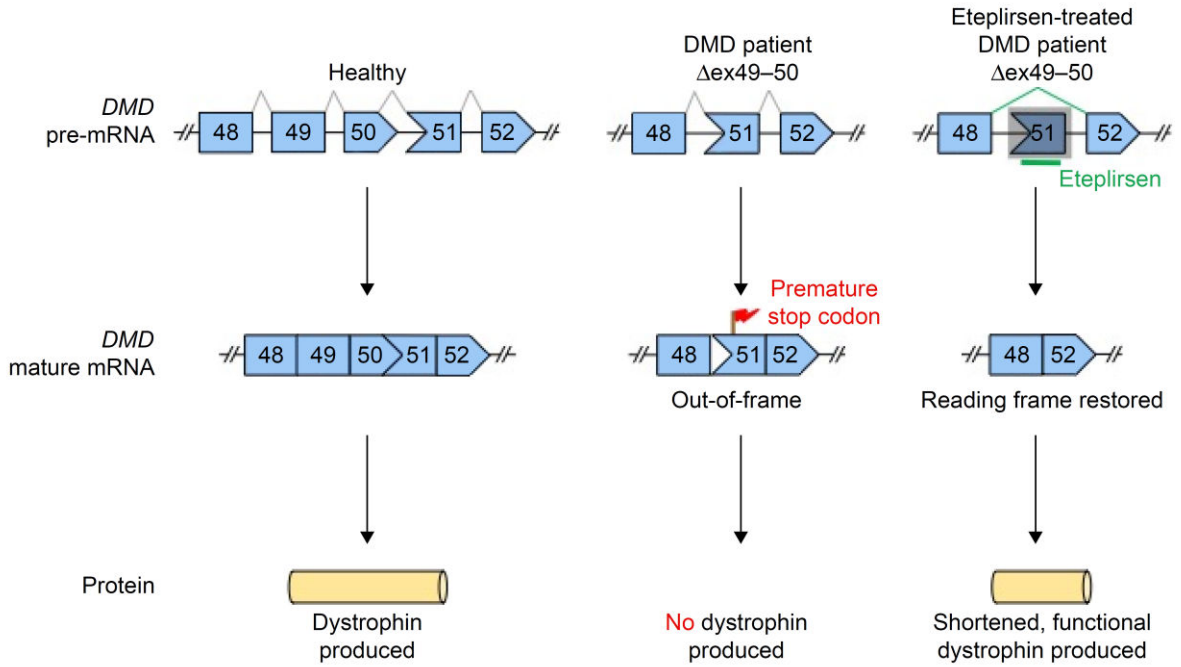


Figure 29 : L'Eteplirsen® (barre verte) reconnaît spécifiquement l'exon 51 du gène DMD. Sa liaison au pré-ARNm *DMD* influence la machinerie d'épissage (le spliceosome) et permet de sauter l'exon 51 lors de la maturation de l'ARNm. Ainsi, le cadre de lecture est restauré, ce qui permet la traduction d'une dystrophine raccourcie. L'exemple concerne un cas où l'Eteplirsen® est utilisé pour traiter un patient DMD ayant une délétion des exons 49 et 50 dans le gène *DMD*. Cette délétion crée un décalage du cadre de lecture qui introduit un codon stop prématuré et donc une non-production de la dystrophine [155].

➢ ***Mini-dystrophines***

Une autre approche thérapeutique consiste en l'utilisation de virus adéno-associé (AAV) modifiés afin d'apporter de la dystrophine exogène pour restaurer la fonction musculaire. Le problème des vecteurs AAV est qu'ils peuvent intégrer à ce jour un matériel génétique d'une

taille maximale de 5 kb. Malheureusement, c'est presque trois fois moins que la taille de l'ADNc de la dystrophine (14 kb) et 2,3 fois moins que la taille de la séquence codante (11.2 kb). Ainsi, l'objectif consiste à développer des micro/mini-dystrophines dont la séquence codante pourrait être incorporée dans des AAV, et qui conserveraient le maximum des propriétés de la dystrophine entière. Ces stratégies sont inspirées entre autres de certaines observations cliniques. Par exemple, dès 1990, il a été rapporté un cas de patient BMD présentant un phénotype peu sévère et étant en fait porteur d'une mutation ayant conduit à la perte de 46% de la dystrophine [156] (Figure 30).

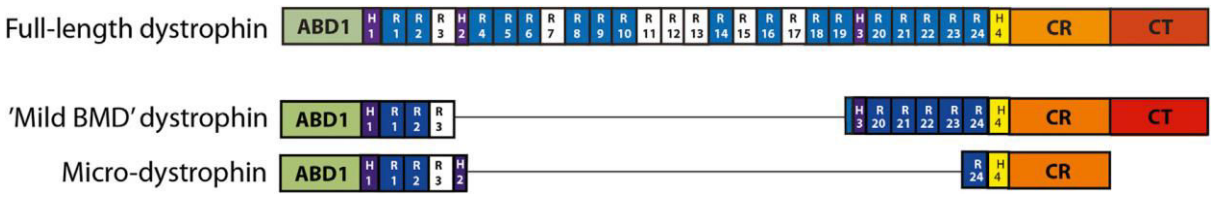


Figure 30 : Comparaison [157] des domaines la dystrophine pleine taille (haut) avec une mini-dystrophine présente chez un patient BMD de phénotype très peu sévère [156] (milieu). Une délétion dans le gène DMD a conduit à l'absence des exons 17 à 48. (Bas) Exemple d'une mini/micro-dystrophine utilisée lors de tests [158].

- ***CRISPR-Cas9***

Le système CRISPR/Cas9 (« clustered regular interspaced short palindromic repeats / CRISPR associated protein 9 ») a été initialement découvert chez les bactéries et s'apparente à l'ancêtre de l'immunité acquise. Il intervient dans la défense des bactéries contre les phages (virus des bactéries). Au vu de son mécanisme, il peut être considéré comme l'ancêtre de l'ARN interférence, un mécanisme de régulation du taux d'ARNm. Le système CRISPR/Cas9 est largement utilisé en routine comme outil de modification de génome, pour invalider un gène par exemple (« knock out »). Entre fin 2015 et début 2016, trois articles issus de trois laboratoires différents, rapportent l'utilisation de la technologie CRISPR/Cas9 pour retirer la portion du gène *DMD* porteur de mutations comme une nouvelle stratégie de thérapie génique en vue du traitement des dystrophinopathies [159–161].

- ***« New and notable » : l'utilisation de lentivirus***

Très récemment, Counsell *et al.* [162] ont rapporté la possibilité d'avoir recours à des lentivirus afin de pouvoir envisager une thérapie générique avec de la dystrophine de pleine longueur. Dans la perspective de leur utilisation en thérapie génique, ces lentivirus possèdent un réel avantage par rapport aux AAV. En effet, il est possible d'y incorporer du matériel

génétique dont la taille est supérieure à 15 kb, soit trois fois plus que pour les AAV classiques. Ces travaux ouvrent de nouvelles stratégies de thérapies géniques dédiées aux traitements des patients DMD et BMD. Les lentivirus, dont fait partie HIV, sont des virus pathogènes, mais sont rendues inoffensifs et ont déjà été utilisés en thérapie génique [163].

1.3.2. Thérapie cellulaire

La deuxième approche majeure à ce jour pour le traitement des dystrophinopathies est la thérapie cellulaire [164], basée sur l'utilisation de cellules souches vivantes. Cette technique consiste à prélever ces cellules, soit chez le patient, soit chez un donneur. Ces cellules souches doivent ensuite être purifiées, multipliées et potentiellement modifiées. L'objectif est d'ensuite implanter ou réimplanter les cellules chez le malade pour remplacer les cellules déficientes. Dans le cas d'une DMD, la transplantation de cellules souches a pour objectif de favoriser la régénération musculaire en vue d'améliorer la fonction motrice. Les premières études menées sont prometteuses [164,165].

2. Le gène *DMD*

La dystrophine est codée par le gène *DMD* [166] qui est le plus long gène humain ($2{,}4.10^6$ paires de bases) (Figure 31). Ce gène comporte sept promoteurs spécifiques selon le type de tissu [167]. La transcription complète de la dystrophine (notée *dp* dans la nomenclature ci-dessous) est contrôlée par trois promoteurs localisés en amont du premier exon. Les promoteurs de muscle (m), cerveau (c) et Purkinje (type de cellules neuronales) (p) sont nommés selon leur principal site d'expression. De plus, il existe quatre promoteurs internes (dp260, dp140, dp116 et dp71) nommés selon le poids moléculaire de la protéine produite [168]. L'isoforme dp260 est principalement retrouvée dans la rétine, dp140 se trouve principalement dans le cerveau, et dp116 dans les cellules de Schwann (un type de cellules gliales formant l'environnement des neurones). Des transcrits dp71 ont été trouvés dans la plupart des tissus autres que le tissu musculaire squelettique, tels que le cerveau, le rein, le foie, les poumons ou encore le muscle cardiaque. La forme dp427m est la principale protéine produite dans le muscle squelettique. Le gène correspondant est sous le contrôle du promoteur du même nom (dp427m) et est composé de 79 exons transcrits en un ARNm de 14 kb après épissage, pour une séquence codante de 11.2 kb.

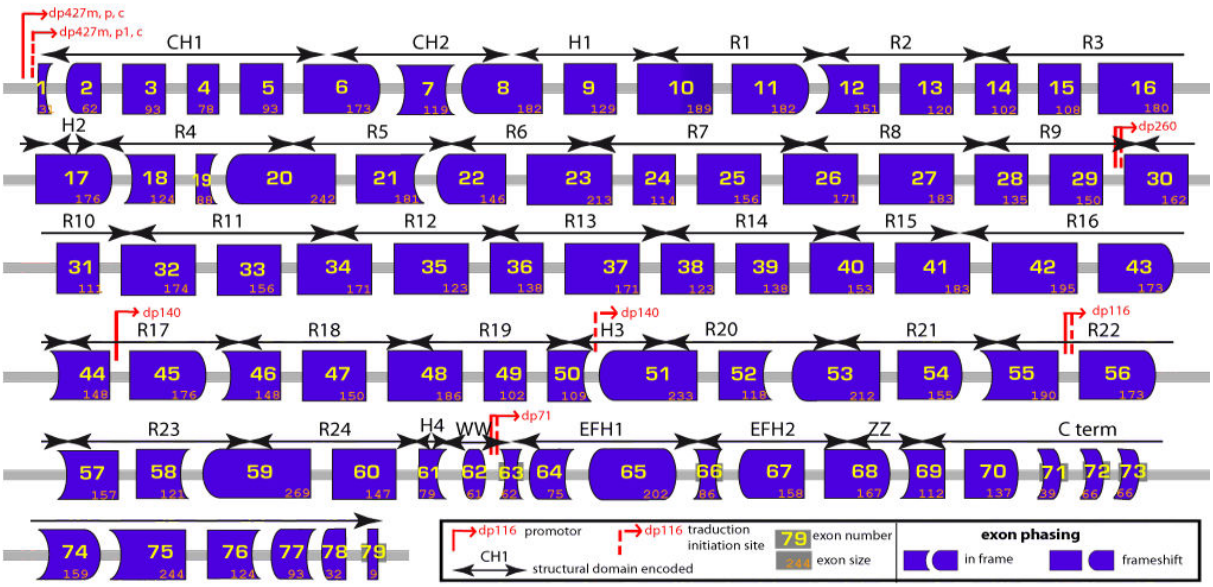


Figure 31 : Représentation du gène *DMD* (*http://edystrophin.genouest.org/*) [169].

L'ARNm est traduit en une grande protéine filamenteuse [170] composée de 3685 acides aminés pour une masse moléculaire de 427 kDa. C'est cette protéine qui fait l'objet de ces travaux de thèse.

3. La dystrophine

3.1. Organisation

La dystrophine est une grande protéine filamenteuse composée de quatre domaines [171] (Figure 32). Les éléments structuraux détaillés seront présentés en **3.2**.

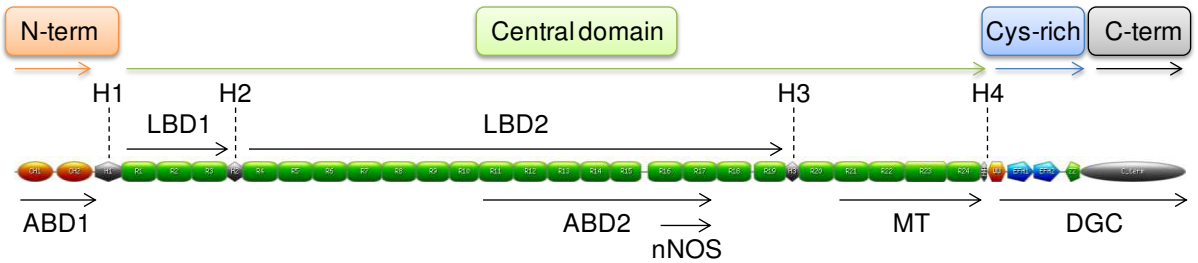


Figure 32 : Représentation schématique de l'organisation de la dystrophine et des interactions avec ses partenaires majeurs. H1-4, charnières (« hinges ») 1-4; ABD, domaine de liaison à l'actine ; LBD : domaine de liaison aux lipides. nNOS : domaine de liaison à la synthase neuronale de l'oxide nitrique MT : domaine de liaison au microtubule, DGC : domaine de liaison au complexe dystrophine et glycoprotéines associées.

➢ ***Le domaine N-terminal (NT)***

Le domaine N-terminal est composé de 240 acides aminés et est structuré en deux sous-domaines CH1 et CH2 (« Calponin Homology »). Il est capable de se lier à l'actine filamenteuse du cytosquelette. Ainsi, il est souvent retrouvé sous l'appellation ABD1 (« actin binding domain 1 »).

➢ ***Le domaine central***

Le domaine central représente 75% de la dystrophine (~ 2800 acides aminés) et est composé de 24 répétitions de type spectrine notées de R1 à R24. Ces répétitions sont organisées en triple faisceau d'hélices α, ou « coiled-coil » (HA, HB et HC), dans lesquels les hélices sont reliées par deux boucles HA/HB et HB/HC. Dans une structure en triple « coiled-coil », l'hélice C d'une répétition est en continuité avec l'hélice A subséquente. Cette région spécifique est appelée « linker » (Figure 33).

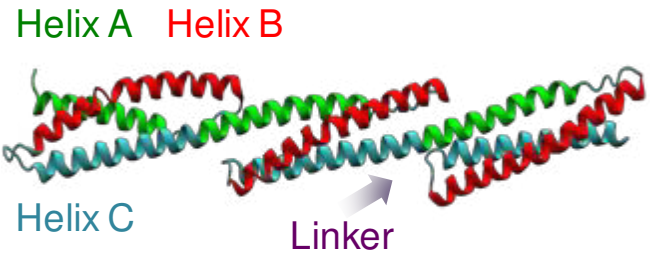


Figure 33 : Représentation de trois répétitions de types spectrine organisées en triple faisceau d'hélices ou « coiled-coil ». PDB: 1U4Q [172].

Chacune des 24 répétitions contient ~ 110 résidus pour une MW d'environ 12 kDa. Bien que partageant une structure supersecondaire hautement conservée, ces répétitions ont des propriétés physico-chimiques propres [173]. Les répétitions du domaine central sont segmentées par quatre régions charnières (« hinges », H1 à H4), donnant lieu à trois sous-domaines compris entre H1-H2 (R1-3), H2-H3 (R4-19) et H3-H4 (R20-24). De plus, une région qualifiée de pseudo-charnière a été localisée entre R15 et R16 [173].

➢ ***Domaine riche en cystéines (Cys-rich, CR)***

Ce domaine tire son nom de la présence de 15 résidus cystéines parmi les 280 qui le composent. Il est composé d'un sous-domaine WW, de deux sous-domaines contenant deux motifs dits main EF et d'un dernier sous-domaine ZZ, capable de fixer le zinc.

➢ ***Le domaine C-terminal (CT)***

Le quatrième et dernier domaine consiste en la partie C-terminale qui compte ~ 330 résidus et forme une plateforme d'interactions avec deux protéines cytoplasmiques, la dystrobrévine et la syntrophine.

3.2.Structure

Pour comprendre et déterminer de façon exhaustive le rôle biologique de la dystrophine, l'équipe SIM utilise une approche structure/fonction. La dystrophine est une très grande protéine (427 kDa). Ainsi, obtenir des informations structurales sur la dystrophine entière est un vrai challenge qui n'a pas encore été relevé avec succès. Cette protéine est structurée en sous-domaines fonctionnels, ce qui permet d'envisager d'obtenir des informations structurales sur des fragments protéiques, tout en gardant une pertinence biologique. A ce jour, seules les structures de trois parties ont été résolues par XRC : le domaine N-terminal [174] (PDB : 1DXX), la répétition R1 du domaine central, cristallisée grâce à un pont disulfure artificiel entre deux monomères pour réduire la flexibilité de la protéine [175] (PDB : 3UUN), et le domaine riche en cystéines [176] (PDB : 1EG3) (Figure 34).

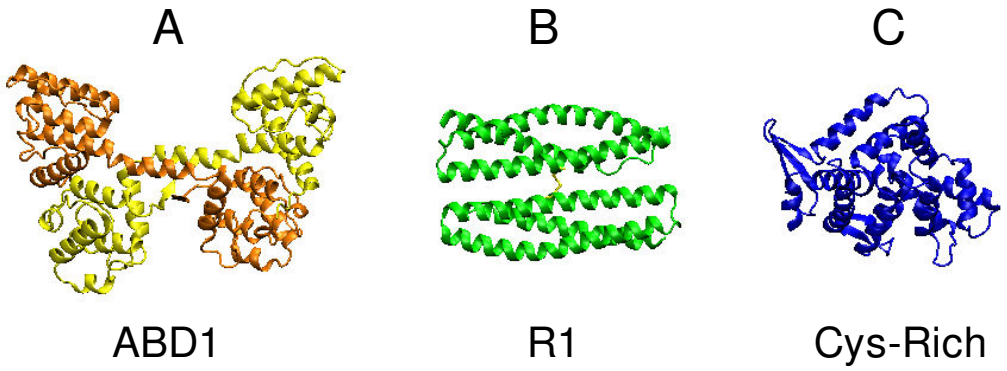


Figure 34 : Structure tridimensionnelle résolue par XRC de parties de la dystrophine. (A) Le domaine N-terminal sous forme dimérique. Les domaines CH (« calponin homology ») adoptent une structure en hélices α. (B) Répétition R1 du domaine centrale structuré en triple faisceau d'hélices α ou « coiled-coil ». Ce fragment protéique a été cristallisé sous forme de dimère à l'aide d'un pont disulfure artificiel (jaune). (C) Partie du domaine CR comprenant les motifs WW, EF1 et EF2.

Le domaine central de la dystrophine représente 75% de la protéine. Pourtant, seule la structure 3D de la répétition R1 a été résolue. En effet, les techniques classiques de biologie structurale, telles que la RMN et la XRC, ne peuvent pas être appliquées au domaine central de

la dystrophine à cause de sa grande taille, de son importante flexibilité et de sa structure répétée en « coiled-coil ». Des méthodes alternatives ont dû être employées pour pallier à ce manque d'information. L'équipe SIM a commencé par générer 21 modèles du domaine central composés de deux répétitions (R1-2, R2-3, … , R22-23 et R23-24) à l'aide du serveur I-TASSER [177]. Ces modèles ont été obtenus par homologie avec la spectrine. Grâce à une approche exclusivement basée sur de la dynamique moléculaire, l'équipe a pu mettre en évidence que chaque répétitions du domaine central possède des propriétés électrostatiques et hydrophobes propres [173]. A partir de ces observations, le long domaine central a pu être décomposé en sept parties arborant des caractéristiques distinctes et pouvant être considérés comme des sous-domaines aux propriétés fonctionnelles et structurales spécifiques [173].

Plus récemment, l'équipe SIM a développé une approche alternative afin d'obtenir des informations structurales sur ce domaine et de proposer des modèles tout atome de fragments du domaine central de la dystrophine. Cette méthode est basée sur la combinaison du SAXS et de la modélisation moléculaire [178]. Molza *et al.*[178] ont d'abord utilisé cette méthode innovante pour les fragments R11-15 et R16-17 : en voici brièvement les étapes. Tout d'abord, un modèle par homologie est obtenu en se basant sur la spectrine, grâce au serveur I-TASSER [177]. Ensuite, ce modèle est converti en modèle gros-grain reliés entre eux par des réseaux de ressorts. En parallèle, une enveloppe SAXS est générée avec le logiciel GASBOR (suite ATSAS, voir partie II) d'après une courbe de diffusion expérimentale. Finalement, les données expérimentales et *in silico* sont combinées grâce à de la dynamique moléculaire interactive (IMD) en utilisant le logiciel BioSpring [179], où le modèle gros-grain est ajusté sous contraintes expérimentales à l'aide d'un bras haptique. Après une étape d'optimisation, des modèles tout-atome des fragments protéines peuvent être proposés (Figure 35). Cette approche à été étendue à d'autres fragments protéiques et la structure du domaine central est ainsi en cours d'élucidation (partie IV, chapitre 1).

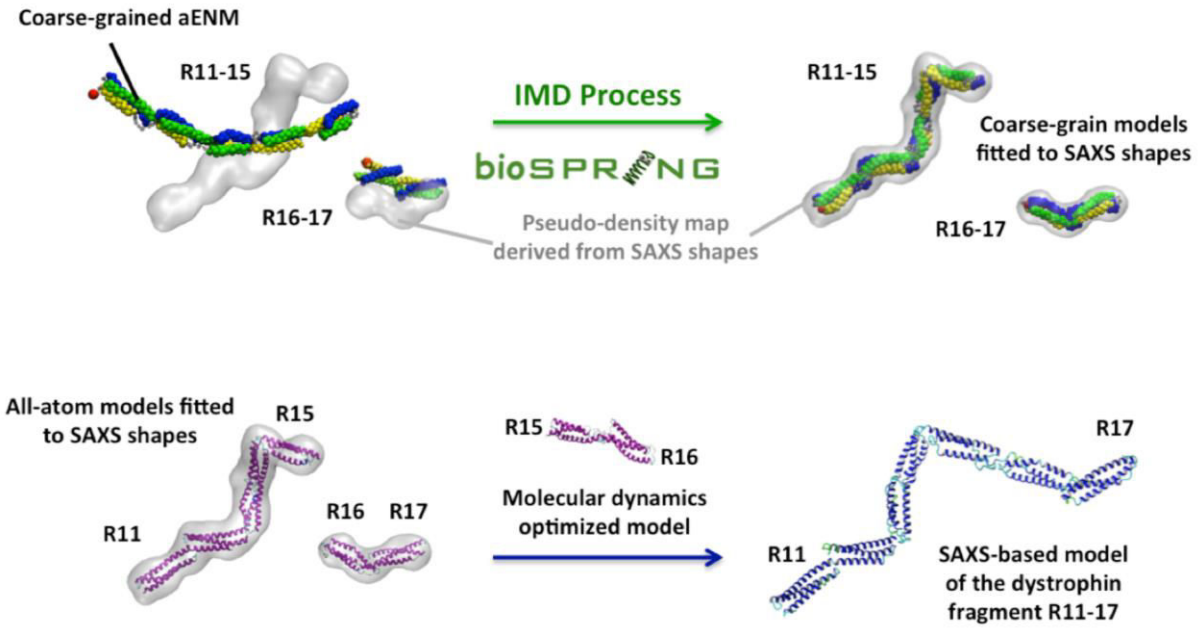


Figure 35 : (Haut) Représentation schématique du protocole de dynamique moléculaire interactive en utilisant le logiciel BioSpring [179], appliqué aux modèles gros-grain des fragments de dystrophine R11-15 etR16-17. aENM : « augmented elastic network model ». Les hélices HA sont représentées en jaune, les hélices HB sont en bleu et les hélices HC en vert. Le grain rouge représente l'extrémité N-terminale. Les enveloppes grises ont été obtenues à partir du programme GASBOR de la suite ATSAS à partir des données de SAXS. (Bas) Modèles tout atome de R11-15 et R16-17 ajustés dans l'enveloppe SAXS et proposition d'un modèle tout-atome du fragment R11-17 après optimisation. Adapté de Molza et al.[178].

3.3.Structure des protéines issues de délétions d'exons

En plus de la caractérisation structurale de domaines de la dystrophine sous forme native, l'équipe SIM s'est penchée sur la structure que pouvait adopter des produits issus de délétion d'exons. En effet, les thérapies dédiées au traitement des patients DMD et/ou BMD s'appuient souvent sur un raisonnement au niveau de l'ARNm. En théorie en cas de saut d'exon, si le cadre de lecture est préservé, la protéine produite doit être fonctionnelle. Cependant, le gène de la dystrophine est organisé d'une manière extrêmement complexe. Par exemple, une répétition du domaine central ne correspond tout simplement pas à un exon (Figure 31). Cette stratégie ne prend pas en compte la structure de la protéine. En 2015, Nicolas *et al.* [180] s'intéressent à l'effet de certaines délétions d'exon sur la structure de la protéine qui en découle et étudient le fragment R16-21 et quatre de ses variants (Δ45–47, Δ45–48, Δ45–49 et Δ45–51) (Figure 36).

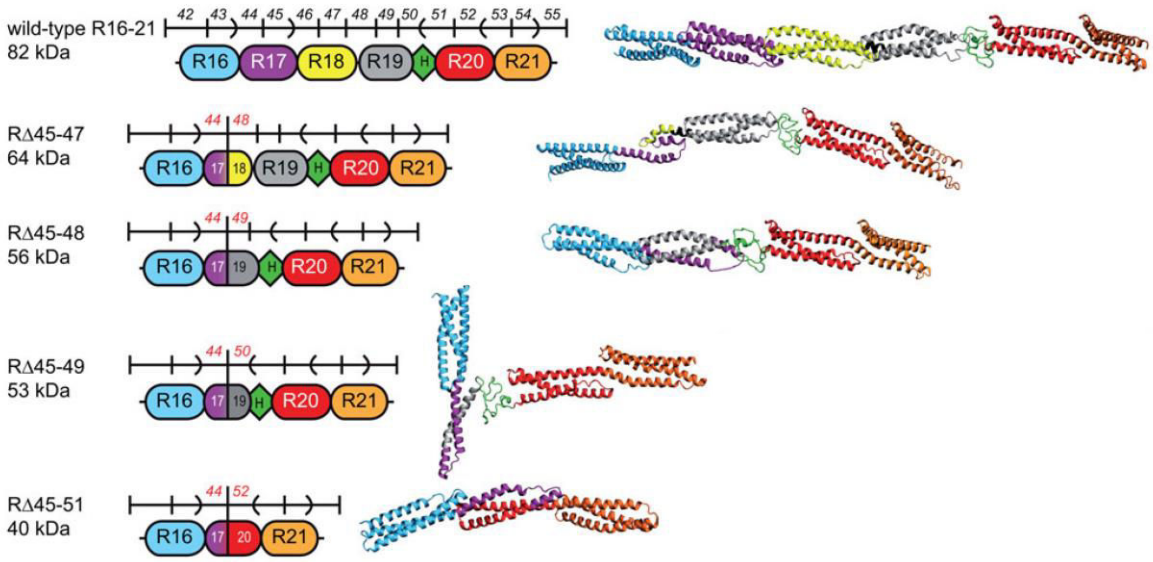


Figure 36 : Conséquences structurales engendrées par des délétions d'exon. Les modèles ont été obtenus par homologie puis simulés par dynamique moléculaire [180].

Ils démontrent ainsi que les délétions d'exons Δ45–48 et Δ45–51 conduisent à la production de protéines possédant une structure similaire à celle de leur équivalent sauvage (répétitions hybrides ou « hybrid repeat »). *A contrario*, les délétions Δ45–47 et Δ45–49 engendrent la production de protéines possédant une structure différentes (répétitions fractionnées ou « fractional repeat »). Ces travaux apportent des évidences que la sévérité des phénotypes des patients BMD est liée à la structure de la dystrophine. De plus, cette étude soulève un point extrêmement important dans le cadre du design de thérapies géniques par saut d'exon. S'assurer du repliement correct des protéines raccourcies apparait comme un point clé dans le succès de ces thérapies.

3.4.Partenaires protéiques majeurs

La dystrophine est une grande protéine filamenteuse possédant de nombreux partenaires identifiés. Elle peut être considérée comme une plateforme d'interactions (Figure 37).

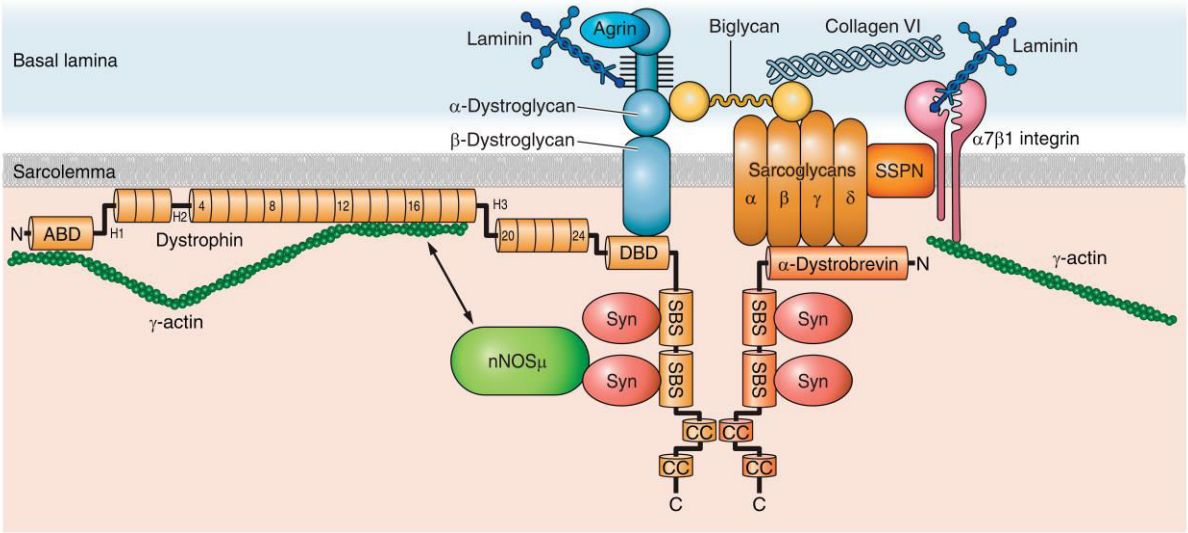


Figure 37 : Représentation schématique de la dystrophine dans le contexte cellulaire avec ses partenaires. Syn, syntrophine ; SSPN, sarcospane ; ABD, domaine de liaison à l'actine ; DBD, domaine de liaison au dystroglycane ; SBS, site de liaison à la syntrophine [181].

➢ ***Actine***

Chez les vertébrés, trois d'isoformes principales d'actine ont été identifiées, α, β et γ. Les isoformes β et γ coexistent dans la plupart des types de cellules en tant que composants du cytosquelette et sont impliquées dans les processus de motilité cellulaire interne. Les α-actines se retrouvent, quant à elles, dans les tissus musculaires et interviennent dans le mécanisme de contraction des fibres musculaires grâce aux interactions qu'elles effectuent avec la myosine. En présence de cations Mg^{2+} et d'ATP, les monomères d'actine globulaire (actine-G) s'associent entre elles pour former des microfilaments (actine-F). L'actine-F est structurée en double hélice droite avec un pas de 37 nm formé par 13 monomères d'actine-G par tour, pour une épaisseur de filament d'environ 7-9 nm. Ce filament est polarisé, c'est-à-dire que la polymérisation est directionnelle. On distingue l'extrémité barbée, siège de la polymérisation, de l'extrémité pointue où a lieu la dépolymérisation. Par analogie avec la concentration micellaire critique, la concentration critique de monomère en équilibre G/F est de l'ordre de 0.1-0.2 µM. C'est par sa région N-terminale (ABD1) et *via* les répétions R11-17 [182] et R11-15 [183] (ABD2) (Figure 38) que la dystrophine établit des interactions avec l'actine-F du cytosquelette (isoforme γ) qui contribuent à son rôle de charpente du sarcolemme (Figure 37).

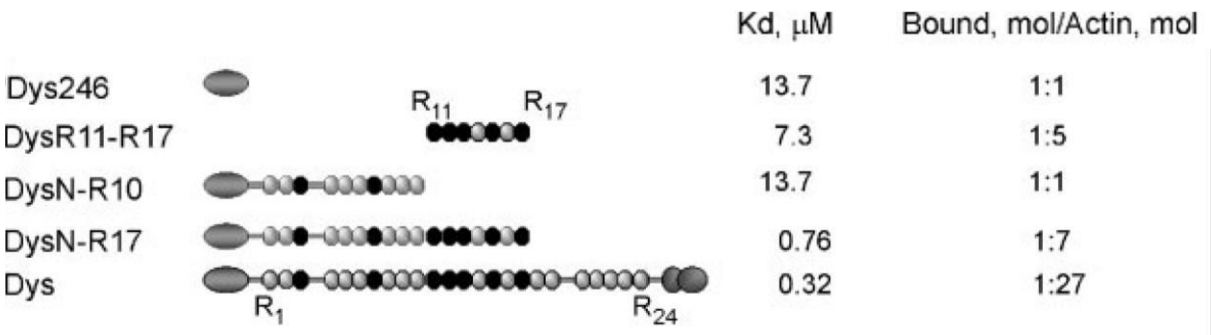


Figure 38 : Interactions dystrophine-actine. Constante de dissociation et stœchiométrie des complexes dystrophine/actine. Séparément, ABD1 (Dys246) et ABD2 ont une constante de dissociation de l'ordre de la dizaine de micromolaire. Lorsque les deux sites de liaison de la dystrophine à l'actine sont présents, l'affinité de la dystrophine pour l'actine est accrue avec une valeur de K_d en-dessous du micromolaire. Adapté de [182].

- ***nNOS***

Les synthases de l'oxyde nitrique (NOS) sont des enzymes impliquées dans différentes voies de signalisation [184]. Dans les cellules musculaires, l'isoforme neuronale (nNOS) est impliquée dans l'interaction avec la dystrophine [185] *via* les répétitions R16 et R17 et plus particulièrement la région « linker » entre les deux répétitions [186]. Dans les tissus musculaires, le produit de synthèse de nNOS, l'oxyde nitrique, induit une vasodilatation locale qui est essentielle pour l'oxygénation normale des muscles [187].

- ***Microtubules***

Tout comme les microfilaments d'actine, les microtubules font partie du cytosquelette. Ils sont composés d'hétérodimères de tubulines α et β qui polymérisent grâce à l'hydrolyse du GTP et s'organisent d'abord en protofilaments qui ensuite s'associent en 13 sous-unités pour former des microtubules. Les microtubules sont des tubes creux ayant un diamètre d'environ 24 nm. Les interactions dystrophine/microtubule se font par l'intermédiaire des répétitions R20-23 avec une affinité plutôt élevée, le K_d étant de l'ordre de 0.5 µM [188]. Comme dans le cas des interactions dystrophine/actine, celles effectuées avec les microtubules contribuent au rôle de charpente de la membrane.

- ***Le complexe dystrophine et glycoprotéines associées (DGC)***

La dystrophine est associée à la membrane par l'intermédiaire de protéines membranaires intégrales et périphériques glycosylées ou non-glycosylées, le DGC [189]. Ce

macrocomplexe inclut l'α et le β-dystroglycane, l'α-dystrobrévine, la syntrophine, la famille des sarcoglycanes (α-, β-, γ- et δ) et finalement le sarcospane. C'est la région C-terminale de la dystrophine qui est impliquée dans la liaison aux protéines du DGC. Le DGC aurait pour fonction de relier le cytosquelette intracellulaire à la matrice extracellulaire environnante et la perte de ces interactions pourrait rendre le sarcolemme plus sujet à être endommagé lors de son exposition aux stress mécaniques lors du processus de contraction/élongation musculaire.

3.5. Interaction avec les lipides membranaires

En plus de ses nombreux partenaires protéiques, la dystrophine est également capable d'interagir avec les lipides membranaires. Des travaux pionniers de De Wolf *et al.* [190] ont montré que la répétition R2 du domaine central de la dystrophine possède une forte affinité pour des phospholipides anioniques (mélange DOPC/DOPS). Ces mêmes auteurs suggéraient déjà que l'ensemble du domaine central de la dystrophine pourrait avoir une forte interaction vis-à-vis de la membrane plasmique. En 2003, Le Rumeur *et al.* [191] approfondissent cette étude et démontrent l'implication des résidus tryptophanes lors des interactions protéine/lipide. Quelques années plus tard, Le Rumeur *et al.* [192] mettent en évidence cette fois-ci que la liaison aux lipides est dépendante du « packing » des phospholipides et prouvent que cette liaison est à la fois due à des forces électrostatiques mais également hydrophobes. Ensuite, Legardinier et *al.*[193] s'intéressent à des fragments plus longs du domaine central de la dystrophine, les fragments R1-3 et R20-24, ces fragments étant bordés par des régions charnières H1-H2 et H3-H4, respectivement. Dans cette étude, il a été mis en évidence que le fragment R1-3 interagit fortement avec les phospholipides anioniques alors que R20-24 ne présente aucune interaction pour les phospholipides, qu'ils soient de types zwitterioniques ou anioniques. Legardinier et *al.* [194] étendent la caractérisation des interactions dystrophine/lipide aux répétitions R4 à R19 du domaine central de la dystrophine en utilisant des fragments de 3, 4 ou 5 répétitions successives se chevauchant plus ou moins (R4-6, R7-9, R12-14, R11-15, R14-17 et R17-19). Aucun fragment court contenant la répétition R10 n'a été obtenu, à cause de la grande instabilité des protéines recombinantes possédant cette partie. Cette restriction mise à part, ces travaux mettent en évidence que, tout comme R1-3, les répétitions R4 à R19 possèdent une force affinité pour des phospholipides anioniques (Figure 39). Ainsi, deux domaines de liaisons aux phospholipides ont été définis, LBD1 et LBD2 (« Lipid Binding Domains 1 et 2 ») (Figures 32 et 37).

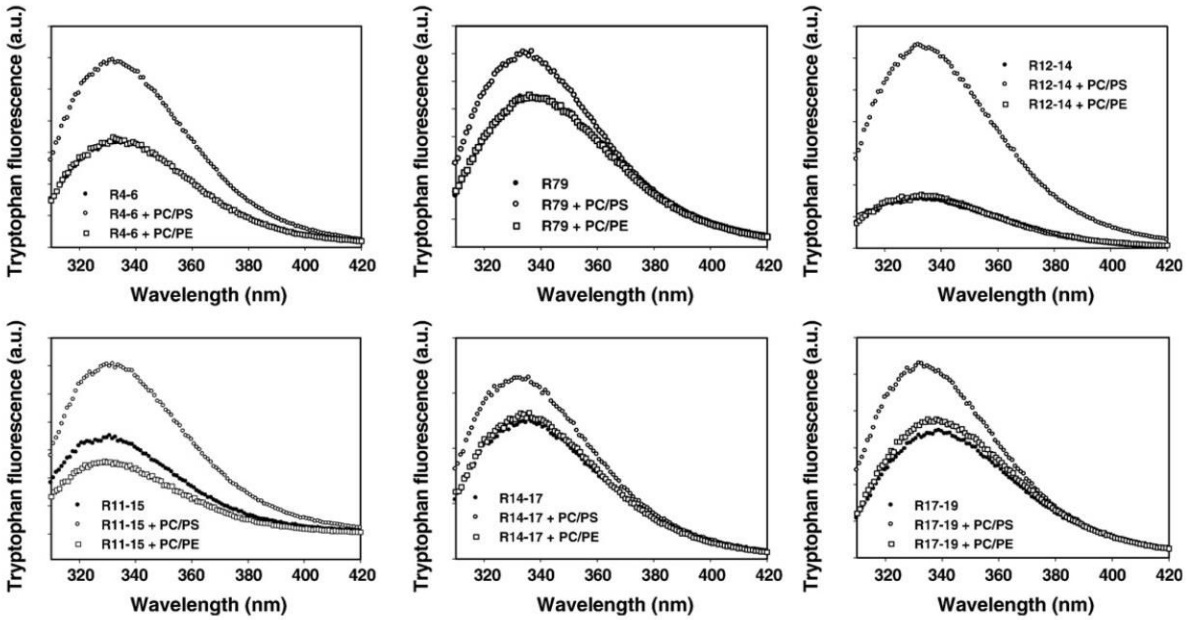


Figure 39 : Mise en évidence des interactions dystrophine/lipides. Variation de fluorescence des résidus tryptophanes des fragments de dystrophine seuls (cercles noirs) ou lors de leur liaison aux liposomes anioniques (DOPC/DOPS, cercles blancs) ou zwitterioniques (DOPS/DOPE, carrés blancs) [194].

Plus tard, Sarkis *et al.* [195] se focalisent sur le fragment R11-15 et montrent que ce dernier est capable de se lier préférentiellement à des liposomes ayant un diamètre relativement faible et donc un fort degré de courbure. En effet, l'interaction protéine/lipide n'a lieu qu'avec des SUV de 50 nm de diamètre et non avec des LUV de 100, 200 ou 400 nm de diamètre. De plus, en utilisant à la fois le modèle des SUV et des monocouches de Langmuir à diverses tensions de surface initiales, il a été mis en évidence que les répétitions R11-15 sont capables de lier à la fois des phospholipides zwitterioniques (DOPC/DOPE) et anioniques (DOPC/DOPS), [195]. Cependant, l'interaction avec les SUV anioniques demeure bien plus forte [195]. Par la suite, Sarkis *et al.* [183] s'interrogent quant à la possibilité du fragment R11-15 d'interagir à la fois avec les lipides membranaires et avec l'actine du cytosquelette (ABD2). Grâce à des études de rhéologie réalisées sur des monocouches de lipides, il a été montré que R11-15 est capable d'interagir avec ses partenaires lipidiques et protéiques simultanément et donc potentiellement *in vivo* de créer un lien entre le sarcolemme et le cytosquelette. Ceci implique que ce sous-domaine assurerait un rôle essentiel dans le maintien de l'intégrité de la membrane plasmique lors du processus de contraction/élongation du muscle. Finalement, Ameziane-Le Hir *et al.* [196] s'intéressent à l'interaction du fragment R16-21 avec des monocouches de phospholipides zwitterioniques ou anioniques dans lesquelles du cholestérol a été incorporé. Ce fragment étant

porteur de deux motifs CRAC (« cholesterol recognition/interaction amino acid consensus sequence ») qui sembleraient nécessaires à la reconnaissance du cholestérol [197]. Dans cette étude, il est montré que R16-21 est plus fortement inséré dans la monocouche dans des conditions de tension de surface physiologique (30 mN/m). Ainsi, il est suggéré qu'*in vivo*, l'ancrage de la dystrophine au feuillet interne du sarcolemme pourrait être assuré par ce sous-domaine *via* l'interaction avec des nanodomaines riches en cholestérol. L'équipe poursuit l'étude sur le fragment R16-21 sous forme native et le compare avec des variants protéiques issus de délétions d'exon : RΔ45−47, RΔ45−48, RΔ45−49, and RΔ45−51 [198]. Ces variants ont été choisis car la région d'ADN du gène *DMD* couvrant les exons 45 à 53 est la plus sujette à des délétions (« hot spot ») [149]. Les auteurs démontrent que les variants possèdent une meilleure affinité que la forme sauvage pour les membranes, ce qui implique que les mutations perturbent le comportement moléculaire de la protéine vis-à-vis du sarcolemme [198]. Un mode d'interaction différent des variants par rapport à la protéine native pourrait être une cause de la variabilité phénotypique observée chez les patients BMD.

En résumé, depuis 2003, l'équipe SIM a largement contribué à l'étude des interactions protéine/lipide impliquant le domaine central de la dystrophine et le sarcolemme. Lors de ces études menées *in vitro*, l'équipe a caractérisé ces interactions par une multitude de techniques et de méthodes en employant divers modèles membranaires. Suite à ces recherches, il découle que les interactions dystrophine/membrane sont régies par une multitude de facteurs. En effet, elles sont modulées à la fois par la région de la protéine impliquée, le « packing » membranaire, ainsi que la nature des phospholipides. Récemment, Zhao *et al.* [199] ont confirmé *in vivo*, tout du moins en partie, ces résultats. Dans cette étude, les auteurs utilisent des modèles murins et canins dans lesquels des fragments de dystrophine couplés à de la GFP ont été insérés dans le tissu musculaire *via* l'utilisation d'AAV. Ils montrent que les sous-domaines R1-3, R10-12, H4-CR et CT sont localisés au niveau de la membrane plasmique (Figure 40) par microscopie à fluorescence et Western blotting. Ces résultats viennent étayer ceux précédemment obtenus par l'équipe *in vitro*. En effet, les interactions du fragment R1-3 avec la membrane étaient déjà bien caractérisées, tout comme celles du fragment R11-15 qui contient deux des trois répétitions (R11 et R12) identifiées par Zhao et *al.* comme étant un sous-domaine liant les lipides *in vivo*.

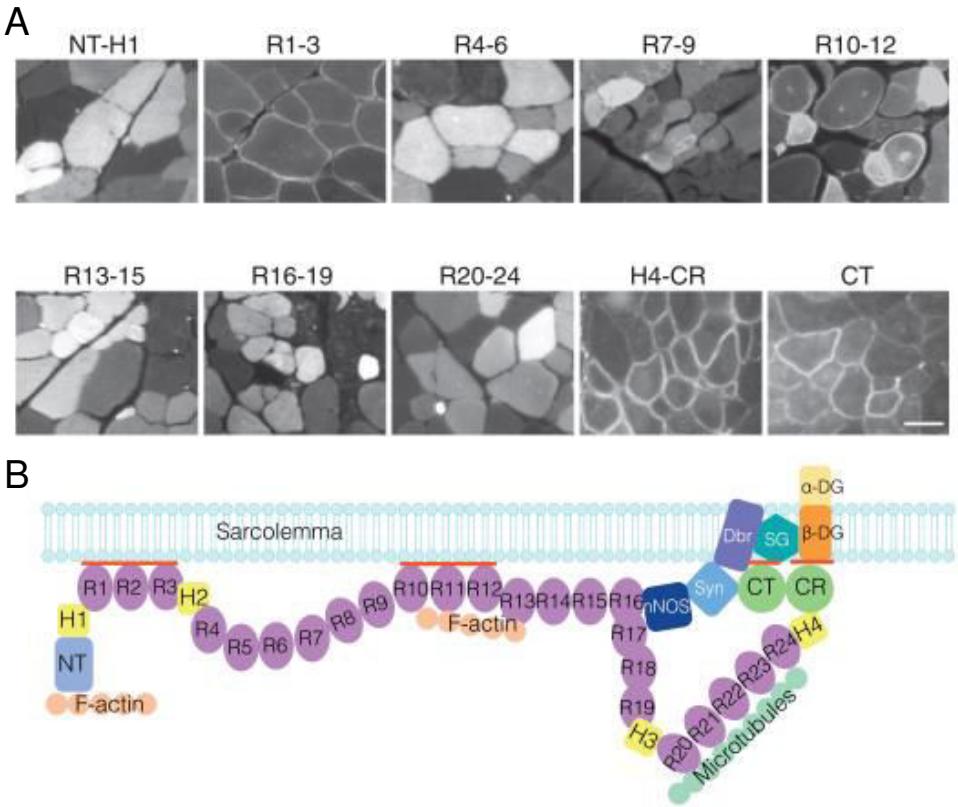


Figure 40 : Evidence *in vivo* d'interactions de certains fragments du domaine central de la dystrophine avec le sarcolemme. (A) Protéines de fusion de la dystrophine humaine, divisée en dix sous-domaines fusionnés à une étiquette GFP, produites individuellement dans le muscle de souris *mdx* (« X-linked muscular dystrophy mouse », un modèle murin pour étudier les DMD) par transfert de gène *via* des AAV. Les sous-domaines R1-3, H4-CR et CT sont exclusivement localisés au sarcolemme alors que R10-12 a également été trouvé dans le cytoplasme. NT : domaine N-terminal, H : « hinge » ou charnière, CR : domaine riche en cystéine, CT : domaine C-terminal. (B) Modèle des interactions dystrophine/sarcolemme mis à jour, où la dystrophine se lie à la membrane plasmique via quatre sous-domaines indépendants (R1-3, R10-12, H4-CR et CT). Adapté de Zhao et al.[199].

En prenant en compte tous les résultats précédents, il est difficile d'obtenir une adéquation parfaite en les travaux effectués *in vitro* et ceux réalisés *in vivo*. Zhao *et al.* ont identifié qu'uniquement R1-3 et R10-12 étaient des domaines de liaisons aux lipides membranaires alors que l'équipe avait identifié une zone allant de R1 à R19. Des études supplémentaires sont nécessaires pour comprendre les discordances entre les résultats obtenus par Zhao *et al.* et ceux obtenus par l'équipe SIM. Cependant, **un accord sur R1-3, et partiellement sur R11-15, conforte la pertinence du choix de ces deux fragments pour ces travaux de thèse.**

PARTIE II :
MATERIEL ET METHODES

Cette deuxième partie du manuscrit a pour but d'apporter des informations complémentaires concernant les techniques et la méthodologie employées lors de ces travaux de thèse, les parties expérimentales étant parfois décrites de manière succincte dans les articles présentés en partie III.

1. Production, extraction et purification des protéines recombinantes

1.1. Généralités

Les protéines étudiées au cours de cette thèse appartiennent au domaine central de la dystrophine humaine (dp427m, uniprot : P11532), qui est composé de 24 répétitions de type spectrine (R1 à R24). Les fragments R1-3 et R11-15 (selon l'alignement de Winder *et al.* [200]) sont produits fusionnés à une étiquette GST (glutathion S-transférase) ou histidine afin de faciliter leur purification. Au début de la thèse, nous disposions uniquement de la construction permettant la production de protéines recombinantes portant une étiquette GST. Au début de l'année 2016, nous avons développé une nouvelle construction permettant la production de protéines avec une étiquette histidine. Les rendements de purification ont été multipliés par un facteur cinq environ.

Tableau 5 : Caractéristiques des protéines étudiées. MW : masse moléculaire ; ε : coefficient d'extinction molaire. Les résidus en gras et en italique (***GS***) sont des résiduels du site de clivage à la thrombine (LVPR/GS) situé entre l'étiquette et la protéine d'intérêt. Les numéros en exposant correspondent à la numérotation de la structure primaire de la dystrophine.

Nom	Séquence	MW (kDa)	ε ($M^{-1}.cm^{-1}$)	Nombre de résidus	Système de production
R1-3	***GS***^{338}EVNLD…QISQA^{668}	38.5	59720	333	*E. coli*
R11-15	***GS***^{1461}FQKPA…LNFAQ^{1973}	60	45950	515	*E. coli*

1.2. Culture et lyse bactérienne

Un clone de la souche bactérienne *Escherichia coli* BL21 (DE3) contenant le plasmide d'intérêt est ensemencé dans 50 mL de milieu de culture LB (Luria-Bertani) avec une concentration finale en antibiotique de 50 µg/mL. Les pré-cultures d'une nuit servent à ensemencer une culture d'1 L, à 37°C. La levée de la répression catabolique s'effectue par ajout d'1 mM d'isopropyl β-D-1-thiogalactopyranoside, un analogue non métabolisable de l'allolactose, lorsque l'absorbance dépasse 0,5 UA. Après 3-4 h de culture, les culots bactériens sont récoltés par centrifugation et congelés à -20°C. Le culot issu d'1 L de culture bactérienne est repris dans 20 mL de tampon PNI (Na phosphate 50 mM – pH 8, NaCl 300 mM, Imidazole 5 mM) ou TN (Tris 20 mM – pH 7,5, NaCl 150 mM) selon le type de chromatographie d'affinité envisagé. Ces tampons sont supplémentés avec de la DNAse (40 µg/mL), du lysozyme (0,5 mg/mL) et un cocktail d'anti-protéases (Roche). La suspension est soumise à des ultrasons (30 s, 50%, P4 ; Sonifier W450, Branson). Le lysat est ensuite centrifugé à 20 000 g pendant 45 min, filtré (0.22 µm) et stocké à 4°C.

1.3. Purification

1.3.1. Chromatographie d'affinité

Deux types de chromatographies d'affinité sont effectuées selon que le fragment protéique porte une étiquette GST ou histidine.

- ***Chromatographie sur complexe GST-GSH***

Les protéines d'intérêt possèdent l'étiquette GST du côté N-terminal, séparée par un site de clivage à la thrombine. La purification est faite grâce à une chromatographie d'affinité sur une résine contenant du glutathion grâce au système Akta Start (GE Healthcare). Le surnageant est filtré à 0,2 µm puis chargé à faible débit (0,5 mL/min) sur une colonne GSTrap (20 mL, GE Healthcare), préalablement équilibrée avec du tampon TN. La colonne est lavée abondamment à débit élevé (4 mL/min) jusqu'à obtenir une absorbance (A_{280nm}) inférieure à 0,02. La protéine est éluée dans du TNE (TN supplémenté avec 0.1 mM d'EDTA ; acide éthylène diamine tétra-acétique) après clivage enzymatique avec de la thrombine directement sur colonne.

- ***Chromatographie d'affinité sur complexe Ni2+ (IMAC)***

Les constructions plasmidiques utilisées permettent la production de la protéine d'intérêt avec une étiquette 6-His en N-terminal, séparées par un site de clivage à la thrombine. La purification est réalisée par chromatographie d'affinité sur une résine de billes d'agarose

couplée à des ions Ni^{2+}, pilotée par un système Akta Start (GE Healthcare). Le surnageant de sonification est filtré à 0,2 µm puis chargé, à faible débit (0,5 mL/min), sur une colonne HisTrap (25 mL, GE Healthcare) préalablement équilibrée avec du tampon PNI. La colonne est lavée abondamment à débit élevé (4 mL/min) jusqu'à obtenir A_{280nm} inférieure à 0,02. La digestion thrombique de l'étiquette se fait directement sur colonne, la nuit à 4°C. Les protéines ne portant plus l'étiquette sont éluées dans du PNI et celles portant encore l'étiquette et liées à la colonne sont éluées par compétition avec du tampon PNI contenant 500 mM d'imidazole. Dans ce dernier cas, la digestion thrombique a lieu ultérieurement.

1.3.2. Chromatographie d'exclusion stérique (SEC)

Toutes les étapes qui suivent sont réalisées à l'aide du système Akta FPLC (GE Healthcare). Quelle que soit la chromatographie d'affinité réalisée, une fraction de l'étape de chromatographie d'affinité est chargée sur colonne d'exclusion stérique HiLoad 16/600 Superdex 200 prep grade (120 mL, GE Healthcare) préalablement équilibrée dans du tampon TNE et collectée par fractions de 2 mL.

1.3.3. Quantification et concentration

La concentration en protéine est déterminée par spectrophotométrie à 280 nm selon la loi de Beer-Lambert. Au besoin, les échantillons sont concentrés en utilisant des concentrateurs (Amicon Ultra-15, 4 ou 0,5, MWCO 10000 ou 3000 kDa, Millipore). Les rendements de purification sont de l'ordre de 3-5 mg/L de culture bactérienne lors de l'utilisation d'une étiquette GST et ils sont supérieurs à 25 mg/L de culture bactérienne en utilisant une étiquette histidine.

1.3.4. Analyse SDS-PAGE

L'électrophorèse en conditions dénaturantes (SDS-PAGE) [201] permet, entre autres, de suivre l'évolution du degré de pureté des protéines au cours de chaque étape de purification. Les protéines sont séparées selon leur taille dans un gel à 4-12% en polyacryamide (SDS-PAGE Precast gels, Expedeon), puis révélées au bleu de Coomassie (InstantBlue, Expedeon). Le degré de pureté des protéines est estimé par densitométrie à l'aide du logiciel ImageQuant TL (GE Healthcare).

1.3.5. Immuno-empreinte (Western-blotting)

Pour les protéines non digérées sur colonne, le retrait de l'étiquette est vérifié par immuno-empreinte. Après SDS-PAGE, un transfert sur membrane de nitrocellulose est réalisé puis les

protéines sont colorées au rouge Ponceau. La membrane est saturée (Tris-buffer-Saline + 5% de lait en poudre) pendant une nuit à 4°C, puis incubée avec l'anticorps primaire anti-polyhistidine (1/1000e, Sigma) pendant 1 h. Après lavages et incubation avec un anticorps secondaire anti-souris (1/5000ème) couplé à de la peroxydase, la révélation est réalisée par chimiluminescence (ImageQuant LAS 4000, GE Healthcare).

2. Préparation des modèles membranaires

2.1. Phospholipides

Les lipides utilisés (Avanti Polar lipids) sont conditionnés dans du chloroforme ou dans un mélange chloroforme/méthanol et sont utilisés sans aucune autre étape de purification. Les mélanges lipidiques sont préparés dans du $CHCl_3$ (ou $CHCl_3$/MeOH) qui est évaporé 12 heures sous vide. Les lipides sont ensuite réhydratés dans du tampon TNE ou dans du tampon d-TNE (Tris d-11 20 mM, NaCl 150 mM, EDTA d-16 0.1 mM, 100% D_2O, pD 7.1) pour le besoin des expériences de RMN et/ou de SANS.

Lipides utilisés :

- DMPC (1,2-dimyristoyl-sn-glycéro-3-phosphocholine, 14:0 PC)
- DHPC (1,2-dihexanoyl-sn-glycéro-3-phosphocholine, 6:0 PC)
- DMPS (1,2-dimyristoyl-sn-glycéro-3-phospho-L-sérine, 14:0 PS)
- DMPC d-67 (1,2-dimyristoyl-d54-sn-glycéro-3-phosphocholine-1,1,2,2-d4-N,N,N-trimethyl-d9)
- DHPC d-35 (1,2-dihexanoyl-d22-sn-glycéro-3-phosphocholine-1,1,2,2-d4-N,N,N-trimethyl-d9)
- DMPS d-54 (1,2-dimyristoyl-d54-sn-glycéro-3-phospho-L-sérine).

2.2. Les liposomes de DMPC

Les lipides secs sont réhydratés dans du tampon TNE et forment des MLV. Les SUV sont obtenus par extrusion en utilisant un « mini extruder » (Avanti Polar Lipids) et des filtres de 50 nm. Au moins 11 allers-retours sont effectués pour obtenir une suspension claire.

2.3.Les micelles de DHPC

Les lipides secs sont simplement repris dans un volume approprié de tampon TNE afin d'obtenir une concentration supérieure à la concentration micellaire critique du DHPC (15 mM), ce qui induit la formation spontanée de micelles.

2.4.Les bicelles

Deux types de bicelles sont préparés : des bicelles zwitterioniques (DMPC/DHPC, 1/1 mol/mol) ou anioniques (DMPC/DMPS/DHPC, 0,67/0,33/1 mol/mol). Les lipides secs sont repris dans du TNE pour une concentration finale d'au moins 200 mM de lipides. Les mélanges lipidiques réhydratés sont ensuite refroidis dans de l'azote liquide (10 s), chauffés à 40°C (10 min), agités au vortex (1 min) et centrifugés (1 min, 6000 rpm). Cette procédure est répétée trois fois pour l'obtention d'une solution claire et homogène de bicelles.

Les phospholipides à chaînes courtes, DHPC dans notre cas, sont en équilibre rapide entre la forme liée et la forme libre. Le ratio mol/mol des deux constituants est défini par q et q_{eff} où :

$$q = \frac{[DMPC\ ou\ DMPS]}{[DHPC]} \qquad \text{Eq 12}$$

$$q_{eff} = \frac{[DMPC\ ou\ DMPS]}{[DHPC]_{totale} - [DHPC]_{libre}} \qquad \text{Eq 13}$$

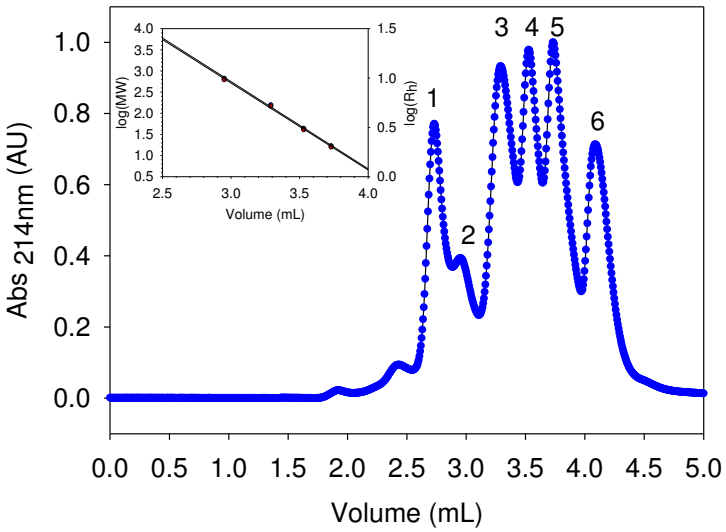
Ainsi, nous prenons en compte la concentration bicellaire critique ($[DHPC]_{libre}$ = CBC = 6 mM) [59], par analogie avec la concentration micellaire critique dans le cas de micelles (CMC = 15 mM). A concentration élevée en lipides, la proportion de DHPC libre est négligeable mais, dans notre cas, les bicelles sont utilisées à une concentration de 50 mM de lipides pour une valeur de q = 1 et ainsi q_{eff} est égal à 1,3. (voir partie I, chapitre 1 pour plus de détails).

3. Caractérisation biophysique des objets d'étude

3.1.Estimation du R_h et de la MW apparente par SEC

La chromatographie d'exclusion stérique est basée sur la séparation des objets selon leur rayon hydrodynamique (R_h). Grâce à la calibration de la colonne avec des protéines de R_h et de masses molaires (MW) connues, il est possible d'estimer ces deux paramètres pour les objets

d'étude, que ce soit pour les protéines ou pour les bicelles (Figure 41). Cependant, il s'agit d'une méthode indirecte.



	Nom	Volume d'élution	MW (kDa)	R_h (nm)
1	Agrégats	2.73	-	-
2	Thyroglobuline	2.95	670	9.6
3	γ-Globuline	3.29	158	5.2
4	Ovalbumine	3.53	44	3.0
5	Myoglobine	3.73	17	2.0
6	Vitamine B12	4.09	-	-

Figure 41 : Chromatogramme HPSEC obtenu avec un mélange de calibrants (BioRad) permettant l'estimation de la MW apparente et de R_h pour les fragments protéiques utilisés lors de cette étude. Insert : Log(MW) ou Log(R_h) = f(volume d'élution).

L'estimation de la MW par ce moyen implique que la protéine d'intérêt possède une structure du même type que les calibrants, c'est-à-dire ici globulaire. Cette approche appliquée aux fragments du domaine central de la dystrophine, possédant une structure tridimensionnelle allongée, est utilisée uniquement d'un point de vue comparatif et la MW obtenue est considérée comme apparente.

3.2.Diffusion statique de la lumière (SLS) multi-angles (MALS)

Contrairement à la précédente technique, la SLS permet de déterminer directement la MW d'un objet, quelle que soit sa forme, selon l'équation de Rayleigh :

$$\frac{KC}{R_\theta} = \left(\frac{1}{MW} + 2A_2C\right)\frac{1}{P_\theta} \qquad \text{Eq 14}$$

où C est la concentration, R_θ le ratio de Rayleigh, A_2 le second coefficient du viriel, MW la masse molaire, K une constante optique et P_θ le facteur de forme. Ces deux derniers paramètres sont définis par :

$$K = \frac{4\pi^2 {n_0}^2}{N_A \lambda_0^4} \left(\frac{dn}{dc}\right)^2 \qquad \text{Eq 15}$$

et

$$P_\theta = 1 + \frac{16\pi^2 n_0^2 R_g^2}{3\lambda_0^2} \sin^2\left(\frac{\theta}{2}\right) \qquad \text{Eq 16}$$

où λ_0 est la longueur d'onde du laser, N_A le nombre d'Avogadro, n_0 l'indice de réfraction du solvant, *dn/dc* l'incrément d'indice de réfraction, R_g le rayon de giration et θ l'angle de mesure. Dans des conditions diluées, ce qui est le cas des expériences de chromatographie d'exclusion stérique couplées à la diffusion multi-angle (SEC-MALS), et pour $P_\theta = 1$ (si $\theta = 0°$), l'Eq. 14 devient :

$$R_\theta = K.C.MW \qquad \text{Eq 17}$$

Les mesures ont été réalisées sur des dispositifs HPLC (Agilent) couplés avec des détecteurs de diffusion de lumière (miniDAWN TREOS ou DAWN HELEOS II, Wyatt). La calibration est réalisée avec du toluène et la normalisation avec de la BSA (Figure 42).

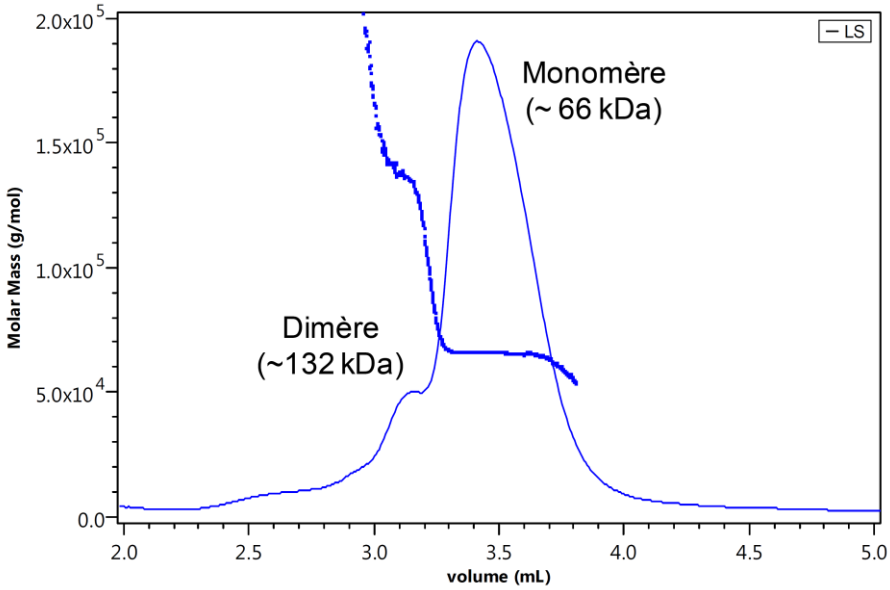


Figure 42 : Exemple d'une expérience SEC-MALS avec de la BSA. Les masses des formes monomériques et dimériques sont aisément distinguables.

3.3.Diffusion dynamique de la lumière (DLS)

Cette technique permet d'estimer la taille des particules présentes dans une suspension en déterminant leur rayon hydrodynamique (R_h). La lumière d'un laser illumine la molécule

d'intérêt et certains photons sont diffusés. L'intensité diffusée est mesurée en fonction du temps. Plus une molécule est de taille élevée, plus la lumière est diffusée intensément mais peu souvent et inversement pour une petite molécule. Un traitement mathématique d'auto-corrélation du signal est ensuite effectué afin de comparer le signal à lui-même sur de courts intervalles de temps. La vitesse de perte de signal (Γ) de la fonction d'auto-corrélation est reliée au coefficient de diffusion (D) de la particule par :

$$\Gamma = DQ^2 \qquad \text{Eq 18}$$

où Q est le vecteur de diffusion et est défini par :

$$Q = \frac{4\pi n \sin\theta}{\lambda} \qquad \text{Eq 19}$$

où n est l'indice de réfraction du milieu, λ la longueur d'onde du laser et 2θ l'angle de diffusion. R_h est déterminé par l'équation de Stokes-Einstein :

$$R_h = \frac{k_B T}{6\pi\eta D} \qquad \text{Eq 20}$$

où k_B est la constante de Boltzmann, T la température et η la viscosité du milieu.

3.4.Spectroscopie RMN ^{1}H et ^{31}P

Les expériences ont été réalisées sur la plateforme PRISM (Rennes) en collaboration avec Arnaud Bondon et Liza Mouret. La RMN est basée sur la propriété qu'ont certains noyaux atomiques d'absorber les ondes radiofréquences. Chaque type de noyau (^{1}H, ^{13}C, ^{15}N, ^{31}P, ...) absorbe ces ondes à une fréquence particulière dite « fréquence de résonance ». L'échantillon est inséré dans un aimant et est soumis à un champ magnétique B_0 très intense. Il en résulte une aimantation macroscopique, parallèle à B_0. L'application d'un champ radiofréquence, perpendiculaire à B_0 et centré sur la fréquence de résonance des noyaux d'intérêt, engendre des transitions énergétiques pour ces noyaux et, à l'échelle macroscopique, un mouvement de bascule de l'aimantation dans le plan transverse. À l'arrêt du champ radiofréquence, le retour à l'équilibre du système génère un signal électrique mesuré par l'appareil puis traduit en fréquence par transformée de Fourier. La fréquence de résonance des noyaux dépend finement de leur environnement chimique et, sur le spectre obtenu, chaque pic correspond à des noyaux situés dans un environnement différent. Lors de ces travaux de thèse, la RMN a été utilisée en tant que contrôle qualité des bicelles afin de vérifier la bonne organisation des phospholipides en bicelles et la proportion correcte des composants.

4. Caractérisation des interactions protéine/lipide

4.1. Spectrofluorimétrie : fluorescence intrinsèque des tryptophanes

Une molécule fluorescente est capable d'absorber de la lumière, ce qui entraine un passage d'un électron de l'état fondamental à un état excité. Le retour à l'état fondamental produit de l'énergie sous forme de lumière à une longueur d'onde plus grande et donc de plus basse énergie. La fluorescence intrinsèque du tryptophane est utilisée pour caractériser les interactions protéine/lipide. Lors de la liaison de la protéine aux lipides, le changement de la polarité de l'environnement des résidus tryptophanes a une influence sur leurs propriétés de fluorescence. Typiquement, le tryptophane possède un maximum d'intensité d'émission de fluorescence à 330 nm dans un environnement apolaire alors qu'il se trouve à 350 nm dans un milieu polaire. Les tryptophanes sont excités à 295 nm pour éviter la contribution des résidus phénylalanines et tyrosines. La fluorescence est recueillie sur une gamme de 310 à 450 nm. Les mesures sont réalisées sur un appareil Fluorolog-3 (Jobin-Yvon).

4.2. Dichroïsme circulaire (CD)

Cette technique est basée sur la propriété rotatoire de certaines molécules, capables d'absorber différemment la lumière polarisée circulairement gauche ou droite. Le spectre dichroïque correspond à la différence d'absorbance ΔA entre les deux composantes A_L et A_R :

$$\Delta A = A_L - A_R = (\varepsilon_L - \varepsilon_R).C.l = \Delta\varepsilon.C.l \qquad \text{Eq 21}$$

où ε_L et ε_R sont les coefficients d'extinction molaire pour la lumière polarisée circulairement gauche ou droite, C la concentration (mol L^{-1}) et l le trajet optique (cm). Le signal mesuré est l'ellipticité θ (mdeg). Pour relier l'ellipticité au dichroïsme circulaire, on fait l'approximation que ce dichroïsme est faible ($\Delta A << 1$) et on définit l'ellipticité molaire $[\theta]$ reliant θ et $\Delta\varepsilon$ selon :

$$[\theta] = \frac{\theta}{C.l} = 32{,}98 * \Delta\varepsilon \qquad \text{Eq 22}$$

Le dichroïsme circulaire est largement utilisé pour déterminer la structure secondaire des protéines [202]. Dans l'UV lointain, l'information structurale est obtenue grâce à la liaison amide formant la liaison peptidique. Cette liaison possède deux transitions électroniques : une centrée autour de 220 nm, de type $n \rightarrow \pi^*$, et une centrée autour de 190 nm, de type $\pi \rightarrow \pi^*$. La structure secondaire des protéines influence l'état d'excitation de la liaison peptidique. Dans une protéine, la présence de feuillets β implique la présence d'un minimum à 210 nm et d'un

maximum à 195 nm. Au contraire, la présence d'hélices α se traduit par deux minima, à 222 et 208 nm, et par un maximum à 193 nm. De plus, le ratio $\theta_{222\ nm}/\theta_{208\ nm}$ permet de mettre en évidence des interactions inter-hélices et donc la présence de faisceaux d'hélices (« coiled-coil ») [203]. Les mesures sont réalisées sur un appareil J-815 (Jasco).

4.3. Thermophorèse micro-échelle (MST)

La thermophorèse micro-échelle (MST) (NanoTemper Technologies) est une technique basée sur la propriété qu'ont les molécules à migrer selon, leur (i) taille, (ii) couche d'hydratation et (iii) charge, dans un gradient de température [204]. Les dispositifs de type Monolith (Figure 43A) permettent d'utiliser la MST afin de caractériser les interactions entre un récepteur et un ligand (protéine/protéine, protéine/ADN ou ARN, protéine/lipide, …). Cette technique permet notamment de déterminer une constante de dissociation K_d. Le récepteur est marqué avec une sonde fluorescente, puis est titré avec le ligand. Une expérience est typiquement réalisée dans 16 capillaires où la sonde fluorescente est excitée par une lampe et où un laser infrarouge (IR) focalisé est utilisé pour chauffer localement le capillaire (Figure 43B). Initialement, les molécules sont réparties de façon homogène et une fluorescence initiale constante est détectée. Après l'activation du laser IR, une décroissance de fluorescence est observée due au changement de température mais surtout à cause du mouvement thermophorétique des molécules marquées. Après l'extinction du laser IR, les molécules retournent vers leur position initiale (Figure 43C).

Le mouvement thermophorétique d'une molécule fluorescente (Figure 41D, courbe noire) change lors de sa liaison à un ligand non fluorescent (Figure 41D, courbe rouge), ce qui entraîne une différence dans le signal MST. Pour l'analyse des données, la modification de la thermophorèse est exprimée en changement de fluorescence normalisée (ΔFnorm), qui est définie par le ratio F_{hot}/F_{cold} (Figure 41D). La titration du ligand non fluorescent entraîne une variation graduelle de la thermophorèse. Une courbe classique de titration est obtenue en traçant ΔFnorm en fonction de la concentration en ligand Figure 41D. Cette courbe est utilisée pour déterminer la constante de dissociation K_d. Dans la partie III, nous présentons les données obtenues pour les complexes R1-3 ou R11-15/bicelles.

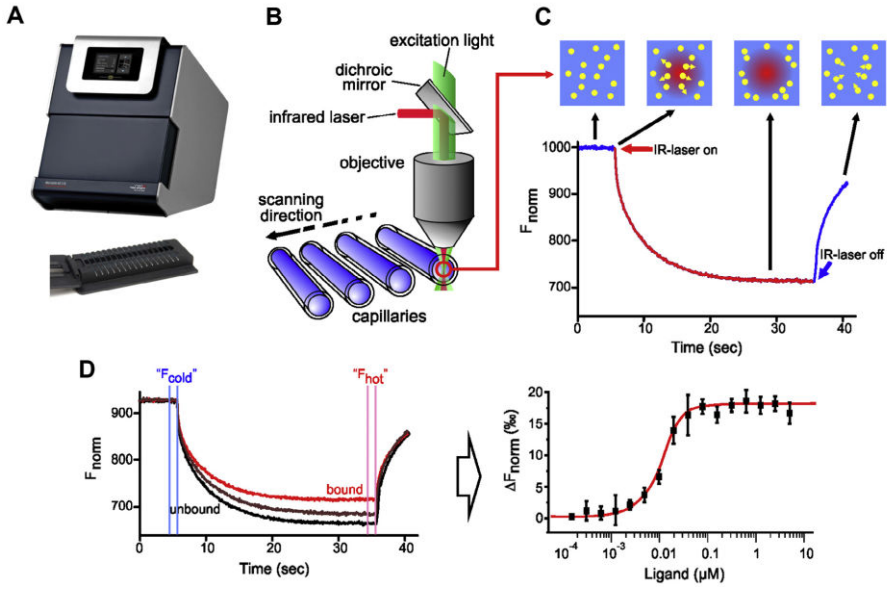


Figure 43 : Principes de la Thermophorèse (A) Appareil Monolith NT.115 de la société NanoTemper Technologies. (B) Représentation schématique du dispositif. (C) Signal typique d'une expérience MST. (D) Exemple de données obtenues lors d'une expérience MST [205].

A notre connaissance, c'est la première fois que les bicelles sont utilisées dans des expériences de MST. D'ailleurs, nous écrirons une note technique avec la société NanoTemper dès que les résultats seront publiés. Les mises au point expérimentales ont été effectuées avec le complexe R1-3/bicelle zwitterionique et la BSA a été utilisée comme contrôle négatif afin de mettre en évidence la spécificité des interactions dystrophine/bicelle.

4.4. Marquage aux billes d'or

Afin de faciliter sa visualisation par microscopie électronique, le fragment protéique R1-3 est marqué avec des nanoparticules (NP) d'or (1,4 nm monomaléimido ou monosulfo-N-hydroxy-succinimido (NHS) Nanogold, Nanoprobes) *via* ses cystéines ou lysines. Le marquage est basé soit sur la réactivité thiol/maléimide et se fait dans du tampon TNE, soit sur la réactivité amine/NHS-esters et s'effectue dans du tampon PBS. Le conjugué R1-3/NP et l'excès de NP libres sont séparés par HPSEC grâce au suivi de l'absorbance des NP à 420 nm. Le dosage protéine/nanoparticule s'effectue par spectrophotométrie à 280 et 420 nm.

4.5. Coloration négative et microscopie électronique en transmission (TEM)

Les expériences ont été réalisées sur la plateforme de microscopie MRIc (Rennes) en collaboration avec Agnès Burel. Pour les expériences de TEM, l'échantillon est déposé sur une grille recouverte d'un film de carbone (30 s). La grille est lavée deux fois avec de l'eau MQ

(30 s) et finalement mise en contact avec le contrastant (2% acétate d'uranyle) (1 min). Les observations sont réalisées avec un microscope électronique JEM 1400 (JEOL).

4.6. Cryo-micro électronique en transmission (Cryo-TEM)

Les expériences ont été réalisées en collaboration avec Marie-Sousai Appavou (JCNS at MLZ, Garching). La préparation des échantillons pour la Cryo-TEM consiste en une congélation rapide de l'échantillon dans l'éthane liquide (-180°C) afin de figer les objets à observer dans de la glace vitreuse. Les observations sont réalisées avec un microscope électronique JEM-FS2200 (JEOL).

4.7. Chimie-click et spectrométrie de masse

➢ ***Protéine/protéine***

Les fragments protéiques ABD1 ou R11-15 (ABD2) sont mis contact avec de l'actine-F et un agent de pontage (« cross-linker ») zéro distance (EDC, (1-éthyl-3-(3-diméthylaminopropyl)carbodiimide hydrochloride) pendant une heure (Figure 44).

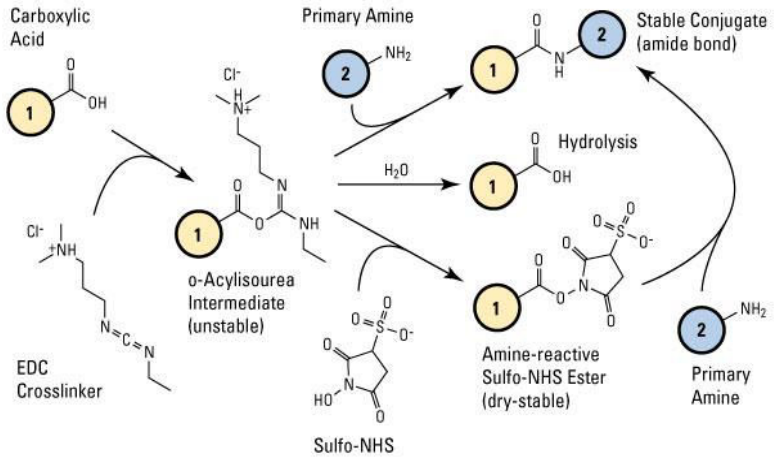


Figure 44 : Résumé de la réaction de pontage (« cross-link ») avec de l'EDC (carbodiimide). Dans le cas de deux protéines, la réaction se fait entre des groupements carboxylates de la protéine 1 (jaune) et les groupements amines de la protéine 2 (bleu). Ce sont les résidus Asp/Glu et Lys qui sont impliqués dans la réaction et forment une liaison amide entre les protéines 1 et 2 [206].

Les échantillons sont ultracentrifugés (100000 g, 30 min, 4°C). Les surnageants et les culots sont séparés puis analysés par SDS-PAGE. Les parties du gel correspondant aux échantillons liés de manière covalente et co-sédimentés sont excisées en vue d'une analyse par spectrométrie de masse.

- ***Protéine/lipide***

Les bicelles ont été marquées avec 5% (mol/mol) de 1-palmitoyl-2- (9- (3-pent-4-ynyl-3-H-diazirine-3-yl)-nonanoyl)-sn-glycéro-3-phosphocholine (pacFA PC) (Avanti Polar Lipids). Ce phospholipide modifié porte sur l'une des queues du phospholipide, à la fois un groupement diazirine photoactivable (réactivité des carbènes suite à un départ de N_2) et un groupement alcyne cliquable, par réaction alcyne-azide [207]. Les complexes protéines/bicelles sont exposés aux UVA (5 x 15 W Stratalinker, GE Healthcare) pendant 5 min. Les échantillons sont ensuite analysés par SDS-PAGE. Les parties du gel correspondant à la protéine portant le pacFA PC sont excisées en vue d'une analyse par spectrométrie de masse.

- ***Cartographie des interactions par spectrométrie de masse***

Les expériences et analyses de spectrométrie de masse ont été réalisées en collaboration avec Mélanie Lagarrigue de la plateforme de protéomique Protim (Rennes). Les pièces de gel sont découpées en petits cubes puis décolorées avec de l'éthanol (30% *v*/*v*) à chaud (70°C). Les échantillons sont déshydratés dans une solution d'acétonitrile pur. Les protéines sont ensuite réduites avec du dithiothréitol, alkylées avec de l'iodoacétamide, puis digérées à la trypsine pendant une nuit à 37°C. L'extrait peptidique est chargé sur un système nano-HPLC (LC Packings Ultimate 3000, Dionex) équipé d'une pré-colonne (5 mm x 300 µm id, 300 µg de taille de pore, Pepmap C18, 5 µm, ThermoScientific) et d'une colonne analytique (15 cm × 75 µm id, 300 Å taille des pores, Pepmap C18, 5 µm, ThermoScientific). Les peptides sont séparés en phase inverse puis élués directement de la colonne nano-HPLC vers la source d'ionisation (nanoESI) d'un spectromètre de masse LTQ-Orbitrap XL (ThermoScientific). Les produits de pontage covalent protéine/protéine sont analysés avec le logiciel StravroX [208]. Pour les interactions protéine/lipide, la modification pacFA PC est définie manuellement avec la prise en compte de la perte du groupement phosphocholine lors de la MS/MS (voir partie III, chapitre 2, article 2 pour plus de détails).

5. Caractérisation structurale des objets d'étude

5.1. Diffusion aux petits angles (small-angle scattering, SAS)

La diffusion des neutrons ou des rayons X aux petits angles (SANS et SAXS respectivement) sont des techniques de choix pour d'obtenir des informations structurales de biomacromolécules en solution, l'échelle de mesure allant de quelques angströms à plusieurs centaines de nanomètres (Figure 45A). Les photons (rayons X) sont diffusés par le cortège

électronique des atomes, tandis que les neutrons sont diffusés par les noyaux des atomes. La capacité d'un atome à diffuser les rayons X augmente avec le numéro atomique alors que celle d'un atome à diffuser un neutron varie selon son état isotopique ainsi que son spin (Figure 45B). Cette propriété s'appelle la longueur de diffusion (« scattering length » ou SL) et est exprimée en unité de longueur. Elle représente la probabilité qu'un photon ou un neutron soit diffusé par l'atome considéré. Dans le cas de la diffusion de neutrons, il existe un terme cohérent et un terme incohérent. Lors d'une expérience de SANS, c'est la diffusion cohérente qui est mesurée et qui permet d'obtenir des informations structurales alors que la diffusion incohérente, dans cette technique, contribue au bruit de fond. Il n'existe pas d'équivalent de bruit de fond dans le cas du SAXS. On définit la densité de longueur de diffusion (« scattering length density » ou SLD), qui correspond à la somme des SL des atomes composant une molécule divisée par son volume moléculaire. Elle est généralement exprimée en 10^{10} cm^{-2} ou 10^{-6} $Å^{-2}$.

$$SLD = \left(\sum_{i=1}^{n} b_{c_i}\right) / V_M \qquad \text{Eq 23}$$

où b_{ci} est la longueur de diffusion cohérente du $i^{ème}$ des n atomes d'une molécule et V_M est le volume moléculaire.

Les techniques de SAXS et de SANS permettent d'obtenir des informations à basse résolution sur des molécules en solution. Un faisceau de rayons X ou de neutrons illumine l'échantillon et certaines particules, diffusées par ce dernier, arrivent sur le détecteur. Après une intégration circulaire (dans le cas d'un échantillon isotrope), une courbe à une dimension est obtenue, donnant l'intensité diffusée en fonction du vecteur de diffusion Q :

$$Q = \frac{4\pi \sin\theta}{\lambda} \qquad \text{Eq 24}$$

où Q est le vecteur de diffusion, λ la longueur d'onde et 2θ l'angle de diffusion. Nous faisons l'hypothèse d'une diffusion élastique, c'est-à-dire qu'il n'y pas d'échange d'énergie entre le neutron incident et le noyau.

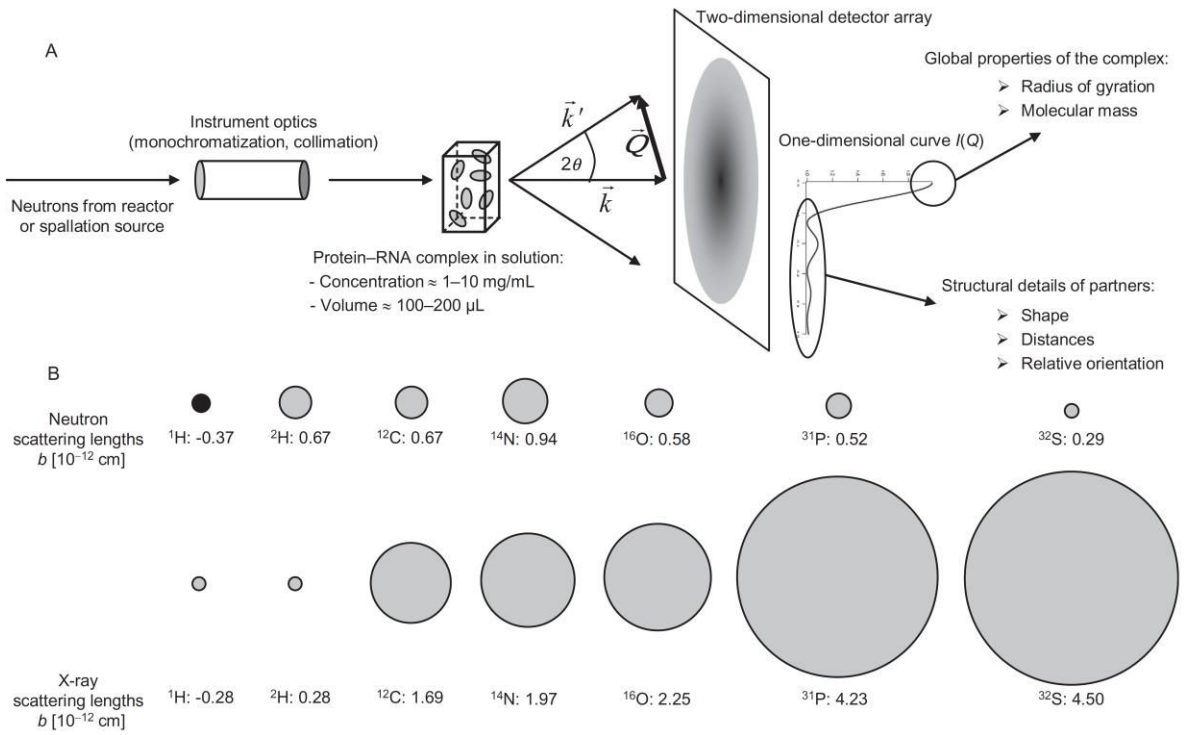


Figure 45 : (A) Représentation simplifiée d'une expérience de SANS ou de SAXS. (B) Comparaison de la capacité (« scattering length », SL, en 10^{-12} cm) de certains atomes à diffuser des neutrons ou des rayon X [209]. Les isotopes ^{1}H et ^{2}H n'ont pas la même capacité à diffuser des neutrons, contrairement aux rayons X. C'est cette propriété qui est exploitée dans les expériences de variation de contraste en diffusion de neutrons.

Pour un système homogène, l'intensité diffusée en fonction du vecteur de diffusion est définie par [210,211] :

$$I\,(Q) = \Delta\rho^2.N.V^2.P(Q).S(Q) \qquad \text{Eq 25}$$

où $\Delta\rho$ est le contraste (différence de SLD avec le solvant), N est le nombre de molécules par unité de volume, V est le volume moléculaire, $P(Q)$ est le facteur de forme et $S(Q)$ est le facteur de structure. $S(Q)$ contient des informations sur les distances entre les particules en solution. Dans un système suffisamment dilué, $S(Q) = 1$. Ainsi, dans ces conditions, l'ordre de grandeur de $I(Q)$ dépend du contraste, de la concentration et surtout de la forme globale de la macromolécule, donnée par le facteur de forme $P(Q)$. Cette hypothèse est systématiquement vérifiée de manière expérimentale avant de procéder au reste de l'analyse.

- ***La technique de variation de contraste***

La technique de variation de contraste est essentiellement basée sur la différence de capacité d'un neutron à être diffusé par un atome d'hydrogène ou par l'un de ses isotopes, le deutérium.

Cette propriété permet de masquer le signal d'un composant d'un système macromoléculaire composé de plusieurs partenaires. Dans notre étude, les partenaires sont les fragments de dystrophine et le modèle membranaire de bicelle. Pour un complexe formé de deux partenaires (1 et 2), le signal SANS peut être décrit par :

$$I\,(Q) = N\left[\Delta\rho_1^2.V_1^2.P_1(Q) + \Delta\rho_2^2.V_2^2.P_2(Q) + 2\Delta\rho_1.V_1\,\Delta\rho_2.V_2.P_{1,2}(Q)\right] \qquad \text{Eq 26}$$

Le dernier terme de l'équation correspond aux termes croisés du complexe [212]. Si le contraste du partenaire 2 avec le solvant est nul ($\Delta\rho_2 = 0$), alors le signal résultant correspond uniquement à la contribution du partenaire 1.

5.2. Analyse des données SAS

Les données SAXS ont été obtenues exclusivement sur la ligne SWING (synchrotron SOLEIL, France) et les données SANS ont été obtenues sur les instruments de SANS suivants : PACE (Laboratoire Léon-Brillouin, LLB, France) KWS-1 et KWS-2 (Heinz Maier-Leibnitz Zentrum, MLZ, Allemagne) et D22 (Institut Laue-Langevin, ILL, France). Les données SAS ont été analysées avec les logiciels de la suite ATSAS[213], développée par l'EMBL et/ou le logiciel Scatter [214].

- ***Etapes de contrôle qualité***

Pour l'acquisition de données SAS, un contrôle drastique des échantillons doit être réalisé en amont afin d'éviter toute interprétation biaisée des données [215]. Pour des échantillons protéiques préparés en laboratoire, une étape de purification de type SEC préparatoire ou semi-préparatoire, suivie d'une analyse SDS-PAGE sont souvent des étapes nécessaires. Pour tout type d'échantillon, une analyse par diffusion de lumière (DLS et/ou SLS) peut révéler rapidement et facilement la qualité de l'échantillon et donc sa compatibilité avec le SAS. En effet, la diffusion des neutrons ou des photons est proportionnelle au carré du volume (Eq 25). Ainsi, dans le cas d'une protéine, une très faible proportion d'agrégats de grande taille affecte fortement les données. Pour pallier à ce problème et s'assurer de la monodispersité de l'échantillon à analyser, certaines lignes SAXS sont équipées d'un système HPLC avec une colonne SEC en amont de la cellule ou du capillaire de détection SAXS. Il est ainsi possible de séparer les agrégats protéiques de l'état d'oligomérisation, souvent monomérique, de l'échantillon d'intérêt. La ligne SWING a été pionnière dans la mise au point de ce dispositif [216], qui est désormais répertorié sous l'appellation SEC-SAXS. L'utilisation d'un tel système sous-entend d'avoir un fort flux de rayons X, ce qui est rendu possible grâce au rayonnement

synchrotron. A l'ILL, grâce à un fort flux de neutrons, un système HPLC est désormais disponible sur l'instrument D22 [217].

- ***Approximations de Guinier***

Une fois la courbe SAS obtenue, *I(Q),* une première étape pour attester de la qualité des données consiste à utiliser l'approximation de Guinier [218]. Pour de petites valeurs de *Q*, cette approximation (valable pour $QR_g < 1.3$) est utilisée pour déterminer l'intensité à angle nul *I(0)* et le rayon de giration R_g, et est définie par [218]:

$$I\ (Q) = I(0)exp\left(-\frac{Q^2R_g^2}{3}\right) \qquad \text{Eq 27}$$

Pour des valeurs de *Q* intermédiaires, pour un objet en forme de cylindre allongé, de longueur *L* et de rayon *R*, l'approximation de Guinier (valable pour $QR_c < 1.3$) permet de déterminer R_c et est définie par [218]:

$$I(Q) = \frac{I(0)}{Q}exp\left(-\frac{Q^2R_c^2}{2}\right) \qquad \text{Eq 28}$$

où

$$R_c^2 = \frac{R^2}{2} \qquad \text{Eq 29}$$

et

$$R_g^2 = \frac{L^2}{12} + \frac{R^2}{2} \qquad \text{Eq 30}$$

Pour des valeurs de *Q* intermédiaires, pour un objet en forme de disque, de rayon *R* et d'épaisseur *t*, l'approximation de Guinier (valable pour $QR_d < 1.3$) permet de déterminer R_d et est définie par [218]:

$$I(Q) = \frac{I(0)}{Q^2}exp(-Q^2R_d^2) \qquad \text{Eq 31}$$

où

$$R_d^2 = \frac{t^2}{12} \qquad \text{Eq 32}$$

et

$$R_g^2 = \frac{R^2}{2} + \frac{t^2}{12} \qquad \text{Eq 33}$$

- ***Détermination de la MW apparente***

Une deuxième étape, pouvant être qualifiée de classique en ce qui concerne l'analyse de données SAS, consiste en la détermination de la masse apparente *MW* de l'échantillon grâce à l'extrapolation de l'intensité à angle nul *I(0)*, déterminée par l'approximation de Guinier.

La masse moléculaire (*MW* en g mol^{-1}) peut être estimée selon l'équation suivante :

$$I(0) = (\Delta\rho V_P)^2 . c . MW / N_A \qquad \text{Eq 34}$$

où $\Delta\rho$ (10^{10} cm^{-2}) est le contraste (différence de SLD) entre les objets et le tampon, V_P est le volume spécifique partiel (cm^3.g^{-1}) des protéines, *c* est la concentration en protéines (g.cm^{-3}) et N_A est la constante d'Avogrado. Le V_P est déterminé à partir de la densité des protéines selon leur séquence d'acides aminés [219]. La faiblesse principale de cette méthode est qu'elle nécessite que l'intensité diffusée soit obtenue en unité absolue (cm^{-1}). Une détermination précise de la concentration de l'échantillon et de sa SLD est également nécessaire. Cette dernière peut être inexactement estimée dans les expériences de SANS à cause de l'échange H/D lorsque les expériences sont réalisées dans du D_2O.

Une alternative très simple pour déterminer la *MW* consiste à utiliser une protéine standard, de *MW* et concentration connues, avec une simple règle de proportionnalité entre la *MW* et *I(0)*.

$$MW_{POI} = \frac{I(0)_{POI}}{c_{POI}} . \frac{MW_{std}}{I(0)_{std}/c_{std}} \qquad \text{Eq 35}$$

La protéine standard utilisée est parfois la BSA. Etonnamment, cette protéine est connue pour former des dimères, voire des trimères. Ainsi, la détermination de *MW*, tout comme celle de *I(0)*, de l'échantillon BSA pourraient être biaisées. De plus, cette approche se base sur l'hypothèse que la protéine d'intérêt et la protéine standard ont le même V_p. Cependant, cela constitue un moyen simple et rapide de déterminer l'état d'oligomérisation.

Une dernière méthode pour déterminer la MW indépendamment de la concentration, du contraste et de l'intensité en unité absolue a été développée par Rambo et Tainer [220]. Elle fait appel au volume de corrélation (V_c) qui est défini par:

$$V_c = \left(\frac{I(0)}{\int Q.I(Q).dQ}\right) \qquad \text{Eq 36}$$

et

$$Q_R = \left(\frac{V_c^2}{R_g}\right) \qquad \text{Eq 37}$$

où Q_R est le paramètre de masse qui suit une loi de puissance en fonction de la masse de la particule considérée selon :

$$\ln(Q_R) = k.\ln(MW) + c \qquad \text{Eq 38}$$

où k et c sont des constantes inhérentes au type de particule. Pour déterminer la MW des fragments protéiques utilisés dans ces travaux de thèse, nous avons utilisé des valeurs de k = 0.973 et c = -1.878 (provenant de la ligne SWING) et considéré l'intégration des données ($Q.I(Q)$ *vs.* Q) jusqu'à 0.2 $Å^{-1}$. Cette méthode peut être utilisée pour d'autres types de macromolécules que les protéines, comme les acides nucléiques par exemple.

- ***Détermination du facteur de forme***

Un moyen d'obtenir des informations plus précises concernant la forme de l'échantillon consiste à ajuster les données SAS avec un modèle analytique, en d'autres termes déterminer le facteur de forme *P(Q)*. Au cours de ces travaux de thèse, nous avons utilisé les modèles « core-shell cylinder » et « core-shell bicelle » pour ajuster les données des bicelles et le modèle « flexible cylinder » pour les protéines (Figure 46).

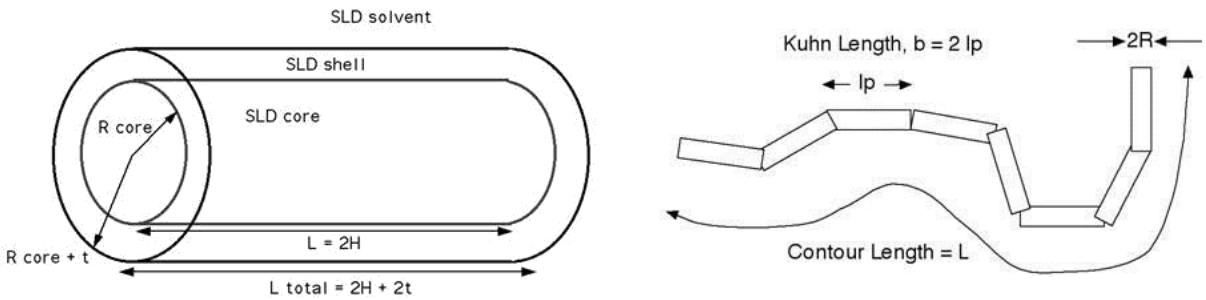


Figure 46 : Description des modèles analytiques [221]. Le modèle « core-shell cylinder » permet de distinguer les SLD de la partie hydrophile (« shell ») et de la partie hydrophobe (« core »), où L est l'épaisseur de la partie hydrophobe, R le rayon de la partie hydrophobe et t l'épaisseur de la partie hydrophile. Notons que le modèle « core-shell bicelle » ne diffère que par la possibilité d'avoir une épaisseur et une SLD de la partie hydrophile du tore différente de celles de la partie bicouche. Cela est utile surtout si les bicelles ne sont pas uniquement formées de phospholipides, comme ceux composé de CHAPSO par exemple dont la SLD de la tête hydrophile est différente de celle d'une tête PC. Dans le cas du modèle « flexible cylinder »

pour analyser les données des protéines étudiées, R est le rayon, L la longueur totale et la longueur de Kuhn (« Kuhn length ») correspond au double de la longueur de persistance.

➢ ***Fonction de distribution des distances***

GNOM est un programme utilisant une transformée de Fourier inverse pour traiter les données SAS. Il permet ainsi de passer de l'espace réciproque à l'espace réel en évaluant la fonction de distribution des distances des particules *P(r)* selon l'équation :

$$I(Q) = 4\pi \int_0^{D_{max}} P(r) \frac{\sin Q(r)}{Qr} dr \qquad \text{Eq 39}$$

Un exemple de l'application de cette méthode à des modèles géométriques simples est présenté Figure 47.

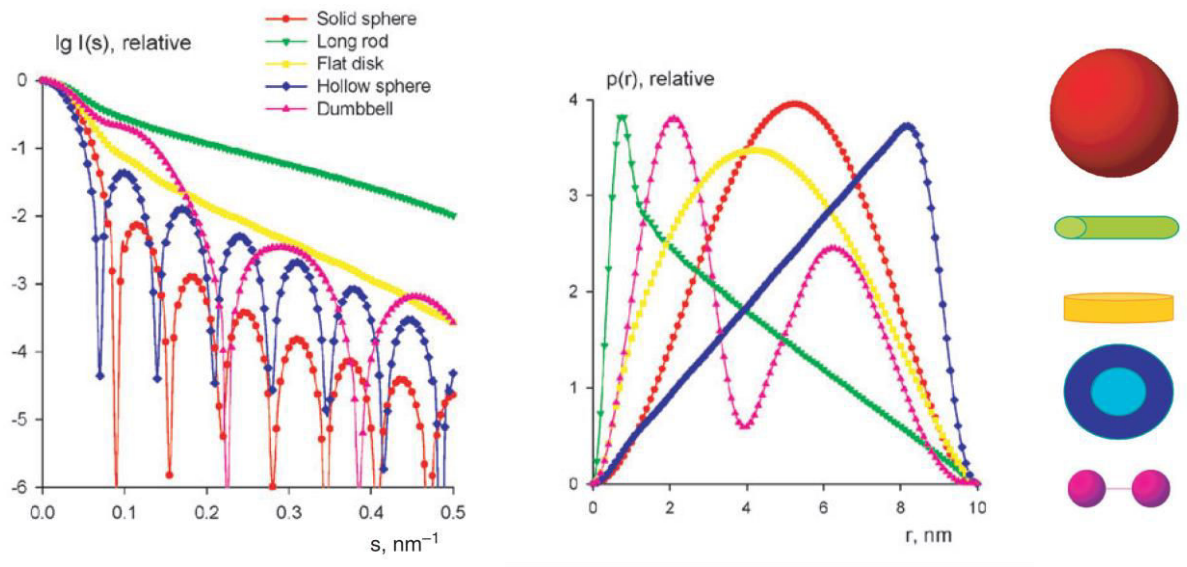


Figure 47 : Exemples de courbes de diffusion et de leurs équivalents en termes de P(r) de différentes formes géométriques. Adapté de [212].

Cette fonction (Eq 39) permet de déterminer la distance maximale au sein de la particule diffusante, D_{max}, et également de déterminer R_g et *I(0)* dans l'espace réel selon :

$$R_g^2 = \frac{\left(\int_0^{D_{max}} r^2 P(r) dr\right)}{\left(\int_0^{D_{max}} P(r) dr\right)} \qquad \text{Eq 40}$$

et

$$I(0) = 4\pi \int_0^{D_{max}} P(r) dr \qquad \text{Eq 41}$$

Cette estimation alternative de R_g et de *I(0)* présente l'avantage d'utiliser toute la courbe de diffusion et est moins sensible à la présence d'une petite fraction d'oligomères ou d'agrégats.

- ***Génération de modèles ab initio***

Dans le domaine de la biologie structurale, les techniques de SAS sont désormais bien connues comme permettant d'obtenir des enveloppes à basse résolution (« shape ») de la protéine d'intérêt. Plusieurs logiciels sont disponibles pour obtenir de telles enveloppes et sont basés sur la génération de modèles *ab initio* par recuit simulé. Les logiciels DAMMIN et GASBOR de la suite ATSAS [213] en sont deux exemples non exhaustifs.

- ***Génération de courbes SAS à partir de modèles***

Une des applications majeures du SAS, et plus particulièrement du SAXS, consiste en la comparaison de la structure cristalline obtenue par XRC à celle obtenue en solution. Ceci permet de vérifier si les conditions de cristallisation ne sont pas artefactuelles et n'ont pas donné lieu à des changements de conformation ou d'état oligomérique de la protéine. Lorsque les données XRC et SAXS sont combinées, elles permettent également d'obtenir des informations sur la position de résidus manquants des données XRC. La combinaison XRC/SAXS n'est pas exclusive. En effet, le SAS est une technique qui peut aussi être combinée avec la RMN, la cryo-EM et la modélisation moléculaire. Un exemple de logiciel permettant la génération de courbes de diffusion théorique à partir de modèles tout-atome sont CRYSOL et CRYSON pour les données SAXS et SANS, respectivement.

6. Réflectivité de neutrons

Les données de réflectivité de neutrons (NR) ont été obtenues sur les instruments FIGARO (ILL), MARIA (MLZ) et REFSANS (MLZ). Les données ont été analysées avec le logiciel Motofit [222]. La réflectivité de neutrons est une technique qui permet de sonder les surfaces et les interfaces. Elle consiste à envoyer un faisceau de neutrons sur une surface avec un certain angle d'incidence et à mesurer la réflectivité (Figure 48) de l'échantillon en fonction de Q_Z selon :

$$Q_Z = \frac{4\pi \sin\theta}{\lambda} \qquad \text{Eq 42}$$

où Q_Z est le vecteur de diffusion, θ l'angle d'incidence et λ la longueur d'onde.

La courbe de réflectivité ainsi obtenue permet de déterminer l'épaisseur et les SLD de chaque couche. Notons que nous parlons ici uniquement de réflectivité spéculaire, par opposition par exemple à la technique SANS à incidence rasante (« grazing-incidence » ou GI-SANS) qui permet d'obtenir des informations dans le plan.

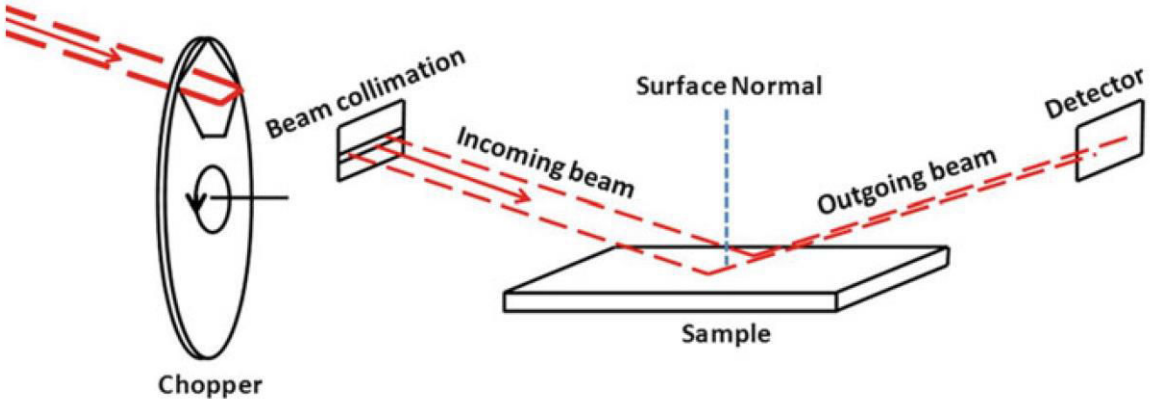


Figure 48 : Représentation simplifiée d'un instrument de réflectivité de neutrons montrant un faisceau de neutrons illuminant une surface et dont certains neutrons sont réfléchis par interaction avec les noyaux de l'échantillon [36].

- ***NR et monocouche de Langmuir***

Plusieurs cuves de Langmuir ont été utilisées avec des volumes variant de 50 à 200 mL. La tension de surface est mesurée par la méthode de Wilhelmy. La cuve est lavée abondamment en alternant EtOH 70% (v/v) et eau MQ. La cuve est remplie par un mélange de tampon h-TNE/d-TNE selon le contraste souhaité. La monocouche lipidique est formée par ajout goutte à goutte du mélange lipidique d'intérêt conditionné dans du $CHCl_3$ ou un mélange $CHCl_3$/MeOH, jusqu'à obtenir une tension de surface de 20 mN/m. Le résiduel de $CHCl_3$ ou $CHCl_3$/MeOH est laissé évaporer et la monocouche est stabilisée pendant 15 min. La tension de surface est ajustée au besoin à l'aide des barrières. Les fragments protéiques sont ensuite injectés dans la sous-phase à une concentration de 0.1 µM. La cinétique d'adsorption est suivie par variation de la tension de surface. Les mesures ont été effectuées à deux angles différents couvrant une gamme en Q de $7\text{-}9.10^{-3}$ à 0,15-0,2 $Å^{-1}$ ($\Delta\lambda/\lambda \approx$ 5-7%) selon l'instrument. Un exemple de données provenant de la littérature de NR obtenues sur une monocouche lipidique est présenté Figure 49.

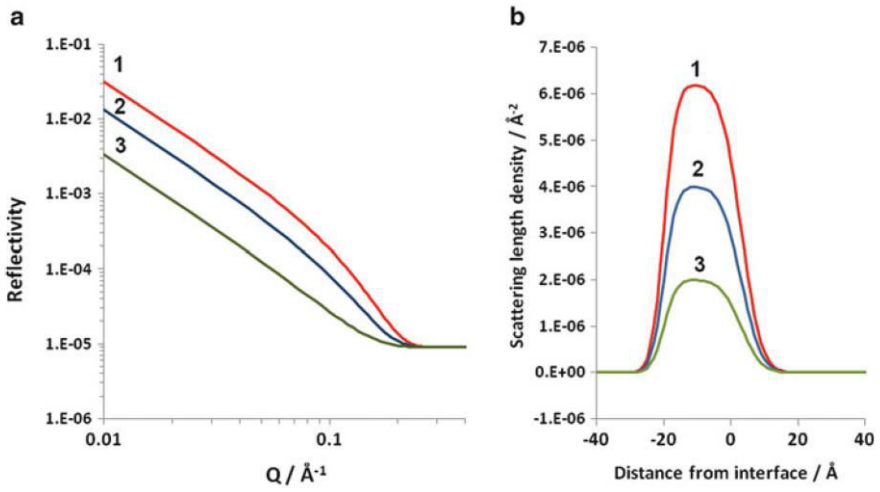


Figure 49 : Courbes de réflectivité de neutrons (a) d'une monocouche de DPPC, dont les queues sont deutérées, à l'interface air/liquide sur une sous-phase aqueuse dont la SLD est 0. Un mélange 8% D_2O/92% H_2O a une SLD de 0 comme celle de l'air. Ce mélange ne présente donc aucun contraste avec l'air et est nommé ACMW « air contrast matched water » ou NRW « null-reflecting water ». Dans cet exemple, la surface occupée par la monocouche représente 32% (3, vert), 64% (2, bleu) ou 100% (1, rouge). (b) Profil de SLD correspondant par rapport à la normale [36].

- ***NR et bicouche supportée (SLB)***

Les différents mélanges de lipides secs sont réhydratés dans du tampon TNE/d-TNE selon les besoins et forment ainsi des MLV. Cette suspension de MLV, contenant ~ 2g/L de phospholipides, est soumise à des ultrasons (Sonics Vibra-Cell VCX130) pendant plusieurs minutes, jusqu'à ce que la solution trouble initiale devienne presque claire, signe de la formation de SUV. L'environnement échantillon ainsi que la préparation des « wafers » de silicium sont décrits dans l'article de Alexandros Koutsioumpas [223]. La formation des bicouches supportées est réalisée par fusion de liposomes [35]. L'excès de SUV est retiré par lavage avec du tampon. Dans les mélanges contenant des phospholipides PS, le tampon est initialement supplémenté avec 2 mM de $CaCl_2$ [224]. La solution protéique est injectée à la seringue pour une concentration finale de 0.1 µM. L'échange de tampon se fait à l'aide d'une pompe à un débit de 1-2 mL/min. Les mesures ont été effectuées à deux longueurs d'onde différentes : 12 Å pour la région à petits Q et 6 Å pour la région à grands Q permettant d'obtenir une gamme en Q de 9.10^{-3} à 0,2 $Å^{-1}$ ($\Delta\lambda/\lambda \approx 10\%$). Un exemple de données provenant de la littérature de NR obtenues sur une SLB de DPPC est présenté Figure 50.

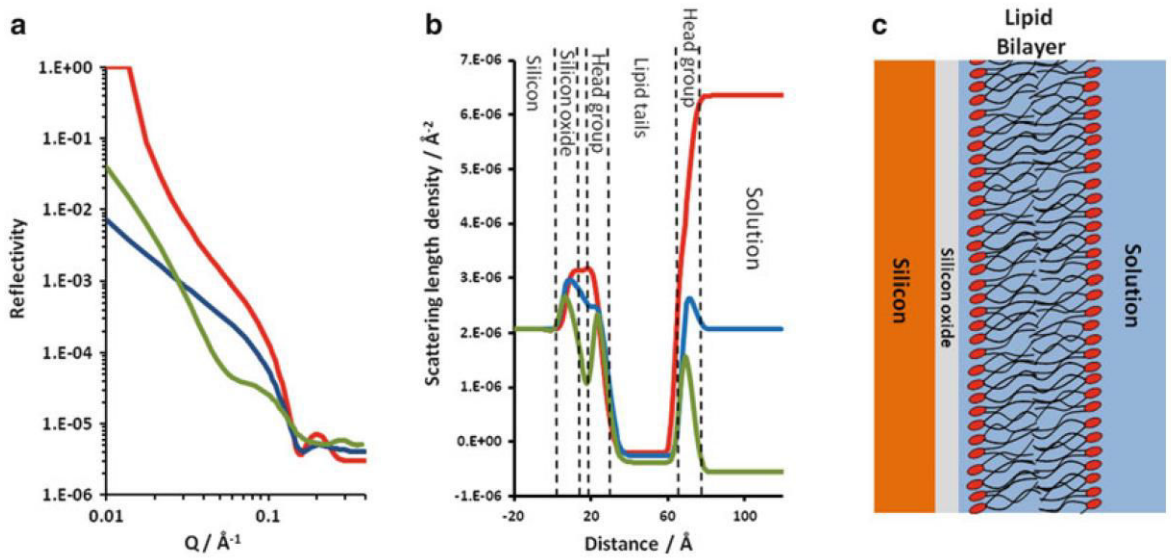


Figure 50 : (a) Courbes de réflectivité de neutrons obtenues pour une bicouche de h-DPPC adsorbée sur un « wafer » de silicium dans trois ratios D_2O/H_2O : 100% de D_2O (rouge), NRW 38% D_2O et 100% H_2O (vert). (b) Profils SLD correspondants avec le même code couleur. (c) Modèle de SLB correspondant aux données [36].

Tableau 6 : Tableau récapitulatif des paramètres utilisés pour l'analyse des données SAS et NR.

Nom	Formule	MW (g mol-1)	Volume (Å3)	SLD SANS (10^{-6} Å^{-2})	SLD SAXS (10^{-6} Å^{-2})
DMPC	C36 H72 N O8 P	678	1090	0.28	9.68
Head	C10 O8 P N H18	312	336	1.79	13.85
Tail	C26 H54	366	754	-0.39	7.89
DMPC-d	C36 H5 N O8 P D67	745	1090	6.65	
Head	C10 O8 P N H5 D13	325	336	5.84	
Tail	C26 D54	420	754	7.02	
DHPC	C20 H40 N O8 P	454	660	0.67	10.54
Head	C10 O8 P N H18	312	336	1.79	13.85
Tail	C10 H22	142	324	-0.49	7.16
DHPC-d	C20 H5 N O8 P D35	489	660	6.20	
Head	C10 O8 P N H5 D13	325	336	5.84	
Tail	C10 D22	164	324	6.56	
DMPS	C34 H65 N O10 P	679	979	0.57	11.04
Head	C8 H11 O10 P N	313	225	4.60	13.56
Tail	C26 H54	366	754	-0.39	7.89
DMPS-d	C34 H8 O10 P N D57	736	979	6.76	
Head	C8 H8 O10 P N D3	316	225	5.22	
Tail	C26 D54	420	754	7.02	

Les volumes moléculaires utilisés proviennent de [84,225]. Le volume de la partie hydrophobe est estimé d'après la règle empirique de Tanford [19]. Les densités sont estimées d'après la MW et le volume moléculaire. Les SLD sont calculés en utilisant le site internet du NIST [226].

7. Modélisation moléculaire

Les expériences *in silico* ont été réalisées au sein de l'équipe SIM essentiellement par Olivier Delalande, Dominique Mias-Lucquin et Anne-Elisabeth Molza. Ils m'ont initié aux bases conceptuelles et méthodologiques de la modélisation moléculaire. Ceci m'a permis de participer au design des expériences.

Un champ de force désigne l'ensemble des équations et des paramètres associés nécessaires pour décrire l'énergie potentielle d'un système moléculaire [227]. On distingue l'énergie liée comprenant la contribution des liaisons covalentes, des angles de valence et des angles dièdres de l'énergie non-liée qui contient la contribution des interactions électrostatiques (Coulomb) et de van der Waals (Figure 51).

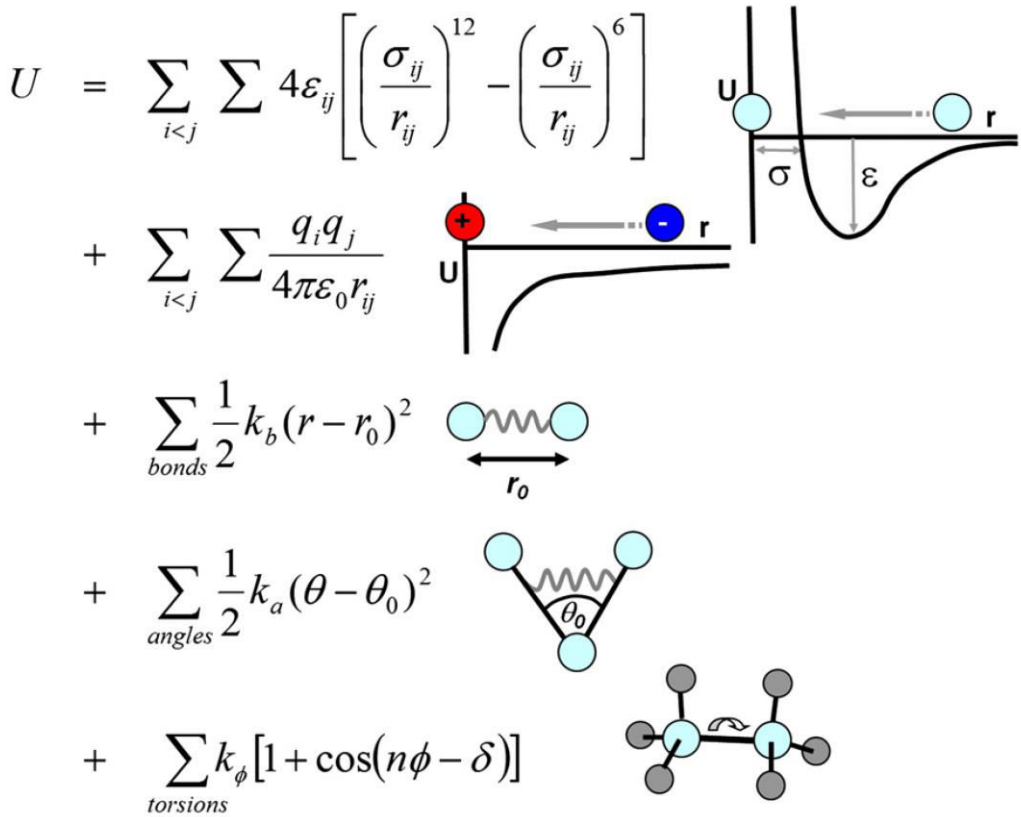


Figure 51 : Energie potentielle des interactions moléculaires selon l'approximation de la mécanique moléculaire. Les molécules sont traitées comme étant constituées de boules (les atomes) reliées par des ressorts (les liaisons). Les deux premiers termes décrivent les interactions de van der Waals et celles de Coulomb, respectivement, pour les atomes i et j. Les autres termes concernent l'énergie liée associée à la liaison covalente, l'angle de liaison et l'angle de torsion des molécules composants le système [228].

Au cours d'une simulation moléculaire, le champ de force contraint le système selon les paramètres imposés. Il existe plusieurs types de champ de force, les champs de force « tout-atome » et « gros-grain ». La dynamique moléculaire gros-grain a pour but de réduire le nombre de particules constituant un système afin de diminuer le temps de calcul des simulations moléculaires, en comparaison au modèle tout-atome. Ces modèles gros-grain regroupent plusieurs atomes, classiquement une fonction chimique, au sein d'un grain afin de conserver les propriétés physico-chimiques inhérentes à chaque type de molécule. Le champ de forces gros-grain MARTINI (Figure 52) a été utilisé lors de ces travaux de thèse. Le champ de forces MARTINI a été initialement conçu pour la simulation par dynamique moléculaire des lipides [229], puis a été étendu aux interactions protéines/lipides [230].

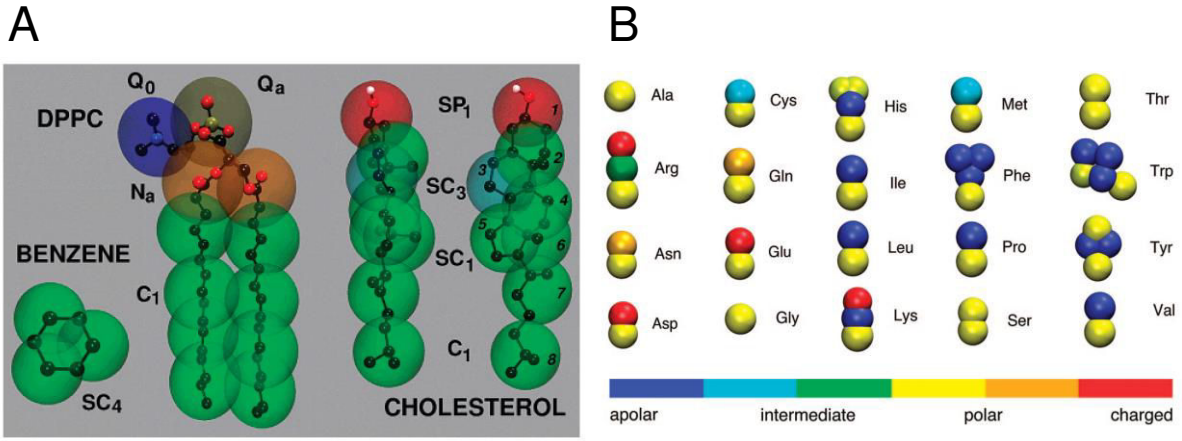


Figure 52 : Représentation du modèle gros-grains MARTINI (A) Comparaison des cartographies des structures chimiques tout-atome et gros-grain pour une molécule de DPPC, de cholestérol et de benzène, où un grain regroupe une fonction chimique [229]. (B) Représentation gros-grain des acides aminés. Les différentes couleurs représentent différents types de particules [230].

PARTIE III : RESULTATS ET DISCUSSION

Contexte de travail et objectif de la thèse

Ces travaux de thèse s'inscrivent dans la problématique générale des relations structure/fonction des protéines. La dystrophine est une protéine essentielle des cellules musculaires au sein desquelles elle assure un rôle de charpente et de soutien du sarcolemme. Ce rôle est réalisé grâce aux diverses interactions qu'elle assure avec ses partenaires protéiques mais également avec les lipides membranaires. A cause de mutations dans le gène *DMD,* l'absence de la dystrophine conduit à la myopathie de Duchenne, tandis que sa production sous forme tronquée plus ou moins fonctionnelle est responsable de la myopathie de Becker, qui présente des phénotypes variés, de très sévères jusqu'à quasiment asymptomatiques. Une compréhension approfondie du rôle biologique de la dystrophine implique une caractérisation fine et précise du mode de liaison impliqué dans ses interactions avec le sarcolemme. L'objectif à long terme de ce projet est de concevoir des thérapies de façon rationnelle ayant pour but le traitement des dystrophinopathies, notamment le design de mini-dystrophines minimales fonctionnelles.

Les interactions protéine/lipide impliquant ces deux fragments protéiques ont été finement caractérisées par l'équipe SIM ces dernières années. Cependant, la structure adoptée par R1-3 et R11-15 en interaction avec la membrane plasmique demeurait inconnue. **L'objectif de cette thèse était de déterminer la conformation adoptée par deux fragments du domaine central de la dystrophine, R1-3 et R11-15, en présence de modèles membranaires**. Pour mener à bien l'objectif de la thèse, nous avons utilisé, entre autres, une approche de SANS/SAXS couplée à de la modélisation moléculaire. Nous avons tiré profit de la technique de variation de contraste en SANS qui permet de sonder le signal d'un composant au sein d'un complexe macromoléculaire à plusieurs partenaires. Dans notre cas, cette technique est la seule permettant d'obtenir uniquement le signal de la protéine lors de sa liaison avec des bicelles *ad hoc*, c'est-à-dire dont le contraste avec le solvant est nul. L'objectif de l'approche par SAXS était d'obtenir des informations sur le complexe entier protéine/bicelle. Une des difficultés consiste alors à obtenir des données SAXS uniquement sur le complexe sans la contribution des fragments protéiques ou des bicelles seuls en solution. Le système SEC-SAXS disponible sur la ligne SWING doit permettre de séparer les bicelles libres des complexes protéines/bicelles. Les

données SAXS des complexes entiers peuvent être comparées avec des modèles obtenus par modélisation moléculaire. L'adéquation entre les données expérimentales et *in silico* devrait permettre de proposer des modèles fiables de complexes de fragments de dystrophine/bicelles pour une meilleure compréhension des bases moléculaires impliquées dans la liaison de la dystrophine au sarcolemme.

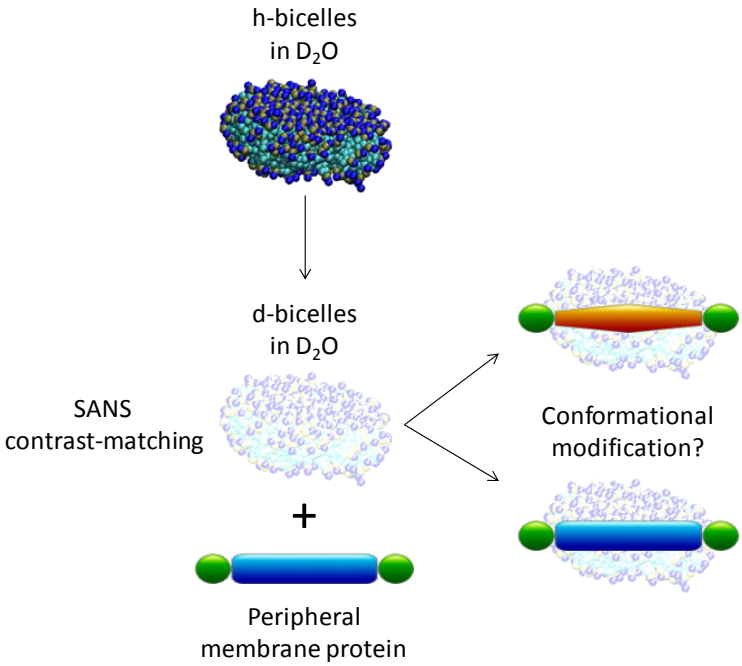
Chapitre 1 : Caractérisation du modèle « bicelle »

Points clés :

- *En utilisant la variation de contraste en SANS, nous montrons que le modèle de bicelle peut être utilisé pour obtenir des informations structurales de protéines membranaires périphériques en interaction avec des lipides membranaires. La preuve de principe de la méthode a été appliquée à R11-15.*

- *Pour des protéines moins stables que R11-15, le modèle de bicelle est compatible avec un dispositif SEC-SANS.*

- *Cette approche est potentiellement applicable aux protéines membranaires intégrales.*



1. Introduction

Nous avons choisi le modèle de bicelle en tant que membranomimétique pour résoudre notre problématique. Ce modèle est beaucoup moins classique que celui de micelles de détergents ou de liposomes. Un moyen simple et rapide de mettre en exergue cette tendance est d'interroger la base de données PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) : ~ 50000 entrées pour le mot clé « liposome », ~ 30000 pour « micelle » et moins de 600 pour « bicelle ». Ainsi, une compréhension exhaustive de ce modèle était nécessaire avant de l'utiliser comme modèle membranaire. Dans cette partie, nous présentons les résultats obtenus sur la caractérisation des bicelles par diverses méthodes.

2. Publication

Une partie de ces travaux a fait l'objet d'un article scientifique publié dans la revue *Langmuir* [231]. L'objet principal de cette publication est la démonstration que des bicelles deutérées « éteintes » peuvent être obtenues par variation de contraste, en utilisant des lipides deutérés dans un tampon 100% D_2O. Ces conditions permettent de sonder uniquement le signal provenant d'une protéine membranaire périphérique en interaction avec des lipides membranaires, ainsi que les potentiels changements conformationnels associés. La faisabilité de la méthode a été appliquée au fragment de dystrophine R11-15.

Article 1: Contrast-matched isotropic bicelles: a versatile tool to specifically probe the solution structure of peripheral membrane proteins using SANS.

LANGMUIR

Article

pubs.acs.org/Langmuir

Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS

Raphael Dos Santos Morais,[†,‡,§,∥] Olivier Delalande,[†,‡] Javier Pérez,[∥] Liza Mouret,[⊥] Arnaud Bondon,[⊥] Anne Martel,[#] Marie-Sousai Appavou,[∇] Elisabeth Le Rumeur,[†,‡] Jean-François Hubert,[†,‡] and Sophie Combet*[,§]

[†]Université de Rennes 1, F-35043 Rennes, France

[‡]CNRS UMR 6290, Institut de Génétique et Développement de Rennes, F-35043 Rennes, France

[§]Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, Gif-sur-Yvette F-91191, France

[∥]SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP48, Saint-Aubin, Gif-sur-Yvette F-91192, France

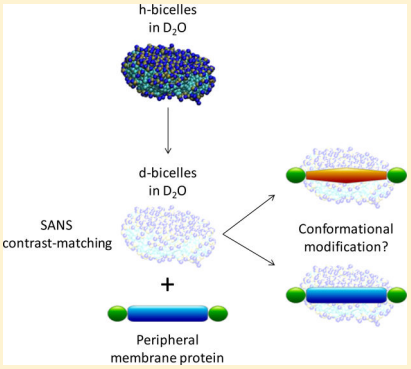
[⊥]CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-350043 Rennes, France

[#]Institut Laue-Langevin, F-38042 Grenoble, France

[∇]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany

Supporting Information

ABSTRACT: Obtaining structural information on integral or peripheral membrane proteins is currently arduous due to the difficulty of their solubilization, purification, and crystallization (for X-ray crystallography (XRC) application). To overcome this challenge, bicelles are known to be a versatile tool for high-resolution structure determination, especially when using solution and/or solid state nuclear magnetic resonance (NMR) and, to a lesser extent, XRC. For proteins not compatible with these high-resolution methods, small-angle X-ray and neutron scattering (SAXS and SANS, respectively) are powerful alternatives to obtain structural information directly in solution. In particular, the SANS-based approach is a unique technique to obtain low-resolution structures of proteins in interactions with partners by contrast-matching the signal coming from the latter. In the present study, isotropic bicelles are used as a membrane mimic model for SANS-based structural studies of bound peripheral membrane proteins. We emphasize that the SANS signal coming from the deuterated isotropic bicelles can be contrast-matched in 100% D_2O-based buffer, allowing us to separately and specifically focus on the signal coming from the protein in interaction with membrane lipids. We applied this method to the DYS-R11−15 protein, a fragment of the central domain of human dystrophin known to interact with lipids, and we were able to recover the signal from the protein alone. This approach gives rise to new perspectives to determine the solution structure of peripheral membrane proteins interacting with lipid membranes and might be extended to integral membrane proteins.



INTRODUCTION

While approximately 26% of the human proteome is predicted to be composed of membrane proteins,[1] the structures of only 685 unique membrane proteins are available.[2] This highlights the difficulty in handling membrane proteins during their solubilization, purification, and crystallization.[3] For even more recalcitrant membrane proteins incompatible with high-resolution methods (nuclear magnetic resonance (NMR) and X-ray crystallography (XRC)), small-angle X-ray and neutron scattering (SAXS and SANS, respectively) are good alternative approaches to obtain the structural information on a protein not only alone in solution but also in interaction with its partner(s).[4,5] In the SANS-based approach, the contrast matching (CM) technique allows the signal coming from one component to be specifically probed by hiding the signal coming from other partners in a multicomponent system. This is achieved by adjusting the deuteration level of the components in an appropriate H_2O/D_2O ratio. The CM technique has been successfully applied to the determination of the structure of one component in protein−protein and protein−DNA complexes.[6] Protein−lipid complexes have also been successfully investigated in the same way with different

Received: April 21, 2017
Revised: June 2, 2017
Published: June 5, 2017

 © 2017 American Chemical Society

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

membrane mimic environments.[7] In an elegant study, Gabel et al.[8] investigated the conformation of the β-barrel transporter FhaC surrounded by a β-octyl glucoside corona by coupling SANS and molecular modeling, obtaining a conformation of the protein similar to that obtained by XRC. Although detergent micelles are considered suitable for maintaining the tridimensional structure of membrane proteins, some authors[9] proposed that the loss of activity evidenced for some membrane proteins could be inherent to the lack of their internal dynamics in detergent micelles compared with bilayered membrane mimics. Indeed, membrane models other than micelles are available and the interest for reconstructing a biological native lipid environment for membrane proteins is growing.[10−12] Among them are the classical liposome model but also relatively more recent ones, such as nanodiscs and bicelles, which can be good alternatives.[10−12] However, all of these membrane mimics have their advantages and drawbacks.[10−12]

Nanodiscs are phospholipid bilayers surrounded by either a membrane scaffolding protein (MSP)[13] or a stryrene maleic acid (SMA) copolymer.[14] By using these membrane mimics, Maric et al.[15] proposed an interesting method to use SANS to analyze integral membrane proteins incorporated into contrast-matched nanodiscs. In the case of peripheral membrane proteins, it is essential to maintain a phosphatidylcholine head all over the membrane mimic to avoid nonspecific interactions that could happen between the scaffold of the nanodisc and the protein of interest. This can be achieved by handling peripheral proteins with fully phospholipid-based bicelles. Moreover, the size of bicelles can be easily adjusted and does not require the use of several MSP truncated forms.[16,17] Last but not least, in addition to the intrinsic highly curved rim region, few curvature-inducing molecules can be incorporated into bicelles in order to modify the planarity of the bilayer part.[18] These properties reinforce the suitability of the bicelle model for mimicking membrane bending, known to be essential for various biological processes.[19] These three points constitute the most important advantages of phospholipid-based bicelles compared to nanodiscs.

Bicelle aggregates are made of long-chain phospholipids arranged in a bilayer surrounded by a torus of short-chain phospholipids or detergents. Pioneering studies describing such objects were performed on lecithin/bile-salt or analogues[20,21] and dipalmitoylphosphatidylcholine/diheptanoylphosphatidylcholine (DPPC/DH_7PC) solution.[22,23] To better mimic the membrane bilayer, unsaturated phospholipids, cholesterol,[24,25] or anionic phospholipids[26] were successfully incorporated into the bilayer part. Recently, bicelles were even constructed from native *Escherichia coli* lipids,[27] but the currently best-described system is made of dimyristoylphosphatidylcholine/dihexanoylphosphatidylcholine (DMPC/DHPC) since its first characterization by solid-state NMR.[28] The molar ratio (DMPC/DHPC), denoted "q", of the two components forming the bicelles is the major parameter governing the bicelle size. This parameter can be easily adjusted according to the desired final size for a required application. Its value distinguishes two types of bicelles. The larger ones ($q \geq 2.5$) are able to align in a magnetic field above the gel-to-fluid phase transition temperature (T_m). They are commonly used in solid state NMR[29,30] and increasingly in XRC,[31−33] while existing as perforated lamella rather than disks in these conditions.[34] The other type, small isotropic bicelles ($q < 2.5$), are commonly used in solution NMR.[29,30] Despite their versatility, the morphology and the size of the bicelles are reported to be highly sensitive to q, temperature, and dilution.[35,36] DMPC/DHPC bicelles were thoroughly investigated with cryo-electronic microscopy, dynamic light scattering (DLS) and NMR,[35−39] but only a few studies were performed with SANS.[40−44] The study by Luchette et al.[40] is, to our knowledge, the only one devoted to SANS characterization of hydrogenated DMPC/DHPC isotropic bicelles.

In the present work, we analyzed both hydrogenated and deuterated isotropic bicelles (h-bicelles and d-bicelles, respectively) in the range of $1 \leq q \leq 1.3$. We demonstrate that the d-bicelles can be contrast-matched in 100% D_2O-buffer regardless of their morphology and temperature. We applied this method to a protein fragment of the dystrophin central domain, DYS R11−15, known to interact with the lipid membranes.[45−47] We highlight that one can specifically probe the signal of a peripheral membrane protein bound to bicelles. In addition, we show that SANS analysis can be properly achieved on bicelles previously submitted to high-pressure size exclusion chromatography.

■ MATERIALS AND METHODS

Materials. 1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC), 1,2-dimyristoyl-d54-*sn*-glycero-3-phosphocholine-1,1,2,2-d4-*N*,*N*,*N*-trimethyl-d9 (DMPC-d67), and 1,2-dihexanoyl-d22-*sn*-glycero-3-phosphocholine-1,1,2,2-d4-*N*,*N*,*N*-trimethyl-d9 (DHPC-d35), conditioned in chloroform, are from Avanti Polar Lipids and were used without any further purification. D_2O, Tris-d11, and EDTA-d16 are from Eurisotop.

Bicelle Preparation. A chloroform solution containing the appropriate amounts of DMPC/DHPC or DMPC-d67/DHPC-d35, to obtain a ratio (mol/mol) $q = 1$, was dried overnight under vacuum. The lipid mixture was then rehydrated in d-TNE buffer solution (20 mM Tris-d11, 150 mM NaCl, and 0.1 mM EDTA-d16, pD 7.5) or TNE buffer to reach a total lipid concentration of at least 200 mM. Once rehydrated, the solution was frozen (10 s in liquid N_2), thawed (10 min at 40 °C), vigorously shaken with a vortex (1 min), and then centrifuged (1.5 min, 6000 rpm, MiniSpin, Eppendorf). This procedure was repeated two more times to homogenize the solution. Stock solutions were diluted in d-TNE or TNE based on needs.

The Ideal Bicelle Model (IBM). Isotropic bicelles are considered to be disk-shaped objects made of long-chain phospholipids forming a bilayer, surrounded by short-chain phospholipids located at the rim (Figure 1A). The molar ratio of the two phospholipids, denoted "q" (not to be confused with "Q" used here for the scattering vector in SANS and DLS experiments), is the main parameter governing the size of the bicelles. The effective molar ratio is denoted "q_{eff}"[35,37] when the proportion of free DHPC is considered, and is defined by

$$q_{eff} = \frac{[\mathrm{DMPC}]}{[\mathrm{DHPC}]_{total} - [\mathrm{DHPC}]_{free}} \tag{1}$$

Several IBMs are proposed to correlate q_{eff} to the radius of the bicelle. These models are based either on headgroup area[35,48] or on phospholipid volume.[49] We consider the latter model to be the most reliable one since the molecular volumes of phospholipids are barely different below and above T_m[50,51] and the equation does not depend on the thickness of the bilayer. The radius of the IBM can be expressed as[38,49] (Figure 1A)

$$R = r_\perp + \frac{r_\perp q_{eff}}{4\Lambda}\left[\pi + \left(\pi^2 + \frac{32\Lambda}{3q_{eff}}\right)^{1/2}\right] \tag{2}$$

where R is the radius of the bicelle, $r_\perp$ is the thickness of the rim (11 Å, the length of a DHPC molecule), Λ is the volume ratio of DHPC to DMPC (0.61), and t is the thickness of the bilayer (assumed to be 40

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

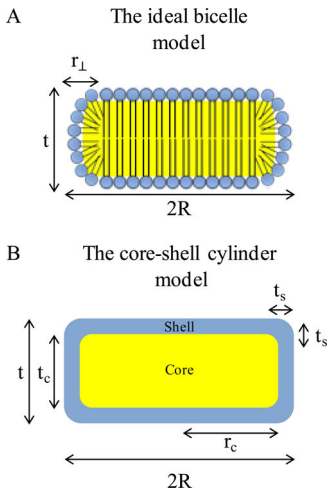


Figure 1. (A) Schematic representation of the ideal bicelle model. For DMPC/DHPC bicelles, R is the radius of the bicelle, $r_\perp$ is the thickness of the rim (11 Å, the length of a DHPC molecule), and t is the thickness of the DMPC bilayer (assumed to be 40 Å).[49] (B) Schematic representation of the "core-shell cylinder" model, where t_c is the thickness of the core, r_c is the radius of the core, and t_s is the thickness of the shell. As described in the Materials and Methods, the calculated scattering length densities of the core and of the shell are, respectively, −0.58 and 2.24 × 10^{-6} $Å^{-2}$ for h-bicelles and 7.39 and 5.05 × 10^{-6} $Å^{-2}$ for d-bicelles.

Å).[49] According to Small,[52] the volumes of DMPC and DHPC are 1090 and 660 $Å^3$, respectively.

^{31}P NMR Spectroscopy. ^{31}P NMR spectra were recorded on a Bruker spectrometer Avance 500 equipped with a 5 mm BBO probe operating at 202.46 MHz. Spectra of 10 kHz spectral width and 32 K data points were acquired with proton decoupling using 128 scans, a 30° flip angle, and a 0.5 s relaxation delay. The data were processed with the TopSpin3.2 software (Bruker). Before applying the Fourier transform, free induction decays were treated with an exponential broadening of 2 Hz. 85% H_3PO_4 was used as an external standard for the ^{31}P chemical shift.

Dynamic Light Scattering. The monodispersity and the hydrodynamic radius (R_h) of the bicelles were estimated by DLS using a Zetasizer instrument (Nano ZS, Malvern Instruments). Measurements were done in a temperature range from 15 to 37 °C in low volume Hellma cells (100 μL). The data were processed with the Zetasizer software v7.11 with default parameters and characterized by the size distribution by intensity. The decay rate, Γ, of the autocorrelation function is linked to the diffusion coefficient D by $\Gamma = DQ^2$, where Q, the momentum transfer, is defined as $Q = \frac{4\pi n \sin\theta}{\lambda}$, where n is the refractive index of the medium, λ the wavelength of the laser, and 2θ the scattering angle. R_h is determined with the Stokes−Einstein equation: $R_h = \frac{k_B T}{6\pi\eta D}$, where k_B is the Boltzmann constant, T the temperature, and η the viscosity of the medium. Data were processed using the viscosity of heavy water according to the temperature.[53]

DYS R11−15 Protein Purification. The DYS R11−15 protein was expressed and purified as previously described.[45] The purity was assessed by SDS-PAGE (Figure S3A) and Coomassie blue staining (InstantBlue, Expedeon). TNE to d-TNE buffer exchange was performed with Amicon Ultra-15 (MWCO 10 kDa). Exactly the same buffer was used for both bicelle rehydration and SANS acquisition to ensure a perfect buffer subtraction in all cases.

Intrinsic Tryptophan Fluorescence. Fluorescence measurements were obtained on a Fluorolog spectrofluorometer (Jobin-Yvon). Tryptophan fluorescence emission spectra were recorded at 20 °C in low volume quartz cuvettes (120 μL) between 310 and 420 nm using an excitation wavelength of 295 nm (bandwidth 2 nm) in TNE buffer. The protein concentration was 10 μM and the lipid concentration was 50 mM.

Chromatographic Coelution. Data were collected using a Bio SEC-5 500 Å column (5 μm, 4.6 mm × 300 mm, Agilent) mounted on an *ÄKTA* Explorer HPLC system (GE Healthcare). The column was equilibrated with TNE buffer supplemented with 6 mM of DHPC to avoid bicelle deformation under diluted concentration. A sample volume of 50 μL was loaded onto the column. For DYS R11−15 alone, the protein concentration was 60 μM and, for the bicelles alone, the lipid concentration was 50 mM. For the lipid−protein complex, the concentrations were 94 μM for the protein and 50 mM for the lipids. The flow rate was 0.2 mL/min.

Small Angle Neutron Scattering Experiments. SANS data were recorded on either the PACE[54] (LLB, Saclay), KWS1[55,56] (MLZ, Garching), or D22[57] (ILL, Grenoble) SANS instruments. According to the predicted size of the bicelles, two or three sample-to-detector distances were used with a wavelength varying from 4.7 to 6 Å, to cover a Q-range from 0.005 to 0.5 $Å^{-1}$ for the largest one, where $Q = \frac{4\pi \sin\theta}{\lambda}$ is the momentum transfer, λ is the wavelength, and 2θ is the scattering angle. All measurements were done in 1 mm Hellma Suprasil quartz (QS) cells.

High-Pressure Size Exclusion Chromatography-SANS. High-pressure size exclusion chromatography (HPSEC)-SANS data were acquired on the D22 instrument (ILL, Grenoble) with the single configuration of 8 m collimation and sample−detector distance, at λ = 6 Å ± 10%, covering a Q-range from 0.008 to 0.17 $Å^{-1}$. The sample environment was described elsewhere.[58] Data were collected using a Bio SEC-5 500 Å column (5 μm, 4.6 mm × 300 mm, Agilent) equilibrated with the d-TNE buffer supplemented with 6 mM of DHPC. A volume of 100 μL of bicelle solution at 50 mM was loaded onto the column. The flow rate was 0.2 mL/min. The SANS signal of the buffer was collected in line before the void volume, allowing a perfect subtraction to be done.

SANS Data Analysis. SANS data were analyzed with the ATSAS suite,[59] following its guidelines unless otherwise indicated. The PRIMUS software[60] was used to estimate the dimensions of the bicelles. At small Q values, the Guinier approximation (for $QR_g < 1.3$) was used to determine the forward intensity $I(0)$ and the radius of gyration R_g and is defined by[61]

$$I(Q) = I(0)\exp\left(-\frac{Q^2R_g^2}{3}\right) \tag{3}$$

At intermediate Q values, for a disk-shaped object, the Guinier approximation allows the cross-sectional radius of gyration R_t to be approximated and is defined by[61]

$$I(Q) = \frac{I(0)}{Q^2}\exp(-Q^2R_t^2) \tag{4}$$

For a disk-shaped object of homogeneous SLD contrast, R_t and R_g are related to the thickness t and the radius R of the disk by[61]

$$R_t^2 = \frac{t^2}{12} \tag{5}$$

and

$$R_g^2 = \frac{R^2}{2} + \frac{t^2}{12} \tag{6}$$

SANS Data Fitting Parameters. For our calculations, we used the volumes of DMPC and DHPC from Small[52] (1090 and 660 $Å^3$ respectively). The volume of the tail, excluding the carbons belonging to the carbonyl functions, was estimated using the empirical rule of Tanford.[62] Then, the volume of the lipid head was obtained simply by subtracting the volume of the tail from the total volume. The densities were estimated from the molecular formula and the volumes. The scattering length density (SLD) was calculated using the NIST Web site.[63] A summary table is available (Table S1). Data were fitted with the "core-shell cylinder" model[64] (Figure 1B), described in the

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

Supporting Information (eq S1, S2, and S3). We fixed the following parameters to fit SANS data: the thicknesses of the shell (t_s) and of the core (t_c) were 6 and 28 Å, respectively. The SLD of the solvent was 6.34×10^{-6} Å^{-2}. The thickness of the core had to be released in order to fit the data of d-bicelles at 37 °C. All data fitting procedures for form factor determination were done using the SASview v3.1.0 software.[65]

■ RESULTS AND DISCUSSION

Characterization of the h-Bicelles. Beaugrand et al.,[37] in analogy with the critical micellar concentration for the micelles, introduced the critical bicellar concentration (CBC), which corresponds to the concentration of free DHPC in a bicelle solution.[35] They also defined the concentration boundaries to maintain the bicelle size for a defined q ratio. Mineev et al.[38] recently proposed to determine the concentration of free DHPC by NMR diffusion, as previously done by Chou et al.,[66] and to dilute the sample in a buffer containing the same DHPC concentration in order to maintain the expected q ratio and thus the bicelle size under diluted conditions. We decided to take advantage of the free DHPC in solution to obtain bicelles with a molar ratio, q_{eff}, ranging from 1 to 1.3 by assuming a CBC value of 6 mM throughout this study.[37] This value is known to be almost constant in our q_{eff}-range and in a wide range of temperatures.[37] Note that all concentrations hereafter indicate the total lipid concentration where the CBC is not subtracted. The q_{eff}-range ($1 \leq q_{eff} \leq 1.3$) was chosen since the morphology of small isotropic bicelles ($q_{eff} < 1$) is still largely debated. Below this threshold, the objects would be present as mixed-micelles[37] or as real disk-shaped objects.[38,39] First, we checked the specific organization in bicelles and the correct molar ratio of our preparation by analyzing a concentrated sample (200 mM, q = 1) by ^{31}P NMR (Figure 2A). The spectrum is characteristic of isotropic bicelles with a narrow peak and a broader one attributed to DHPC and DMPC, respectively.[28] The molar ratio of 1, calculated from the integrals of the two peaks, is in line with the q = 1 expected value. Since the morphology and the size of the bicelles may be affected by q_{eff},[35,36] less concentrated samples diluted with the d-TNE buffer, at 100, 75, and 50 mM of lipids (Figure 2B), were analyzed by SANS at 20 °C, below the gel-to-fluid phase transition temperature of DMPC (24 °C).[67] For these three concentrations, a plateau is reached, which is characteristic of objects with a finite size. The radii of gyration (R_g) obtained for the three concentrations are 25.5 ± 0.2, 28.0 ± 0.3, and 33.5 ± 0.4 Å. As expected, the bicelle size increases with q_{eff} (Figure 2B). On the other hand, the intermediate radius of gyration (R_t) remains constant at 12.0 ± 0.1 Å. For a disk-shaped object of homogeneous SLD contrast, the R_g values correspond to geometric radii of approximately 32, 36, and 44 Å, while the R_t value corresponds to a thickness (t) of approximately 41.5 Å, in line with the dimension of a DMPC bilayer in gel phase.[68] We also analyzed the stock solution (265 mM) and observed a structure peak (Figure S2) associated with an interparticle spacing previously observed by Luchette et al.[40] in similar conditions. This peak is no longer observable for the more diluted samples, down to 50 mM. More diluted samples were also analyzed (25 and 10 mM) and, as expected, the samples are no longer disk-shaped objects at these concentrations since q_{eff} increases strongly (data not shown). Finally, to confirm that the bicelles at 50 mM are effectively discoidal objects, data obtained at this concentration were fitted with the "core-shell cylinder" form factor model.[64] This two-step SLD cylinder model allows the tails and the heads of the phospholipids to be distinguished. We assume that the SLD for DMPC and DHPC heads are identical since they have the same chemical function. This affirmation might not be absolutely true because differences in the level of hydration between the bilayer and the rim could modify the SLD. Indeed, the areas occupied by the heads of both phospholipids are known to be different in the case of a DMPC bilayer in gel (47 Å^2)[68] or in fluid phase (60 Å^2)[69] or for DHPC micelles (100 Å^2).[70] We also fixed the thickness of the core and of the shell to 28 and 6 Å, respectively, to reduce the number of free parameters during the fitting procedure. Despite these approximations, good fits were obtained (Figure 2B) using this simplified model for h-bicelles, giving overall dimensions of 40 Å for the thickness (t) and 42 Å for the total radius (R). These results are in line with the first Guinier approximations and the IBM proposed by Triba et al.[49] Altogether, these data show that the experimental conditions used for the h-bicelles led to disk-shaped objects. Therefore, the same experimental conditions were used in the following experiments with the d-bicelles.

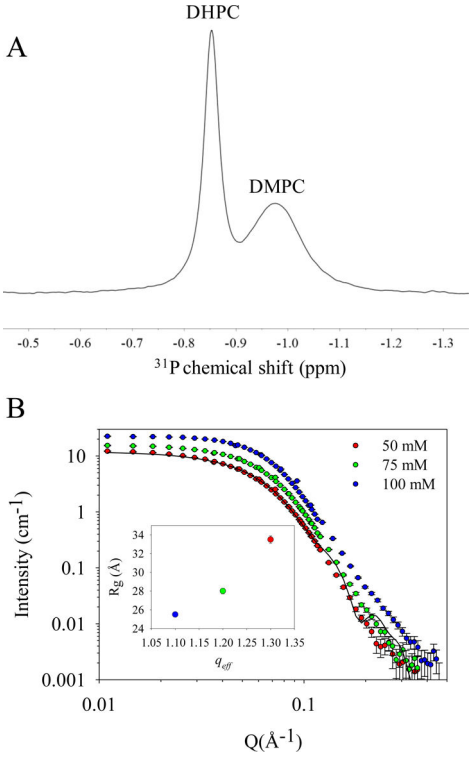


Figure 2. (A) ^{31}P NMR spectrum of q = 1 h-bicelles at 200 mM, typical of a bicelle organization. (B) SANS intensities measured for q = 1 h-bicelles at 20 °C for a total DMPC/DHPC concentration of 100 (q_{eff} = 1.1) (blue), 75 (q_{eff} = 1.2) (green), and 50 (q_{eff} = 1.3) (red) mM in 100% D_2O d-TNE buffer. Data at 50 mM were fitted (black line) with the "core-shell cylinder" model. The thickness (t) is 40 Å and the radius (R) is 42 Å, in line with the dimensions of a disk. Insert: R_g as a function of q_{eff} showing the increased bicelle size with q_{eff}.

Deuterated Bicelles Are Contrast-Matched in 100% D_2O Buffer. We used almost fully deuterated DMPC and DHPC lipids, commercially available, to form the d-bicelles (five hydrogen atoms remain near the glycerol group). The d-bicelles at 50 mM were analyzed at three contrasts: 42, 70, and 100% D_2O/H_2O ratios (Figure 3). First, to probe the expected disk-shaped morphology, data from the d-bicelles at 42% D_2O (corresponding to the classical contrast-matching point of hydrogenated proteins) was used for further analysis after

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

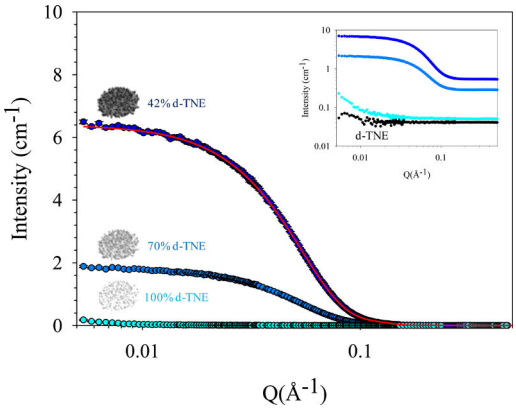


Figure 3. SANS intensities measured for q_{eff} = 1.3 d-bicelles at 50 mM lipids in 42% (dark blue), 70% (blue), and 100% (cyan) D_2O in d-TNE buffer (black) at 20 °C. The d-bicelles in 42% d-TNE were fitted with the "core-shell cylinder" model (red line), giving 40 and 42 Å for the thickness (t) and the radius (R), respectively. The incoherent SANS signal was subtracted. Insert: the same data, without incoherent subtraction in a log/log scale, highlighting the small residual signal of d-bicelles at very small Q-values and compared to d-TNE buffer prepared in 100% D_2O (black).

incoherent subtraction (Figure 3). We used the Guinier approximation to estimate the overall dimensions of these disk-shaped objects. We obtained R_g and R_t values of 31.5 ± 0.1 and 12.5 ± 0.1 Å, respectively, corresponding to a radius (R) of approximately 41 Å and a thickness (t) of 43 Å. Then, the data for d-bicelles were fitted with the same model and assumptions as for the h-bicelles. A good fit was obtained (Figure 3) by using the "core-shell cylinder" model, giving overall dimensions of 40 Å for the thickness (t) and 44 Å for the total radius (R), in line with the Guinier approximation and the theoretical IBM.

According to our calculations, the SLD for DMPC-d67 and DHPC-d35 are 6.65 and 6.20 × 10^{-6} $Å^{-2}$, respectively, and are close to the SLD of 100% D_2O-based buffer of 6.34 × 10^{-6} $Å^{-2}$. Therefore, considering SANS resolution, d-bicelles should be nearly contrast-matched in 100% d-TNE buffer. As expected, d-bicelles at 100% D_2O are virtually contrast-matched in SANS, with a small residual signal at small angles, which is negligible (Figure 3, insert). The signal of the d-bicelles in 70% D_2O is more than 20-fold higher and almost 2 orders of magnitude higher in 42% than in 100% D_2O. By plotting the square-root of SANS intensity as a function of the percentage of D_2O, we obtained for the data above a match-point of 104% D_2O (Figure S1A). To compare with our experimental data, theoretical curves were generated for a Q-range from 2.5.10^{-2} to 0.5 $Å^{-1}$ with the SASview software, using the above parameters obtained for the data set at 42% D_2O, in 0, 20, 42, 70, 90, and 100% D_2O (Figure S1C). The theoretical match-point of the d-bicelles was determined with the same plot as above (Figure S1B) and is around 99% D_2O, in line with the experimental determination. Considering experimental and theoretical data, 100% D_2O is assumed to be a reliable match-point value. We concluded that the d-bicelles are virtually contrast-matched in 100% d-TNE buffer, whose salt concentration and pH are relevant for biological applications.[46] These first analyses were performed at 20 °C. The gel-to-fluid transition temperatures of DMPC and DMPC-d54 are 24 and 20 °C,[71,72] respectively, and the acute control of the bicelle size and morphology is a key parameter in these studies. Therefore, we decided to explore structural modification of the isotropic bicelles depending on temperature since this was previously achieved in other systems.[34,36]

Thermal Effects on Bicelle Morphology and Contrast-Matching. To ensure that the studied membrane mimic objects were disk-shaped, we performed a screening of temperature effects on both the h- and d-bicelles using DLS. Measurements were performed from 15 to 37 °C (Figure 4A).

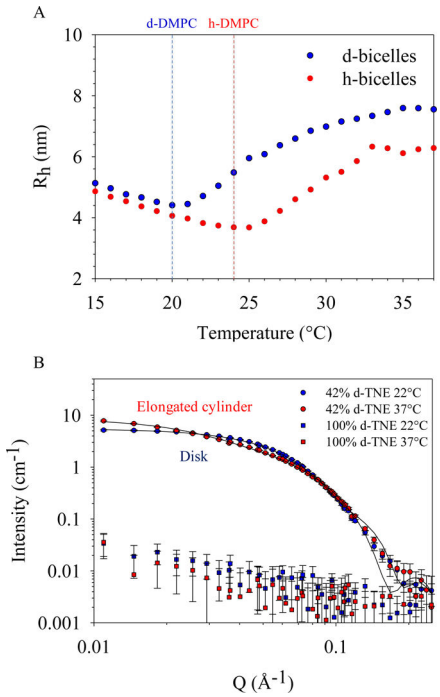


Figure 4. (A) DLS analysis of h- (red) and d-bicelles (blue) with q_{eff} = 1.3 bicelles at 50 mM from 15 to 37 °C in d-TNE. Dotted lines indicate the temperature transition for d- and h- DMPC (blue and red, respectively). (B) SANS intensities measured for q_{eff} = 1.3 d-bicelles (50 mM) at 22 (blue) and 37 °C (red) in 42% (circles) and 100% (squares) d-TNE. The data were fitted (black lines) with the "core-shell cylinder" model. At 22 °C, the thickness (t) is 40 Å and the radius (R) is 42.5 Å, in line with the dimensions of a disk, while at 37 °C, the radius is 26 Å and the thickness is 192 Å, corresponding to an elongated cylinder.

The figure highlights the hydrodynamic radius (R_h) variation of the bicelles according to temperature. For the h-bicelles, R_h decreases from 4.9 to 3.7 nm when the temperature rises from 15 to 24 °C (T_m) and then progressively increases up to a plateau value of 6.2 nm at temperatures above 30 °C. The d-bicelles exhibited the same qualitative behavior, with R_h values of 5.1, 4.4, and 7.0 nm at 15, 20 (T_m), and above 30 °C. These results clearly highlight that structural modifications occur above T_m, as previously reported in other studies.[34,36,39] Next, we explored by SANS the effect of the temperature on d-bicelle morphology in 42% d-TNE at 22 °C, two degrees above the gel-to-fluid transition temperature, and at 37 °C, which is the physiological temperature (Figure 4B). In correlation with our DLS data, the morphology of d-bicelles is clearly modified with temperature. Not surprisingly, this change occurs close to T_m. At 22 °C, good fits were obtained with the "core-shell cylinder" model with a radius (R) of 42 Å and a thickness (t) of 40 Å (Figure 4B) and we can conclude that disk-shaped objects remain present at least two degrees above T_m. On the other hand, no suitable fits with the same parameters were obtained

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

for data acquired at 37 °C. Instead, at 37 °C, good fits were obtained by keeping the same model but by changing the initial parameters and unfixing the thickness of the core in order to design an elongated cylinder (Figure 4B). The calculated values for the cylinder were 26 Å for the total radius (R) and 192 Å in thickness (t). The same observations were done for the h-bicelles (data not shown). We conclude that elongated mixed-micelles do appear at this temperature, as reported in previous studies.[36,43]

As shown in Figure 4B, the SANS signal observed for the d-bicelles in 100% d-TNE is at least 2 orders of magnitude lower than that in 42% TNE. Thus, despite the change in morphology, deduced from data of d-bicelles in 42% d-TNE, we conclude that full contrast-matching is obtained for d-bicelles in 100% d-TNE.

To sum up, whatever the form adopted by the bicelles, we show that they can be contrast-matched. This novel approach may be designed to specifically probe the SANS signal of a protein in interaction with membrane lipids. The morphological modifications should, however, be considered for any further investigations, particularly if one wants to use molecular dynamics simulations to model the entire protein/bicelle complex.

Experimental Application: DYS-R11−15 in Interaction with d-Bicelles. To experimentally validate this approach and to specifically probe the signal coming from a protein bound to membrane lipids, we applied the method to a protein−lipid complex. This complex is made of an amphipathic peripheral protein bound to a bicelle. For the protein part, we used a protein fragment belonging to the dystrophin central domain DYS R11−15 (from the 11th to the 15th spectrin-like repeats) for which both stability alone in solution and strong interaction with lipids have been thoroughly described.[45−47] Although the three-dimensional structure of R11−15 alone is not accessible by NMR and XRC due to its size (60 kDa) and its flexibility, an all-atom model was recently proposed by coupling SAXS and molecular modeling.[73] For the bicellar part of the complex, we used DHPC/DMPC h-bicelles (q_{eff} = 1.3, 50 mM lipids). We examined the interaction of the protein with the h-bicelles, in the same experimental conditions as for further SANS analysis, by performing intrinsic tryptophan fluorescence measurements and chromatographic coelution (Figures S3B and S3C). Then, we analyzed the DYS R11−15 protein alone or in interaction with the d-bicelles by SANS (Figure 5). We used exactly the same buffer and the same batch of bicelles in the samples of bicelles alone and DYS R11−15/bicelle complexes. Data were recorded at 93 μM (5.6 g/L) of DYS R11−15 in 100% d-TNE buffer at 18 °C. The specific SANS data were obtained following buffer or d-bicelles signal subtraction for the protein alone or in the presence of d-bicelles, respectively, and after incoherent background subtraction. In experiments designed to probe three-dimensional modifications of peripheral proteins bound to lipids, it is fundamental to avoid any signal coming from the peripheral proteins alone in solution. Within the concentrations used in our experiments presented in Figure 5, each protein in solution is bound to at least one bicelle, as concluded from chromatographic coelution (Figure S3C). SANS raw data of the dystrophin in absence and presence of bicelles are perfectly superimposed, as shown in the inset on Figure 5. From the Guinier approximation, the radii of gyration are 56.2 ± 0.8 and 55.9 ± 0.8 Å for the protein alone and for the protein in the presence of contrast-matched d-bicelles, respectively. In the case of large conformational changes due to lipid binding, R_g would have been affected and the scattering curves would be different. Under the conditions of this experiment, we conclude that the native tertiary structure of the protein is maintained when bound to bicelles or that no conformational modification is detectable by SANS. A study of conformational modification of dystrophin fragments bound to bicelles in other conditions is beyond the scope of the present paper. To determine whether the method would be suitable for experiments requiring higher amounts of bicelles, in additional experiments, more concentrated samples of d-bicelles alone were also analyzed. We observed that their residual signal remains negligible (Figure S2). According to the IBM and using a CBC of 6 mM, bicelles with q_{eff} = 1.3 at 50 mM, q_{eff} = 1.1 at 100 mM, and q_{eff} = 1 at 225 mM of lipids correspond approximately to 130, 350, and 920 μM of bicellar objects, respectively. These concentrations of bicellar objects seem large enough to ensure that all proteins are bound to at least one contrast-matched bicelle in the classical protein concentration range used in SANS (5−10 g/L).

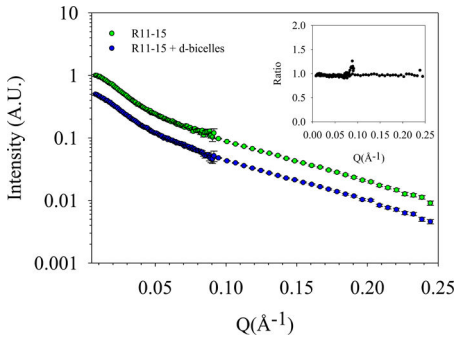


Figure 5. Normalized SANS intensities measured for DYS R11−15 at 93 μM (5.6 g/L) alone (green) or in the presence of the d-bicelles (blue) in 100% D_2O d-TNE at 18 °C. Insert: the ratio of the two scattering curves. The blue curve is shifted for clarity since the two curves superimpose, as shown by the ratio of the raw data in the inset.

Going Further in Bicelle Analysis: Bicelles in HPSEC-SANS. All the SANS data presented above were acquired in a standard sample environment (Hellma cell) since the DYS-R11−15 is a quite stable protein. Some less stable membrane proteins tend to form aggregates that would dramatically affect the SANS signal. To overcome this, we attempted to explore the behavior of bicelles in a size exclusion chromatography system, with the aim of targeting the SANS acquisition on the appropriate membrane mimic system, either alone or putatively in interaction with a partner. The results in Figure 6 show that the DMPC/DHPC h-bicelles are compatible with an online HPLC system such as the one available on D22[58] (ILL, Grenoble) since they are eluted as a single peak. The elution of bicelles can be followed thanks to their absorbance at 210 nm and does not need the incorporation of a probe as previously described with liquid chromatography (LC) experiments performed on DMPC/DPC (dodecylphosphocholine) bicelles[18] (the new ILL-D22 SEC-SANS set up enables to record absorbance at up to 4 different wavelengths). Interestingly, SANS data acquired on samples from the top of the elution peak were perfectly fitted with the "core-shell cylinder" model,[64] highlighting that bicelles maintain their disk-shaped conformation after traveling through the HPSEC column. We could not analyze the d-bicelles in the same way because of the prohibitive cost of the d-DHPC-supplemented elution buffer.

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

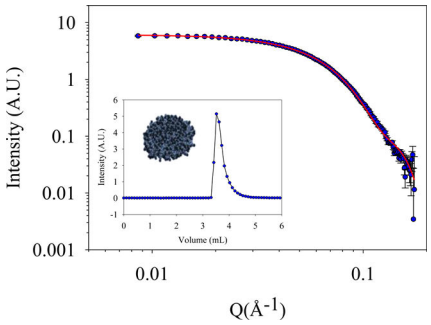


Figure 6. HPSEC-SANS intensities measured for h-bicelles fitted (red line) with the expected "core-shell cylinder" model as for the same sample in the Hellma cell environment. Insert: the associated scattergram (SANS detector) as a function of elution volume, showing one unique peak corresponding to h-bicelles.

Using h-DHPC instead does not allow the d-bicelles to remain contrast-matched due to the rapid exchange between the d-DHPC from the bicelle rim and the free h-DHPC present in the buffer (data not shown). Nevertheless, supplementing the buffer with d-DHPC should allow the d-bicelles to remain contrast-matched throughout an HPSEC-SANS measurement. Thus, this system could be applied to the characterization of other peripheral proteins and even integral membrane proteins in a bicellar environment, as previously done in NMR and XRC.

■ CONCLUSION

In the present paper, we show that isotropic hydrogenated and deuterated bicelles can be characterized by SANS and finely controlled in size for biological applications. We demonstrate that the isotropic d-bicelles can be contrast-matched in SANS experiments using a 100% D_2O-based buffer at physiological pH and salt concentrations. Moreover, we highlight that the signal coming from a peripheral protein bound to the d-bicelles can be separately and specifically probed by SANS. As a reminder, deuterated phospholipids used in the present study are commercially available enabling an easy preparation of d-bicelles with a determined size. Thereby, our method may represent an advantage compared to the deuterated nanodiscs, for which MSP purification in deuterated form might be tricky. These contrast-matched d-bicelles give rise to a wide range of biological applications to separately and specifically probe the solution structure of peripheral proteins and even fully integral membrane proteins in interaction with membrane lipids.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.langmuir.7b01369.

Summary table of the parameters used for the fitting procedure; equations describing the"core-shell cylinder" model; experimental and theoretical matchpoints, theoretical SANS intensities of the "core-shell cylinder" model in various % D_2O; SANS data of h- and d-bicelles at several concentrations; SDS-PAGE analysis of DYS R11−15, tryptophan fluorescence spectra of the protein with or without h-bicelles and chromatographic coelution analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sophie.combet@cea.fr.

ORCID

Raphael Dos Santos Morais: 0000-0002-8500-876X

Anne Martel: 0000-0002-1232-8519

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the Association Française contre les Myopathies−Telethon, the Conseil Régional de Bretagne, the Laboratoire Léon-Brillouin (LLB, Saclay, France), and the Synchrotron SOLEIL (Saint-Aubin, France) for funding this project. The Laboratoire Léon-Brillouin, the Jülich Centre for Neutron Science (JCNS) at the Heinz Maier-Leibnitz Zentrum (MLZ, Garching, Germany), and the Institut Laue-Langevin (ILL, Grenoble, France) neutron facilities are acknowledged for beamtime allocation and support on, respectively, PACE, KWS-1, and D22 SANS instruments. SANS experiments at JCNS-MLZ have been supported by the European Commission under the seventh Framework Programme through the "Research Infrastructures" action of the "Capacities" Programme, NMI3-II Grant number 283883. The authors acknowledge, at the University of Rennes 1, the Centre de Purification des Protéines of the UMR-CNRS 6290, the PRISM platform for NMR facilities, and BIOSIT for spectrofluorimetry facilities. The authors are also grateful to Giulio Gambarota (Université de Rennes 1) for his reading of the manuscript. Schemes of the proteins in the graphical abstract were drawn from MyDomains—Image Creator. The bicelle models were generated by molecular modeling.[74,75]

■ REFERENCES

(1) Fagerberg, L.; Jonasson, K.; von Heijne, G.; Uhlén, M.; Berglund, L. Prediction of the Human Membrane Proteome. *Proteomics* **2010**, *10* (6), 1141−1149.

(2) Membrane Proteins of Known Structure, http://blanco.biomol.uci.edu/mpstruc/.

(3) Seddon, A. M.; Curnow, P.; Booth, P. J. Membrane Proteins, Lipids and Detergents: Not Just a Soap Opera. *Biochim. Biophys. Acta, Biomembr.* **2004**, *1666* (1−2), 105−117.

(4) Petoukhov, M. V.; Svergun, D. I. Analysis of X-Ray and Neutron Scattering from Biomacromolecular Solutions. *Curr. Opin. Struct. Biol.* **2007**, *17* (5), 562−571.

(5) Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. A Brief Survey of State-of-the-Art BioSAXS. *Protein Pept. Lett.* **2016**, *23* (3), 217−231.

(6) Heller, W. T. Small-Angle Neutron Scattering and Contrast Variation: A Powerful Combination for Studying Biological Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (11), 1213−1217.

(7) Breyton, C.; Gabel, F.; Lethier, M.; Flayhan, A.; Durand, G.; Jault, J.-M.; Juillan-Binard, C.; Imbert, L.; Moulin, M.; Ravaud, S.; et al. Small Angle Neutron Scattering for the Study of Solubilised Membrane Proteins. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2013**, *36* (7), 71.

(8) Gabel, F.; Lensink, M. F.; Clantin, B.; Jacob-Dubuisson, F.; Villeret, V.; Ebel, C. Probing the Conformation of FhaC with Small-Angle Neutron Scattering and Molecular Modeling. *Biophys. J.* **2014**, *107* (1), 185−196.

(9) Frey, L.; Lakomek, N.-A.; Riek, R.; Bibow, S. Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics. *Angew. Chem., Int. Ed.* **2017**, *56* (1), 380−383.

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

(10) Dörr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schäfer, M.; van Walree, C. A.; Killian, J. A. The Styrene-Maleic Acid Copolymer: A Versatile Tool in Membrane Research. *Eur. Biophys. J.* **2016**, *45* (1), 3−21.
(11) Vestergaard, M.; Kraft, J. F.; Vosegaard, T.; Thøgersen, L.; Schiøtt, B. Bicelles and Other Membrane Mimics: Comparison of Structure, Properties, and Dynamics from MD Simulations. *J. Phys. Chem. B* **2015**, *119* (52), 15831−15843.
(12) Warschawski, D. E.; Arnold, A. A.; Beaugrand, M.; Gravel, A.; Chartrand, É.; Marcotte, I. Choosing Membrane Mimetics for NMR Structural Studies of Transmembrane Proteins. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808* (8), 1957−1974.
(13) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2* (8), 853−856.
(14) Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y.-P.; Dafforn, T.; Overduin, M. Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *J. Am. Chem. Soc.* **2009**, *131* (22), 7484−7485.
(15) Maric, S.; Skar-Gislinge, N.; Midtgaard, S.; Thygesen, M. B.; Schiller, J.; Frielinghaus, H.; Moulin, M.; Haertlein, M.; Forsyth, V. T.; Pomorski, T. G.; et al. Stealth Carriers for Low-Resolution Structure Determination of Membrane Proteins in Solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70* (Pt 2), 317−328.
(16) Wang, X.; Mu, Z.; Li, Y.; Bi, Y.; Wang, Y. Smaller Nanodiscs Are Suitable for Studying Protein Lipid Interactions by Solution NMR. *Protein J.* **2015**, *34* (3), 205−211.
(17) Grinkova, Y. V.; Denisov, I. G.; Sligar, S. G. Engineering Extended Membrane Scaffold Proteins for Self-Assembly of Soluble Nanoscale Lipid Bilayers. *Protein Eng., Des. Sel.* **2010**, *23* (11), 843−848.
(18) Draney, A. W.; Smrt, S. T.; Lorieau, J. L. Use of Isotropically Tumbling Bicelles to Measure Curvature Induced by Membrane Components. *Langmuir* **2014**, *30* (39), 11723−11733.
(19) Zimmerberg, J.; Kozlov, M. M. How Proteins Produce Cellular Membrane Curvature. *Nat. Rev. Mol. Cell Biol.* **2006**, *7* (1), 9−19.
(20) Mazer, N. A.; Benedek, G. B.; Carey, M. C. Quasielastic Light-Scattering Studies of Aqueous Biliary Lipid Systems. Mixed Micelle Formation in Bile Salt-Lecithin Solutions. *Biochemistry* **1980**, *19* (4), 601−615.
(21) Sanders, C. R.; Prestegard, J. H. Magnetically Orientable Phospholipid Bilayers Containing Small Amounts of a Bile Salt Analogue, CHAPSO. *Biophys. J.* **1990**, *58* (2), 447−460.
(22) Bian, J.; Roberts, M. F. Phase Separation in Short-Chain Lecithin/Gel-State Long Chain Lecithin Aggregates. *Biochemistry* **1990**, *29* (34), 7928−7935.
(23) Lin, T. L.; Liu, C. C.; Roberts, M. F.; Chen, S. H. Structure of Mixed Short-Chain Lecithin/Long-Chain Lecithin Aggregates Studied by Small-Angle Neutron Scattering. *J. Phys. Chem.* **1991**, *95* (15), 6020−6027.
(24) Minto, R. E.; Adhikari, P. R.; Lorigan, G. A. A 2H Solid-State NMR Spectroscopic Investigation of Biomimetic Bicelles Containing Cholesterol and Polyunsaturated Phosphatidylcholine. *Chem. Phys. Lipids* **2004**, *132* (1), 55−64.
(25) Triba, M. N.; Devaux, P. F.; Warschawski, D. E. Effects of Lipid Chain Length and Unsaturation on Bicelles Stability. A Phosphorus NMR Study. *Biophys. J.* **2006**, *91* (4), 1357−1367.
(26) Struppe, J.; Whiles, J. A.; Vold, R. R. Acidic Phospholipid Bicelles: A Versatile Model Membrane System. *Biophys. J.* **2000**, *78* (1), 281−289.
(27) Liebau, J.; Pettersson, P.; Zuber, P.; Ariöz, C.; Mäler, L. Fast-Tumbling Bicelles Constructed from Native Escherichia Coli Lipids. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858* (9), 2097−2105.
(28) Sanders, C. R.; Schwonek, J. P. Characterization of Magnetically Orientable Bilayers in Mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphosphatidylcholine by Solid-State NMR. *Biochemistry* **1992**, *31* (37), 8898−8905.
(29) Dürr, U. H. N.; Gildenberg, M.; Ramamoorthy, A. The Magic of Bicelles Lights up Membrane Protein Structure. *Chem. Rev.* **2012**, *112* (11), 6054−6074.
(30) Dürr, U. H. N.; Soong, R.; Ramamoorthy, A. When Detergent Meets Bilayer: Birth and Coming of Age of Lipid Bicelles. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *69*, 1−22.
(31) Poulos, S.; Morgan, J. L. W.; Zimmer, J.; Faham, S. Bicelles Coming of Age: An Empirical Approach to Bicelle Crystallization. *Methods Enzymol.* **2015**, *557*, 393−416.
(32) Ujwal, R.; Bowie, J. U. Crystallizing Membrane Proteins Using Lipidic Bicelles. *Methods* **2011**, *55* (4), 337−341.
(33) Ujwal, R.; Abramson, J. High-Throughput Crystallization of Membrane Proteins Using the Lipidic Bicelle Method. *J. Visualized Exp.* **2012**, No. No. 59, e3383.
(34) van Dam, L.; Karlsson, G.; Edwards, K. Morphology of Magnetically Aligning DMPC/DHPC Aggregates-Perforated Sheets, Not Disks. *Langmuir* **2006**, *22* (7), 3280−3285.
(35) Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R. Structural Evaluation of Phospholipid Bicelles for Solution-State Studies of Membrane-Associated Biomolecules. *Biophys. J.* **2001**, *81* (4), 2163−2171.
(36) van Dam, L.; Karlsson, G.; Edwards, K. Direct Observation and Characterization of DMPC/DHPC Aggregates under Conditions Relevant for Biological Solution NMR. *Biochim. Biophys. Acta, Biomembr.* **2004**, *1664* (2), 241−256.
(37) Beaugrand, M.; Arnold, A. A.; Hénin, J.; Warschawski, D. E.; Williamson, P. T. F.; Marcotte, I. Lipid Concentration and Molar Ratio Boundaries for the Use of Isotropic Bicelles. *Langmuir* **2014**, *30* (21), 6162−6170.
(38) Mineev, K. S.; Nadezhdin, K. D.; Goncharuk, S. A.; Arseniev, A. S. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir* **2016**, *32*, 6624.
(39) Ye, W.; Lind, J.; Eriksson, J.; Mäler, L. Characterization of the Morphology of Fast-Tumbling Bicelles with Varying Composition. *Langmuir* **2014**, *30* (19), 5488−5496.
(40) Luchette, P. A.; Vetman, T. N.; Prosser, R. S.; Hancock, R. E.; Nieh, M. P.; Glinka, C. J.; Krueger, S.; Katsaras, J. Morphology of Fast-Tumbling Bicelles: A Small Angle Neutron Scattering and NMR Study. *Biochim. Biophys. Acta, Biomembr.* **2001**, *1513* (2), 83−94.
(41) Nieh, M.-P.; Glinka, C. J.; Krueger, S.; Prosser, R. S.; Katsaras, J. SANS Study on the Effect of Lanthanide Ions and Charged Lipids on the Morphology of Phospholipid Mixtures. Small-Angle Neutron Scattering. *Biophys. J.* **2002**, *82* (5), 2487−2498.
(42) Nieh, M.-P.; Raghunathan, V. A.; Glinka, C. J.; Harroun, T. A.; Pabst, G.; Katsaras, J. Magnetically Alignable Phase of Phospholipid "Bicelle" Mixtures Is a Chiral Nematic Made Up of Wormlike Micelles. *Langmuir* **2004**, *20* (19), 7893−7897.
(43) Li, M.; Morales, H. H.; Katsaras, J.; Kučerka, N.; Yang, Y.; Macdonald, P. M.; Nieh, M.-P. Morphological Characterization of DMPC/CHAPSO Bicellar Mixtures: A Combined SANS and NMR Study. *Langmuir* **2013**, *29* (51), 15943−15957.
(44) Nieh, M.-P.; Raghunathan, V. A.; Pabst, G.; Harroun, T.; Nagashima, K.; Morales, H.; Katsaras, J.; Macdonald, P. Temperature Driven Annealing of Perforations in Bicellar Model Membranes. *Langmuir* **2011**, *27* (8), 4838−4847.
(45) Legardinier, S.; Raguénès-Nicol, C.; Tascon, C.; Rocher, C.; Hardy, S.; Hubert, J.-F.; Le Rumeur, E. Mapping of the Lipid-Binding and Stability Properties of the Central Rod Domain of Human Dystrophin. *J. Mol. Biol.* **2009**, *389* (3), 546−558.
(46) Sarkis, J.; Hubert, J.-F.; Legrand, B.; Robert, E.; Chéron, A.; Jardin, J.; Hitti, E.; Le Rumeur, E.; Vié, V. Spectrin-like Repeats 11−15 of Human Dystrophin Show Adaptations to a Lipidic Environment. *J. Biol. Chem.* **2011**, *286* (35), 30481−30491.
(47) Sarkis, J.; Vié, V.; Winder, S. J.; Renault, A.; Le Rumeur, E.; Hubert, J.-F. Resisting Sarcolemmal Rupture: Dystrophin Repeats Increase Membrane-Actin Stiffness. *FASEB J.* **2013**, *27* (1), 359−367.
(48) Vold, R. R.; Prosser, R. S. Magnetically Oriented Phospholipid Bilayered Micelles for Structural Studies of Polypeptides. Does the Ideal Bicelle Exist? *J. Magn. Reson., Ser. B* **1996**, *113* (3), 267−271.

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

(49) Triba, M. N.; Warschawski, D. E.; Devaux, P. F. Reinvestigation by Phosphorus NMR of Lipid Distribution in Bicelles. *Biophys. J.* **2005**, *88* (3), 1887−1901.
(50) Petrache, H. I.; Tristram-Nagle, S.; Gawrisch, K.; Harries, D.; Parsegian, V. A.; Nagle, J. F. Structure and Fluctuations of Charged Phosphatidylserine Bilayers in the Absence of Salt. *Biophys. J.* **2004**, *86* (3), 1574−1586.
(51) Träuble, H.; Haynes, D. H. The Volume Change in Lipid Bilayer Lamellae at the Crystalline-Liquid Crystalline Phase Transition. *Chem. Phys. Lipids* **1971**, *7* (4), 324−335.
(52) Small, D. M. *The Physical Chemistry of Lipids: From Alkanes to Phospholipids*; Plenum Press, 1986.
(53) Cho, C. H.; Urquidi, J.; Singh, S.; Robinson, G. W. Thermal Offset Viscosities of Liquid H2O, D2O, and T2O. *J. Phys. Chem. B* **1999**, *103* (11), 1991−1994.
(54) Auvray, L.; Lairez, D. Small Angle Neutron Scattering facility PACE; http://www-llb.cea.fr/spectros/pdf/pace-llb.pdf.
(55) Feoktystov, A. V.; Frielinghaus, H.; Di, Z.; Jaksch, S.; Pipich, V.; Appavou, M.-S.; Babcock, E.; Hanslik, R.; Engels, R.; Kemmerling, G.; et al. KWS-1 High-Resolution Small-Angle Neutron Scattering Instrument at JCNS: Current State. *J. Appl. Crystallogr.* **2015**, *48* (1), 61−70.
(56) Frielinghaus, H.; Feoktystov, A.; Berts, I.; Mangiapia, G. KWS-1: Small-Angle Scattering Diffractometer. *J. Large-Scale Res. Facil. JLSRF* **2015**, DOI: 10.17815/jlsrf-1-26.
(57) Instrument layout (D22), https://www.ill.eu/instruments-support/instruments-groups/instruments/d22/description/instrument-layout/.
(58) Jordan, A.; Jacques, M.; Merrick, C.; Devos, J.; Forsyth, V. T.; Porcar, L.; Martel, A. SEC-SANS: Size Exclusion Chromatography Combined in Situ with Small-Angle Neutron Scattering. *J. Appl. Crystallogr.* **2016**, *49* (6), 2015−2020.
(59) Petoukhov, M. V.; Franke, D.; Shkumatov, A. V.; Tria, G.; Kikhney, A. G.; Gajda, M.; Gorba, C.; Mertens, H. D. T.; Konarev, P. V.; Svergun, D. I. New Developments in the *ATSAS* Program Package for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2012**, *45* (2), 342−350.
(60) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *PRIMUS*: A Windows PC-Based System for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2003**, *36* (5), 1277−1282.
(61) Glatter, O.; Kratky, O. *Small Angle X-Ray Scattering*; Academic Press, 1982.
(62) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; Wiley, 1980.
(63) Neutron Activation Calculator, https://www.ncnr.nist.gov/resources/activation/.
(64) Kline, S. R. Reduction and Analysis of SANS and USANS Data Using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39* (6), 895−900.
(65) SasView, http://www.sasview.org/.
(66) Chou, J. J.; Baber, J. L.; Bax, A. Characterization of Phospholipid Mixed Micelles by Translational Diffusion. *J. Biomol. NMR* **2004**, *29* (3), 299−308.
(67) Nagle, J. F.; Wilkinson, D. A. Lecithin Bilayers. Density Measurement and Molecular Interactions. *Biophys. J.* **1978**, *23* (2), 159−175.
(68) Tristram-Nagle, S.; Liu, Y.; Legleiter, J.; Nagle, J. F. Structure of Gel Phase DMPC Determined by X-Ray Diffraction. *Biophys. J.* **2002**, *83* (6), 3324−3335.
(69) Kučerka, N.; Liu, Y.; Chu, N.; Petrache, H. I.; Tristram-Nagle, S.; Nagle, J. F. Structure of Fully Hydrated Fluid Phase DMPC and DLPC Lipid Bilayers Using X-Ray Scattering from Oriented Multilamellar Arrays and from Unilamellar Vesicles. *Biophys. J.* **2005**, *88* (4), 2626−2637.
(70) Gabriel, N. E.; Roberts, M. F. Short-Chain Lecithin/Long-Chain Phospholipid Unilamellar Vesicles: Asymmetry, Dynamics, and Enzymatic Hydrolysis of the Short-Chain Component. *Biochemistry* **1987**, *26* (9), 2432−2440.
(71) Guard-Friar, D.; Chen, C. H.; Engle, A. S. Deuterium Isotope Effect on the Stability of Molecules: Phospholipids. *J. Phys. Chem.* **1985**, *89* (9), 1810−1813.
(72) Wang, G. X.; Chen, C. H. Thermodynamic Elucidation of Structural Stability of Deuterated Biological Molecules: Deuterated Phospholipid Vesicles in H2O. *Arch. Biochem. Biophys.* **1993**, *301* (2), 330−335.
(73) Molza, A.-E.; Férey, N.; Czjzek, M.; Le Rumeur, E.; Hubert, J.-F.; Tek, A.; Laurent, B.; Baaden, M.; Delalande, O. Innovative Interactive Flexible Docking Method for Multi-Scale Reconstruction Elucidates Dystrophin Molecular Assembly. *Faraday Discuss.* **2014**, *169*, 45−62.
(74) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812−7824.
(75) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435−447.

■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on the Web on June 16, 2017, with errors in the Figure 1 caption. The corrected version was reposted on June 20, 2017.

DOI: 10.1021/acs.langmuir.7b01369
Langmuir 2017, 33, 6572−6580

Supporting information

Contrast-matched isotropic bicelles: a versatile tool to specifically probe the solution structure of peripheral membrane proteins using SANS.

Raphael Dos Santos Morais[1,2,3,4], Olivier Delalande[1,2], Javier Pérez[4], Liza Mouret[5], Arnaud Bondon[5], Anne Martel[6], Marie-Sousai Appavou[7], Elisabeth Le Rumeur[1,2], Jean-François Hubert[1,2], and Sophie Combet[3,].*

[1]Université de Rennes 1, F-35043 Rennes, France. [2]CNRS UMR 6290, Institut de Génétique et Développement de Rennes, F-35043 Rennes, France. [3]Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, Gif-sur-Yvette CEDEX F-91191, France. [4]SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP48, Saint-Aubin, Gif-sur-Yvette, F-91192, France. [5]CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-350043 Rennes, France. [6]Institut Laue-Langevin, F-38042 Grenoble, France. [7]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany.

Table S1: Summary of the parameters used for the fitting procedure.

Name	Formula	MW (g/mol)	Volume (Å^3)	SANS SLD (10^{-6} Å^{-2})	SANS contrast in 100% D_2O (10^{-6} Å^{-2})	SANS contrast in 42% D_2O (10^{-6} Å^{-2})
h-DMPC	C36 H72 N O8 P	678	1090	0.28	-6.06	
Head	C10 O8 P N H14	308	336	2.24	-4.10	
Tail	C26 H58	370	754	-0.58	-6.92	
h-DHPC	C20 H40 N O8 P	454	660	0.67	-5.67	
Head	C10 O8 P N H14	308	336	2.24	-4.10	
Tail	C10 H26	146	324	-0.95	-7.29	
d-DMPC	C36 H5 N O8 P D67	745	1090	6.65		4.31
Head	C10 O8 P N H5 D9	317	336	5.05		2.71
Tail	C26 D58	428	754	7.39		5.05
d-DHPC	C20 H5 N O8 P D35	489	660	6.20		3.86
Head	C10 O8 P N H5 D9	317	336	5.05		2.71
Tail	C10 D26	172	324	7.38		5.04

The SLD of 42% (corresponding to the SANS matching point of most hydrogenated proteins) and 100% (corresponding to the SANS matching point of d-bicelles) D_2O are 2.34 and 6.34.10^{-6} Å^{-2} respectively.

Description of the "core-shell cylinder" form factor model:

$$P(Q,\alpha) = \frac{scale}{V_s} f^2(Q) + bkg \qquad \text{Eq. S1}$$

where

$$f(Q) = \frac{2(\rho_c - \rho_s)V_c \sin\left[Qt_c \cos\frac{\alpha}{2}\right]}{\left[Qt_c \cos\frac{\alpha}{2}\right]\frac{J_1(qr_c \sin\alpha)}{(Qr_c \sin\alpha)}} + \frac{2(\rho_s - \rho_{solv})V_s \sin\left[Q(t_c + 2t_s)\cos\frac{\alpha}{2}\right]}{\left[Q(t_c + 2t_s)\cos\frac{\alpha}{2}\right]\frac{J_1(Q(r_c + t_s)\sin\alpha)}{(Q(r_c + t_s)\sin\alpha)}} \qquad \text{Eq. S2}$$

and

$$V_s = \pi(r_c + t_s)^2(t_c + 2t_s) \qquad \text{Eq. S3}$$

where α is the angle between the axis of the cylinder and the Q-vector, V_s is the volume of the outer shell (*i.e.* the total volume, including the shell), V_c is the volume of the core, t_c is the thickness of the core, r_c is the radius of the core, t_s is the thickness of the shell, ρ_c, ρ_s, and ρ_{solv} are the scattering length densities of the core, the shell, and the solvent, respectively, and bkg is the background level. The outer radius of the shell is given by $r_c + t_s$ and the total length of the outer shell is given by $t_c + 2t_s$. J_1 is the first order Bessel function.

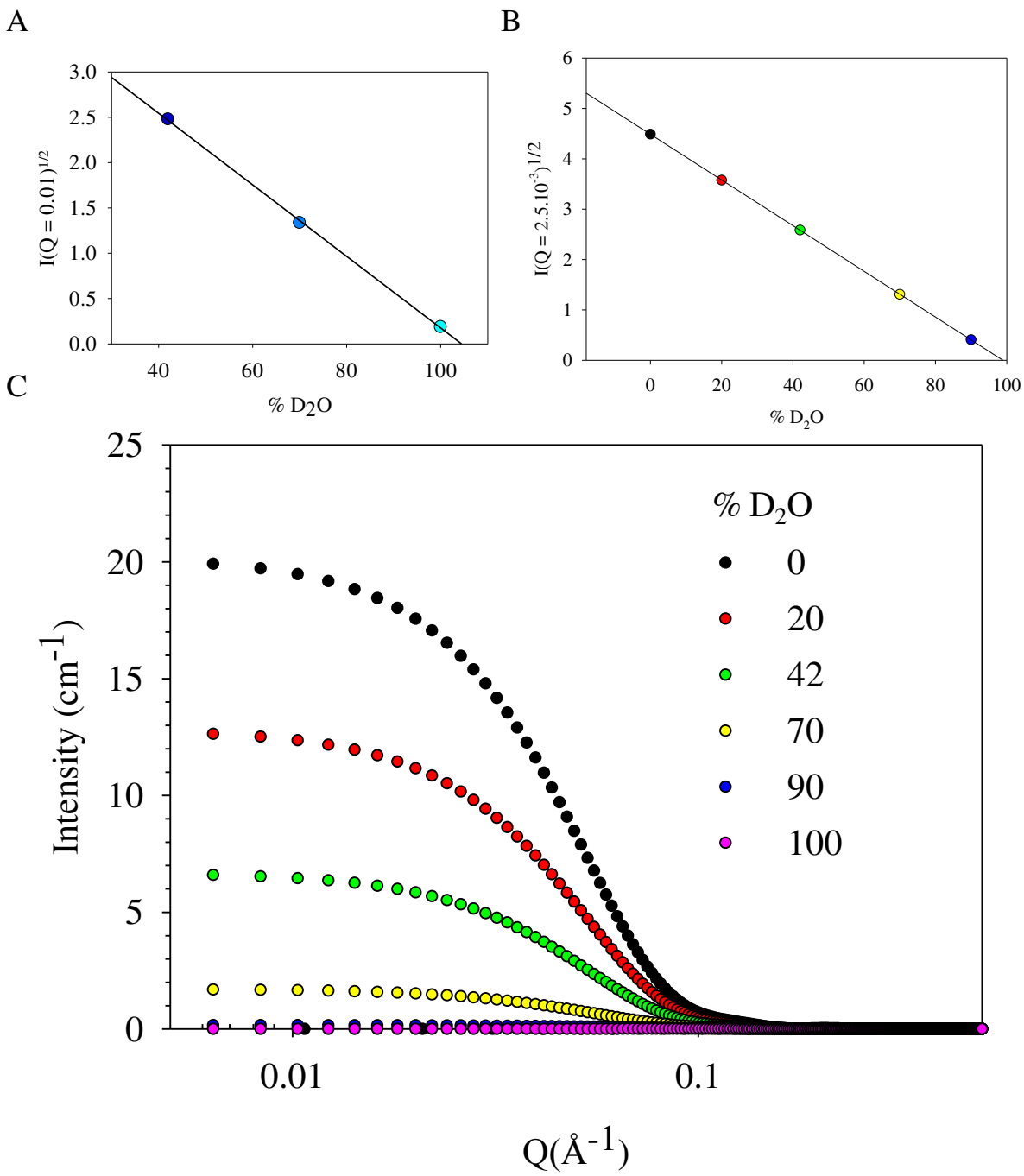


Figure S1: (A) Square root of I(Q = 0.01 Å^{-1}) *versus* the percentage of D_2O plot obtained with the experimental data of d-bicelles in 42, 70, and 100% D_2O (see Figure 3). Experimental match-point is 104% D_2O. (B) Square root of I(Q = 2.5 10^{-3} Å^{-1}) *versus* the percentage of D_2O plot obtained with theoretical data generated with the "core-shell cylinder" model. Theoretical match-point is 99% D_2O. (C) Simulated SANS intensities of the "core-shell cylinder" model in 0 (black), 20 (red), 42 (green), 70 (yellow), 90 (dark blue), and 100 (magenta) % D_2O.

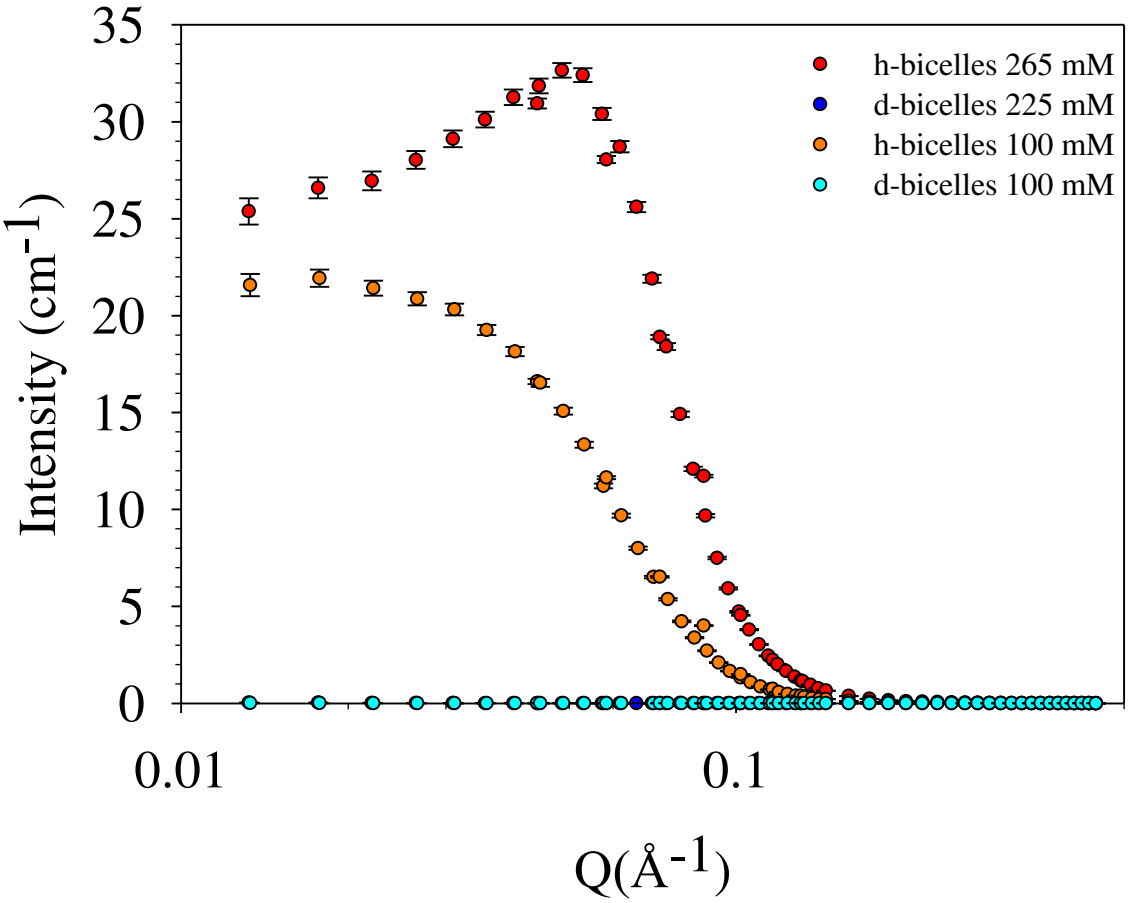


Figure S2: SANS intensities measured for q = 1 h-bicelles at 20°C for a total DMPC/DHPC concentration of 265 (q_{eff} = 1.05) (red) and 100 (q_{eff} = 1.1) (orange) mM and for q = 1 contrast-matched d-bicelles at 225 (q_{eff} = 1.05) (dark blue) and 100 (q_{eff} = 1.1) (cyan) mM in 100% D_2O d-TNE buffer. At 265 mM, the structure peak is observed for h-bicelles. The contrast-matched d-bicelles are superimposed and close to zero intensity.

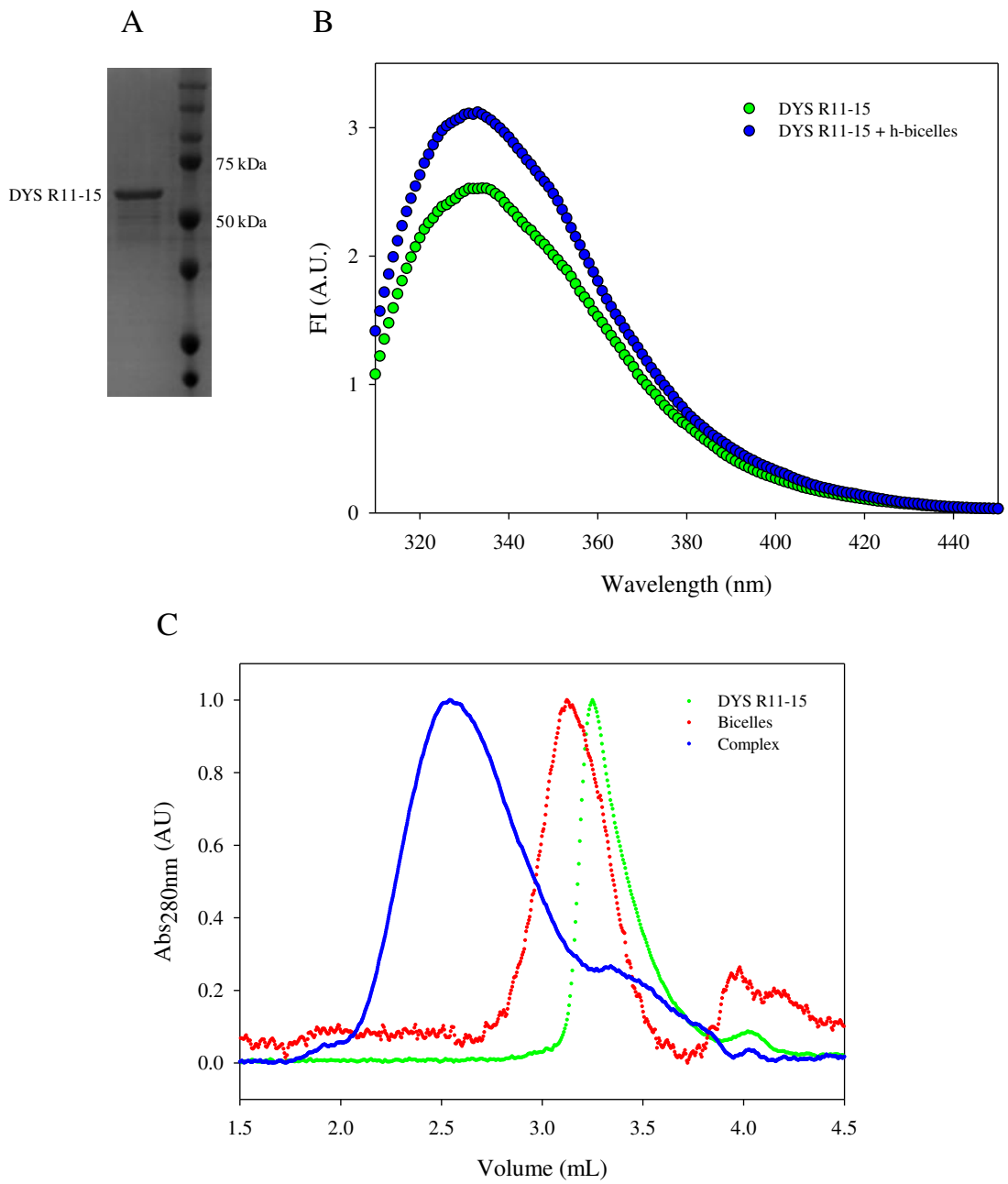


Figure S3: (A): SDS-PAGE analysis of DYS-R11-15, assessing the purity and the expected molecular weight of the protein. (B): Tryptophan intrinsic fluorescence variation of DYS-R11-15 alone (green) or in the presence of h-bicelles (blue). Fluorescence intensity variation is attributed to a modification in the tryptophan environment due to lipid binding. (C) Chromatographic (HPLC) co-elution (Abs_{280nm} detector) of DYS-R11-15 bound to h-bicelles (blue) versus the elutions of h-bicelles (red) or DYS R11-15 (green) alone. Top of the peaks was normalized to the unity for clarity.

3. Résultats complémentaires

3.1. Micelles de DHPC et liposomes de DMPC

Avant l'utilisation exclusive des bicelles comme modèle membranaire, nous nous sommes intéressés aux structures que peuvent adopter les deux composants lipidiques (DHPC et DMPC) des bicelles, seuls en solution, à savoir la formation de micelles ou de liposomes. Les micelles de DHPC ont une CMC de 14-15 mM [76]. Au-dessus de cette concentration, les molécules de DHPC s'organisent en micelles en forme d'ellipsoïdes avec un nombre d'agrégation d'environ 35 [76]. Nous avons réalisé des expériences de SAXS et SANS pour vérifier la morphologie adoptée par les micelles. Pour les deux jeux de données, le plateau atteint pour de petites valeurs de Q est caractéristique d'objets de taille finie, ce qui permet d'exclure la présence d'agrégats de micelles (Figure 53A). Le rayon de giration obtenu est de 12,5 Å pour les données SANS et de 20 Å pour les données SAXS. Cette différence s'explique par la différence de contraste (Tableau 6). Le fit des données avec un modèle « core-shell » d'ellipsoïde permet d'obtenir des dimensions en adéquation avec celles précédemment rapportées dans la littérature [232,233]. Les micelles de détergents sont largement utilisées pour solubiliser des protéines membranaires intégrales et ainsi former des complexes protéine/détergents. Cependant, le modèle « micelle » de DHPC n'est pas adapté à notre problématique. En effet, bien que possédant des têtes PC, ces objets sont de trop petite taille par rapport à la taille des deux fragments protéiques ciblés dans cette étude et ont l'inconvénient de ne pas présenter une bicouche lipidique.

Les SUV de DMPC ont été obtenues par extrusion de MLV à travers des pores de 50 nm de diamètre. Les données SANS ont été obtenues pour une concentration de 10 mM de lipides. Contrairement aux données obtenues pour les micelles, celles des SUV ne permettent pas de déterminer une valeur de R_g dans la gamme en Q utilisée. Les données sont typiques de SUV obtenues pour les mêmes lipides [234] (Figure 53B). Le système de liposomes est l'un des modèles les plus utilisés dans l'étude des interactions protéine/lipide. Il ne permet cependant pas de répondre aux attentes soulevées dans le cadre de notre problématique. Bien qu'ayant l'avantage de posséder une bicouche lipidique, ce système présente, en effet, un degré de polydispersité trop élevé pour permettre une utilisation en SAS. De plus, la taille relativement importante des SUV implique un grand nombre de lipides et donc un grand nombre d'atomes, ce qui n'est pas à ce jour compatible avec des temps de calculs raisonnables de simulation de complexes protéine/lipide, même dans le cadre d'une dynamique moléculaire gros-grain.

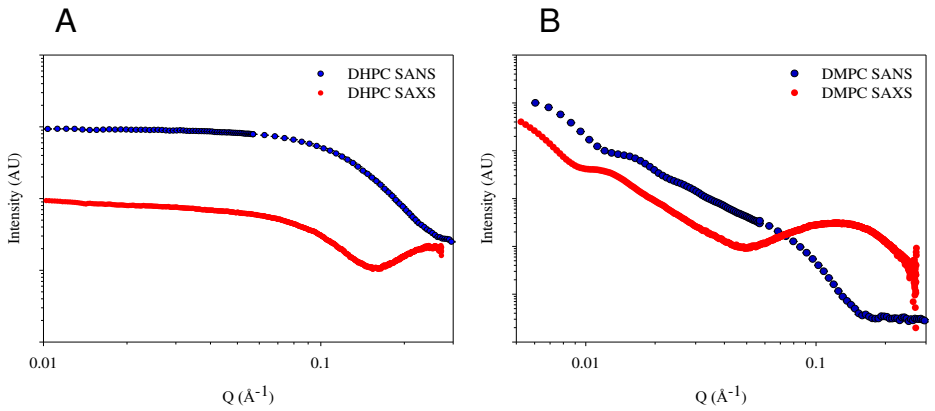


Figure 53 : Données SANS (en bleu) et SAXS (en rouge) obtenues pour des micelles de DHPC (A) et des SUV de DMPC (B).

3.2. Analyse complémentaire des bicelles par SANS

En complément de l'article 1, nous avons analysé des bicelles de DMPC/DHPC à différents ratios pour une même concentration (Figure 54A). Comme attendu, la taille des bicelles augmente selon q_{eff}. En effet, *I(0)* varie en fonction de ce ratio, ce qui indique que les bicelles augmentent de taille. Nous présentons également ici les données SANS obtenues entre 10 et 25 mM de lipides totaux pour q = 1, en complément de celles obtenues dans l'article 1 où la morphologie des bicelles commence à être modifiée entre 25 et 50 mM (Figure 54B).

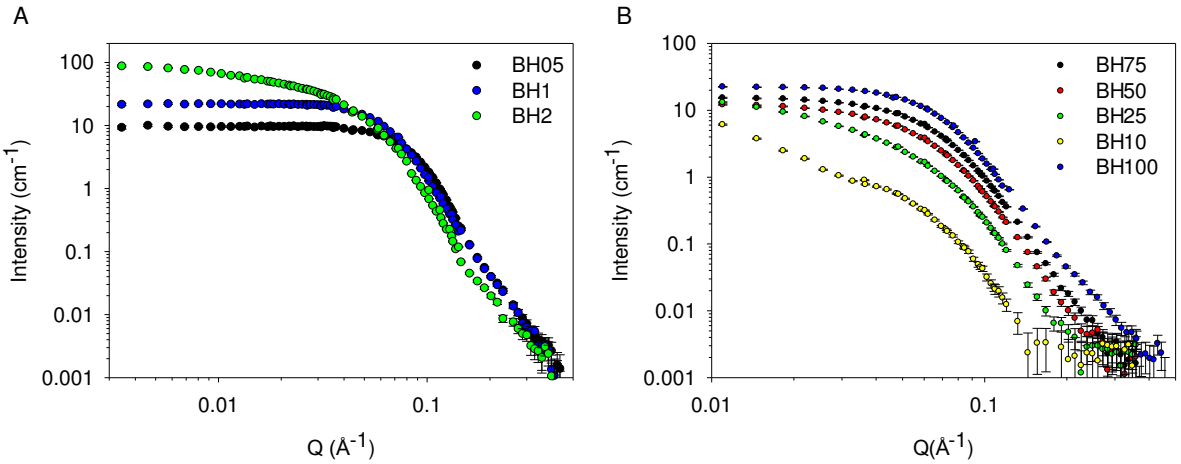


Figure 54 : Analyse par SANS. (A) Variation de la taille des bicelles hydrogénées (BH) selon q. Données SANS obtenues pour 100 mM de lipides totaux pour un ratio q = 0.5 (noir), 1 (bleu) et 2 (vert). (B) Modification de la taille et de la morphologie des bicelles en fonction de q_{eff}. La concentration en lipides est de 100 (bleu), 75 (noir), 50 (rouge), 25 (vert) et 10 (jaune) mM. Entre 25 et 50 mM, la morphologie en forme de disque n'est plus présente et l'apparition de vésicules à lieu à partir de 10 mM.

En complément des études en température menées dans l'article (Figure 4 B, article 1), nous avons obtenu des données sur les bicelles zwitterioniques (DMPC/DHPC) à trois concentrations en lipides totaux (100, 75 et 50 mM) pour q = 1. Nous avons étendu l'analyse en température pour les bicelles deutérées anioniques (contenant de la DMPS) pour une concentration de 50 mM de lipides. Pour tous les systèmes considérés, la morphologie discoïdale des bicelles est maintenue pour une température d'au moins 2°C au-dessus du *Tm* de la DMPC. De plus, comme présenté dans l'article (Figure 4B, article 1), à 37°C, la morphologie obtenue par ajustement des données est celle de cylindres allongés.

Comme cela a été décrit dans l'article 1, les bicelles sont compatibles avec un dispositif HPLC-SEC. A notre connaissance, l'article de Draney *et al.*[77] est le seul utilisant la SEC pour caractériser les bicelles. Cependant, dans cet article, les expériences ont été réalisées sur des bicelles DMPC/DPC à basse pression. Dans notre cas, en complément des expériences de SEC-SANS réalisées sur les bicelles zwitterioniques présentées dans l'article (Figure 6, article 1), nous avons également analysé des bicelles anioniques par HPLC-SEC. Une morphologie discoïdale est obtenue avec ce type de bicelles, qui sont donc compatibles avec le système HPLC-SEC (Figure 55A). De plus, nous avons vérifié les conséquences de ne pas ajouter 6 mM de DHPC dans le tampon d'élution pour l'analyse des bicelles par SEC. Comme attendu, la dilution induite par la colonne conduit à une forte augmentation de q_{eff} et donc à l'apparition de très gros objets lipidiques de types MLV (Figure 55B).

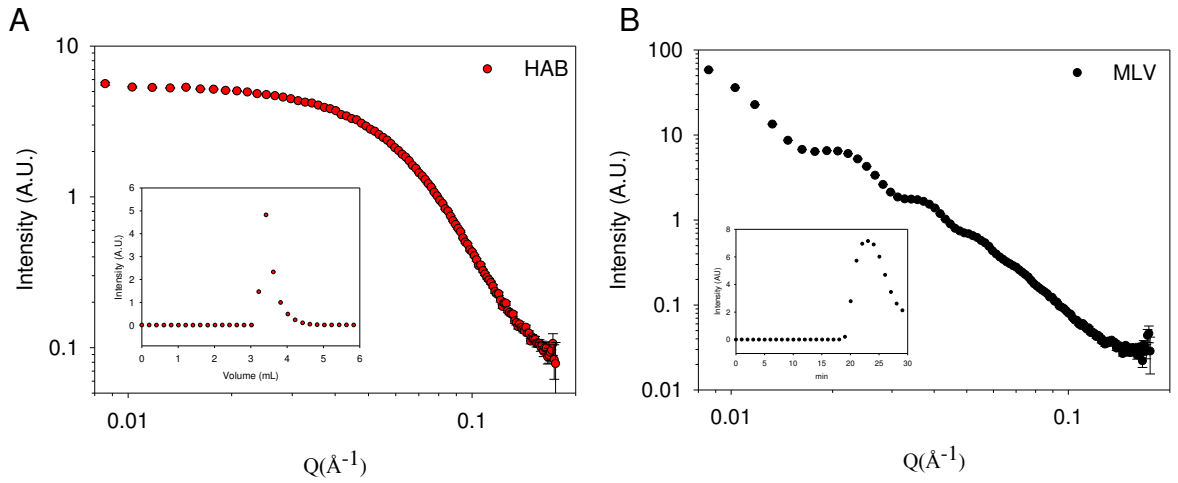


Figure 55 : (A) Analyse SEC-SANS des bicelles anioniques montrant le maintien de la morphologie discoïdale. (B) Modification de la morphologie des bicelles lorsque la colonne SEC n'est pas équilibrée avec du tampon contenant 6 mM de DHPC. Insert : Chromatogrammes associés avant exposition aux neutrons.

3.3. Analyse complémentaire des bicelles par SAXS

Des expériences de SAXS ont été réalisées sur les bicelles hydrogénées zwitterioniques (HZB) ou anioniques (HAB). Les expériences ont été réalisées en injection directe (pipetage par un robot) ou en utilisant le système HPLC disponible sur la ligne SWING [216]. Contrairement aux données SANS, aucun plateau n'est atteint dans la gamme en *Q* utilisée. La présence d'une remontée pour des petites valeurs de *Q* est significative de la présence de gros objets (Figure 56B). Pourtant, le chromatogramme en amont de la cellule SAXS atteste que les bicelles ne sont pas de gros objets puisqu'elles sont éluées dans la gamme de séparation de la colonne et ne se trouvent pas dans le volume d'exclusion (Figure 56A).

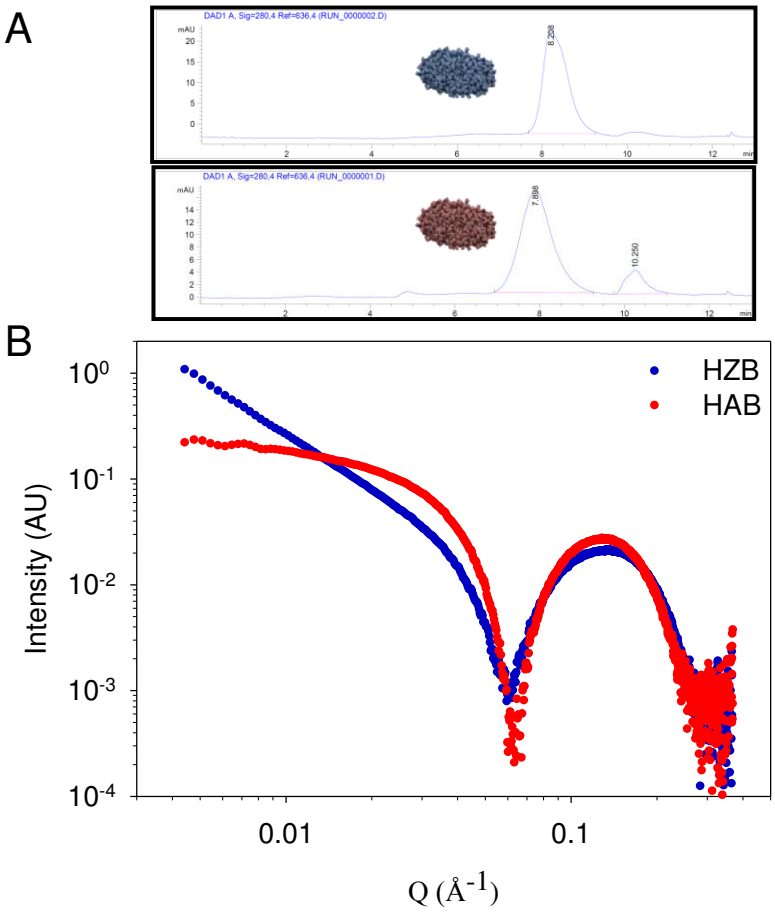


Figure 56 : (A) Profils de SEC des HZB (bleu) ou HAB (rouge) en sortie de colonne HPLC (détecteur UV). (B) Courbes SEC-SAXS correspondantes.

Le système HPLC ne peut pas être mis en cause, puisque nous avons démontré que les bicelles sont compatibles avec le système SEC-SANS disponible sur l'instrument D22 [217] (ILL). De plus, dans l'article 2 (voir chapitre suivant), nous avons montré que les bicelles peuvent être caractérisées par SEC-MALS et que la MW obtenue est en accord avec l'IBM. Afin de

comprendre à quoi est due la présence de cette remontée du signal, plusieurs hypothèses ont été testées sur les mesures de SAXS/SEC-SAXS : température (10, 15, 20 et 25°C), concentration en lipides, débit de pipetage, radicaux libres (utilisation de DTT pour les piéger), matière du capillaire (quartz ou Kapton® (polyimide)), puissance du faisceau, surface d'illumination, temps d'exposition et également un effet possible entre l'utilisation de D_2O ou de H_2O. De plus, pour certaines expériences, l'échantillon restait collé au capillaire, au moins en partie. Ce problème de salissure du capillaire a parfois empêché une bonne soustraction du pseudo-tampon. Notre dernière investigation pour comprendre ce phénomène a été menée en utilisant un détecteur DLS *in situ*. Afin de mieux comprendre ce phénomène, des expériences de DLS ont été réalisées avant (SEC-DLS-SAXS), puis après (SEC-SAXS-DLS), exposition aux rayons X afin de vérifier l'intégrité des bicelles (Figure 57).

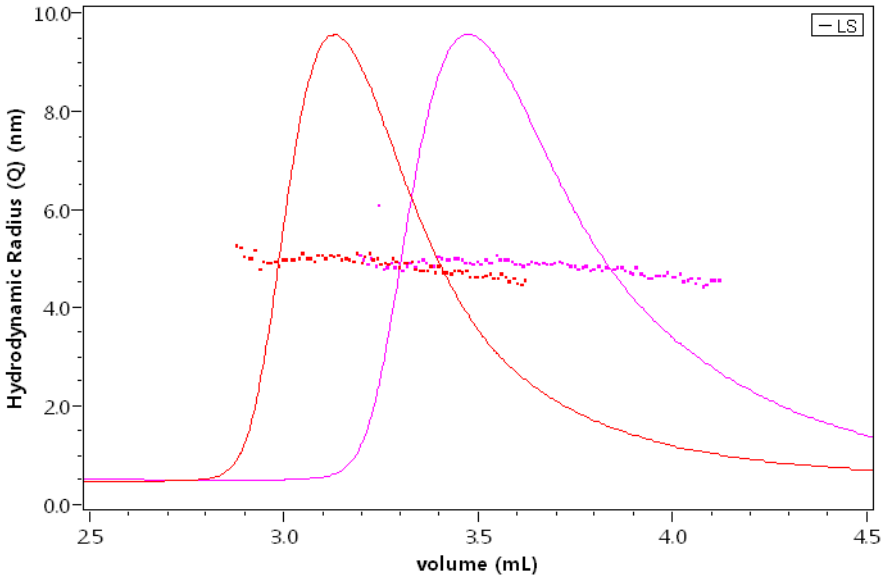


Figure 57 : Analyse SEC-DLS des HAB en configuration DLS-SAXS (rouge) et SAXS-DLS (rose). Ces données montrent que les bicelles possèdent un R_h constant le long du pic d'élution avant ou après exposition aux rayons X. Les mêmes résultats ont été obtenus pour les HZB.

Etonnamment, les bicelles possèdent un R_h similaire (~ 5 nm) avant et après exposition aux rayons X. Le décalage des pics n'est dû qu'à la différence de la longueur de la tubulure et l'élargissement du pic est dû à la dilution qui en découle. Ces résultats suggèrent que, s'il y a un changement de forme des bicelles, celui-ci a lieu uniquement lors de l'exposition aux rayons X et qu'il est réversible. Malgré les nombreux obstacles rencontrés, certaines données SAXS obtenues sur les bicelles sont exploitables et riches d'informations. Par exemple, pour les données SAXS présentées dans la Figure 58A, le décalage du minimum vers les petites valeurs

de Q indique une augmentation de la taille des objets bicelles en fonction de la valeur de q. De plus, ces données ont pu être ajustées avec le modèle « core-shell bicelle », qui permet de différencier la valeur des SLD entre la partie « shell » du tore et la bicouche, ainsi que leur épaisseur respective. Mise à part une augmentation du rayon des bicelles, les paramètres d'ajustement des données SAXS nous apportent une information très intéressante. En effet, les valeurs d'épaisseur trouvées pour la partie « shell » du tore sont d'environ 1 à 2 Å lorsque les SLD sont fixées. *A contrario*, la SLD du tore se rapproche de celle de l'eau (9.47 10^{-6} Å^2) lorsque l'épaisseur du « rim » est fixée. Une explication plausible est une forte hydratation de cette partie, qui diminue le contraste avec le tampon. En effet, l'aire moléculaire d'une tête PC est sensiblement supérieure dans une micelle de DHPC (100 Å^2) [235] de l'aire occupée par une tête PC dans une bicouche de DMPC en phase gel (47 Å^2) [236] ou fluide (60 Å^2) [237]. Ceci implique que le « packing » des lipides est plus ou moins important et de telles différences impliquent une différence d'hydratation.

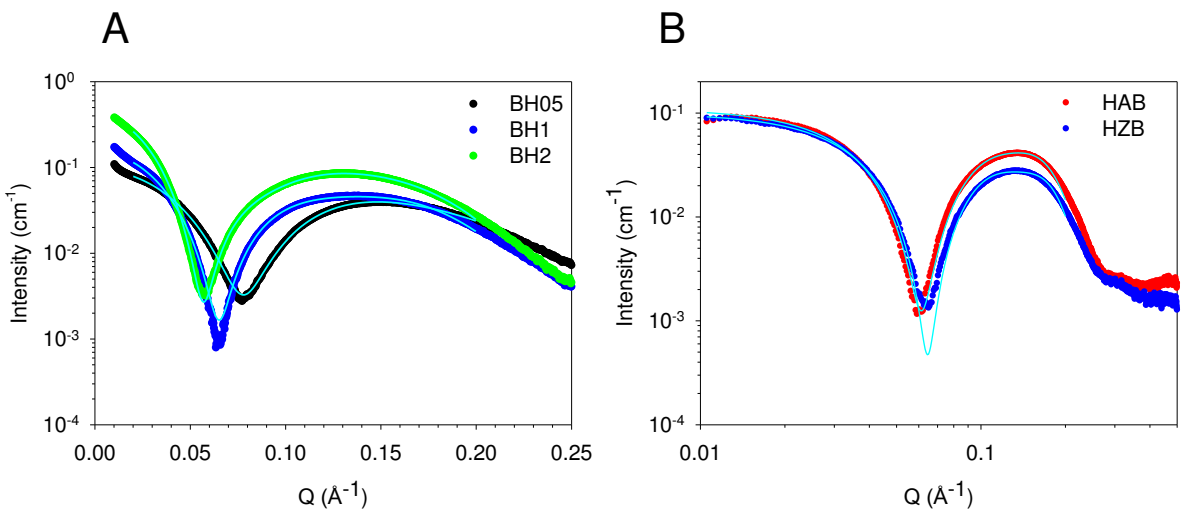


Figure 58 : (A) Données SAXS obtenues à 75 mM de lipides totaux pour des ratios q = 0.5 (noir), 1 (bleu) et 2 (vert) (q_{eff} = 0,6, 1,2 et 2,6). (B) Données SAXS obtenues pour des bicelles hydrogénées zwitterioniques (HZB) ou anioniques (HAB) à 50 mM de lipides et q_{eff} = 1,3. Les données ont été ajustées avec le modèle « core-shell bicelle » (cyan).

Les toutes dernières acquisitions SAXS sur les bicelles (Figure 58B) sont l'un des meilleurs jeux de données. Pour ces bicelles, les données ont été obtenues dans un environnement échantillon différent de celui des précédents, à savoir un simple capillaire rempli d'une solution de bicelles qui a été stabilisée pendant environ 15 min avant exposition aux rayons X. Cet environnement échantillon semble le plus adapté pour obtenir des données SAXS de bicelles

de bonne qualité. Ainsi, des valeurs de R_g de bonne qualité ont pu être obtenues : 46 Å pour les deux types de bicelles.

Une extension aux bicelles du logiciel Memprot[127], initialement dédié aux PDC, est en cours de réalisation, dans le cadre d'un projet mené par Maciej Baranowski et Javier Pérez. Ce projet a pour but d'obtenir des modèles gros-grain de bicelles seules sous contraintes SAXS puis potentiellement de caractériser des complexes IMP/bicelle.

4. Conclusion

Dans ce chapitre, nous avons finement étudié le modèle de bicelle en vue de son utilisation en tant que modèle membranaire pour caractériser les interactions protéine/lipide. De plus, nous avons démontré la possibilité de l'utiliser en tant que modèle membranaire pertinent pour l'obtention d'informations structurales de PMP lors de leur liaison aux lipides membranaires dans une approche de variation de contraste en SANS. En effet, nous avons mis en évidence que le signal SANS de ce modèle membranaire, en version deutérée, peut être masqué dans un tampon 100% D_2O à concentration en sels et pH physiologiques. Ces conditions permettent de travailler à un contraste élevé en utilisant une protéine hydrogénée dans du tampon 100% D_2O et permet d'éviter le recours à des mélanges de tampon H_2O/D_2O qui réduiraient le contraste et donc l'intensité du signal.

Un modèle alternatif à celui des bicelles est celui de nanodisque. Ce modèle est plus stable que les bicelles. Cependant, la taille des nanodisques à base de MSP (voir partie I chapitre 1) dépend de la longueur des protéines utilisées. Ainsi, travailler avec des nanodisques de plusieurs tailles requiert un long travail de purification de protéines, alors que la modulation de la taille des bicelles nécessite uniquement l'utilisation d'un ratio q approprié. De plus, dans le cas de leur utilisation avec des PMP, la nature du tore des nanodisques peut conduire à des interactions non-spécifiques. Cette même remarque peut être étendue à des bicelles n'étant pas exclusivement formées de phospholipides (par exemple du CHAPSO).

Le modèle « bicelle » présente de nombreux avantages. Des précautions sont cependant à prendre pour son utilisation. La première est qu'il faut absolument prendre en considération la $[DHPC]_{libre}$ (CBC) afin d'éviter tout biais d'interprétation dû à la variation de morphologie des bicelles. Une des possibilités consiste à utiliser des phospholipides à chaîne courte ou des détergents possédant une CMC plus faible que celle de la DHPC afin de pouvoir travailler dans

des conditions de dilutions plus importantes. Cependant, pour les besoins du SANS, leur actuelle non-disponibilité en version deutérée peut être un frein à cette alternative. La morphologie des bicelles est sensible à la température. Pour être certain de travailler avec des objets en formes de disque, 2°C au-dessus de *Tm* est l'écart maximum tolérable, dans le cadre des expériences de diffusion. Au-delà de cette valeur, la morphologie discoïdale des bicelles n'est plus rigoureusement conservée : les bicelles DMPC/DHPC à 37°C adoptent par exemple une morphologie en forme de cylindres allongés. Pour travailler à des températures plus élevées tout en maintenant une morphologie en forme de disque, une alternative est d'utiliser un phospholipide ayant une longueur de chaîne hydrophobe de taille supérieure, comme le DPPC, et donc d'avoir une valeur de *Tm* plus élevée. Cependant, ce choix soulève un autre problème : celui de travailler au dessus de 41°C pour que les lipides soient en phase fluide, ce qui peut poser des problèmes liés à la thermostabilité des protéines et à la signification biologique de résultats. Un dernier point de discussion concerne une faiblesse des techniques de diffusion en général. La présence, même en très faible proportion, d'un gros objet affecterait très fortement le signal de diffusion. Des approches permettant une distribution de population par nombre, de type TEM et/ou cryo-EM, pourrait être utiles afin d'affiner les diagrammes de phase des bicelles.

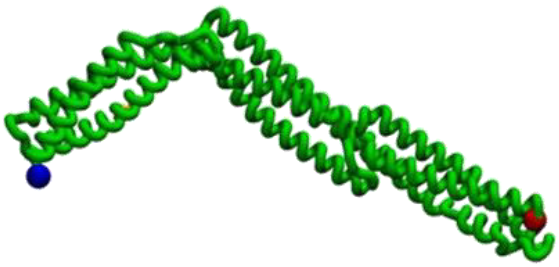
Chapitre 2 : Analyse des interactions dystrophine/bicelle

Points clés :

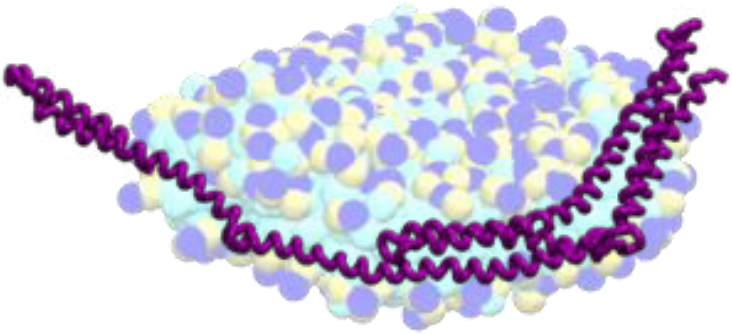
- *R1-3 et R11-15 conservent leur conformation lors de leur interaction avec des bicelles zwitterioniques. A contrario, R1-3 et R11-15 adoptent une structure étendue en présence de bicelles anioniques.*

- *Dans le cas de R1-3, ce changement de structure est attribué à une ouverture du faisceau d'hélices α (« coiled coil ») de la répétition R1, ouverture qui pourrait être nécessaire à l'ancrage au sarcolemme.*

SANS-based model of a dystrophin fragment free in solution



Contrast-matched anionic bicelle



SANS-based model of a dystrophin fragment bound to anionic bicelle

1. Introduction

L'équipe a finement caractérisé les interactions dystrophine/lipides depuis environ 15 ans. Pour cela, les modèles de SUV et de monocouche de Langmuir ont largement été employés. Dans l'article présenté ci-après, nous avons utilisé le modèle de bicelles. Nous avons étudié les interactions protéine/bicelles en utilisant deux types de bicelles, zwitterioniques et anioniques, afin de prendre en compte les propriétés de charge de la membrane plasmique. Nous avons également obtenu des données à des températures inférieures ou supérieures à la *Tm* pour prendre en compte la fluidité de la membrane. Ces interactions protéine/bicelle ont notamment été mises en évidence par MST. Enfin, nous avons étoffé et renforcé notre approche de SANS en la couplant à des analyses de dynamique moléculaire gros-grain et de chimie-click/spectrométrie de masse. Ceci nous a permis de proposer un modèle tout-atome de R1-3 en interaction avec des bicelles anioniques.

2. Publication

Les résultats sont présentés sous forme d'un article scientifique soumis à publication et intitulé:

Article 2 : Dystrophin-membrane interactions: evidence of structural changes upon binding to anionic phospholipids.

Dystrophin-membrane interactions: evidence of conformational modification upon binding to anionic phospholipids

Raphael Dos Santos Morais[1,2,3,4*], Olivier Delalande[1,2], Javier Pérez[4], Dominique Mias-Lucquin[1,2], Mélanie Lagarrigue[1,5], Anne Martel[6], Anne-Elisabeth Molza[1,2,§], Angélique Chéron[1,2], Thomas Chenuel[1,2], Arnaud Bondon[7], Marie-Sousai Appavou[8], Elisabeth Le Rumeur[1,2], Sophie Combet[3*], and Jean-François Hubert[1,2*].

[1]Université de Rennes 1, F-35043 Rennes, France. [2]Institut de Génétique et Développement de Rennes, CNRS UMR 6290, F-35043 Rennes, France. [3]Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, F-91191Gif-sur-Yvette CEDEX, France. [4]SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP 48, Saint-Aubin, F-91192 Gif-sur-Yvette, France. [5]Inserm U 1085, Protim-Plate-forme Protéomique. [6]Institut Laue-Langevin, F-38042 Grenoble, France. [7]CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-35043 Rennes, France. [8]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany.

ABSTRACT. Dystrophin is a filamentous peripheral membrane protein supporting the plasma membrane of muscular cells. Its absence due to gene mutations leads to severe Duchenne muscular dystrophy. Most of the dystrophin protein consists of a central domain made of 24 spectrin-like coiled-coil repeats. Using small-angle neutron scattering (SANS) and the contrast variation technique, we specifically probed the solution structure of two fragments of the dystrophin central domain, both known to interact with membrane lipids. Their solution structures were compared to the structures adopted in the presence of phospholipid-based bicelles. SANS data for the protein/lipid complexes were obtained under various conditions to probe the role of electrostatic interactions and membrane fluidity. When bound to anionic bicelles, large modifications of the protein 3D-structures were detected. One complex was further analyzed by coarse-grained molecular dynamic simulation. The final models of the protein bound to the membrane lipids are totally in agreement with the experimental SANS and click chemistry/mass spectrometry data. We conclude that the membrane anchoring that occurs during the contraction/elongation process of muscles could be ensured by a coiled-coil opening in the first spectrin-like repeat. Understanding these structural changes may help in the design of rationalized shortened dystrophins for gene therapy. Finally, our approaches can be extended to other peripheral and integral membrane proteins.

INTRODUCTION

Determining the 3D-structure of membrane proteins is very challenging due to the difficulties associated with the handling of these proteins[1]. However, considerable progress has been made since the first high-resolution structure determination of an integral membrane protein, the photosynthetic reaction center, in 1985, using X-ray crystallography (XRC)[2]. One of the most noticeable areas of progress is the combined use of X-ray free electron laser and iododetergent-based bicelles for phasing[3]. Nevertheless, only ~ 700 membrane protein structures have been made available (http://blanco.biomol.uci.edu/mpstruc/), and this number is considerably less than the numerous structures obtained for soluble proteins[4]. The particular case of peripheral membrane proteins (PMPs) could be even trickier, since some of these proteins are partially soluble in an aqueous environment but are also able to interact with membrane lipids[5]. For some of the PMPs, membrane anchoring is carried out with a post-translational modification such as myristoylation, palmitoylation, or glypiation. In the specific case of part-time docked PMPs, interactions with membrane lipids can be achieved *via* a hydrophobic loop, an amphipathic α-helix, or electrostatic interactions. For the study of PMPs, solid and/or liquid nuclear magnetic resonance (NMR) is classically employed to obtain the 3D-structure of the protein either free in solution and/or bound to a membrane mimic such as detergent micelles, liposomes, or bicelles[6]. Nevertheless, the main drawback when using NMR is the limited size of the protein[7]. For membrane proteins not compatible with XRC and/or NMR, small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) are alternative techniques of choice to obtain structural information on biological macromolecules, either free in solution, or bound to partners[8,9], and for membrane proteins[10,11]. The SANS-based contrast variation technique[12] allows one to specifically obtain low-resolution structures of proteins involved in macromolecular complexes and has been successfully applied to protein-protein[12,13], protein-polymer[14], protein-DNA/RNA[12,15], and protein-lipid/detergent complexes[16,13]. We recently demonstrated that phospholipid-based bicelles can be contrast-matched virtually and employed in SANS experiments, as a relevant membrane mimic to probe conformational modification of peripheral membrane proteins when interacting with membrane lipids[17]. In the present work, this approach was used to explore the structure-function relationship of a biological system associated with severe genetic diseases.

Cytoskeletal proteins protect muscle cell membranes against shear stresses. Among them, dystrophin plays a key role, and its absence leads to the severe Duchenne muscular dystrophy (DMD)[18], whereas its deficiency causes the heterogeneous Becker muscular dystrophy (BMD). Dystrophin is a large (427 kDa), monomeric, amphipathic, and filamentous peripheral membrane

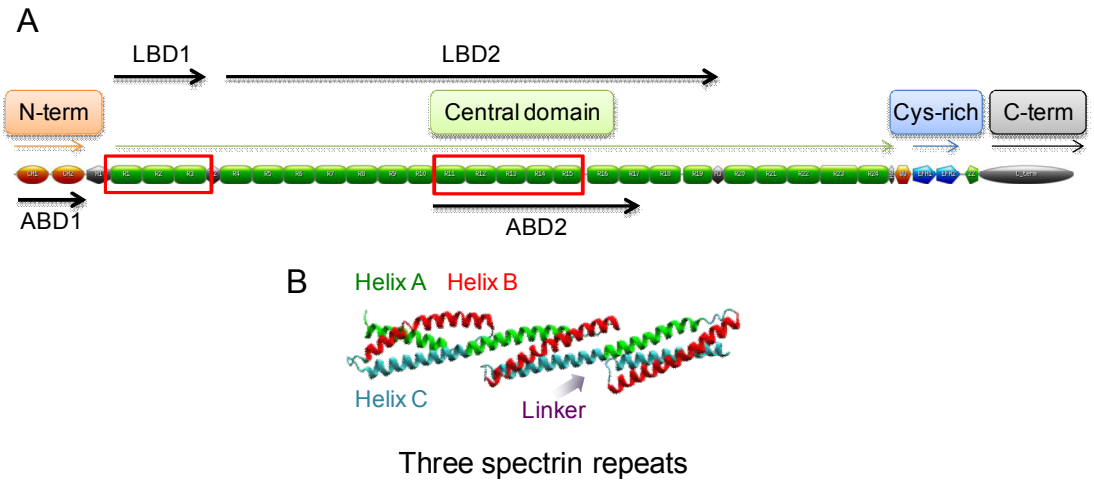


Figure 1: (A) Schematic representation of dystrophin and its four domains, including the central domain composed of 24 spectrin-like repeats (R). The R1-3 and R11-15 protein fragments used in this study are framed in red and can both interact with membrane phospholipids. The R11-15 fragment can also interact with F-actin. LBD: lipid binding domain. ABD: actin binding domain. (B) Representation of the 3D-structure of three spectrin repeats folded in a triple coiled-coil (1U4Q[19]). The helix C of one repeat is in continuity with the helix A of the subsequent repeat.

protein composed of four main domains[20]: the N-terminal domain (NT), the long central domain separated by four hinges, the Cys-rich domain (CR), and the C-terminal domain (CT) (Figure 1A). The central domain accounts for 75% of the dystrophin sequence and is made of 24 spectrin-like repeats folded in a triple coiled-coil (Figure 1B) (denoted R1 to R24). Although these repeats have a highly conserved coiled-coil structure, they possess specific surface properties making them unique[21]. Dystrophin is a scaffolding protein, forming a platform for interactions with various partners. Indeed, F-actin interacts with NT and R11-17[22] or R11-15[23] domains, neuronal nitric oxide synthase with R16-17[24], β-dystroglycan with CR[25], and finally the microtubules with R20-23[26]. In addition to these interactions with protein partners, dystrophin is also able to interact with the lipid part of the sarcolemma[20,27]. *In vitro* experiments showed that the region from R1 to R19 is able to bind anionic phospholipids[28–30], while R11-15 is able to bind both anionic and zwitterionic phospholipids[31] (Figure 1A). Moreover, this region is also the only region to bind both F-actin and membrane lipids simultaneously[23] (Figure 1A). Recently, *in vivo* experiments performed on GFP-dystrophin fragment fusion proteins highlighted that R1-3, CR, and CT are exclusively localized at the sarcolemma, while R10-12 is found both in the cytosol and at the sarcolemma[27]. The combination of these observations led to the conclusion that interactions of dystrophin with membrane lipids must be considered for understanding its crucial scaffolding role in muscle cells. In particular, during the contraction process of muscles, the sarcolemma bends, and areas containing invaginations are formed[32]. These regions present a high degree of curvature leading to a local lipid packing default. The influence of this phenomenon on the dystrophin/lipid interactions has been investigated with liposomes of various sizes[31]. Our previous work highlighted that the interfacial properties of dystrophin are modulated according to the region of the protein involved, the nature of lipids, and lipid packing, which all play key roles in the physiology of muscle cells[28–31]. High-resolution techniques of structural biology such as NMR and XRC failed to solve the structure of the central domain of dystrophin, mainly due to its large size, flexibility, and repeat-based organization. The structures of only three parts of the protein were solved by XRC: the NT domain[33] (PDB: 1DXX), the CR[34] (PDB: 1EG3), and the first coiled-coil repeat R1[35] (PDB: 3UUN). The latter was crystallized as a dimer through the formation of an artificial disulfide bond to reduce its flexibility[35]. Despite the impossibility of using XRC or NMR, we were able to propose solution structures of fragments of the central domain, through the combined use of SAXS and molecular modeling, first for R11-15 and R16-17[36] and more recently for other fragments covering the whole dystrophin central domain (Delalande *et al.*, submitted). These dystrophin SAXS-derived models were successfully deposited in the small-angle scattering biological data bank (SASBDB[37]). Our results highlight that the linear structure admitted for the dystrophin central domain should be reconsidered, since kinks are present all along this rod-shaped but not straight domain (Delalande *et al.*, submitted). Therefore, we obtained evidence of the conformation adopted by the protein fragments free in solution, in the absence of a lipid environment. However, the structure of these protein fragments in interactions with the membrane remains unknown. In the present paper, we took advantage of the SANS contrast variation method to investigate the solution structure of two dystrophin fragments of the central domain (R1-3 and R11-15). We analyzed their structures when free in solution and when in an interaction with zwitterionic (dimyristoylphosphatidylcholine / dihexanoylphospha tidylcholine, DMPC / DHPC) or anionic (DMPC / dimyristoylphosphatidylserine (DMPS) / DHPC) bicelles. We show that the tertiary structure of R1-3 and R11-15 is maintained during the binding to zwitterionic phospholipids, whereas in the presence of anionic phospholipids, R11-15 and, to a larger extent, R1-3, exhibit conformational modifications. The R1-3/bicelle complex was further investigated using classical coarse-grained molecular dynamics (CG-MD) and interactive CG-MD (CG-IMD) simulations. Moreover, we were able to accurately map the protein/lipid interactions by coupling click chemistry with mass spectrometry. We propose an all-atom model of R1-3 in interaction with membrane phospholipids. The tridimensional modification of R1-3 can be attributed to an opening of the coiled-coil of repeat 1. The biological and physiological relevance of this observation is discussed.

EXPERIMENTAL SECTION

Intrinsic tryptophan (Trp) fluorescence. Fluorescence measurements were carried out on a Fluorolog spectrofluorometer

(Jobin-Yvon) at 22 and 26°C. Tryptophan fluorescence emission spectra were recorded in low volume quartz Hellma cells (120 µL) between 310 and 420 nm, using an excitation wavelength of 295 nm (bandwidth of 2 nm). Blanks corresponding to the buffer or bicelles alone were subtracted from the respective spectra. The protein concentration was 20 µM, and the total lipid concentration was 50 mM (~ 150 µM of bicellar objects).

Far-UV circular dichroism (CD) analysis. Far-UV CD spectra of the proteins were recorded on a Jasco J-815 spectropolarimeter, at 22 and 26°C, from 200 to 240 nm in 0.02 cm path length Hellma cells. Blanks corresponding to the buffer or bicelles alone were subtracted from the respective spectra. The ratio of the ellipticity at 222 to 208 nm ($\theta_{222}/\theta_{208}$) was used as an indicator of the presence of a coiled-coil[29,38]. The protein concentration was 20 µM and the total lipid concentration was 50 mM (~ 150 µM of bicellar objects).

Microscale thermophoresis (MST). Bicelles were labeled with 1 µM of 1,2-dioleoyl-sn-glycero-3-phospho-L-serine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl) for a final concentration of 15 mM of lipids (~ 50 µM of bicellar objects) at $q_{eff} = 1.3$ (see Supporting Information) and the [DMPC]/[DMPS] (mol/mol) ratio into the bilayer part was maintained equal to 2. The bicelles were titrated with a 0.75-fold serial dilution of proteins from 338 to 4.5 µM for R1-3 and 184 to 2.5 µM for R11-15. The 16 solutions were then loaded into Monolith NT premium capillaries (NanoTemper) and thermophoresis was measured with a Monolith NT.115 instrument (NanoTemper) at 22 and 26°C. Instrument parameters were as follows: 30% LED power, 40% MST power, and 5/30/5 laser off/on/off. Data were analyzed with the NT MO Affinity Analysis software v2.1.3 (Nano Temper).

Small-angle neutron scattering (SANS) experiments. Preparatory works were carried out at KWS-2 instrument (MLZ, Garching) while final investigations were done at either PACE (LLB, Saclay), D22 (ILL, Grenoble), or KWS-1 (MLZ, Garching) instruments. Two to three sample-to-detector distances were used with a wavelength varying from 4.7 to 7 Å ($\Delta\lambda/\lambda \sim 10\%$), to cover a Q-range from 0.008 to 0.5 Å^{-1} for the largest one, where $Q = \frac{4\pi \sin\theta}{\lambda}$ is the momentum transfer, λ is the wavelength, and 2θ is the scattering angle. All measurements were performed in 1 mm thickness Hellma *QS* cells, and the intensities obtained are in absolute units (cm^{-1}). For SANS acquisitions, exactly the same buffer that had been used for bicelle rehydration was used to prepare protein/bicelle samples to guarantee a perfect buffer subtraction.

SANS data analysis. SANS data were analyzed with the ATSAS suite[39] and the Scatter software (http://www.bioisis.net/), following guidelines unless otherwise indicated. The PRIMUS software was used to determine the intensity at zero angle (I_0) and the radius of gyration (R_g). These values are defined at small Q-values ($QR_g < 0.8$-1.1) by the Guinier approximation (Eq. 1)[40]:

$$I(Q) = I(0)exp\left(-\frac{Q^2R_g^2}{3}\right) \qquad Eq.\ 1$$

The apparent molecular weight (MW, in g·mol^{-1}) of the proteins was estimated with the following equation:

$$I(0) = (\Delta\rho V_P)^2.c.MW/N_A \qquad Eq.\ 2$$

where $\Delta\rho$ (cm^{-2}) is the contrast (*i.e.,* the difference in the neutron scattering length densities, SLDs) between R1-3 or R11-15 and d-TNE buffer, V_P is the partial specific volume (cm^3·g^{-1}) of the protein fragments, c is the protein concentration (g·cm^{-3}), and N_A is the Avogadro number (mol^{-1}). The V_P were determined from the density of the proteins according to their amino acid sequence (http://psldc.isis.rl.ac.uk/Psldc/). Fitting procedures were performed using the SASview v3.1.0 software (http://www.sasview.org/). Protein and bicelle data were fitted with the "flexible cylinder" and the "core-shell cylinder" form factor models[43], respectively. The GNOM module from the ATSAS suite was used to determine the pair-distribution function $P(r)$, the maximum distance D_{max}, and the R_g in real space (denoted R_g^{real}) from $P(r)$. Twenty *ab initio* models were generated with the GASBOR software considering the data up to $Q = 0.25$ Å^{-1}. The models obtained were aligned and averaged with the DAMAVER software. The presented *ab initio* models correspond to the GASBOR model with the smallest normalized spatial discrepancy (NSD) surrounded by the corresponding DAMAVER model. A summary table showing all this information is available (Table S1).

Click-chemistry and liquid chromatography coupled with mass spectrometry (LC-MS/MS). HZB and HAB were labeled with 5% (mol/mol) of 1-palmitoyl-2-(9-(3-pent-4-ynyl-3-H-diazirin-3-yl)-nonanoyl)-sn-glycero-3-phosphocholine (pacFA PC) (Avanti Polar Lipids) bearing both photoactivable diazirine and clickable alkyne groups on one tail of the phospholipid[44] (Figure S10A). The modified bicelles were mixed with R1-3 to reach a final concentration of 50 mM of lipids and 15 µM of protein. Then, the protein/lipid complexes were exposed to UVA radiation (5 x 15 W, Stratalinker, GE Healthcare) for 5 min, just before SDS-PAGE. The staining was performed using Coomassie blue (InstantBlue, Expedeon). The smeared parts of the gel (Figure S10B), just above the band corresponding to R1-3, were cut into small pieces. Proteins contained in the gel pieces were reduced, alkylated, and then digested with trypsin (Promega), and the resulting peptides were extracted as previously described[45]. The peptide extract was loaded (10 µL) on a nano-HPLC system (LC Packings Ultimate 3000, Dionex) equipped with a trapping precolumn (5 mm × 300 µm id, 300 Å pore size, Pepmap C18, 5 µm, ThermoScientific) and an analytical column (15 cm × 75 µm id, 300 Å pore size, Pepmap C18, 5 µm, ThermoScientific). Reversed-phase separation was performed with the same gradient as described previously[46]. Peptides were directly eluted from the nano-HPLC column to the nanoelectrospray ion source of a LTQ-Orbitrap XL (ThermoScientific) mass spectrometer operating in data-dependent mode by automatically switching between full MS scan and MS/MS acquisitions on the 15 most intense precursor ions as described previously[46]. MS data were saved in the RAW file format with Xcalibur 2.0.7 and subjected to a database search for protein identification using Mascot Distiller 2.6.1.0 and Mascot Server 2.5.1 with its automatic decoy database search. The database searched contained the *E. coli* reference proteome from Uniprot (UP000000625, April 2017), the contaminants database from Mascot and the sequence of the R1-3 protein fragment (785 sequences, 14595443 residues). Mass tolerance was set to 10 ppm for precursors and to 0.5 Da for fragments. Trypsin was selected as enzyme with two missed cleavages allowed. Protein modifications were: fixed carbamidomethylation of cysteines and variable oxidation of methionine, and variable pacFA PC. The pacFA PC modification ($C_{39}H_{72}NO_8P$, 713.4996 Da) was manually defined in the Mascot configuration editor with a neutral loss of phosphocholine ($C_5H_{14}NO_4P$, 183.066 Da). Since the photoactivation of diazirine forms carbene intermediates that can react with any amino acid side chain or peptide backbone, all amino acids were selected as possible modified sites. Proline Studio 1.4 was used for identification validation (peptide rank = 1; false discovery rate < 1% at the peptide spectrum match level)[47].

Coarse-grained molecular dynamics (CG-MD) simulations. CG-MD simulations were run using the GROMACS 5.0 program[48] with the Martini coarse-grained force field[49,50]. Two protein/lipid systems were simulated, *i.e.,* the R1-3 dystrophin

fragment with either hydrogenated zwitterionic bicelles (HZB) or hydrogenated anionic bicelles (HAB). In a manner similar to a previously improved protocol[51], bicelles ($q_{eff} = 1.3$) were built through the generation of a DMPC/DHPC or DMPC/DMPS/DHPC bilayer with the insane tool[52] by defining the number of lipids within the ideal bicelle model (IBM)[53]. This step was followed by a second step consisting of the enhancement of the size of the cubic simulation boxes up to 250 $Å^3$ by the addition of water and NaCl. All CG-MDs were performed at a constant temperature of 303 K and under NPT (constant number, pressure, and temperature) conditions. Final CG-MD trajectories of 20 µs were recorded and first analyzed by focusing on the period of relaxation of the lipid assembly and on the stability of the bicelle organization. Finally, the conversion of CG protein/lipid complexes to final atomic models was performed using the Backward tool[54]. Theoretical SANS curves were generated with default parameters (except 50 harmonics) from 0.008 to 0.18 $Å^{-1}$, with the CRYSON program from the ATSAS suite. The theoretical curves were smeared using a "resolution file" according to the D22 instrument setup.

Interactive coarse-grained molecular dynamic (CG-IMD) simulations. CG-IMD simulations were performed on the final R1-3/HAB model obtained from classical CG-MD using the GROMACS 5.0 implementation based on our previously published approach[55]. Four different scenarios of coiled-coil opening were tested through five independent simulations for each (see Supporting Information). CG-IMD trajectories were recorded during 200 ps after checking that the system returns well to equilibrium. A classical CG-MD of 2 ns was finally run from the open CG-IMD final model. Atomic reconstruction and theoretical SANS curves were processed as for classical CG-MD results.

RESULTS AND DISCUSSION

Protein and h-bicelle quality control. To obtain good and relevant SANS data, a drastic upstream sample quality control is mandatory[56]. The R1-3 and R11-15 protein fragments were selected since their protein/lipid interactions were finely demonstrated both *in vitro*[28–31] and *in vivo*[27] and since they can be considered as independent subdomains[21]. Indeed, R11-15 is an actin and lipid binding domain separated from the subsequent repeat R16 by a pseudo hinge[21], and R1-3 is known to be a lipid binding domain bordered by two non-structured hinges[21] (Figure 1). These protein fragments were obtained with a high degree of purity, as assessed by SDS-PAGE analysis (Figure S1A). Prior to SANS analysis, the molecular weights (MWs) of R1-3 and R11-15 were determined by high-pressure size-exclusion chromatography coupled with multi-angle light scattering (HPSEC-MALS, see Supporting Information). The obtained MWs of 38.2 ± 0.4 and 62.9 ± 0.4 kDa are in line with the expected values of 38.5 and 60 kDa for R1-3 and R11-15, respectively, assessing the monomeric state of both proteins (Figure S1B). These data were obtained in TNE buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 0.1 mM EDTA) supplemented with 6 mM of DHPC, a value that corresponds to the critical bicellar concentration[57] of the bicelles used hereafter and that allows the bicellar shape to be maintained in HPLC experiments[17]. The bicelles employed in this study are zwitterionic (HZB, DMPC/DHPC, 1:1) or anionic (HAB, DMPC/DMPS/DHPC, 0.67:0.33:1). The phospholipid composition was chosen because phosphatidylcholine (PC) is the most prevailing phospholipid headgroup in eukaryotic cells, while phosphatidylserine (PS) is the most abundant charged headgroup and is known to be preferentially located in the internal leaflet of the plasma membrane[58]. Both are well represented in the sarcolemma[59] (~ 45% PC and 18% PS of total phospholipids). The quality control of the bicelles was first assessed by ^{31}P NMR to ensure a proper bicellar organization of the phospholipids, as well as to guarantee the expected molar ratio of each component. The NMR spectra (Figure S2A) are typical of an organization in bicelles[57], and the ratio between the integrals of the long-chain phospholipid peak and the DHPC peak corresponds to the expected value of $q = 1$. The monodispersity of the bicelles was then verified with HPSEC-dynamic light scattering (HPSEC-DLS) (Figure S2B). The hydrodynamic radius (R_h) is constant all along the peak for both types of bicelles and is 4.9 ± 0.1 nm for HZB and 5.1 ± 0.1 nm for HAB. The MW of the bicelles was estimated by HPSEC-MALS giving a value, at the top of the peak, of approximately 200 kDa (Figure S2C), in line with the expected value according to the IBM[53]. Interestingly, we show here that HPSEC-MALS can be used to estimate the MW of bicelles, as previously shown for liposomes[60], detergent micelles, or protein/detergent complexes[61], opening new possibilities if one wants to go deeper into the characterization of the bicelle membrane mimic by this method, for instance at different q ratios (see Supporting Information). All these data allowed us to conclude that both the protein fragments and the bicelles exhibit the expected size and monodispersity and therefore can be further used for the characterization of protein/lipid interactions.

Characterization of protein/h-bicelle interactions. Dystrophin/lipid interactions were previously extensively studied *in vitro* with either liposomes or Langmuir monolayers as membrane mimics. R1-3 was demonstrated to be able to bind anionic (DOPC/DOPS, 2:1) monolayers[30] or anionic small unilamellar vesicles (SUVs) and, to a smaller extent, to interact with zwitterionic (DOPC:DOPE, 2:1) SUVs[28]. However, R11-15 is able to bind anionic as well as zwitterionic SUVs and Langmuir monolayers[31]. Despite a solid characterization of the dystrophin/lipid interactions with vesicle and monolayer models, in the present study we had to verify the interactions of R1-3 and R11-15 with the bicelle membrane mimic model. To characterize these interactions, we employed zwitterionic and anionic bicelles to probe the influence of the electrostatic properties of the membrane. We also obtained data at 22 and 26°C (corresponding to the T_m (24°C) of DMPC ± 2°C) to take into account the effect of membrane fluidity on the protein/lipid interactions. First, the protein/bicelle interactions were highlighted by Trp intrinsic fluorescence measurements. Concerning R11-15, the Trp intrinsic fluorescence of the protein fragment in the presence of HZB or HAB increases slightly but significantly in the presence of HZB or HAB at the two temperatures studied (Figures 2D and S4D). For R1-3 in the presence of HZB or HAB, Trp intrinsic fluorescence increases significantly whatever the temperature (Figures 2A and S4A). In all cases, Trp intrinsic fluorescence is attributed to a modification in the Trp environment due to the interaction of the protein fragments with the phospholipid-based bicelles[31]. To go deeper into the characterization of these protein/bicelle interactions, we used microscale thermophoresis (MST)[62] to determine the dissociation constant (K_d) of the complexes. Bicelles have already been used successfully in isothermal titration calorimetry (ITC) studies[63], but their use in MST experiments has not been previously reported. In conventional MST experiments, the protein of interest is labeled and titrated with a ligand, as we previously did with the R16-21/SUV system[64]. In such a protocol, the size and the morphology of the bicelle model would be affected by dilution[65], compromising the data interpretation. Instead, we labeled the bicelles with fluorescent phospholipids, and then we performed

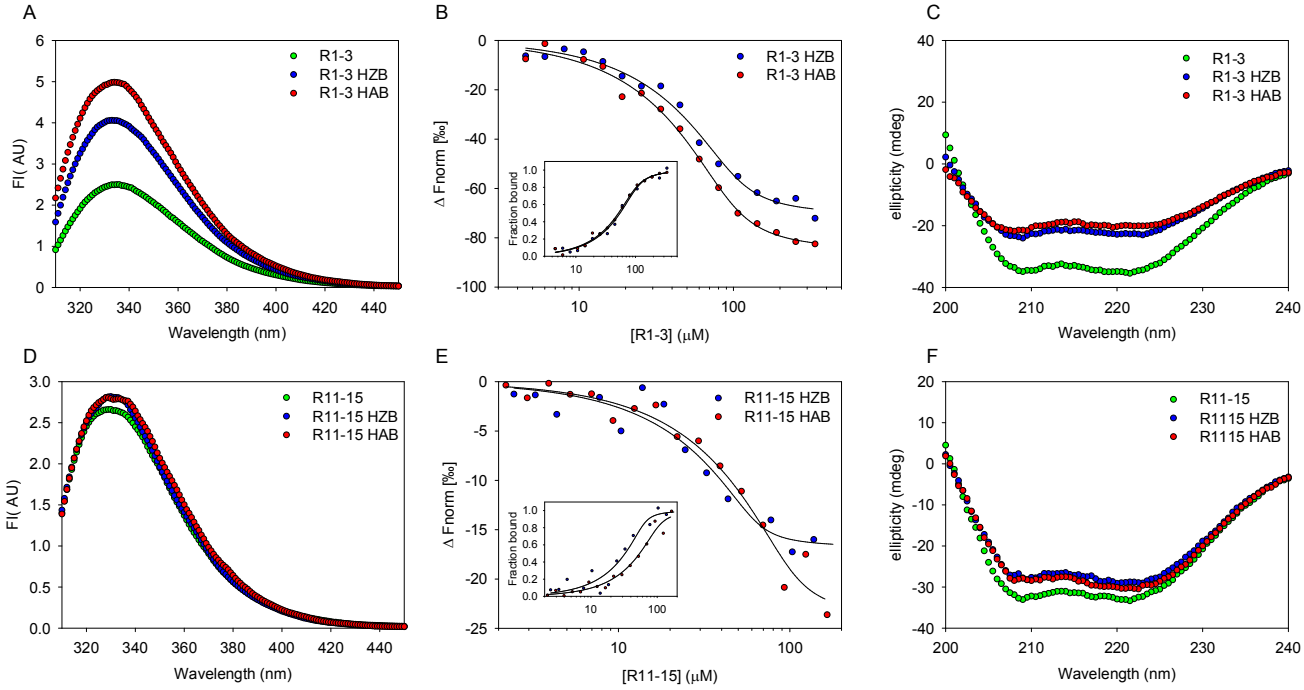


Figure 2: Characterization of the protein/lipid interactions at 26°C for R1-3 (*A*, *B*, *C*) and R11-15 (*D*, *E*, *F*) dystrophin fragments. (*A* and *D*) Tryptophan fluorescence intensity (FI) of each protein in the absence (green) or in the presence of zwitterionic (HZB, blue) or anionic (HAB, red) bicelles. FI variation is due to a modification in the Trp environment attributed to lipid binding. (*B* and *E*) Microscale thermophoresis (MST) data: the dissociation constant (K_d) was determined to be approximately 10 µM for each protein fragment/bicelle complex. (*C* and *F*) Far-UV CD spectra with the same legend as for FI (*A* and *D*). CD spectra highlight that the α-helical folding of both protein fragments is maintained during lipid binding.

the titration with the protein of interest. To avoid using a large quantity of protein and to ensure the saturation of the bicelles with the R1-3 and R11-15 proteins, we determined the minimal concentration of lipids for which the disk shape and size of the bicelles persist. A stock solution of bicelles at $q = 0.5$ was diluted to 15 mM of lipids giving $q_{eff} = 1.3$. The morphology of the bicelles was determined by SANS, and the data were fitted with the "core-shell cylinder" model[43] (Figure S3). The dimensions obtained were 40 Å thickness for both types of bicelles and 48 and 52 Å radius for HZB and HAB, respectively. These dimensions are in line with the expected values for bicelles at $q_{eff} = 1.3$. As shown by SANS data (Figure S3), under these experimental conditions, the morphology of the bicelles is preserved. Finally, to ratiocinate in bicellar objects instead of lipid concentration, we used the IBM[53], as previously done by other authors in ITC experiments[63]. For bicelles with $q_{eff} = 1.3$ at 15 mM of lipids, the concentration of bicellar object is approximately 50 µM since the aggregation number is ~ 300 lipids according to the IBM and the molecular volumes of phospholipids we used (see Supporting Information). The K_d values obtained from fitting the MST data are in the range of 10 µM for R1-3/HZB and R1-3/HAB, whatever the temperature (Figures 2B and S4B). The same values were obtained for R11-15/HAB at 22 and 26°C and for R11-15/HZB at 26°C (Figures 2E and S4E). Unfortunately, for the R11-15/HZB system at 22°C, no K_d could be determined in the protein concentration range that we used. Nevertheless, MST data at the lowest and the highest R11-15 concentration are significantly different, meaning that an interaction occurs (Figure S4E) but is weaker than for the anionic bicelles, as already suggested in our previous studies with other membrane mimics[29,31]. Increasing the concentration of R11-15 above 200 µM (12 g/L) to determine a K_d could have led to protein aggregation, which would have biased the data. In addition, from the fit of the MST data for the other datasets, the concentration of the binding site ranges from 50 to 100 µM. Although close to the predicted number of bicelles according to the IBM, this value can be affected by the bicelle aggregation number and by the value of q_{eff}. Indeed, working at 15 mM of lipids tends to be close to the $[DHPC]_{free}$ that could drastically affect q_{eff} even for a small misestimate of the total lipid concentration. Albeit a 1:1 stoichiometry appears the most likely, we cannot exclude the possibility of two protein fragments being bound to one bicelle. Nevertheless, under all conditions, except for the R11-15/HZB at 22°C, the protein/lipid affinity is approximately 10 µM. For a comparison with other described systems, the affinities of R1-3 and R11-15 to the membrane are lower the apolipoprotein-A1[66] with lipids. Nevertheless, this affinity is in the range of the affinity of cytochrome C for PC membranes with cardiolipin[67]. Then, we used circular dichroism (CD) to probe potential secondary structure modifications of the proteins during their interactions with both types of bicelles (Figure 2, C and F). Our first observation is that the CD spectra of the proteins alone are typical of an overall α-helical folding with the presence of two minima at 222 and 208 nm. These data confirm that both recombinant proteins are properly folded[28,29]. Moreover, the $\theta_{222}/\theta_{208}$ ratio is close to or above the unity, which is an evidence of a coiled-coil supersecondary structure[38]. For R11-15, the CD spectra of the protein alone or in the presence of both types of bicelles, and at the two temperatures studied, are very similar (Figures 2F and S4F). These results show that the secondary structure of R11-15 is maintained during binding to bicelles, as it is following binding to SUV[29] and zwitterionic bicelles[17]. For R1-3 data obtained in the presence of HZB or HAB, the same first observations as for R11-15 can be made, *i.e.*, the α-helical folding is preserved during the interaction with both types of bicelles. Interestingly, the spectra of R1-3 bound to

bicelles are qualitatively different from the spectra of the protein alone. Indeed, the$\theta_{222}/\theta_{208}$ ratio is below the unity, and a loss of ellipticity is observed at 208 nm, plausibly due to a loss of helicity (Figures 2C and S4C). These results suggest that R1-3 could at least partially lose its coiled-coil structure when bound to bicelles. To sum up, all these data show that (*i*) both R1-3 and R11-15 interact with zwitterionic and anionic bicelles, (*ii*) no significant modification of R11-15 secondary structure is detectable upon binding to zwitterionic or anionic bicelles, and (*iii*) large modifications of the supersecondary structure occur for R1-3 in the presence of zwitterionic bicelles and even more modifications occur when bound to anionic phospholipids.

SANS analysis of deuterated anionic bicelles (DABs). In a previous study, we highlighted that contrast-matched deuterated zwitterionic bicelles (DZBs) can be employed to specifically probe the SANS signal of a protein in an interaction with a lipid membrane[17], using the R11-15 protein as proof of principle. In the present work, to take into account the electrostatic properties of the membrane, we enlarged the scope of the study to anionic bicelles containing DMPS[68]. Therefore, we checked whether such modified bicelles remain contrast-matched in d-TNE buffer (TNE in deuterated version, see Supporting Information) and preserve their disk-shaped morphology under our experimental conditions, even if a drastic upstream quality control was already performed (Figure S2). The DABs were analyzed in 100 and 42% d-TNE D_2O-based buffer (42% D_2O is the classical protein SANS match-point) at 22°C. As expected, DABs are contrast-matched in 100% d-TNE buffer (Figure S5, inset). Their disk-shaped morphology, at least two degrees above the T_m of DMPC-d54 (20°C)[69], is supported by the fit of SANS data obtained at 42% d-TNE buffer (Figure S5). When using contrast-matched bicelles, an important point is to use exactly the same batch for the bicelles alone and for the protein/bicelle complex to ensure a perfect subtraction of the contrast-matched bicelles. Thus, we measured the batch of DZBs used in the present study under the same conditions as those used for the DABs (Figure S5). As expected, the DZBs are contrast-matched (Figure S5, inset) in 100% D_2O-based buffer, as we previously demonstrated[17].

SANS analysis of the R1-3 and R11-15 protein fragments free in solution. The present study aims to obtain structural information about dystrophin fragments in interaction with membrane lipids. Thus, a prior analysis of R1-3 and R11-15 free in solution is mandatory. The proteins fragments were first analyzed by SANS in the absence of bicelles. First, the Guinier approximation allowed us to determine the extrapolations of SANS intensity at zero-angle ($I(0)$) to assess that the proteins were monomeric. The calculated MWs are 45 ± 9 kDa and 55 ± 3 kDa for R1-3 and R11-15, respectively (Table S1). The experimental MWs are from three independent measurements included this study, where the protein concentrations were 103 µM (4.2 g/L) and 93 µM (5.6 g/L), for R1-3 and R11-15, respectively. The differences in expected *versus* experimental MWs can easily be explained by an improper estimation of the SLD of the protein fragments due to an inaccurate estimation of H/D exchange and/or by a small variation in the protein concentration inherent to the measurement technique. Nevertheless, these values are in good agreement with a monomeric state of the proteins, consistent with HPSEC-MALS data (Figure S1B). We then determined the radii of gyration (R_g) again thank to the Guinier approximation. R_g values are 42 ± 1 Å and 62 ± 2 Å for R1-3 and R11-15, respectively (Figure S6A and Table S1). Then, Kratky and Porod-Debye plots allowed us to highlight the presence of well-folded and compact proteins (Figure S6, B and C). The data of the proteins in solution in the absence of bicelles were fitted with the "flexible cylinder" form factor model[43]. For both proteins, the radius is approximately 10 Å. The Kuhn length (twice the persistence length) and the total length are 111 and 156 Å for R1-3, respectively, while they are 120 and 276 Å for R11-15 (Table S1). Then, the pair-distance distribution function algorithm enables the determination of D_{max} and R_g^{real} for both proteins. The R_g^{real} values are in line with the Guinier approximation (Table S1). Concerning the $P(r)$, for both fragments, they are typical of elongated proteins with a peak observed at approximately 20 Å and a smooth fall to D_{max} (Figure 3, C and D). The value of 20 Å corresponds to the thickness of a coiled-coil belonging to the dystrophin central domain, as previously determined for R1 using XRC data[35]. The obtained D_{max} are 177 Å for R1-3 and 274 Å for R11-15 (Figure 3, C and D, and Table S1). Similar results were obtained at 18°C (Table S1), meaning that the conformation of the protein fragments is similar at 18 and 22°C. All these results are in line with the previously obtained SAXS data for R1-3 and R11-15 in solution (Delalande et al., *submitted*) and serve as a reference in the data analysis of the subsequent SANS experiments. To consider the electrostatic properties of the membrane, as well as the importance of its fluidity, we recorded SANS data of R1-3 and R11-15, at exactly the same concentration as for the protein fragments free in solution, in the presence of zwitterionic or anionic bicelles at 18 and 22°C (corresponding to T_m (20°C) of DMPC-d54 ± 2°C[69]). The results are described hereafter.

SANS analysis of the R11-15/d-bicelle complexes. The SANS signal of the R11-15/DZB complex at 22°C is almost superimposed with that of R11-15 without DZB, as we previously showed for the same dataset obtained at 18°C[17] (Figures 3B and S7B). By fitting the data with the "flexible cylinder" model (Table S1), no significant difference in the parameters compared to R11-15 free in solution can be exploited, *i.e.*, the protein does not undergo large tridimensional modification. These results mean that even at 22°C, no conformational modification for R11-15 occurs or that they are not detectable by SANS. We confirm that the tertiary structure of R11-15 is maintained in the presence of zwitterionic bicelles. However, for the R11-15/DAB data, R_g values increase from 62 ± 2 Å without bicelles to ~ 80 Å (78 ± 2 Å (18°C) and 81 ± 4 Å (22°C)) in the presence of DAB (Table S1). This increase is also observed for the R_g^{real} determined from $P(r)$ (Table S1). The D_{max} values obtained confirmed the R_g analysis since the values increase from 274 Å for R11-15 alone to ~ 300 Å in the presence of DAB for the two temperatures studied (Figure 3D and S7D, Table S1). The same observations were made from the fit of the data (Table S1). Moreover, the calculated MWs (~ 50 kDa) of R11-15 in the presence of the both types of bicelles exclude the possibility of the dimerization upon lipid biding (Table S1). To sum up, we conclude that no important conformational modifications occur when R11-15 proteins are bound to zwitterionic phospholipid-based bicelles. However, R11-15 adopts a more elongated conformation when bound to anionic bicelles. To support this hypothesis, we generated *ab initio* models with the GASBOR software to obtain low resolution shapes of R11-15 free in solution and when bound to bicelles (Figure 3B). The low-resolution model of R11-15 in the presence of DAB is clearly more extended than the models found for R11-15 free in solution or in interaction with DZB.

SANS analysis of the R1-3/d-bicelle complexes. In the case of the R1-3/DZB complex, the SANS signal is different from that of the protein alone at both 18 and 22°C, but R_g remains constant (~ 40 Å) (Table S1). For the proteins alone, we fitted the data with the "flexible cylinder" model. For DYS-R1-3/DZB, the radius is approximately 11 Å and the total length is 157 Å at both temperatures (Table S1). The only difference when compared to R1-3 alone is that the Kuhn length increases from 111 to 193 Å at 22°C (Table S1). This increase could be attributed to a loss of protein flexibility when bound to the zwitterionic bicelles. In the

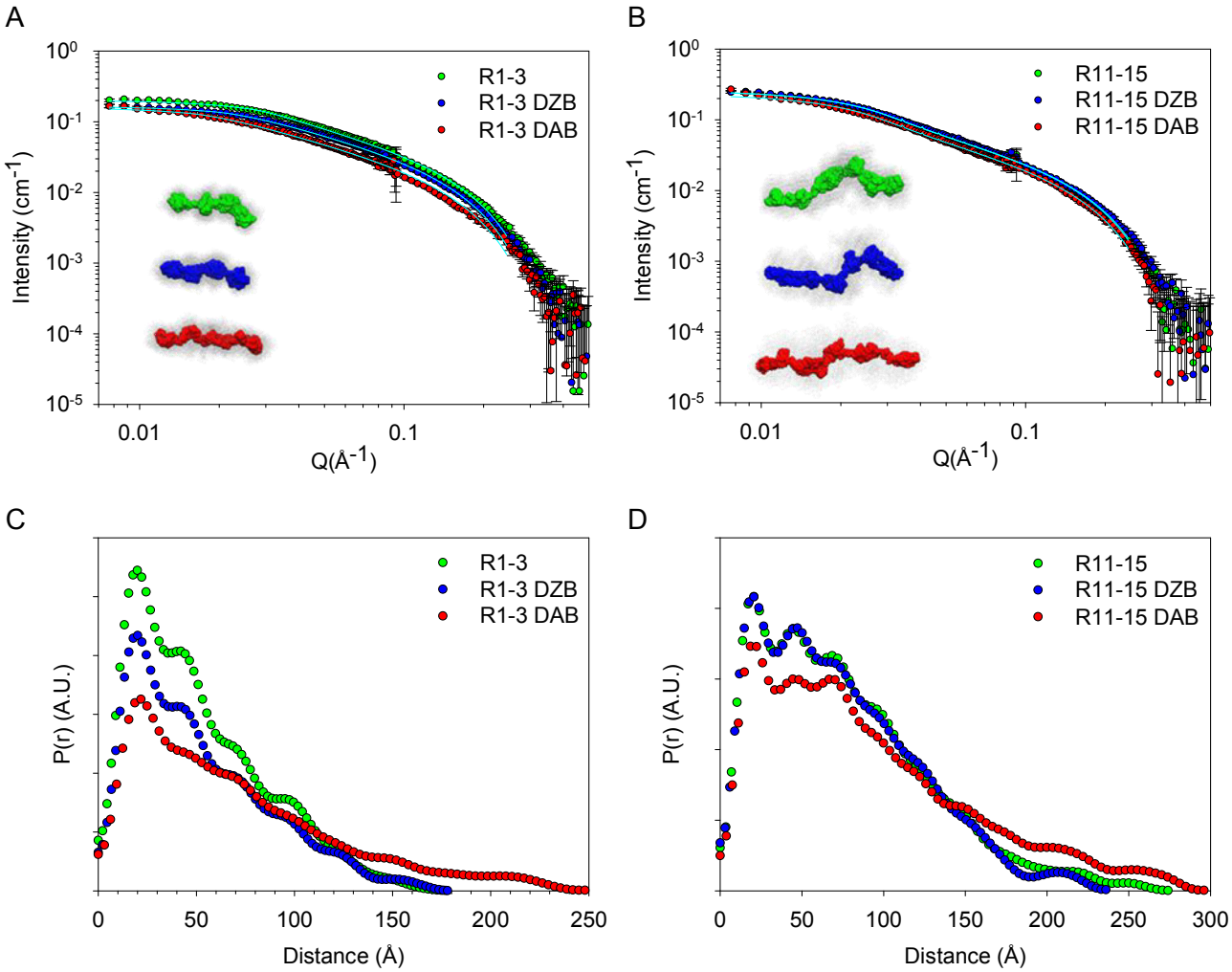


Figure 3: (*A* and *B*) SANS intensities measured at 22°C for R1-3 (*A*) and R11-15 (*B*) dystrophin fragments alone in solution (green) or in interaction with contrast-matched deuterated zwitterionic (blue) or anionic (red) bicelles. SANS data are fitted with the "flexible cylinder" model (cyan). *Inset*: *ab initio* shapes of the dystrophin fragments free in solution (green) or in interactions with DZB (blue) or DAB (red), corresponding to the GASBOR models obtained with the smallest normalized spatial discrepancy and surrounded by the corresponding DAMAVER model (gray). (*C* and *D*) Pair-distribution function $P(r)$ with the same color code showing an increase of D_{max} for both R1-3 and R11-15 when the protein fragments are in interactions with anionic bicelles (red).

presence of DAB, at 18°C the R_g of R1-3 increases from 42 ± 2 Å for the protein alone to 49 ± 2 Å (Table S1). In addition, R_g increases to 62 Å with DAB at 22°C (Table S1). At 18°C, no significant difference in D_{max} can be exploited (Table S1), whereas at 22°C, for R1-3 with or without DAB, the D_{max} values are 177 and 248 Å, respectively. Such increases of R_g and D_{max} indicate a large and significant elongation of the tertiary structure of R1-3 when bound to anionic bicelles, especially in the fluid phase (22°C). As for R11-15, the calculated MWs (~ 45 kDa) of R1-3 in the presence of bicelles allowed us to exclude a potential dimerization upon lipid binding (Table S1). A plausible explanation for these observations could be attributed to a modification of the coiled-coil tridimensional structure, in line with the CD data (Figure 2C). To explain these results, we hypothesize the opening of the coiled-coil of one spectrin-like repeat. Again, we generated *ab initio* models and observed that the envelope of R1-3 in the presence of DAB is clearly different compared to that of the protein free in solution or in the presence of DZB (Figure 3A).

Localization of R1-3 and mapping of the protein/lipid interactions. Only the data of R1-3 in interactions with both types of bicelles were considered for further characterization since they highlight the largest differences compared to the data of the protein free in solution. In addition, R1-3 is bordered by two non-structured hinges that support its relevance as an independent and functional protein domain[21,27]. Finally, the R1, R2, and R3 repeats are in whole or in part highly conserved in the mini-dystrophins designed in gene therapy[70,71]. To obtain accurate information about R1-3 conformational modifications in interactions with bicelles, we carried out *in silico* experiments. We performed CG-MD simulations with the Martini force field[49,50], which is particularly relevant for describing protein/lipid interactions. First, to be closest to the experimental conditions, we rationally designed the bicelles according to the IBM (see the Supporting Information) (Figure S8). The radial distribution function of the lipids indicated that all simulations present stable self-organized bicelles after 20 to 30 ns of CG-MD (Figure S9), which is a long time before the recruitment of the protein fragment observed at 440 and 320 ns for R1-3/HZB and R1-3/HAB complexes,

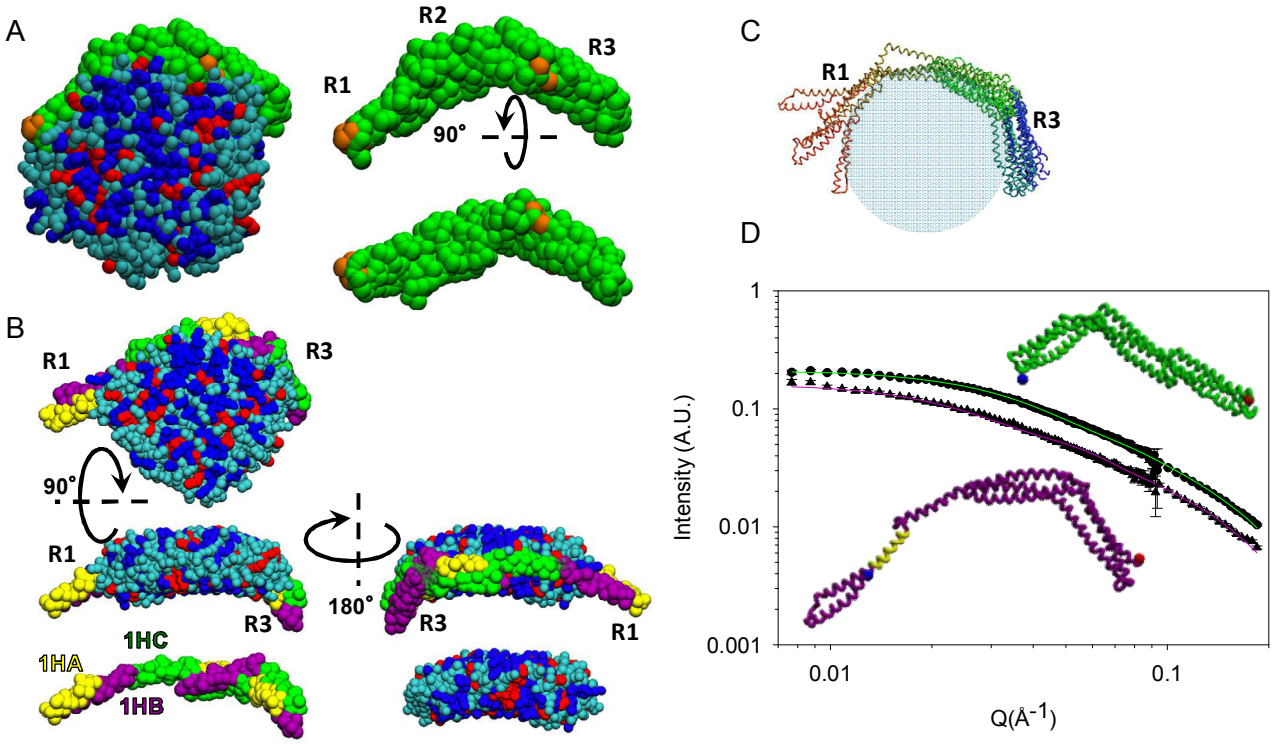


Figure 4: (*A*) Final structural model obtained by CG-MD of R1-3 bound to anionic bicelles (HAB). The protein is shown in green with orange spots corresponding to the pacFA PC cross-linked regions identified by mass spectrometry (see text). Lipids are colored in blue for DMPC, red for DMPS, and cyan for DHPC. (*B*) CG final model for the R1-3/anionic bicelle complex. HA, HB, and HC helices of each repeat are shown in yellow, purple, and green, respectively. (*C*) Collection of the five protein all-atom models obtained after opening the 1HA and 1HB helical regions of the R1 repeat through CG-IMD simulations. (*D*) Experimental SANS intensities (circles and triangles for R1-3 alone or bound to anionic bicelles, respectively) fitted with the theoretical CRYSON curves generated from the R1-3 model free in solution (green) and the R1-3 model bound to anionic bicelles (purple). *Inset*: The corresponding all-atom models are represented with the same color code. Blue and red dots represent the N- and the C-termini, respectively. The putative ALPS motif is colored in yellow (see text).

respectively (Figure S9). R1-3/HZB and R1-3/HAB CG models, obtained after 20 µs of MD simulation, were considered reliable because the relative deviation of the protein particle positions remains globally stable during the last half of the trajectories (Figure S9). Whereas the model proposed for R1-3 in association with HZB is in line with the SANS data, the model computed for R1-3 in association with HAB does not show any structural clue concerning the R_g variation observed experimentally. Surprisingly, in both simulations, the R1-3 dystrophin fragments are spontaneously recruited to the bicelle surface at its torus part (Figure 4A), which presents a high level of default lipid packing, whatever the nature of the bicelle (HZB or HAB). To check these *in silico* results experimentally, we labeled the bicelles with a bifunctional phospholipid[44] (pacFA PC) bearing a photoactivatable diazirin group and a clickable alkyne group (Figure S10A). This modified phospholipid is expected to be found in the bilayer part of the bicelles and to be able to diffuse throughout this part. Its photoactivable group is located close to the fatty acid end of one tail of the phospholipid. Thus, the cross-link should occur only if R1-3 is close to this diazirin group, *i.e.,* if R1-3 is indeed located at the torus as suggested by CG-MD data. The R1-3/bicelle complexes were exposed to UVA and then analyzed by SDS-PAGE (Figure S10B). The smeared profile indicates that the phospholipid is linked to R1-3 for both HZB and HAB upon photoactivation (Figure S10B). To ensure this assumption, the R1-3/lipid complexes were cross-linked at the alkyne group of the pacFA PC after the addition of biotin-azide *via* an azide-alkyne reaction. As expected, the western blot analysis, revealed with an HRP-streptavidin fusion protein, indicates the presence of the modified phospholipids at approximately 40 kDa (Figure S10C) assessing the covalent bond with R1-3. Finally, the smeared parts of the gel (Figure S10A) were submitted to a gel-digestion prior to LC-MSMS analysis. Thanks to this method, we could identify three peptides of R1-3 (R1-3 [77-95], R1-3 [229-243], and R1-3 [305-314]) linked to the phospholipid (Table S2 and Figure S11) corresponding to the protein/lipid interaction areas located in the R1 and R3 repeats (Figure S10D) for both types of bicelles, as suggested by CG-MD. Another possibility that would allow the click reaction to occur is that R1-3 is located inside the bilayer part of the bicelles. To discard this unlikely scenario, we used SANS to measure the R1-3/DZB and R1-3/DAB complexes where R1-3 was contrast-matched, *i.e.*, in 42% d-TNE buffer (Figure S12). The regular disk-shaped morphology of the bicelles is preserved, suggesting that R1-3 is not located inside the bilayer, which would have modified the morphology of the bicelles (Figure S12). To reinforce these results, we performed transmission electron microscopy (TEM) analysis of the R1-3/HAB complex, which presents the highest conformational modifications. To facilitate the localization of R1-3, the protein fragment was labeled with gold nanoparticles prior to TEM analysis. Representative micrographs of R1-3/HAB negatively stained complexes clearly highlight that R1-3 is found at the torus of the bicelles (Figure S13). To sum up, these results provide evidence concerning both the localization of R1-3 at the torus of the bicelles and an accurate

mapping of the protein/lipid interactions at the R1 and R3 coiled-coil repeats of the R1-3 dystrophin fragment.

Probing the conformational modifications of R1-3 in interactions with anionic bicelles. SANS data of the R1-3/DAB complex (unlike the R1-3/DZB complex) suggest that the conformation of R1-3 in the presence of DAB is significantly more extended. This last dataset was thus chosen for further investigation by CG-MD to understand such a conformational modification. The model of R1-3 bound to anionic bicelles from classic CG-MD is not in line with both the experimental SANS data and does not perfectly match MS mapping (Figure 4A). An opening of the coiled-coil could explain the R_g and D_{max} increase of R1-3 obtained in the corresponding SANS experiments (Table S1), as well as the extended *ab initio* envelope (Figure 3A). Moreover, the opening of the coiled-coil at one of the two extremities (R1 or R3) could support MS mapping, discarding the R2 opening. CG-IMD simulations enabled the opening of the dystrophin coiled-coil, thanks to a force applied interactively during CG-MD of the R1-3/HAB system. Reaching spontaneously and/or exploring exhaustively such a protein open state could be difficult in reasonable times of classical simulation. Therefore, modeling an extended protein was achieved through four different CG-IMD scenarios to systematically explore the plausible conformational modifications (Figure S14 and Video S1): opening, either R1, or R3, spectrin-like repeats by user-driven constraints applied to their first or first two helices (1HA or 1HA-1HB in R1) or to their last or last two helices (3HB-3HC or 3HC in R3), respectively (Figures 1B and S14). Only two scenarios were in line with an increase of the R_g value compatible with the SANS data: (*i*) R1 opening through the swinging of the first two helices (1HA-1HB) and (*ii*) R3 opening through the swinging of the last two helices (3HB-3HC). Five models were generated from these two scenarios (Figure 4C) and compared with the experimental data. Among them, the model having the best overall agreement with MS contact mapping (Figure S15), a suitable R_g (59 Å), and the best fit of the SANS data (lowest χ^2 value) was considered the most reliable structural model for R1-3 interacting with an HAB bicelle (Figure 4D), *i.e.*, opening the 1HA-1HB helices of the R1 repeat. Interestingly, based on sequence similarity, a putative amphipathic lipid packing sensor (ALPS) motif was found in the helix 1HB of the R1 repeat of human dystrophin[72]. This motif was first identified in the protein ArfGAP1 (ADP-ribosylation factor GTPase-Activating Protein 1)[73]. The ALPS motifs are made up of 15-20 amino acids and are known to be unfolded in solution and to adopt an α-helical folding when they are in interaction with membrane lipids presenting default packing[74]. The putative ALPS, a motif identified in dystrophin, is close to the area that we hypothesize to open during the interaction with anionic bicelles (Figure 4D) and could be involved in the dystrophin/lipid interaction. To confirm this hypothesis, further experiments such as amino acid substitution are required. However, confirmation of this hypothesis is beyond the scope of the present paper.

CONCLUSIONS

We have characterized the protein/lipid interactions of two protein fragments, R1-3 and R11-15. These fragments belong to the central domain of dystrophin, a protein involved in severe genetic diseases. Obtaining structural information about the central domain of this peripheral membrane protein, free in solution or in the presence of lipids, is not accessible by XRC and/or NMR. We highlighted that R1-3 and R11-15 can interact with zwitterionic bicelles and even more with anionic bicelles, notably by employing microscale thermophoresis. By using the contrast-matched bicelle tool[17] in a SANS approach, we have demonstrated that R1-3 undergoes a large conformational modification when bound to bicelles containing anionic phospholipids, such R11-15 does but in a lesser extent. The R1-3 protein fragment is bordered by two non-structured hinges and is found in numerous therapeutic shortened dystrophins, making it a relevant protein subdomain. We used click-chemistry and mass spectrometry coupled with *in silico* CG-MD to determine a reliable mapping of the R1-3/lipid interactions (Figure 4A). Finally, using the innovative CG-IMD method, we proposed a relevant all-atom model of R1-3 in interactions with anionic bicelles, in total agreement with the experimental SANS data (Figure 4D). These models provide evidence that the R1 repeat of the dystrophin central domain could undergo an opening of its coiled-coil structure. This opening may be necessary for the membrane anchoring of dystrophin into the curved parts of the sarcolemma to firmly support the plasma membrane of muscle cells. Such an anchoring may indeed be involved in the scaffolding of the sarcolemma role played by dystrophin in muscle cells during the muscle elongation/contraction process. Therefore, our work may contribute to the design of rationalized mini-dystrophin as a gene therapy for the treatment of DMD and BMD patients.

Finally, in addition to their use in the SANS contrast variation technique, we demonstrate that bicelles are a versatile tool usable in MST, HPSEC-MALS, and CG-IMD. Therefore, our approaches can confidently be employed for the characterization of protein/lipid interactions as well as the structural analysis of other peripheral and even integral membrane proteins incompatible with high resolution structural methods.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental details about R1-3 and R11-15 protein purification, bicelle preparation, description of the ideal bicelle model (IBM), ^{31}P spectroscopy, high-pressure size-exclusion chromatography coupled with multi-angle and dynamic light scattering, Western blotting, nanogold labeling, transmission electron microscopy (TEM).

Protein and bicelle quality control, SANS data of bicelles used in MST experiments, protein/lipid interactions at 18°C, SANS data of DZB and DAB in 42 and 100% d-TNE buffer, Guinier, Kratky and Porod-Deby plots of R1-3 and R11-15, SANS data at 18 and 22°C, summary table of SANS parameters at 18 and 22°C, bicelle formation by CG-MD, radial distribution function of the CG-MD trajectories, chemical structure of the pacFA PC, SDS-PAGE and Western blotting of the click-chemistry experiment, identified protein/lipid areas, SANS data of the bicelles bound to contrast-matched R1-3, TEM of the R1-3/HAB complexes, CG-IMD scenarios and energy plots, contact frequency of the protein/lipid interactions.

AUTHOR INFORMATION

Corresponding Authors

*E-mail:
raphael.dos-santos-morais@univ-rennes1.fr
sophie.combet@cea.fr
jean-francois.hubert@univ-rennes1.fr

PresentAddresses

[§]present address: CNRS UPR 9080, Université Paris-Diderot, Sorbonne Paris Cité, F-75005 Paris, France.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

The authors thank the Association Française contre les Myopathies-Telethon, the Conseil Régional de Bretagne, the Laboratoire Léon-Brillouin (LLB, Saclay, France), and the Synchrotron SOLEIL (Saint-Aubin, France) for funding this project. The LLB, the Institut Laue-Langevin (ILL, Grenoble, France), and the Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ, Garching, Germany) neutron facilities are acknowledged for beamtime allocation and support on PACE, D22, KWS-1, and KWS-2 SANS instruments, respectively. The authors thank the Grand Equipement National de Calcul Intensif-GENCI program (projet DARI DYSIM). The authors acknowledge, at the University of Rennes 1, Christophe Tascon from the Centre de Purification des Protéines of the UMR-CNRS 6290, the PRISM platform for NMR facilities, BIOSIT for spectroscopy facilities, and Agnès Burel from the MRic plateform for electron microscopy experiment. We gratefully thank Gabriel David and Blandine Pineau from SOLEIL for their support in HPSEC-MALS experiments.

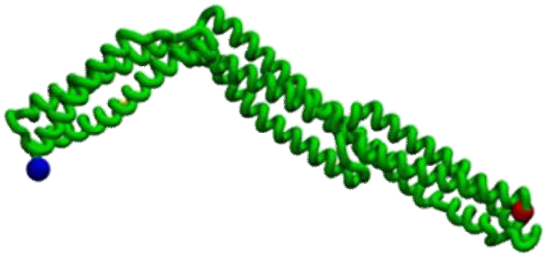
REFERENCES

(1) Seddon, A. M.; Curnow, P.; Booth, P. J. *Biochim. Biophys. Acta* **2004**, *1666* (1–2), 105.
(2) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.
(3) Nakane, T.; Hanashima, S.; Suzuki, M.; Saiki, H.; Hayashi, T.; Kakinouchi, K.; Sugiyama, S.; Kawatake, S.; Matsuoka, S.; Matsumori, N.; Nango, E.; Kobayashi, J.; Shimamura, T.; Kimura, K.; Mori, C.; Kunishima, N.; Sugahara, M.; Takakyu, Y.; Inoue, S.; Masuda, T.; Hosaka, T.; Tono, K.; Joti, Y.; Kameshima, T.; Hatsui, T.; Yabashi, M.; Inoue, T.; Nureki, O.; Iwata, S.; Murata, M.; Mizohata, E. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (46), 13039.
(4) Hendrickson, W. A. *Nat. Struct. Mol. Biol.* **2016**, *23* (6), 464.
(5) Moravcevic, K.; Oxley, C. L.; Lemmon, M. A. Struct. Lond. Engl. 1993 2012, 20 (1), 15.
(6) Dürr, U. H. N.; Gildenberg, M.; Ramamoorthy, A. *Chem. Rev.* **2012**, *112* (11), 6054.
(7) Kaptein, R.; Wagner, G. *J. Biomol. NMR* **2015**, *61* (0), 181.
(8) Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. *Protein Pept. Lett.* **2016**, *23* (3), 217.
(9) Petoukhov, M. V.; Svergun, D. I. *Curr. Opin. Struct. Biol.* **2007**, *17* (5), 562.
(10) Pérez, J.; Koutsioubas, A. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71* (1), 86.
(11) Berthaud, A.; Manzi, J.; Pérez, J.; Mangenot, S. *J. Am. Chem. Soc.* **2012**, *134* (24), 10080.
(12) Heller, W. T. *Acta Crystallogr. Sect. D* **2010**, *66* (11), 1213.
(13) Gabel, F.; Lensink, M. F.; Clantin, B.; Jacob-Dubuisson, F.; Villeret, V.; Ebel, C. *Biophys. J.* **2014**, *107* (1), 185.
(14) Le Cœur, C.; Combet, S.; Carrot, G.; Busch, P.; Teixeira, J.; Longeville, S. *Langmuir ACS J. Surf. Colloids* **2015**, *31* (30), 8402.
(15) Gabel, F. *Methods Enzymol.* **2015**, *558*, 391.
(16) Breyton, C.; Gabel, F.; Lethier, M.; Flayhan, A.; Durand, G.; Jault, J.-M.; Juillan-Binard, C.; Imbert, L.; Moulin, M.; Ravaud, S.; Härtlein, M.; Ebel, C. *Eur. Phys. J. E Soft Matter* **2013**, *36* (7), 71.
(17) Dos Santos Morais, R.; Delalande, O.; Pérez, J.; Mouret, L.; Bondon, A.; Martel, A.; Appavou, M.-S.; Le Rumeur, E.; Hubert, J.-F.; Combet, S. *Langmuir* **2017**, *33* (26), 6572.
(18) Monaco, A. P.; Neve, R. L.; Colletti-Feener, C.; Bertelson, C. J.; Kurnit, D. M.; Kunkel, L. M. *Nature* **1986**, *323* (6089), 646.
(19) Kusunoki, H.; Minasov, G.; Macdonald, R. I.; Mondragón, A. *J. Mol. Biol.* **2004**, *344* (2), 495.
(20) Le Rumeur, E.; Winder, S. J.; Hubert, J.-F. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2010**, *1804* (9), 1713.
(21) Legrand, B.; Giudice, E.; Nicolas, A.; Delalande, O.; Le Rumeur, E. *PloS One* **2011**, *6* (8), e23819.
(22) Rybakova, I. N.; Humston, J. L.; Sonnemann, K. J.; Ervasti, J. M. *J. Biol. Chem.* **2006**, *281* (15), 9996.
(23) Sarkis, J.; Vié, V.; Winder, S. J.; Renault, A.; Le Rumeur, E.; Hubert, J.-F. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2013**, *27* (1), 359.
(24) Lai, Y.; Thomas, G. D.; Yue, Y.; Yang, H. T.; Li, D.; Long, C.; Judge, L.; Bostick, B.; Chamberlain, J. S.; Terjung, R. L.; Duan, D. *J. Clin. Invest.* **2009**, *119* (3), 624.
(25) Campbell, K. P.; Kahl, S. D. *Nature* **1989**, *338* (6212), 259.
(26) Belanto, J. J.; Mader, T. L.; Eckhoff, M. D.; Strandjord, D. M.; Banks, G. B.; Gardner, M. K.; Lowe, D. A.; Ervasti, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (15), 5723.
(27) Zhao, J.; Kodippili, K.; Yue, Y.; Hakim, C. H.; Wasala, L.; Pan, X.; Zhang, K.; Yang, N. N.; Duan, D.; Lai, Y. *Hum. Mol. Genet.* **2016**.
(28) Legardinier, S.; Hubert, J.-F.; Le Bihan, O.; Tascon, C.; Rocher, C.; Raguénès-Nicol, C.; Bondon, A.; Hardy, S.; Le Rumeur, E. *Biochim. Biophys. Acta* **2008**, *1784* (4), 672.
(29) Legardinier, S.; Raguénès-Nicol, C.; Tascon, C.; Rocher, C.; Hardy, S.; Hubert, J.-F.; Le Rumeur, E. *J. Mol. Biol.* **2009**, *389* (3), 546.
(30) Vié, V.; Legardinier, S.; Chieze, L.; Le Bihan, O.; Qin, Y.; Sarkis, J.; Hubert, J.-F.; Renault, A.; Desbat, B.; Le Rumeur, E. *Biochim. Biophys. Acta* **2010**, *1798* (8), 1503.
(31) Sarkis, J.; Hubert, J.-F.; Legrand, B.; Robert, E.; Chéron, A.; Jardin, J.; Hitti, E.; Le Rumeur, E.; Vié, V. *J. Biol. Chem.* **2011**, *286* (35), 30481.
(32) García-Pelagio, K. P.; Bloch, R. J.; Ortega, A.; González-Serratos, H. *J. Muscle Res. Cell Motil.* **2011**, *31* (0), 323.
(33) Norwood, F. L.; Sutherland-Smith, A. J.; Keep, N. H.; Kendrick-Jones, J. *Struct. Lond. Engl. 1993* **2000**, *8* (5), 481.
(34) Huang, X.; Poy, F.; Zhang, R.; Joachimiak, A.; Sudol, M.; Eck, M. J. *Nat. Struct. Biol.* **2000**, *7* (8), 634.
(35) Muthu, M.; Richardson, K. A.; Sutherland-Smith, A. J. *PloS One* **2012**, *7* (7), e40066.
(36) Molza, A.-E.; Férey, N.; Czjzek, M.; Le Rumeur, E.; Hubert, J.-F.; Tek, A.; Laurent, B.; Baaden, M.; Delalande, O. *Faraday Discuss.* **2014**, *169*, 45.
(37) Valentini, E.; Kikhney, A. G.; Previtali, G.; Jeffries, C. M.; Svergun, D. I. *Nucleic Acids Res.* **2015**, *43* (Database issue), D357.
(38) Mehboob, S.; Luo, B. H.; Patel, B. M.; Fung, L. W. *Biochemistry (Mosc.)* **2001**, *40* (41), 12457.
(39) Franke, D.; Petoukhov, M. V.; Konarev, P. V.; Panjkovich, A.; Tuukkanen, A.; Mertens, H. D. T.; Kikhney, A. G.; Hajizadeh, N. R.; Franklin, J. M.; Jeffries, C. M.; Svergun, D. I. *J. Appl. Crystallogr.* **2017**, *50* (4), 1212.

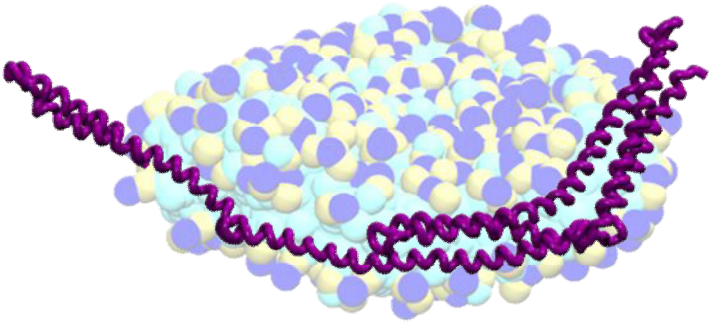
(40) Glatter, O.; Kratky, O. *Small Angle X-ray Scattering*; Academic Press, 1982.
(43) Kline, S. R. *J. Appl. Crystallogr.* **2006**, *39* (6), 895.
(44) Haberkant, P.; Raijmakers, R.; Wildwater, M.; Sachsenheimer, T.; Brügger, B.; Maeda, K.; Houweling, M.; Gavin, A.-C.; Schultz, C.; van Meer, G.; Heck, A. J. R.; Holthuis, J. C. M. *Angew. Chem. Int. Ed Engl.* **2013**, *52* (14), 4033.
(45) Lavigne, R.; Becker, E.; Liu, Y.; Evrard, B.; Lardenois, A.; Primig, M.; Pineau, C. *Mol. Cell. Proteomics MCP* **2012**, *11* (2), M111.012682.
(46) Jumeau, F.; Com, E.; Lane, L.; Duek, P.; Lagarrigue, M.; Lavigne, R.; Guillot, L.; Rondel, K.; Gateau, A.; Melaine, N.; Guével, B.; Sergeant, N.; Mitchell, V.; Pineau, C. *J. Proteome Res.* **2015**, *14* (9), 3606.
(47) Carapito, C.; Lane, L.; Benama, M.; Opsomer, A.; Mouton-Barbosa, E.; Garrigues, L.; Gonzalez de Peredo, A.; Burel, A.; Bruley, C.; Gateau, A.; Bouyssié, D.; Jaquinod, M.; Cianferani, S.; Burlet-Schiltz, O.; Van Dorsselaer, A.; Garin, J.; Vandenbrouck, Y. *J. Proteome Res.* **2015**, *14* (9), 3621.
(48) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435.
(49) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111* (27), 7812.
(50) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4* (5), 819.
(51) Vácha, R.; Frenkel, D. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (15), 4229.
(52) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2015**, *11* (5), 2144.
(53) Triba, M. N.; Warschawski, D. E.; Devaux, P. F. *Biophys. J.* **2005**, *88* (3), 1887.
(54) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. *J. Chem. Theory Comput.* **2014**, *10* (2), 676.
(55) Delalande, O.; Férey, N.; Grasseau, G.; Baaden, M. *J. Comput. Chem.* **2009**, *30* (15), 2375.
(56) Jeffries, C. M.; Graewert, M. A.; Blanchet, C. E.; Langley, D. B.; Whitten, A. E.; Svergun, D. I. *Nat. Protoc.* **2016**, *11* (11), 2122.
(57) Beaugrand, M.; Arnold, A. A.; Hénin, J.; Warschawski, D. E.; Williamson, P. T. F.; Marcotte, I. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (21), 6162.
(58) Leventis, P. A.; Grinstein, S. *Annu. Rev. Biophys.* **2010**, *39* (1), 407.
(59) Fiehn, W.; Peter, J. B.; Mead, J. F.; Gan-Elepano, M. *J. Biol. Chem.* **1971**, *246* (18), 5617.
(60) Wattraint, O.; Saadallah, I.; Silva-Pires, V.; Sonnet, P.; Sarazin, C. *Int. J. Pharm.* **2013**, *454* (2), 621.
(61) Gimpl, K.; Klement, J.; Keller, S. *Biol. Proced. Online* **2016**, *18*.
(62) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. *J. Mol. Struct.* **2014**, *1077*, 101.
(63) Situ, A. J.; Schmidt, T.; Mazumder, P.; Ulmer, T. S. *J. Mol. Biol.* **2014**, *426* (21), 3670.
(64) Ameziane-Le Hir, S.; Paboeuf, G.; Tascon, C.; Hubert, J.-F.; Le Rumeur, E.; Vié, V.; Raguénès-Nicol, C. *Biochemistry (Mosc.)* **2016**, *55* (29), 4018.
(65) van Dam, L.; Karlsson, G.; Edwards, K. *Biochim. Biophys. Acta* **2004**, *1664* (2), 241.
(66) Derksen, A.; Gantz, D.; Small, D. M. *Biophys. J.* **1996**, *70* (1), 330.
(67) Salamon, Z.; Tollin, G. *Biophys. J.* **1996**, *71* (2), 848.
(68) Struppe, J.; Whiles, J. A.; Vold, R. R. *Biophys. J.* **2000**, *78* (1), 281.
(69) Guard-Friar, D.; Chen, C. H.; Engle, A. S. *J. Phys. Chem.* **1985**, *89* (9), 1810.
(70) Meng, J.; Counsell, J. R.; Reza, M.; Laval, S. H.; Danos, O.; Thrasher, A.; Lochmüller, H.; Muntoni, F.; Morgan, J. E. *Sci. Rep.* **2016**, *6*, 19750.
(71) Wang, B.; Li, J.; Xiao, X. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (25), 13714.
(72) Drin, G.; Casella, J.-F.; Gautier, R.; Boehmer, T.; Schwartz, T. U.; Antonny, B. *Nat. Struct. Mol. Biol.* **2007**, *14* (2), 138.
(73) Antonny, B.; Huber, I.; Paris, S.; Chabre, M.; Cassel, D. *J. Biol. Chem.* **1997**, *272* (49), 30848.
(74) Antonny, B. *Annu. Rev. Biochem.* **2011**, *80*, 101.

Insert Table of Contents artwork here

SANS-based model of a dystrophin fragment free in solution



Contrast-matched anionic bicelle



SANS-based model of a dystrophin fragment bound to anionic bicelle

Supporting Information

Dystrophin-membrane interactions: evidence of conformational modification upon binding to anionic phospholipids

Raphael Dos Santos Morais[1,2,3,4*], Olivier Delalande[1,2], Javier Pérez[4], Dominique Mias-Lucquin[1,2], Mélanie Lagarrigue[1,5], Anne Martel[6], Anne-Elisabeth Molza[1,2,§], Angélique Chéron[1,2], Thomas Chenuel[1,2], Arnaud Bondon[7], Marie-Sousai Appavou[8], Elisabeth Le Rumeur[1,2], Sophie Combet[3*], and Jean-François Hubert[1,2*].

[1]Université de Rennes 1, F-35043 Rennes, France. [2]Institut de Génétique et Développement de Rennes, CNRS UMR 6290, F-35043 Rennes, France. [3]Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, F-91191Gif-sur-Yvette CEDEX, France. [4]SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP 48, Saint-Aubin, F-91192 Gif-sur-Yvette, France. [5]Inserm U 1085, Protim-Plate-forme Protéomique. [6]Institut Laue-Langevin, F-38042 Grenoble, France. [7]CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-35043 Rennes, France. [8]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany.

EXPERIMENTAL SECTION

R1-3 and R11-15 purification. R1-3 and R11-15 are designed according to the alignment of Winder[1].

Name	Sequence	MW (kDa)	Number of residues	ε ($M^{-1}.cm^{-1}$)
R1-3	*GS*EVNLD…QISQA	38.5	333	59,720
R11-15	*GS*FQKPA…LNFAQ	60.1	515	45,950

Residues in italic are from thrombin cleavage site.
Human dystrophin (Uniprot P11532).

R1-3 and R11-15 were expressed and purified as previously described with small modifications[2,3]. Briefly, the His-tagged proteins were produced in *E. coli* BL21 (DE3) strain and purified by immobilized metal affinity chromatography (IMAC) on Ni-Sepharose column (HisTrap, GE Healthcare) according to the manufacturer's instructions. The tag was removed by thrombin cleavage and the proteins were further purified with a size exclusion chromatography column (HiLoad 16/600 Superdex 200 prep. grade, GE Healthcare) equilibrated with TNE buffer (20 mM Tris, 150 mM NaCl, and 0.1 mM EDTA, pH 7.5). The purity was assessed by SDS-PAGE stained with Coomassie blue (InstantBlue, Expedeon) and the concentration was determined spectrophotometrically using molar extinction coefficients at 280 nm of 59,720 and 45,950 $M^{-1}.cm^{-1}$ for R1-3 (38.5 kDa) and R11-15 (60 kDa), respectively[2,3]. Buffer exchange, for SANS experiments, was performed by three successive diafiltrations with Amicon Ultra-15 or Ultra-4 (Millipore, MWCO 10 kDa), followed by an ultimate step with a desalting column (NAP5, GE Healthcare) to ensure a perfect buffer exchange.

Bicelle preparation. 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dimyristoyl-sn-glycero-3-phospho-L-serine (DMPS), 1,2-dihexanoyl-sn-glycero-3-phosphocholine (DHPC), 1,2-dimyristoyl-d54-sn-glycero-3-phosphocholine-1,1,2,2-d4-N,N,N-trimethyl-d9 (DMPC-d67), 1,2-dimyristoyl-d54-sn-glycero-3-[phospho-L-serine] (DMPS-d54), and 1,2-dihexanoyl-d22-sn-glycero-3-phosphocholine-1,1,2,2-d4-N,N,N-trimethyl-d9 (DHPC-d35), conditioned in chloroform or chloroform/methanol, were purchased from Avanti Polar Lipids and used without any further purification. D_2O, Tris-d11, and EDTA-d16 were from Eurisotop. Chloroform or chloroform/methanol solutions containing the appropriate amounts of either zwitterionic (DMPC/DHPC, 1:1) or anionic (DMPC/DMPS/DHPC, 0.67:0:33:1) lipid mixtures were dried overnight under vacuum. The lipids were then rehydrated in TNE or d-TNE buffer solution (20 mM Tris-d11, 150 mM NaCl, and 0.1 mM EDTA-d16, pD 7.5) to reach a total lipid concentration of at least 200 mM. Then, the solution was frozen in liquid N_2, thawed (10 min at 40°C), vigorously shaken with a vortex (1 min), and then centrifuged (1.5 min, 6,000 rpm,

MiniSpin, Eppendorf). This procedure was repeated twice, to obtain a clear suspension of zwitterionic or anionic bicelles in hydrogenated or deuterated version, denoted HZB/DZB and HAB/DAB, respectively. Stock solutions were diluted in TNE or d-TNE depending on the experiments.

The ideal bicelle model (IBM). The molar ratio of long-to short-chain phospholipids is denoted "q". It is the main parameter governing the size of the bicelles and is denoted "q_{eff}"[4,5], for effective molar ratio, when the proportion of free DHPC is considered and is defined by the equation Eq. S1.

$$q_{eff} = \frac{[\mathrm{LCP}]}{[\mathrm{DHPC}]_{\mathrm{total}} - [\mathrm{DHPC}]_{\mathrm{free}}} \qquad \text{Eq. S1}$$

where [LCP] is the concentration of long-chain phospholipids, *i.e.* DMPC and DMPS in the present study, and [DHPC] is the concentration of short-chain phospholipids. The model proposed by Triba *et al.*[6], based on the volumes of the phospholipids, was used to estimate the radius of the bicelles by the equation Eq. S2[6,7].

$$\mathrm{R} = \mathrm{r}_{\perp} + \frac{\mathrm{r}_{\perp}\mathrm{q}_{\mathrm{eff}}}{4\Lambda}\left[\pi + \left(\pi^2 + \frac{32\Lambda}{3\mathrm{q}_{\mathrm{eff}}}\right)^{1/2}\right] \qquad \text{Eq. S2}$$

where R is the radius of the bicelle, $r_{\perp}$ is the thickness of the rim (11 Å, the length of a DHPC molecule), and Λ is, either the volume ratio of DHPC to DMPC (0.60), or to the weighted average volume of DMPC and DMPS (0.62). The molecular volumes of DMPC, DMPS, and DHPC are 1,101[8], 979[9], and 660[10] Å^3, respectively. See Dos Santos Morais *et al.*[11] for further information.

^{31}P NMR spectroscopy. ^{31}P NMR spectra of the HZB and HAB were recorded on a Bruker spectrometer Avance III HD 500 equipped with a 5-mm BBO probe operating at 202.46 MHz. Data were recorded at 20°C in d-TNE buffer at a concentration of 240 mM of lipids, with $q = 1$ in 3 mm capillary (200 µL). Spectra of 10 kHz spectral width and 32 K data points were acquired with proton decoupling using 128 scans, a 30° flip angle, and a 0.5 s relaxation delay. The data were processed with the TopSpin3.2 software (Bruker). Before applying the Fourier transform, free induction decays were treated with an exponential broadening of 2 Hz. 85 % H_3PO_4 was used as an external standard for the ^{31}P chemical shift.

High-pressure size-exclusion chromatography coupled with multi-angle light scattering (HPSEC-MALS). HPSEC-multi-angle light scattering (MALS) experiments were performed with an HPLC system (Agilent) equipped with UV detector module coupled with light scattering (miniDAWN TREOS, Wyatt) and refractive index (Viscotek, Malvern) detectors. To determine the molecular weight, the refractive index increment values (dn/dc) were 0.185 mL/g for the

proteins and 0.16 mL/g for the phospholipids[12]. The SEC column (5 µm, 4.6 mm x 300 mm, Agilent) was equilibrated with TNE buffer supplemented with 6 mM of DHPC, corresponding to the $[DHPC]_{free}$, also known as the critical bicellar concentration[5], to avoid bicelle deformation under diluted conditions[13]. The flow rate was 0.3 mL/min and data were processed with the Astra software v6.1 (Wyatt).

High-pressure size-exclusion chromatography coupled with dynamic light scattering (HPSEC-DLS). HPSEC-DLS were recorded with another HPLC system (Agilent) coupled with a light scattering detector (DAWN HELEOS II, Wyatt) in the same conditions as for HPSEC-MALS.

Western blotting. The protein fragment of the protein/lipid complex was precipitated with $CHCl_3$/MeOH (*v*/*v*) and then centrifuged for 5 min at 13,400 rpm (Minispin, Eppendorf). The supernatant containing the solubilized lipids composing the bicelles was removed. The protein pellet was resolubilized overnight at room temperature in AAS buffer (50 mM ammonium acetate,pH 7.0, 0.2%SDS). The protein solution was combined with 1 mM TCEP, 1 mM $CuSO_4$ and 1 mM biotin azide (Lumiprobe) in AAS buffer and incubated for 1h. The sample was loaded into a precast gel (Expedeon) for SDS-PAGE and then transferred onto a nitrocellulose membrane colored with Ponceau. The revelation is performed by electrochemiluminescence using a streptavidin-HRP fusion protein and read with an ImageQuant LAS 4000 device (GE Healthcare).

Gold nanoparticle labeling, negative staining, and transmission electron microscopy (TEM). The R1-3 protein fragment was labeled with gold nanoparticle (1.4 nm monomaleimido Nanogold, Nanoprobes) *via* cysteines according to the manufacturer's instructions. The excess of free nanoparticles was removed by HPSEC with the same column than above HPSEC experiments, equilibrated with TNE buffer. The protein-Nanogold conjugate was added to HAB for a final concentration of 1 µM of conjugate and 15mM of lipids. The sample was deposed onto a carbon grid (30 s), then the grid was washed twice with MQ water (30 s), and finally stained with uranyl acetate (2%) (1 min). The grid was visualized with a JEM 1400 electron microscope (JEOL).

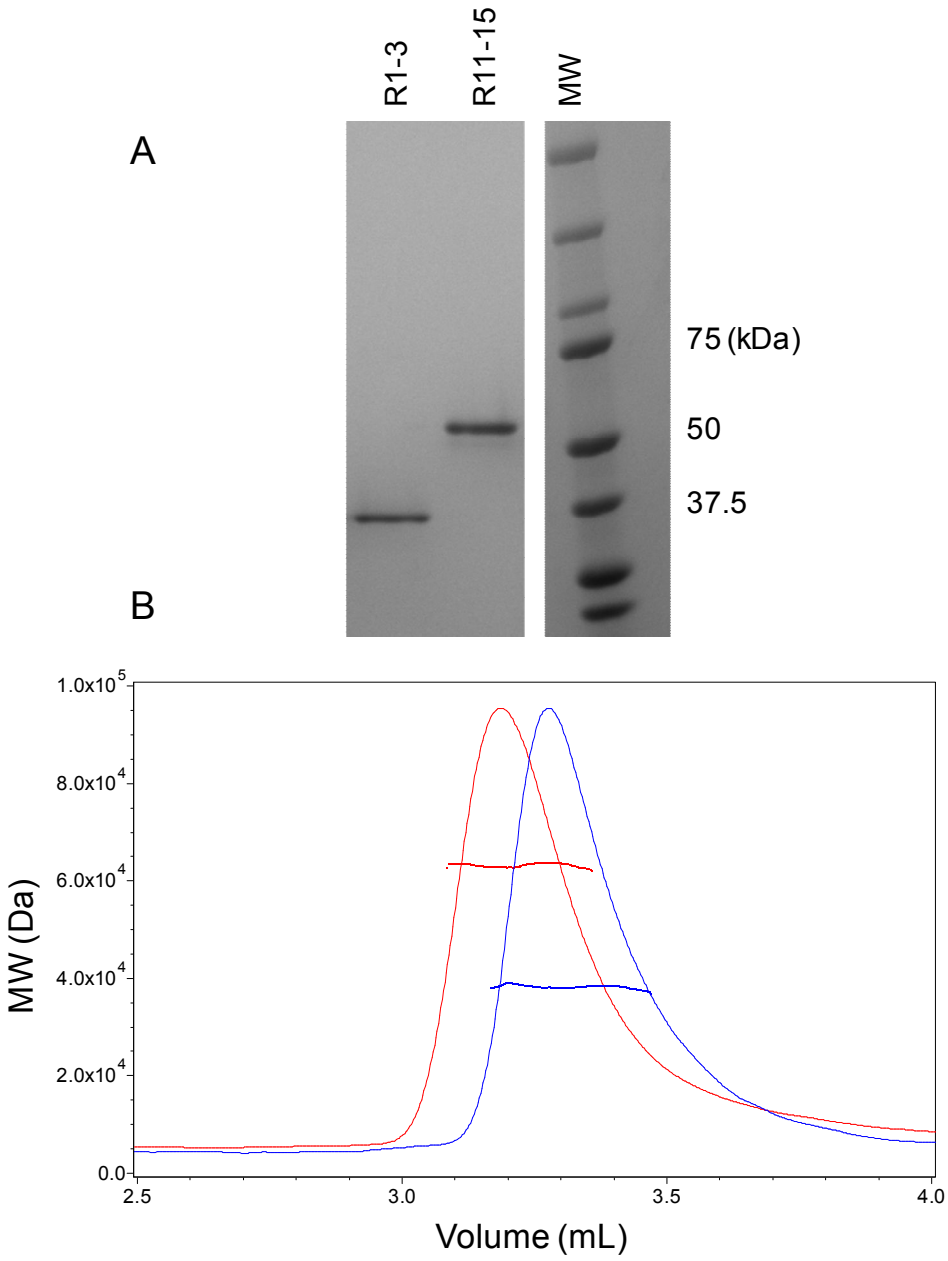


Figure S1: Quality control of the protein fragments. (*A*) SDS-PAGE analysis revealed by Coomassie Blue staining. (*B*) HPSEC-MALS chromatograms of the R1-3 (blue) and R11-15 (red) protein fragments. SDS-PAGE and SEC-MALS chromatograms demonstrate the purity, the monodispersity, as well as the monomeric state of the two protein fragments.

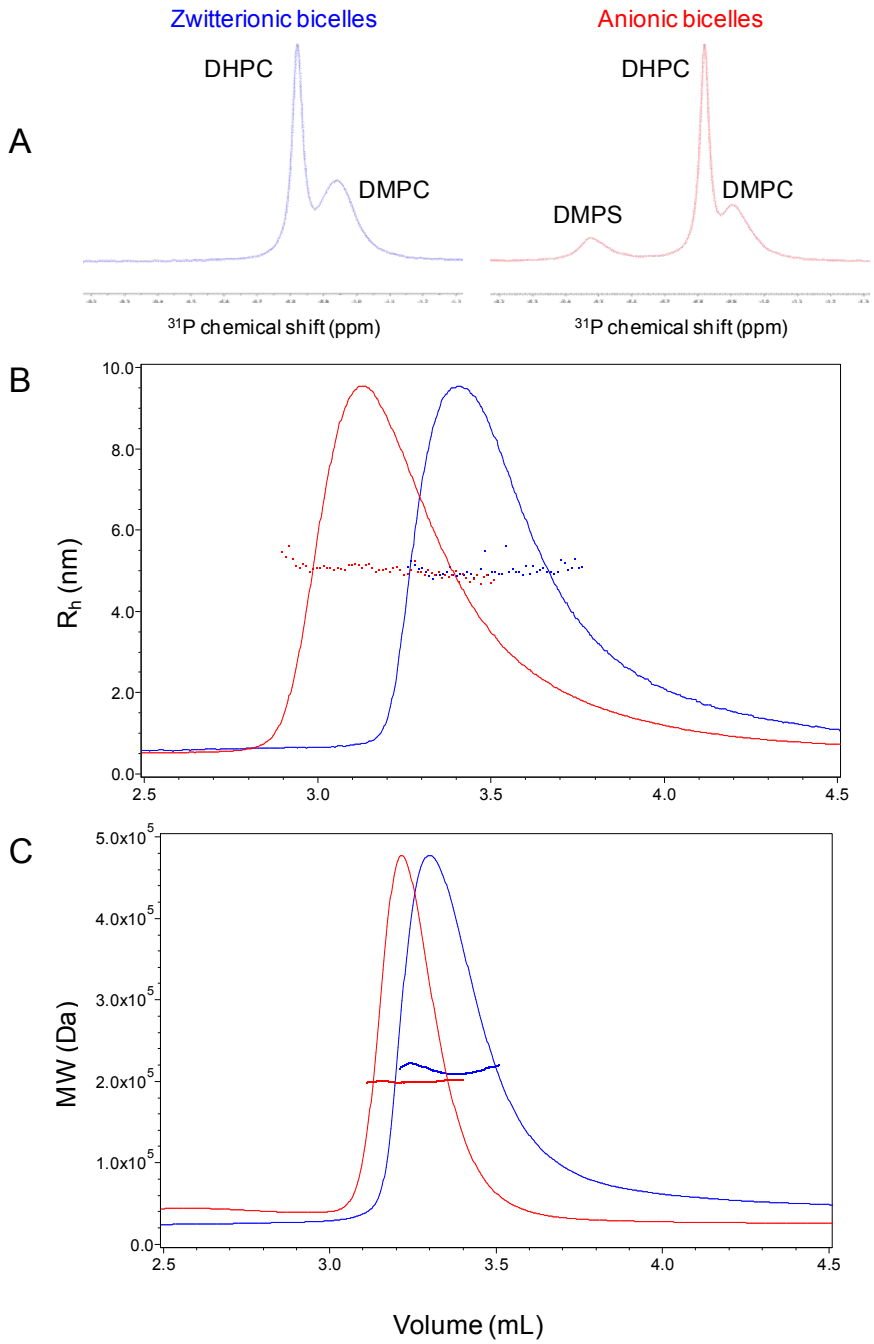


Figure S2: Quality control of the zwitterionic (blue) and anionic (red) bicelles. (*A*) ^{31}P NMR spectra assess the proper organization in bicelles with a *q* ratio equal to 1. (*B*) HPSEC-DLS chromatograms highlight the monodispersity of the bicelles with a constant R_h of 4.9 ± .1 and 5.0 ± .1 nm for the zwitterionic and anionic bicelles, respectively. (*C*) HPSEC-MALS chromatograms show a constant MW of approximately 200 kDa for both types of bicelles all along the peak, in line with the aggregation number predicted by the IBM[6].

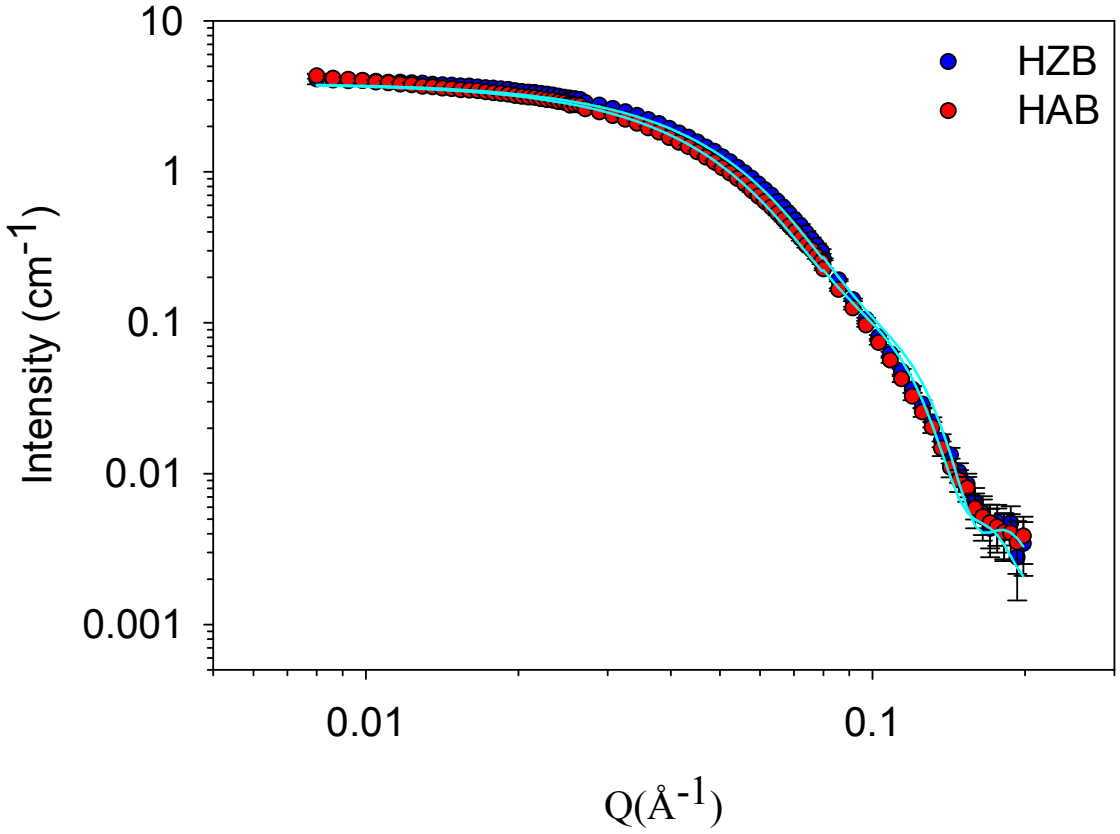


Figure S3: Structural characterization of the HZB (blue) and HAB (red) bicelles at 15 mM of lipids (q_{eff} = 1.3) at 22°C, as used in the MST experiments. SANS data, measured in 100% d-TNE buffer, were fitted (cyan) with the "core-shell cylinder" model, giving the expected dimensions of 43 Å thickness for both types of bicelles and 50 and 54 Å radius for HZB and HAB, respectively.

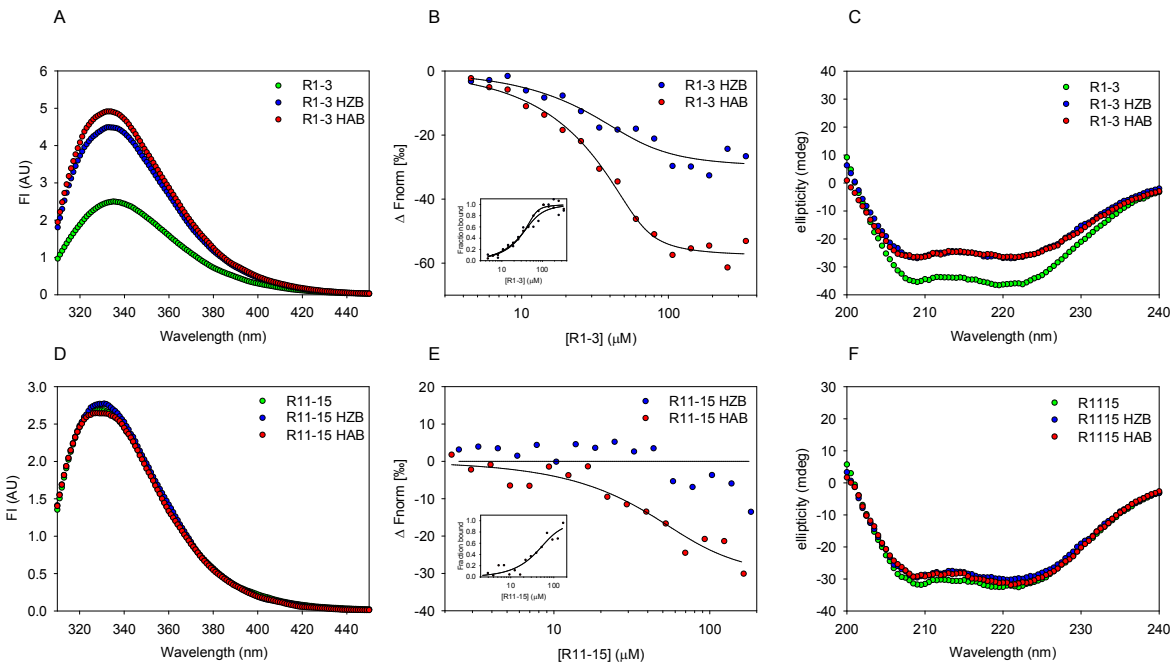


Figure S4: Characterization of the protein/lipid interactions at 18°C for R1-3 (*A*, *B*, *C*) and R11-15 (*D*, *E*, *F*) dystrophin fragments. (*A* and *D*) Tryptophan fluorescence intensity (FI) of each protein in the absence (green) or in the presence of zwitterionic (HZB, blue) or anionic (HAB, red) bicelles. FI variation is due to a modification in Trp environment attributed to lipid binding. (*B* and *E*) Microscale thermophoresis (MST) data: the dissociation constant (K_d) was determined to be around 10 µM for each protein fragment/bicelle complex except for R11-15/HZB. (*C* and *F*) Far-UV CD spectra with the same legend as for FI (*A* and *D*). CD spectra highlight that the α-helical folding of both protein fragments is maintained during lipid binding.

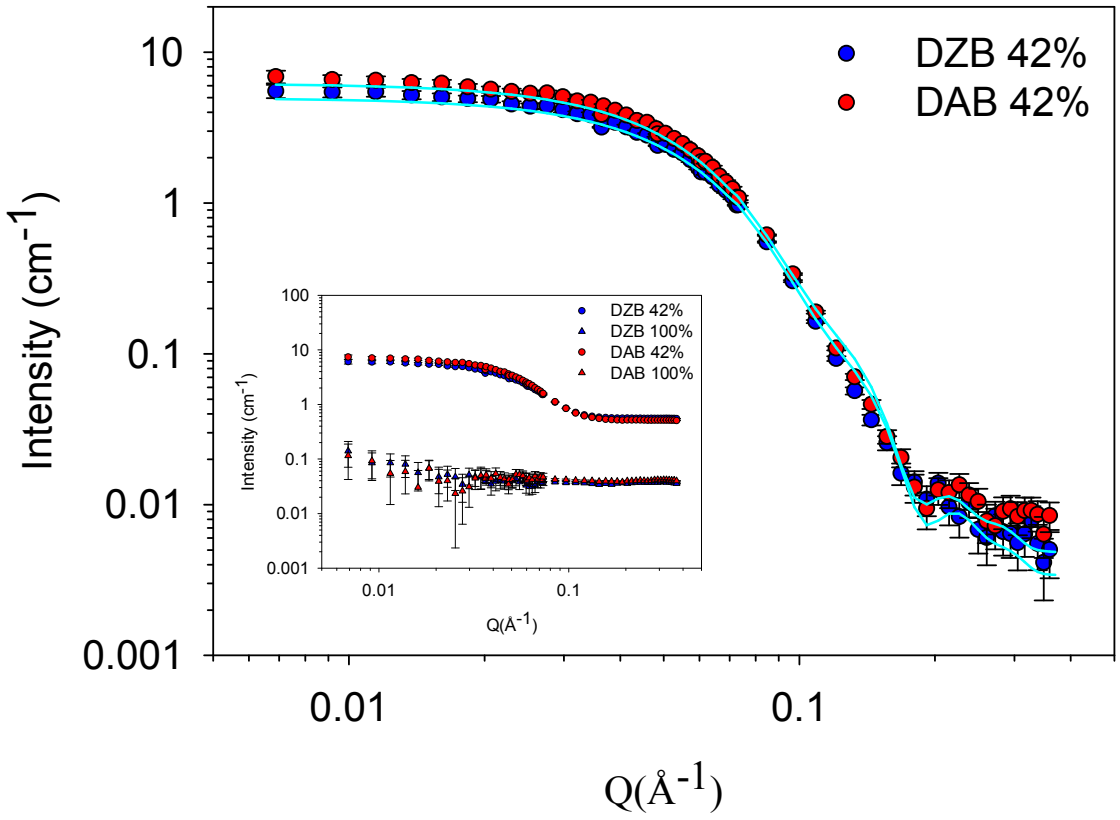


Figure S5: SANS intensities for deuterated zwitterionic (DZB) (blue) and anionic (DAB) (red) bicelles used in the present study at 50 mM of lipids (q_{eff} = 1.3) in 42% (circles) d-TNE buffer, at 22°C. SANS data are fitted with the "core-shell cylinder" model. *Inset*: non-subtracted SANS data in 42% (circles) and 100% (triangles) d-TNE buffer. Both types of bicelles are contrast-matched in 100% d-TNE buffer.

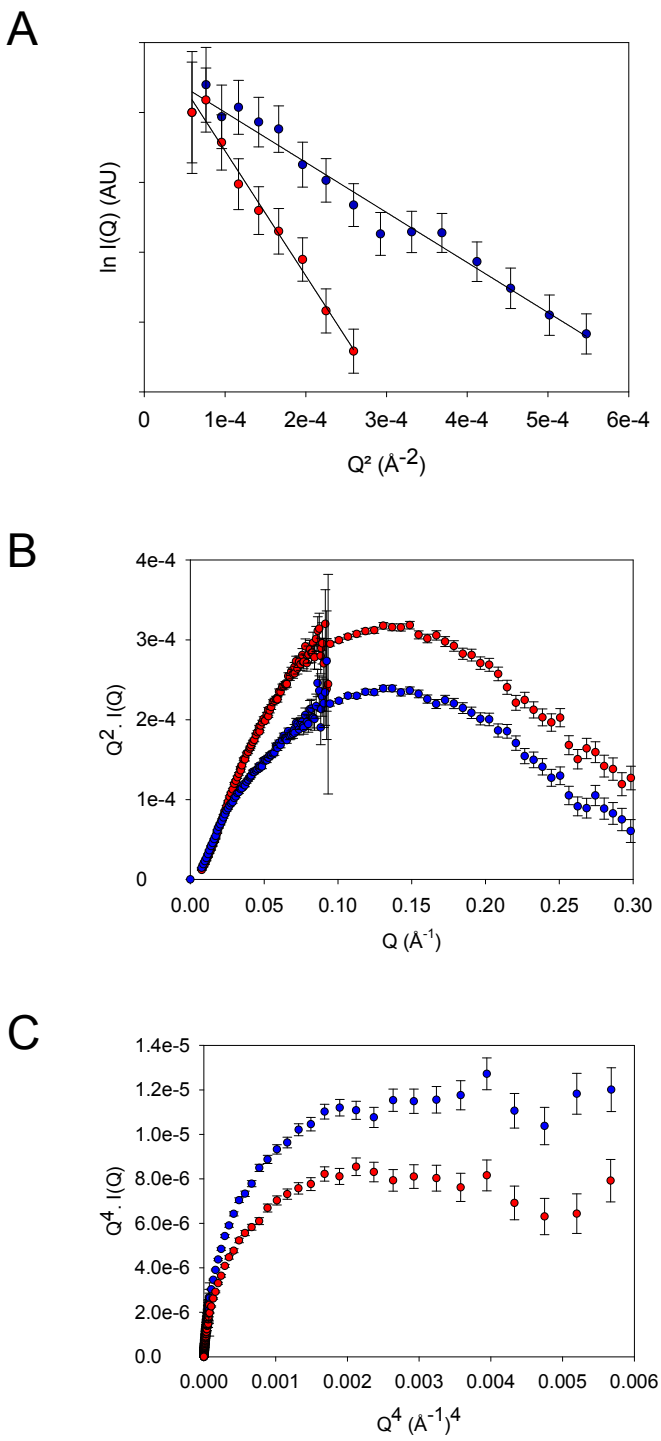


Figure S6: Guinier plots (*A*) for the R1-3 (blue) and R11-15 (red) protein fragments showing the absence of aggregates: the R_g found are 42 ± 1 and 62 ± 2 Å for R1-3 and R11-15, respectively. (B) Kratky and (C) Porod-Debye plots of R1-3 (blue) and R11-15 (red) highlight the presence of well-folded and compact protein fragments.

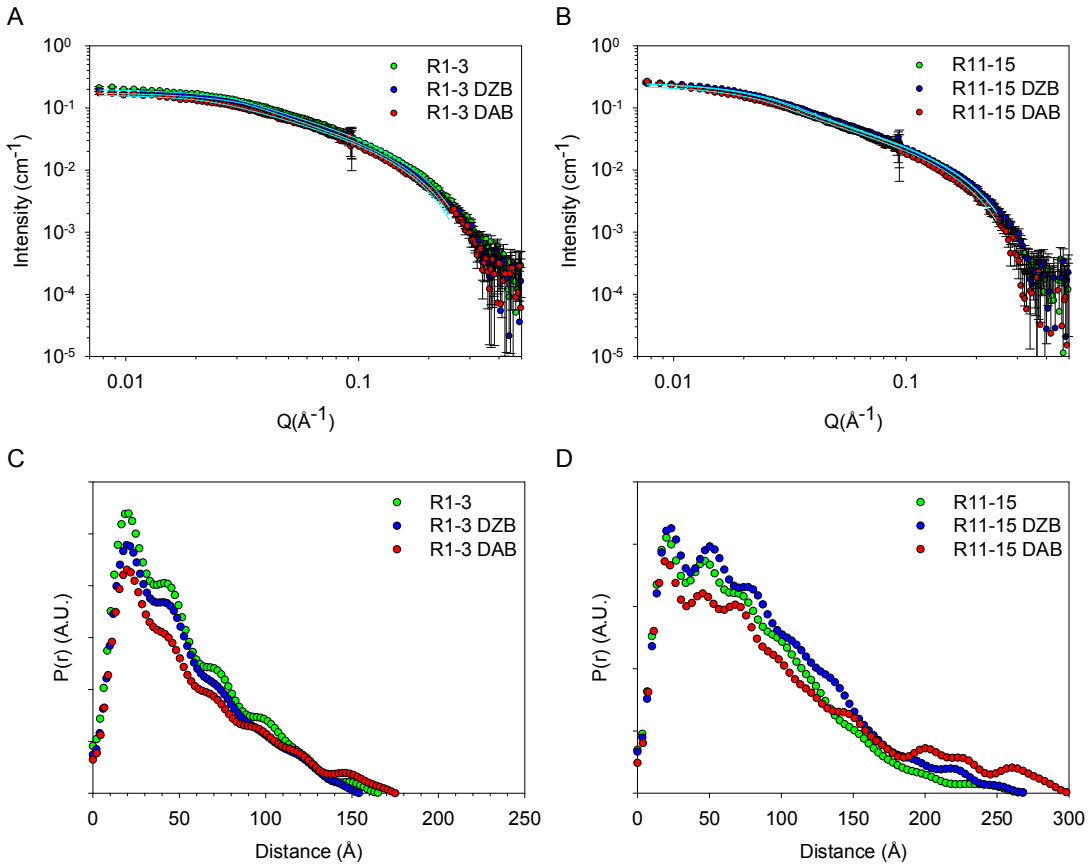


Figure S7: SANS intensities measured at 18°C for R1-3 (*A*) and R11-15 (*B*) alone in solution (green) or in interaction with contrast-matched zwitterionic (blue) or anionic (red) bicelles. (*C* and *D*) Pair-distribution function *P*(*r*), with the same color code, shows an increase of D_{max} when the protein fragments are in interaction with anionic bicelles (red). Figure S7B[13] is reprinted with permission from Dos Santos Morais, R.; Delalande, O.; Pérez, J.; Mouret, L.; Bondon, A.; Martel, A.; Appavou, M.-S.; Le Rumeur, E.; Hubert, J.-F.; Combet, S. *Langmuir* **2017**. Copyright 2017 American Chemical Society.

Table S1: SANS parameters of the R1-3 and R11-15 protein fragments, alone or in interaction with bicelles, at 18 and 22°C.

	R1-3	R1-3/ DZB	R1-3/ DAB	R11-15	R11-15/ DZB	R11-15/ DAB	R1-3	R1-3/ DZB	R1-3/ DAB	R11-15	R11-15/ DZB	R11-15/ DAB
	18°C						***22°C***					
Guinier approximation:												
[¥]QR_g range	0.32-0.97(16)	0.30-0.92 (16)	0.37-0.98 (13)	0.32-0.97(16)	0.32-0.97 (16)	0.47-0.99 (9)	0.32-0.97 (16)	0.32-0.97 (16)	0.47-0.99 (13)	0.47-0.99 (9)	0.48-0.99 (9)	0.62-1.1 (7)
R_g (Å)	42 ± 1	39 ± 1	49 ± 2	60 ± 2	59 ± 2	78 ± 2	42 ± 1	41 ± 1	62 ± 4	62 ± 2	62 ± 2	81 ± 4
R_g^{real} (Å) 1	43	42	46	63	62	79	43	44	61	65	61	77
D_{max} (Å)	165	156	175	268	242	302	177	178	248	274	281	296
$MW^{experimental}$ (kDa)	45 ± 9[§] (56)	49	47	55 ± 3[§] (52)	53	55	45 ± 9[§] (56)	45	46	55 ± 3[§] (52)	52	53
"Flexible cylinder" model:												
Radius (Å)	11	11	11	10	10	11	10	11	11	10	10	10
Length (Å)	158	159	157	281	274	277	156	157	195	276	274	296
Kuhn length (Å)	201	98	176	117	111	128	111	193	162	120	120	125

[¥] Values in brackets indicate the number of points used to determine R_g. [§]The MWs of the protein fragments were determined from three independent experiments; the value in brackets is from the present experiment. The SLDs are 3.11 and 3.12 (10^{-6} $Å^{-2}$) for R1-3 and R11-15, respectively. The V_P are 0.72 and 0.73 (cm^3 g^{-1}) for R1-3 and R11-15, respectively.

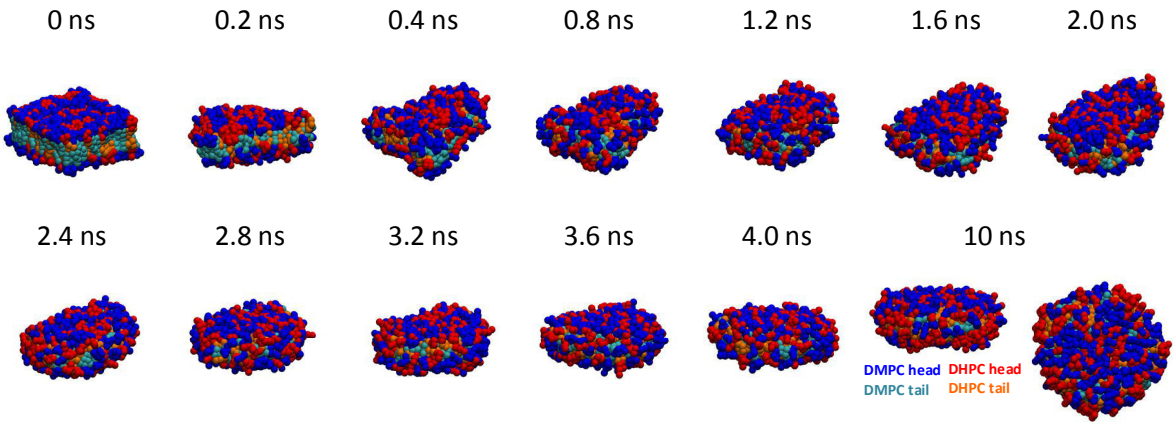


Figure S8: Equilibration of zwitterionic bicelle (HZB) spontaneously formed after 10 ns of CG-MD simulation, from a randomly generated phospholipid bilayer composed of DMPC and DHPC ($q = 1$), by using insane.py script provided for the GROMACS[14] software. Phospholipid head and tail are shown, respectively, in blue and cyan for DMPC and in red and orange for DHPC.

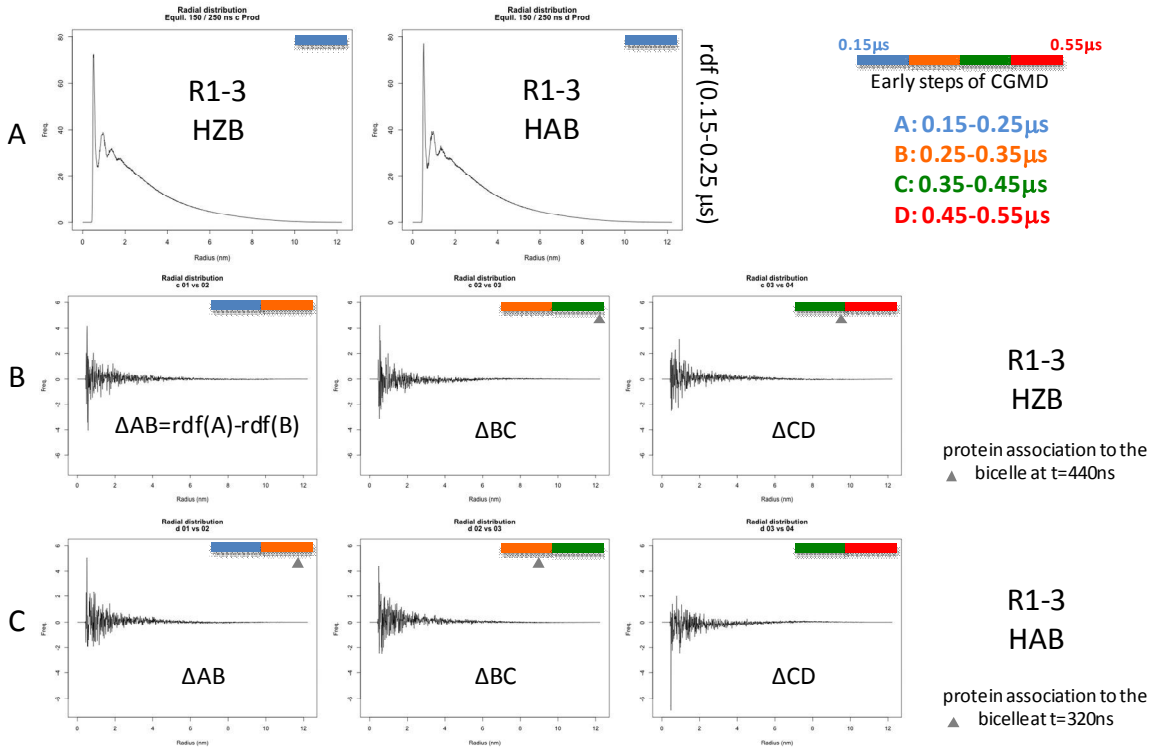


Figure S9: Radial distribution function (rdf) of lipids in the early steps of the CG-MD trajectory (*A*), including system states before and after R1-3 recruitment to HZB or HAB (*B* and *C*). (*A*) Plots of rdf for the 0.15-0.25 µs-range of the trajectory show a well-stabilized bicelle organization before the recruitment of R1-3. (*B* and *C*) The plots corresponding to the difference between two rdf are measured on four short trajectory ranges, as indicated in the schematic colored representation. Small amplitudes of lipid rdf for the system before and after interaction of the R1-3 protein fragment (shown by a grey triangle) with the bicelle indicate that only a light deformation of the bicelle is observed upon association with the protein.

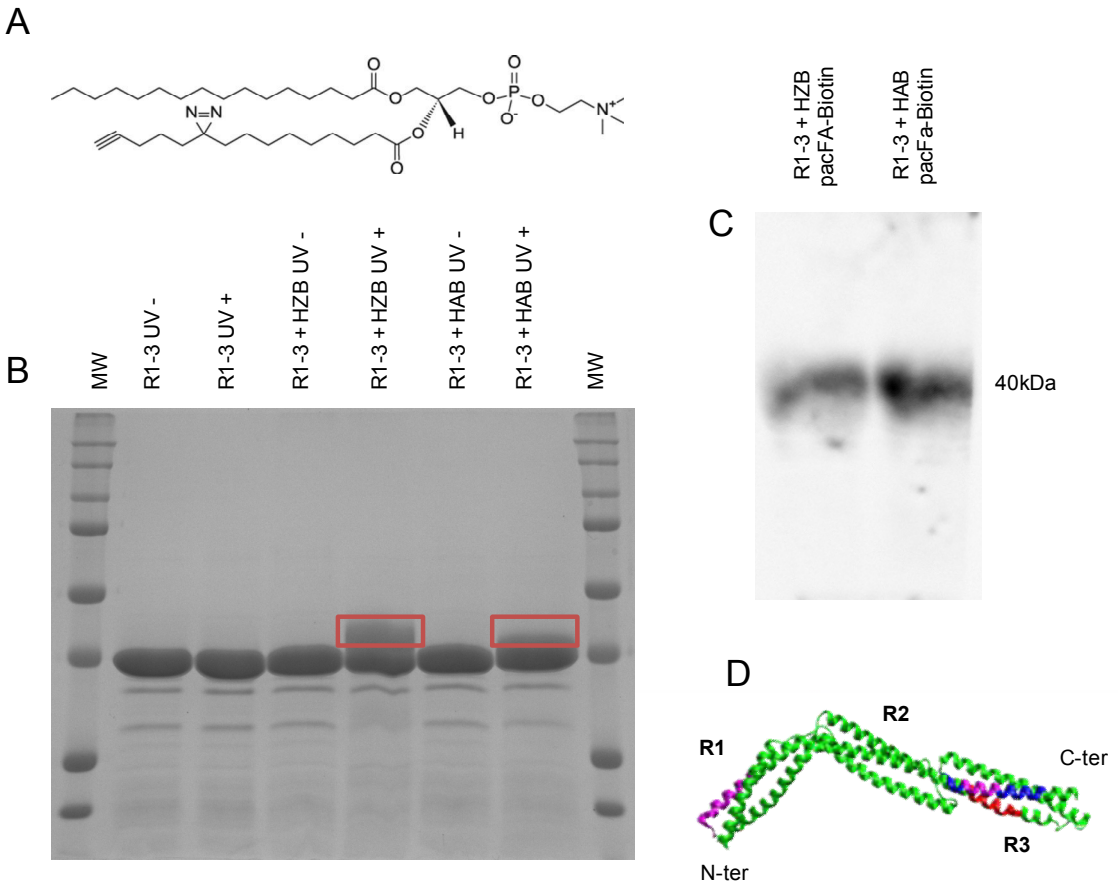


Figure S10: (*A*) Chemical structure of the bifunctional phospholipid (pacFA PC) bearing a photoactivable diazirine group and a clickable alkyne group on the tail. (*B*) SDS-PAGE analysis of R1-3 in the presence of bicelles labeled with pacFA PC, with or without UVA exposure. The gel was revealed by Coomassie Blue staining. Red frames correspond to the parts of the gel analyzed by MS (*C*) Western blot analysis of the R1-3/HZB or HAB samples after labeling with biotin-N_3 revealed with a streptavidin-HRP fusion protein. (*D*) R1-3 model (green) with the identified protein/lipid areas after binding with HZB (blue), HAB (red), or both (magenta).

Table S2: List of peptides identified by LC-MS/MS for the R1-3 dystrophin fragment. Peptides linked to the phospholipid (pacFA PC) are indicated in bold (corresponding to the annotated MS/MS spectra presented in Figure S11).

Protein	Score	Sequence coverage	Number of peptides
R1-3 fragment of dystrophin	2176.8	80.18 %	45

Peptide sequence	Modifications	Ion Score	Start	Stop	Calculated Mass	Exp. m/z	z	Error (ppm)	Missed Cleavage	Rank
LSEDEETEVQEQMNLLNSR	**pacFA PC (D4-E8); pacFA PC (D4-E8)**	**27.67**	**77**	**95**	**3690.03**	**923.52**	**4**	**3.14**	**0**	**1**
LSEDEETEVQEQMNLLNSR	**pacFA PC (E3); pacFA PC (E8)**	**22.8**	**77**	**95**	**3690.03**	**923.52**	**4**	**2.26**	**0**	**1**
LSEDEETEVQEQMNLLNSR	**pacFA PC (E3); pacFA PC(T7)**	**20.84**	**77**	**95**	**3690.03**	**923.52**	**4**	**2.6**	**0**	**1**
LTEEQCLFSAWLSEK	**pacFA PC (T2 or E3); Carbamidomethyl (C6)**	**23.69**	**229**	**243**	**2553.37**	**852.13**	**3**	**1.04**	**0**	**1**
TEAWLDNFAR	**pacFA PC (T1 or E2)**	**18.22**	**305**	**314**	**1935.08**	**646.03**	**3**	**0.88**	**0**	**1**
GSEVNLDR		53.75	1	8	888.43	445.22	2	1.2	0	1
YQTALEEVLSWLLSAEDTLQAQGEISNDVEVVK		35.49	9	41	3676.84	920.22	4	1.42	0	1
DQFHTHEGYMMDLTAHQGR		60.02	42	60	2272.97	758.66	3	-1.62	0	1
DQFHTHEGYMMDLTAHQGR	Oxidation (M10)	63.75	42	60	2288.97	764	3	0.9	0	1
DQFHTHEGYMMDLTAHQGR	Oxidation (M11)	43.4	42	60	2288.97	763.99	3	-3.81	0	1
DQFHTHEGYMMDLTAHQGR	Oxidation (M10); Oxidation (M11)	22.98	42	60	2304.96	769.33	3	-0.08	0	1
VGNILQLGSK		61.47	61	70	1027.6	514.81	2	-0.07	0	1
LSEDEETEVQEQMNLLNSR		78.96	77	95	2263.03	1132.52	2	-0.5	0	1
LSEDEETEVQEQMNLLNSR	Oxidation (M13)	55.37	77	95	2279.02	1140.52	2	2.26	0	1
LSEDEETEVQEQMNLLNSRWECLR	Oxidation (M13); Carbamidomethyl (C22)	17.79	77	100	3023.36	1008.79	3	-0.94	1	1
VLMDLQNQK		49.46	113	121	1087.57	544.79	2	-2.6	0	1
VLMDLQNQK	Oxidation (M3)	62.12	113	121	1103.56	552.79	2	2.23	0	1
LKELNDWLTK		48.47	122	131	1258.69	630.35	2	-0.71	1	1
ELNDWLTK		37.74	124	131	1017.51	509.76	2	2.04	0	1
KMEEEPLGPDLEDLK		66	138	152	1741.84	871.93	2	-0.66	1	1
KMEEEPLGPDLEDLK	Oxidation (M2)	17.94	138	152	1757.84	879.93	2	-0.33	1	1
KMEEEPLGPDLEDLKR		95.63	138	153	1897.95	633.65	3	-1.48	2	1
KMEEEPLGPDLEDLKR	Oxidation (M2)	53.18	138	153	1913.94	638.99	3	-3.08	2	1
MEEEPLGPDLEDLK		61.63	139	152	1613.75	807.88	2	1.11	0	1

MEEEPLGPDLEDLK	Oxidation (M1)	47.93	139	152	1629.74	815.88	2	1.22	0	1
MEEEPLGPDLEDLKR		51.84	139	153	1769.85	885.93	2	0.94	1	1
MEEEPLGPDLEDLKR	Oxidation (M1)	55.01	139	153	1785.85	893.93	2	-0.21	1	1
VLQEDLEQEQVR		82.91	160	171	1484.75	743.38	2	-1.41	0	1
VNSLTHMVVVVDESSGDHATAALEEQLK		66.64	172	199	2978.47	993.83	3	0.79	0	1
VNSLTHMVVVVDESSGDHATAALEEQLK	Oxidation (M7)	45.54	172	199	2994.46	749.62	4	1.06	0	1
WTEDRWVLLQDILLK		35.99	211	225	1927.06	964.54	2	0.7	1	1
WVLLQDILLK		56.59	216	225	1239.76	620.89	2	1.07	0	1
LTEEQCLFSAWLSEK	Carbamidomethyl (C6)	111.02	229	243	1839.87	920.95	2	2.26	0	1
LTEEQCLFSAWLSEKEDAVNK	Carbamidomethyl (C6)	51.65	229	249	2496.18	833.07	3	0.75	1	1
IHTTGFK		29.52	250	256	802.43	402.22	2	-0.22	0	1
IHTTGFKDQNEMLSSLQK		48.43	250	267	2076.03	1039.02	2	-1.02	1	1
IHTTGFKDQNEMLSSLQK	Oxidation (M12)	55.47	250	267	2092.03	1047.02	2	4.47	1	1
DQNEMLSSLQK		61.23	257	267	1291.61	646.81	2	0.94	0	1
DQNEMLSSLQK	Oxidation (M5)	54.15	257	267	1307.6	654.81	2	2.92	0	1
LAVLKADLEK		38.76	268	277	1098.66	550.34	2	-5.03	1	1
LYSLKQDLLSTLK		38.8	285	297	1520.88	507.97	3	-2.6	1	1
QDLLSTLK		35.41	290	297	916.52	459.27	2	0.12	0	1
TEAWLDNFAR		45.13	305	314	1221.58	611.8	2	0.58	0	1
CWDNLVQK	Carbamidomethyl (C1)	40.58	315	322	1061.5	531.75	2	-1.06	0	1
CWDNLVQKLEK	Carbamidomethyl (C1)	25.83	315	325	1431.72	478.25	3	-1.89	1	1

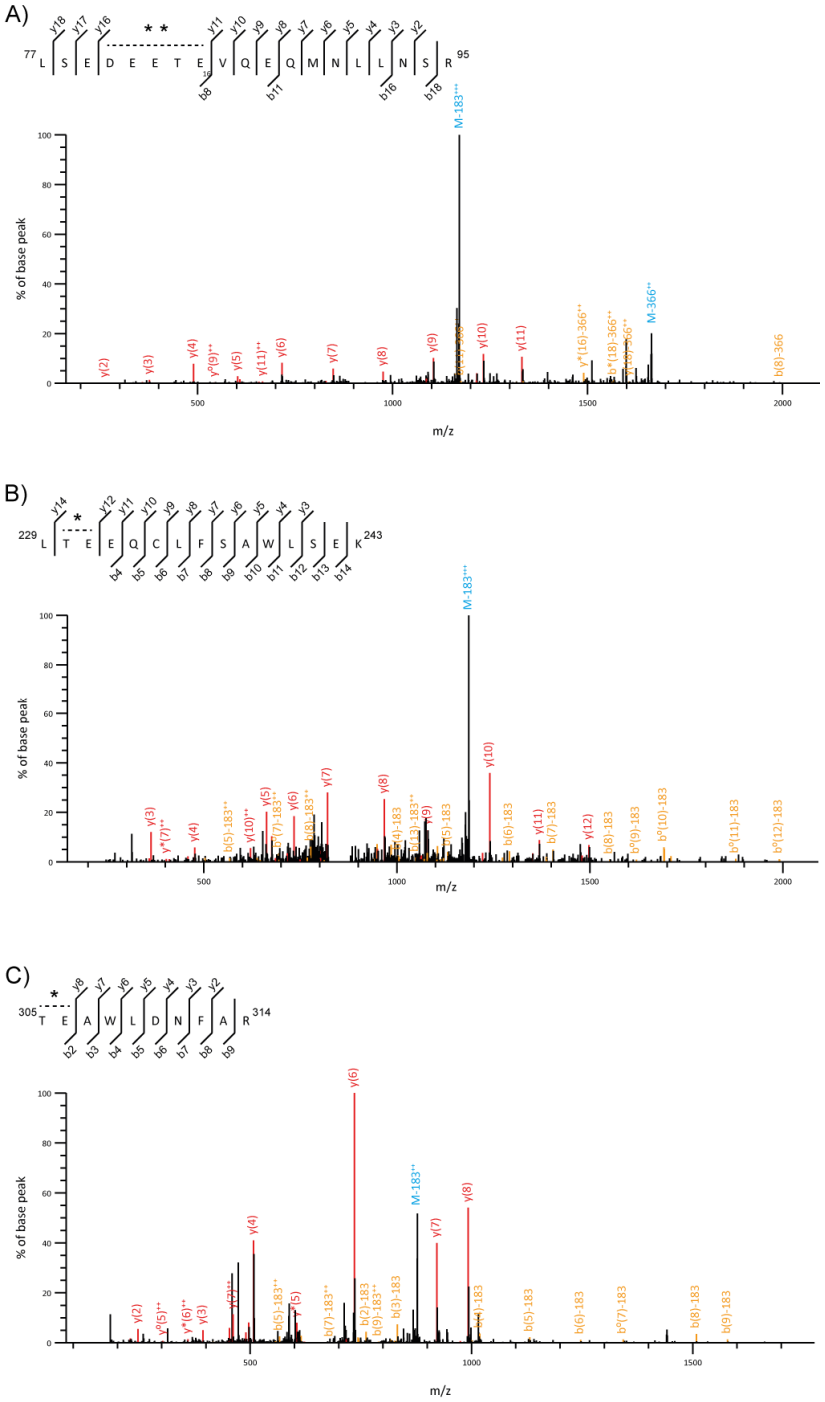


Figure S11: Annotated MS/MS spectra corresponding to the three peptides identified with a pacFA PC modification: (*A*) R1-3 [77-95], (*B*) R1-3 [229-243], and (*C*) R1-3 [305-314]. The presence of the phospholipid modification is indicated by an asterisk on the peptide sequence. Dashed lines indicate the amino acids that can bear the phospholipid modification according to MS/MS spectra.

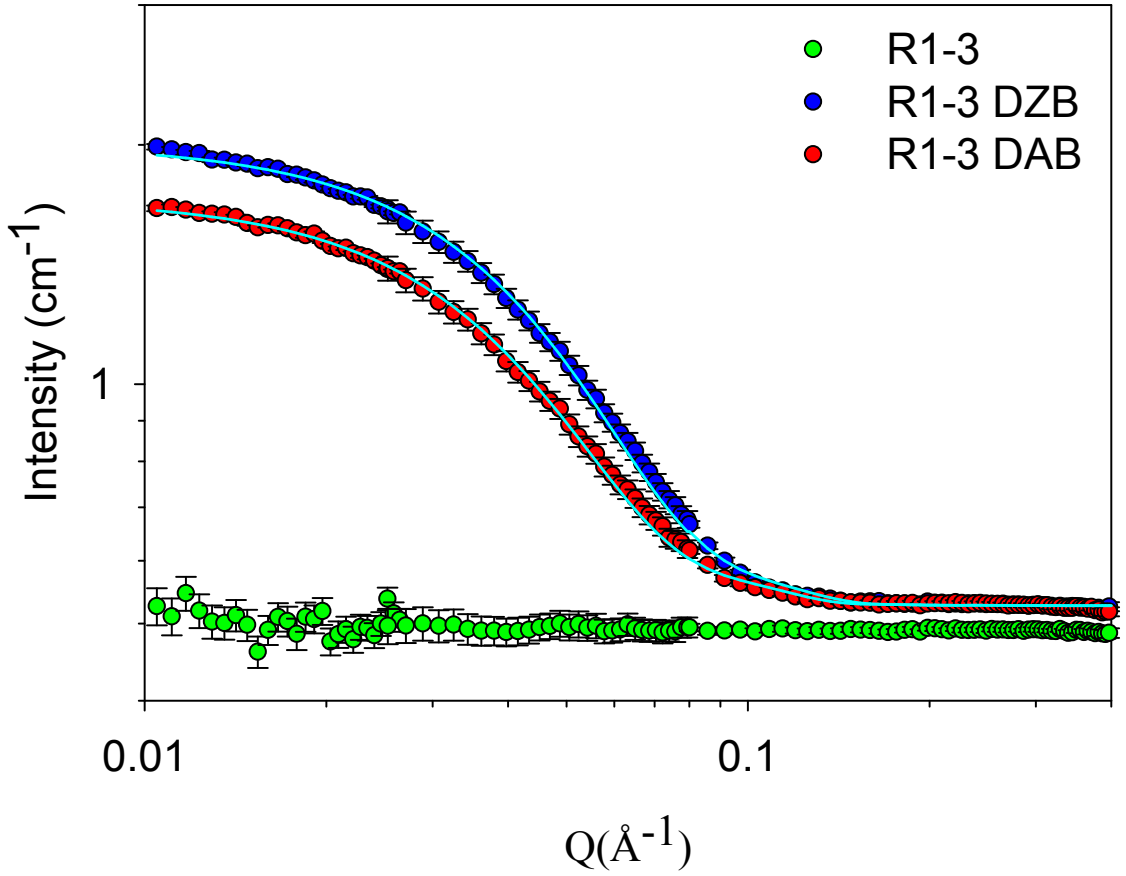


Figure S12: SANS data for DZB (blue) or DAB (red) in the presence of contrast-matched R1-3 compared to contrast-matched R1-3 alone (green) in 42% d-TNE buffer, at 18°C. SANS data were fitted (cyan) with the "core-shell cylinder" model, giving the expected dimensions of 40 Å thickness for both types of bicelles and 48 and 52 Å radius for HZB and HAB, respectively.

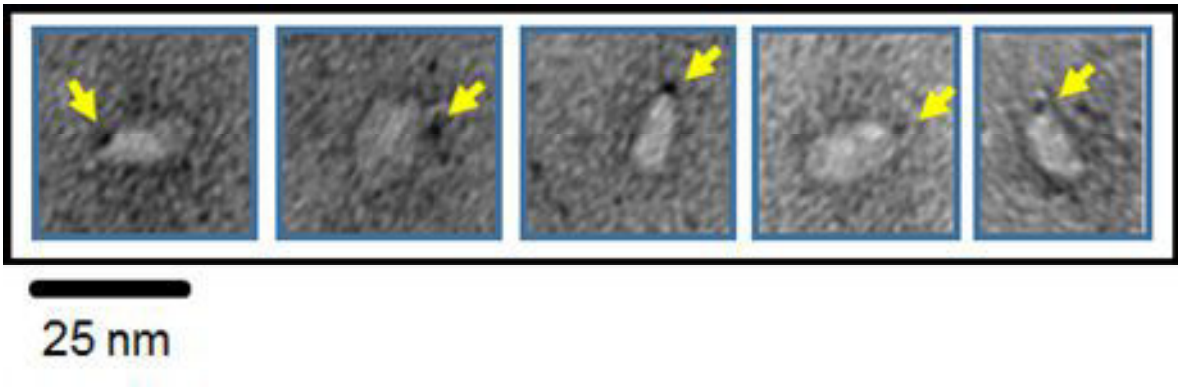


Figure S13: Representative negative stain TEM micrographs of the R1-3/HAB complex. The R1-3 protein fragment is labeled with gold nanoparticle (indicated with yellow arrows) and is found only at the torus part of HAB. Nanogold labeling was done with 1.4 nm monomaleimido Nanogold (Nanoprobes) *via* cysteines according to the manufacturer's instructions.

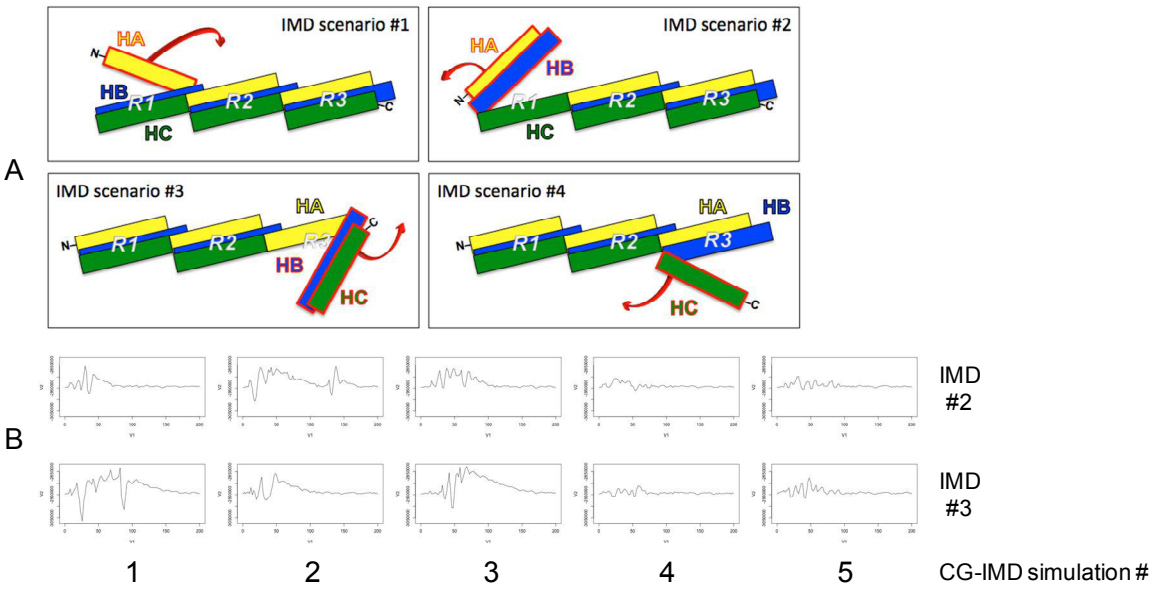


Figure S14: (*A*) Four different scenarios for CG-IMD simulations to explore the conformational changes of the R1-3 protein fragment highlighted by the SANS data. Opening of the coiled-coil fold is performed by the application of an additional force to one or two helice(s) of a repeat (red frames and arrows). (*B*) Fluctuations of the system total energy followed throughout the five independent CG-IMD simulations computed for scenarios #2 and #3. Scenarios #1 and #4 were excluded because they are not compatible with an increase of R_g. Opening of the coiled-coil filament of R1-3 dystrophin fragment occurred for a moderate energy fluctuation, quickly relaxed to the whole simulation box. Therefore, the loss of energy upon unfolding of the coiled-coil is compensated by protein/lipid non-bonding interactions

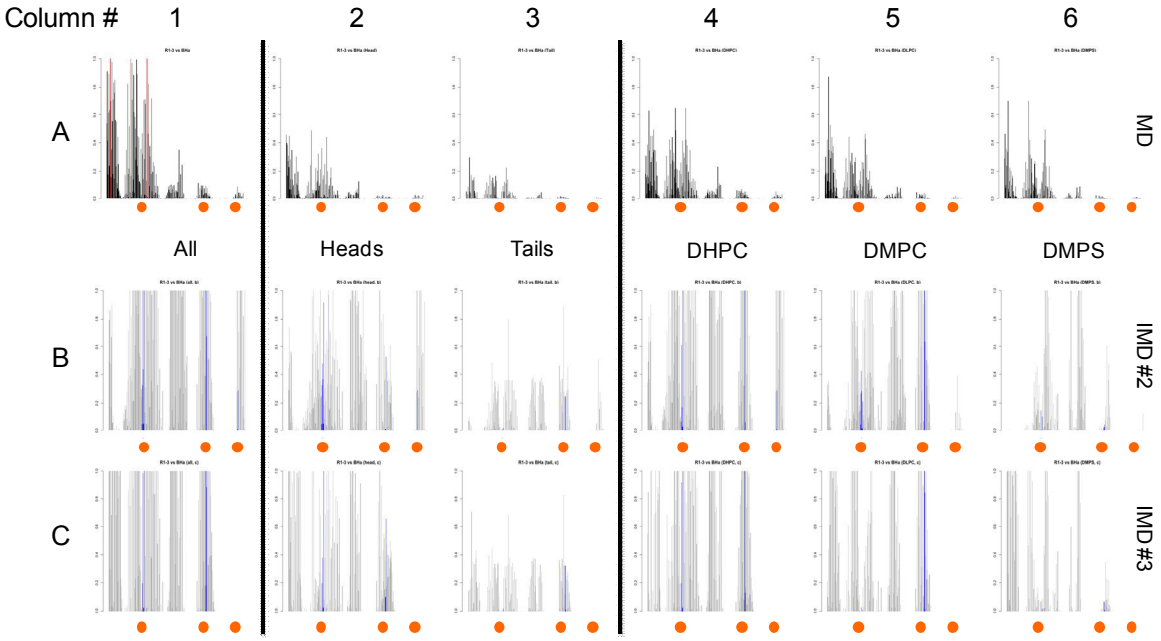
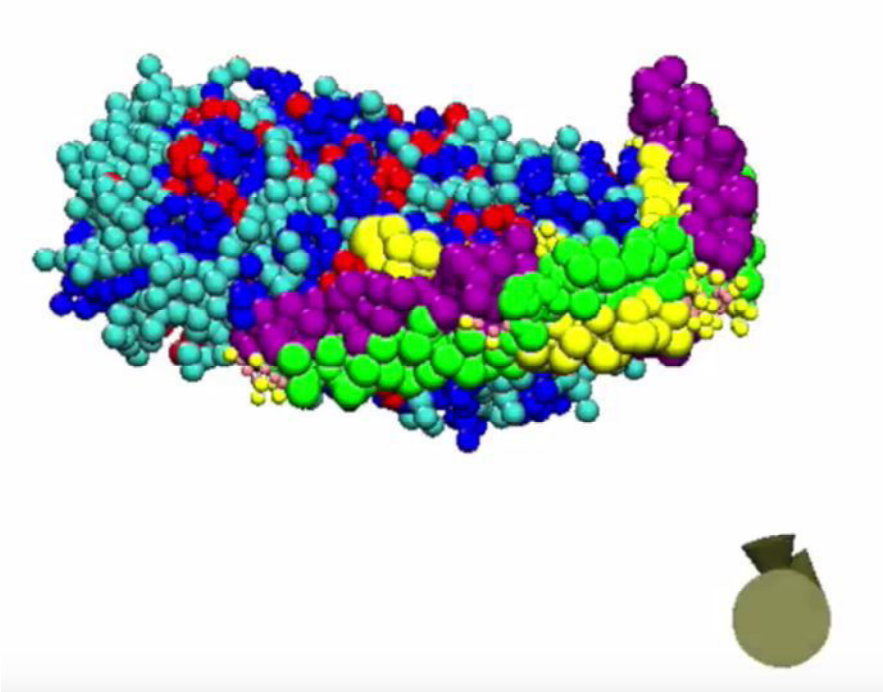


Figure S15: Frequencies of protein/lipid contacts measured during classical CG-MD before (*A*) and after (*B* and *C*) CG-IMD simulations. Protein primary sequence of R1-3 is represented on the *x*-axis. First column concerns the contact frequencies of R1-3 with phospholipids. The second and third columns show the contact frequencies with, respectively, heads and tails of HAB. The three last columns show protein/lipids contact frequencies given each type of phospholipid (DHPC, DMPC, or DMPS). The contact maps obtained from the simulations present the best accordance with mass spectrometry experimental mapping (orange dots) in the case of the CG-IMD #2 scenario, after opening of repeat R1 through the unfolding of its two first helices (1HA and 1HB). The corresponding models also provide the best fits to SANS experimental profiles (see the CRYSON simulated scattering curve in Figure 4D).

Video S1: A movie showing the CG-IMD process.
(https://www.youtube.com/watch?v=F2yTnmVJwIs&feature=em-share_video_user).



REFERENCES

(1) Winder, S. J.; Gibson, T. J.; Kendrick-Jones, J. *FEBS Lett.* **1995**, *369* (1), 27.
(2) Vié, V.; Legardinier, S.; Chieze, L.; Le Bihan, O.; Qin, Y.; Sarkis, J.; Hubert, J.-F.; Renault, A.; Desbat, B.; Le Rumeur, E. *Biochim. Biophys. Acta* **2010**, *1798* (8), 1503.
(3) Sarkis, J.; Hubert, J.-F.; Legrand, B.; Robert, E.; Chéron, A.; Jardin, J.; Hitti, E.; Le Rumeur, E.; Vié, V. *J. Biol. Chem.* **2011**, *286* (35), 30481.
(4) Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R. *Biophys. J.* **2001**, *81* (4), 2163.
(5) Beaugrand, M.; Arnold, A. A.; Hénin, J.; Warschawski, D. E.; Williamson, P. T. F.; Marcotte, I. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (21), 6162.
(6) Triba, M. N.; Warschawski, D. E.; Devaux, P. F. *Biophys. J.* **2005**, *88* (3), 1887.
(7) Mineev, K. S.; Nadezhdin, K. D.; Goncharuk, S. A.; Arseniev, A. S. *Langmuir ACS J. Surf. Colloids* **2016**.
(8) Petrache, H. I.; Tristram-Nagle, S.; Nagle, J. F. *Chem. Phys. Lipids* **1998**, *95* (1), 83.
(9) Petrache, H. I.; Tristram-Nagle, S.; Gawrisch, K.; Harries, D.; Parsegian, V. A.; Nagle, J. F. *Biophys. J.* **2004**, *86* (3), 1574.
(10) Small, D. M. *The Physical Chemistry of Lipids: From Alkanes to Phospholipids*; Plenum Press, 1986.
(11) Dos Santos Morais, R.; Delalande, O.; Pérez, J.; Mouret, L.; Bondon, A.; Martel, A.; Appavou, M.-S.; Le Rumeur, E.; Hubert, J.-F.; Combet, S. *Langmuir* **2017**, *33* (26), 6572.
(12) Refractive index increment dndc for proteins, polymers SLS http://www.materials-talks.com/blog/2013/06/18/refractive-index-increment-dndc-values/
(13) Dos Santos Morais, R.; Delalande, O.; Pérez, J.; Mouret, L.; Bondon, A.; Martel, A.; Appavou, M.-S.; Le Rumeur, E.; Hubert, J.-F.; Combet, S. *Langmuir* **2017**.
(14) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435.

3. Résultats complémentaires

3.1. Stœchiométrie du complexe R1-3/bicelle anionique

Un point important dans la caractérisation du complexe R1-3/bicelle anionique consiste à déterminer la stœchiométrie du système. A cette fin, nous avons obtenu des données SEC-MALS avec un système de triple détecteurs (UV, LS et RI) [238]. Au sommet des pics, la MW des bicelles anioniques est de 200 kDa et celle de R1-3 est de 39 kDa (Figure 59A). Le premier pic à gauche correspond à une entité de R_h plus élevé que celui de chacun des partenaires seuls en solution. Nous en concluons qu'il s'agit du complexe R1-3/HAB. Une MW constante est observée le long du pic d'élution, ce qui souligne l'homogénéité de taille du complexe, ce qui est un prérequis obligatoire pour les analyses en SAS. Pour rappel, la MW dépend de l'incrément d'indice de réfraction (*dn/dc*)². Les deux espèces qui composent ce complexe protéine/lipide possèdent des *dn/dc* différents. Grâce aux deux détecteurs de concentrations (UV et RI), connaissant les *dn/dc* des deux objets et leur coefficient d'extinction massique (ε/MW), il est possible de déterminer la contribution de chaque partenaire au signal de diffusion de lumière. Le logiciel Astra (Wyatt) permet de déconvoluer les signaux et ainsi d'obtenir la stœchiométrie du système [238]. Au sommet du pic (Figure 59B), les MW des parties phospholipidiques (« modifier ») et protéique (« protein ») calculées sont de 155 et 30 kDa, respectivement, ce qui donne un complexe de 185 kDa (« total »). Les MW des deux composés sont inférieures d'un facteur ~ 1,3 à celles déduites des deux partenaires seuls en solution. Dans le cas du complexe, il est probable que le restant de R1-3 non co-éluée vienne affecter le signal (Figure 59A). Une stœchiométrie 1:1 au sein du complexe semble donc la plus probable. Cette conclusion est en accord avec celles des expériences de MST et de microscopie électronique. Lors des expériences SEC, l'échantillon est *de facto* dilué. Si les concentrations effectives sont proches de la valeur de K_d, la proportion de R1-3 libre peut augmenter. Ceci expliquerait pourquoi toute la protéine n'est pas co-éluée avec les bicelles (Figure 59A). Les mêmes expériences ont été réalisées pour R11-15. Cependant, à cause d'une saturation du détecteur UV, due à une plus forte concentration en protéine, il n'a pas été possible de déconvoluer le signal et déterminer la stœchiométrie était impossible. Néanmoins, en se basant sur l'ensemble des résultats (SEC et MST), nous pouvons émettre raisonnablement l'hypothèse que tous les autres complexes étudiés ont également une stœchiométrie 1:1.

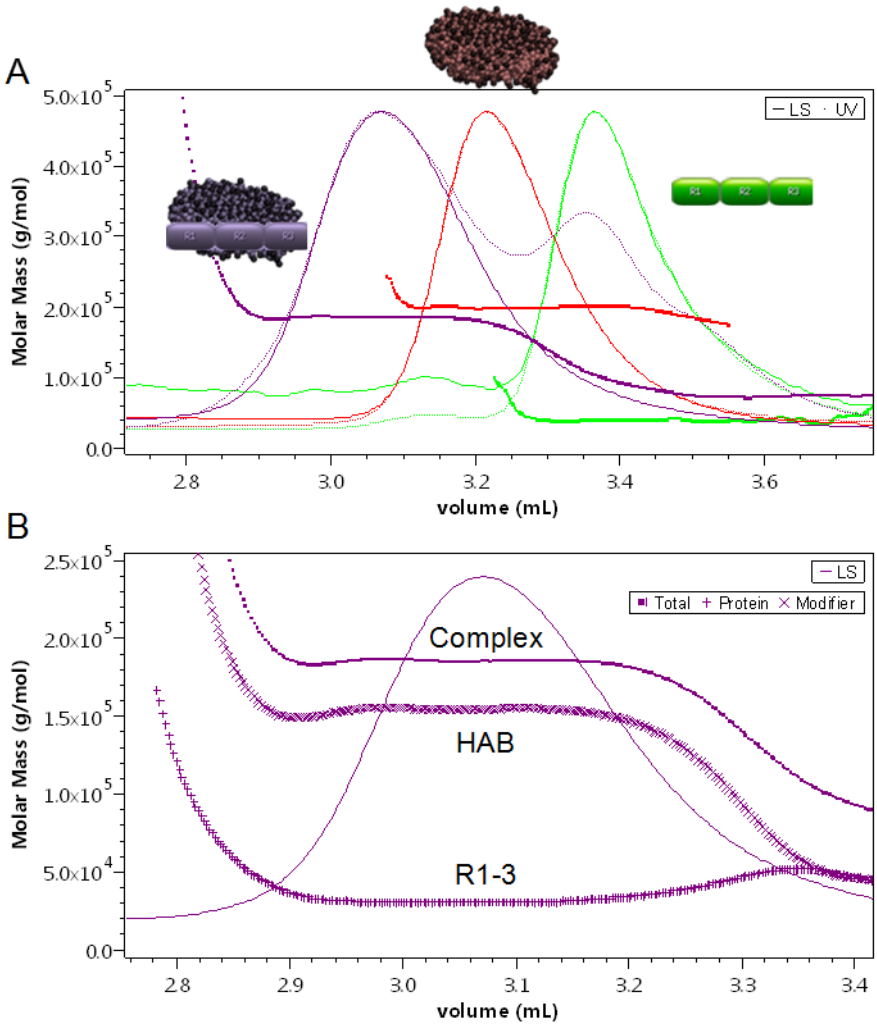


Figure 59 : Analyse SEC-UV/MALS/RI du complexe R1-3/HAB. (A) Analyse SEC-MALS de R1-3 seule (vert), de HAB seule (rouge) et du complexe R1-3/HAB (violet), montrant une co-élution des deux partenaires, ce qui atteste de leur interaction. Détecteurs de diffusion de lumière (LS, trait plein) et UV (pointillé). (B) Déconvolution du signal LS du complexe R1-3/HAB présenté en (A) permettant de déterminer la stœchiométrie du système. « Protein » : R1-3, « Modifier » : HAB.

3.2.R1-3/bicelles anionique : approche par cryo-EM

En complément des résultats de TEM présentés dans l'article 2, nous avons caractérisé les complexes R1-3/bicelle anioniques par cryo-EM (Figure 60). Les bicelles sont présentées vues de dessus. La faible résolution et l'absence de traitement d'images ne permettent pas de visualiser directement la protéine. Cependant, les billes d'or associées à R1-3 sont localisées sur la périphérie des bicelles. Comme nous en avons conclu dans l'article 2 par TEM (Figure S13), R1-3 est localisée sur le côté de la bicelle. De plus une seule bille d'or est retrouvée sur chaque bicelle. Ceci conforte les résultats de stœchiométrie TEM, MST et MALS.

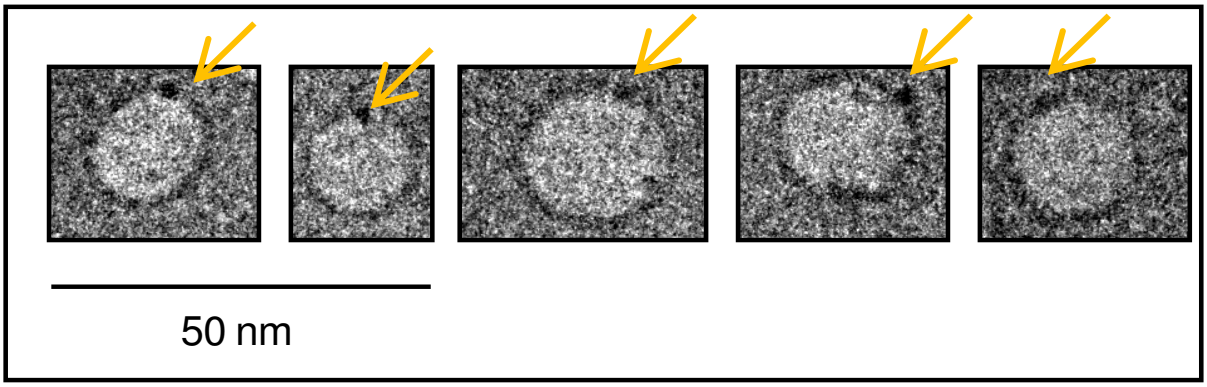


Figure 60 : Micrographes de cryo-EM représentatifs des complexes R1-3/HAB. R1-3 a été marquée avec une bille d'or (flèches orange). Le marquage avec les billes d'or est réalisé *via* les résidus lysines (1.4 nm Mono-Sulfo-NHS-Nanogold®, Nanoprobes).

3.3.R11-15/bicelle : analyse par CG-MD

Dans l'article précédent, nous avons uniquement présenté les résultats CG-MD obtenus pour R1-3. Or, nous avons également réalisé des expériences *in silico* sur les complexes R11-15/bicelles entier dans les mêmes conditions de simulation. Tout comme pour R1-3, R11-15 est spécifiquement recrutée sur la partie toroïdale de la bicelles, qu'elles soient zwitterioniques ou anioniques. Avec la même démarche que celle utilisée pour R1-3, nous avons déduit de nos analyses de chimie-click/spectrométrie de masse que cette interaction s'opère sur le côté de la bicelle. En effet, comme pour R1-3, l'analyse SDS-PAGE après pontage photo-induit révèle une partie présentant un « smear » pour R11-15 en présence des deux types de bicelles, suggérant une liaison covalente avec le pacFA PC. A la suite de l'analyse MSMS, nous avons pu identifier un peptide de R11-15 portant le phospholipide modifié (^{209}AEEWLNLLLEYQK^{221}) uniquement pour le complexe R11-15/HAB. Cette observation est compatible avec le modèle R11-15/bicelle obtenu par CG-MD (Figure 61). Etonnamment, aucun peptide portant le pacFA PC n'a pu être identifié pour le complexe R11-15/HZB. Pourtant, la présence d'un « smear » suggère que le phospholipide en interaction avec R11-15 est bien modifié. Une explication simple et plausible est que les peptides portant le phospholipide ne s'ioniseraient pas correctement lors de leur exposition à la source d'ionisation du spectromètre de masse. Nous n'avons pas effectué d'expérience de CG-MD interactive sur ces complexes pour le moment car le nombre de particules est plus élevé que pour le système R1-3/bicelles (donc le temps de calcul est significativement plus long). Ces analyses sont cependant envisagées ultérieurement au laboratoire.

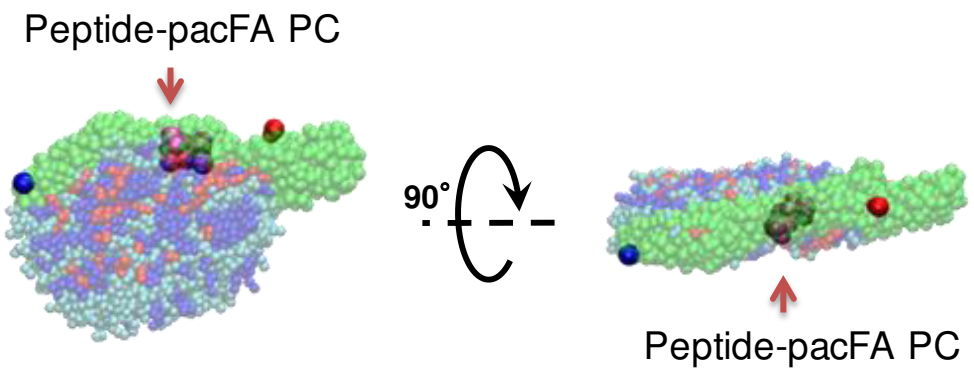


Figure 61 : Modèle gros-grain du complexe R11-15/HAB après 20 µs de simulation. Les lipides sont représentés en bleu (DMPC), cyan (DHPC) et rouge (DMPS). R11-15 est représentée en vert et les extrémités N-ter et C-ter correspondent aux boules bleu et rouge, respectivement. Le peptide portant le pacFA PC identifié en MS est coloré en rose.

3.4.Digestion ménagée

Un moyen simple et rapide de savoir si la liaison aux bicelles conduit à un changement de conformation consiste à digérer les échantillons avec une enzyme, typiquement la trypsine, et de comparer ensuite les profils de digestion par SDS-PAGE. Nous avons réalisé ce type d'expériences sur nos systèmes protéine/bicelle. La première observation est que les protéines seules possèdent une conformation les rendant extrêmement résistantes à une digestion trypsique (Figure 62). Les profils de digestion de R11-15 et de R11-15/HZB sont très similaires. Ceci est en accord avec les données SANS obtenues pour ce complexe, à savoir que la protéine ne subit pas de changement de conformation majeur lors de sa liaison aux bicelles zwitterioniques. Dans le cas du complexe R11-15/HAB, le profil de digestion de la protéine liée est différent de celui de la protéine seule en solution et permet de conclure à un changement de conformation. Pour les complexes R1-3/HZB ou HAB, les profils de digestion sont différents de celui de la protéine seule. Ces résultats impliquent un changement conformationnel de R1-3 lors de sa liaison aux bicelles. Ce changement structural semble plus prononcé dans le cas de l'interaction avec des bicelles anioniques puisque la digestion trypsique est plus rapide. Les résultats obtenus par cette technique simple et rapide sont en accord avec ceux obtenus avec les techniques plus complexes décrites dans l'article 2.

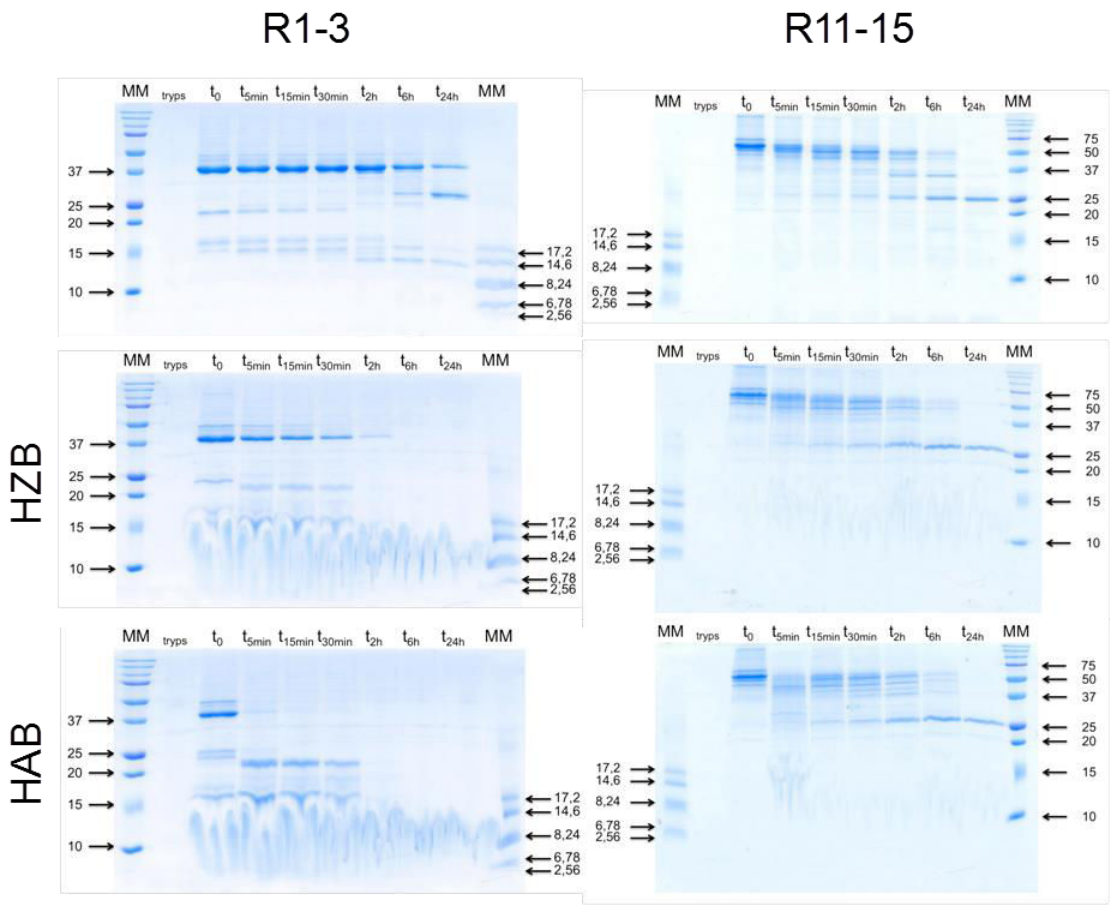


Figure 62 : Analyse SDS-PAGE des profils de digestion ménagée à la trypsine de R1-3 et R11-15 en absence et présence de bicelles zwitterioniques (HZB) ou anioniques (HAB) en fonction du temps.

3.5. Complexe protéine/bicelle étudié par SAXS

En complément des résultats présentés ci-dessus, nous avions comme objectif d'obtenir des données SAXS sur les complexes protéine/bicelle. Nous avons tout d'abord vérifié que la structure de R1-3 et R11-15 n'est pas affectée par la présence de 6 mM de DHPC dans le tampon TNE. Les résultats sont très similaires à ceux précédemment obtenus par l'équipe (partie IV, chapitre 1, article 3), ce qui est le signe que la conformation des fragments protéiques n'est pas affectée par la présence de DHPC libre à la concentration de 6 mM (Figure 63).

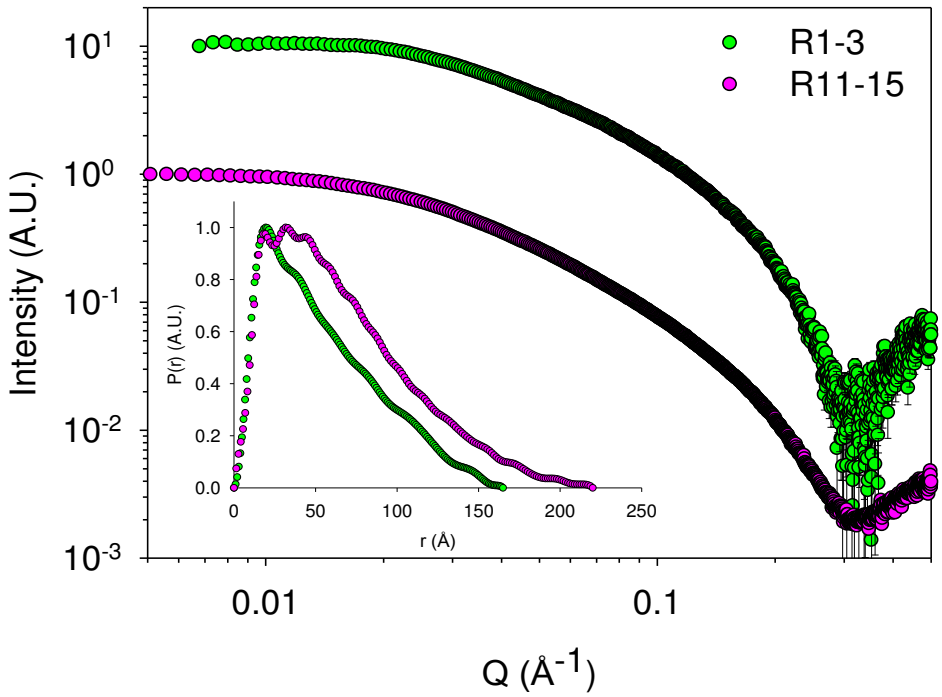


Figure 63 : Courbes SEC-SAXS et équivalents en P(r) obtenus pour R1-3 (vert) et R11-15 (magenta) dans du tampon TNE supplémenté avec 6 mM de DHPC. Ces données sont similaires à celles obtenues sans DHPC.

Notre objectif était d'obtenir des données sur le complexe entier et ensuite de coupler ces données à celles de modélisation moléculaire afin de proposer des modèles tout-atome des complexes protéine/lipide. De nombreux obstacles expérimentaux ont été rencontrés lors de l'utilisation des bicelles en SAXS (partie III, chapitre 1). Néanmoins, certains jeux de données obtenus récemment sur des complexes protéine/bicelle sont prometteurs. En analysant les données SEC-SAXS, la conclusion majeure est que le signal des bicelles seules est significativement différent de celui des complexes protéine/bicelle (Figure 64). Le signal de la protéine n'est donc pas écrasé par celui des bicelles. Il est important de relever que la remontée à petits angles, non attendue et encore inexpliquée pour les objets bicelles (partie III, chapitre 1), est atténuée dans le cas des complexes protéine/bicelle (Figure 64).

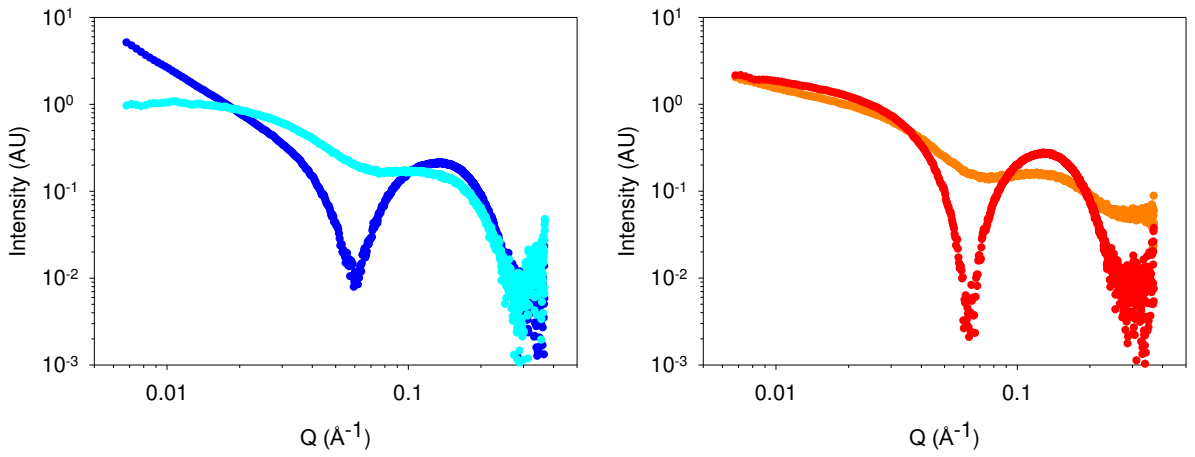


Figure 64 : Intensités SEC-SAXS mesurées pour les HZB (bleu) et les HAB (rouge) seules ou en présence de R1-3 (cyan et orange, respectivement).

Au début de la thèse, nous ne disposions pas des données MST pour les complexes protéine/bicelle. Grâce à ces données, nous savons désormais quelles sont les concentrations des deux composants permettant d'obtenir uniquement des complexes protéines/bicelles en absence quasiment complète des partenaires sous forme libre. Les expériences de SAXS envisagées pour la suite des travaux pourront donc être entreprises sans avoir recours à un système HPLC.

3.6. Thermostabilité de la dystrophine en présence de phospholipides

En vue de poursuivre la caractérisation des interactions protéine/lipide, nous nous sommes intéressés à la stabilité thermique des fragments protéiques lors de leur interaction avec les bicelles. Une variation de la température de demi-dénaturation thermique (*Tm* dans tout ce paragraphe) serait le signe d'un changement de stabilité. Nous avons utilisé tout d'abord le CD pour mettre en évidence un éventuel changement de thermostabilité. L'ellipticité à 222 nm a été suivie en fonction de la température, sur une gamme en température de 20 à 80°C et avec une vitesse d'incrémentation de 1°C/min. Les spectres CD initiaux ont été obtenus avant chauffage. Les *Tm* calculées pour R1-3 et R11-15 en l'absence de bicelles sont 56 et 58°C, respectivement. Ces valeurs ont été obtenues après ajustement des données à un modèle à deux états (natif N, dénaturé D), avec un état D non réversible [239]. Dans le cas des complexes R11-15/bicelles, les courbes de dénaturation thermique de la protéine sont significativement différentes de celle de la protéine seule, que ce soit en présence de HZB ou HAB. Les données ont été ajustées avec un modèle à trois états [239] permettant l'obtention de deux *Tm* pour R11-15 en présence de HZB et de deux autres *Tm* en présence de HAB. Comme dans le cas de la

protéine seule, un précipité est apparu à 80°C dans la cuve de mesure, ce qui montre que la dénaturation thermique n'est pas réversible.

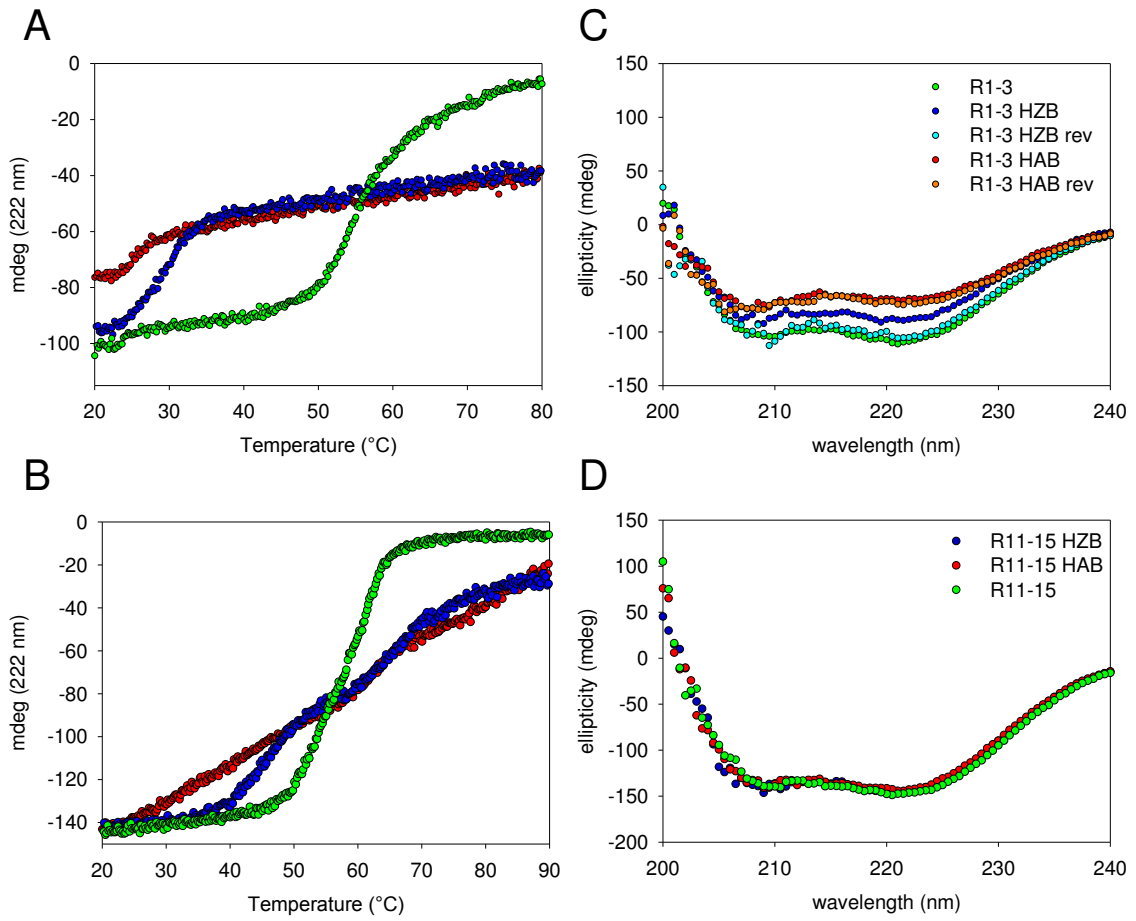


Figure 65 : Evaluation de la thermostabilité de R1-3 (A) et R11-15 (B) mesurée en CD. Les données ont été obtenues pour les fragments protéiques seuls (vert) ou en présence de HZB (bleu) ou HAB (rouge). L'ellipticité a été mesurée à 222 nm sur une gamme de température de 20 à 80°C. Spectres CD dans l'UV proche avec chauffage de R1-3 (C) et R11-15 (D) avec le même code couleur. Pour R1-3 en présence de HZB (vert clair) ou HAB (orange), les spectres CD ont également été mesurés après retour à la température initiale (« rev »).

En ce qui concerne R1-3 en présence de bicelles, les courbes de dénaturation sont très différentes de celle de R1-3 seule. Pour les deux types de complexes, les données suggèrent que *Tm* est fortement diminuée et donc que la protéine est beaucoup moins stable que lorsqu'elle est seule en solution. De manière très surprenante, aucun précipité n'est visible en fin de mesure. De plus, les spectres CD, après retour à la température initiale, sont très similaires à ceux obtenus avant chauffage. Ces données montrent que, impliquée dans un complexe, R1-3 est résistante à la précipitation thermo-induite jusqu'à une température de 80°C. Pour vérifier que ce comportement est spécifique aux conséquences des interactions R1-3/bicelles, nous avons

réalisé les mêmes expériences, tout d'abord en complétant le tampon TNE avec 6 mM de DHPC, puis en présence de SUV de DOPC/DOPS (2:1), ce dernier modèle membranaire étant l'un des plus utilisés par l'équipe pour la caractérisation des interactions protéine/lipide. Dans les deux cas, le *Tm* de R1-3 n'est que très faiblement affecté et, surtout, la dénaturation thermique n'est pas réversible. Ceci confirme que ce sont uniquement les interactions entre R1-3 et les bicelles qui confèrent une thermo-résistance à la protéine. Une question se pose néanmoins à propos de ces interactions. Il est évident que la morphologie des bicelles a été affectée par la température [39,42]. Cependant, le spectre CD des bicelles n'est pas affecté (résultat non montré) et, bien que changeant de morphologie, les bicelles restent un assemblage de phospholipides. Des expériences complémentaires doivent être réalisées afin de mieux comprendre ce phénomène.

Pour compléter les résultats obtenus par CD, nous avons utilisé la technologie proposée par la société NanoTemper : le dispositif Prometheus. Cette technique est basée sur la fluorescence des résidus tryptophanes. Le ratio de fluorescence 330 nm/350 nm (apolaire/polaire) est suivi en fonction de la température et permet de déterminer *Tm*. Nous avons donc analysé la thermostabilité des fragments protéiques en présence de bicelles en suivant la fluorescence. Pour R1-3 et R11-15 seules, cette technologie a permis de mettre en évidence un *Tm* de 51,2°C pour R1-3 et non plus un mais deux *Tm* pour R11-15 de 48,3 et 58,6°C . Pour R11-15, le premier *Tm* peut être attribué à la perte de la structure en « coiled-coil », alors que le deuxième, proche de celui déterminé par CD, pourrait correspondre à une dénaturation globale des hélices α de la protéine. Pour les données en présence de bicelles, nous observons de nouveau des variations de *Tm*. Pour R11-15/HZB, les *Tm* sont de 39,7 et 61,9 °C et de 27,8 et 74,5 °C pour R11-15/HAB. Pour R1-3/HZB, le *Tm* est de 23,5°C (résultats non montrés). L'expérience n'a pas pu être réalisée pour le complexe R1-3/HAB à cause d'une trop forte fluorescence. En effet, l'expérience se déroule en capillaires (48 positions) et la fluorescence des autres échantillons mesurés en même temps a dû être ajustée afin de détecter le signal, conduisant à la saturation du détecteur pour l'échantillon R1-3/HAB.

Tout comme pour les données de CD, des valeurs de *Tm* en-dessous de 30°C ne sont pas forcément le signe d'une dénaturation de la protéine en présence de bicelles, mais plutôt d'un changement de conformation de la protéine lors de l'interaction avec les bicelles en fonction de la température.

4. Conclusion

Dans ce chapitre, nous avons caractérisé les interactions protéine/lipide de deux fragments protéiques du domaine central de la dystrophine, R1-3 et R11-15, qui peuvent être considérés comme deux sous-domaines fonctionnels [173]. De plus, R1-3 est retrouvée dans de nombreuses mini-dystrophines et R11-15 est le seul domaine de la dystrophine capable de lier à la fois les lipides membranaires et l'actine-F du cytosquelette [183]. Ces points rendent l'étude de ces deux fragments extrêmement importante et pertinente. Nous avons mis en évidence que R1-3 et R11-15 peuvent interagir avec des bicelles zwitterioniques et encore plus fortement avec des bicelles anioniques, notamment en utilisant la MST.

Nous rappelons qu'obtenir des informations structurales sur le domaine central de la dystrophine, que ce soit libre en solution ou en présence de lipides, n'est pas accessible par RMN et/ou XRC. En « éteignant » les bicelles en SANS, nous avons démontré que R1-3 et R11-15, ce dernier dans une moindre mesure, subissent un changement de conformation lors de leur interaction avec des bicelles contenant des phospholipides anioniques.

Nous avons utilisé la chimie-click et la spectrométrie de masse couplées à une approche *in-silico* de CG-MD pour déterminer une cartographie des interactions R1-3 / bicelles. En utilisant la méthode innovante de CG-MD interactive, nous avons alors pu proposer un modèle fiable de R1-3 en interaction avec des bicelles anioniques, modèle en adéquation avec les données SANS. Ce modèle montre que la répétition R1 du domaine central de la dystrophine pourrait subir une ouverture de sa structure en « coiled-coil ». Cette ouverture pourrait être nécessaire à l'ancrage de la dystrophine dans des parties incurvées du sarcolemme. Un tel ancrage peut en effet être impliqué dans le rôle de charpente assuré par la dystrophine au sein des cellules musculaires pendant le processus d'élongation/contraction.

Les dystrophies de Duchenne et Becker restent actuellement des maladies incurables. Une des stratégies dédiées au traitement de ces dystrophinopathies est la production de mini-dystrophines minimales fonctionnelles. Actuellement, les vecteurs AAV ne peuvent contenir du matériel génétique qu'à hauteur de 5 kb, permettant potentiellement la production d'une protéine de ~ 180 kDa. Il est donc nécessaire de déterminer quelles sont les parties de la dystrophine indispensable pour assurer sa fonction. Ainsi, nous espérons que le présent travail pourra contribuer ä la conception de mini-dystrophines en tant que thérapie génique pour le traitement des patients atteints de DMD ou de BMD.

Enfin, nous avons démontré que les bicelles sont un outil polyvalent utilisable dans des expériences de MST, SEC-MALS et CG-IMD, en plus de leur utilisation en SANS en s'appuyant sur la technique de variation de contraste. Nous pouvons affirmer avec un haut degré de confiance que notre approche peut être utilisée pour la caractérisation des interactions protéine/lipide, ainsi que pour l'analyse d'autres protéines membranaires périphériques et intégrales dont les études structurales sont incompatibles avec d'autres méthodes.

Chapitre 3 : Degré d'enfouissement de la dystrophine dans le sarcolemme

1. Introduction

Dans les deux chapitres précédents, nous avons caractérisé les interactions en solution du complexe dystrophine/membrane avec le modèle membranaire « bicelle ». L'objectif principal était de déterminer quelle est la conformation des deux fragments de dystrophine, R1-3 et R11-15, lors de leur interaction avec la membrane. Dans ce troisième chapitre, nous nous sommes intéressés au comportement de ces fragments de dystrophine aux interfaces, c'est-à-dire ici en interaction avec des monocouches de lipides à l'interface air/eau ou au contact de membranes lipidiques supportées. Ce thème avait déjà été abordé par l'équipe ces dernières années, notamment en collaboration avec Véronique Vié et Gilles Pabœuf de l'Institut de Physique de Rennes (IPR), en utilisant la microscopie à force atomique et des mesures en cuve de Langmuir [195,240].

L'objectif était de déterminer si les fragments de dystrophine s'insèrent dans la membrane et, le cas échéant, quel est leur degré d'insertion. Nous nous sommes appuyés ici sur la technique de réflectivité de neutrons en utilisant, soit des monocouches de phospholipides en cuve de Langmuir comme modèle membranaire à l'interface air/eau, soit des bicouches de phospholipides supportées sur bloc de silicium. Nous présentons les résultats préliminaires obtenus qui s'avèrent encourageants.

2. Mesures NR en monocouches de Langmuir

Dans un premier temps, nous nous sommes penchés sur le modèle de monocouche de Langmuir en utilisant de la DMPC, afin de conserver les mêmes lipides que pour le modèle de bicelles. Nous avons mis à profit la technique de variation de contraste inhérente aux expériences de neutrons afin de caractériser les liaisons protéine/lipide à l'interface air/eau. Trois contrastes différents ont été utilisés lors de ces expériences de réflectivité de neutrons (1. d-lipides/D_2O ; 2. h-lipides/D_2O ; 3. d-lipides/ACMW, d et h signifiant « deutérés » et « hydrogénés », respectivement et ACMW signifiant « air contrast matched water »). Toutes les mesures ont été réalisées à 20°C dans du tampon TNE ou dans un mélange TNE/d-TNE

selon le contraste souhaité. Pour chaque contraste, nous avons suivi la cinétique d'adsorption à la monocouche des fragments protéiques grâce à la variation de tension de surface ($\Delta\pi$). Dans chaque cas, la tension de surface initiale (π_i) était de 20 mN/m. A la fin de la cinétique d'adsorption des fragments protéiques, $\Delta\pi$ était ~5 mN/m pour R1-3 et ~2.5 mN/m pour R11-15, quel que soit le contraste considéré. Tous les résultats de NR présentés correspondent à l'état du système en fin de cinétique, lorsque la tension de surface est stable.

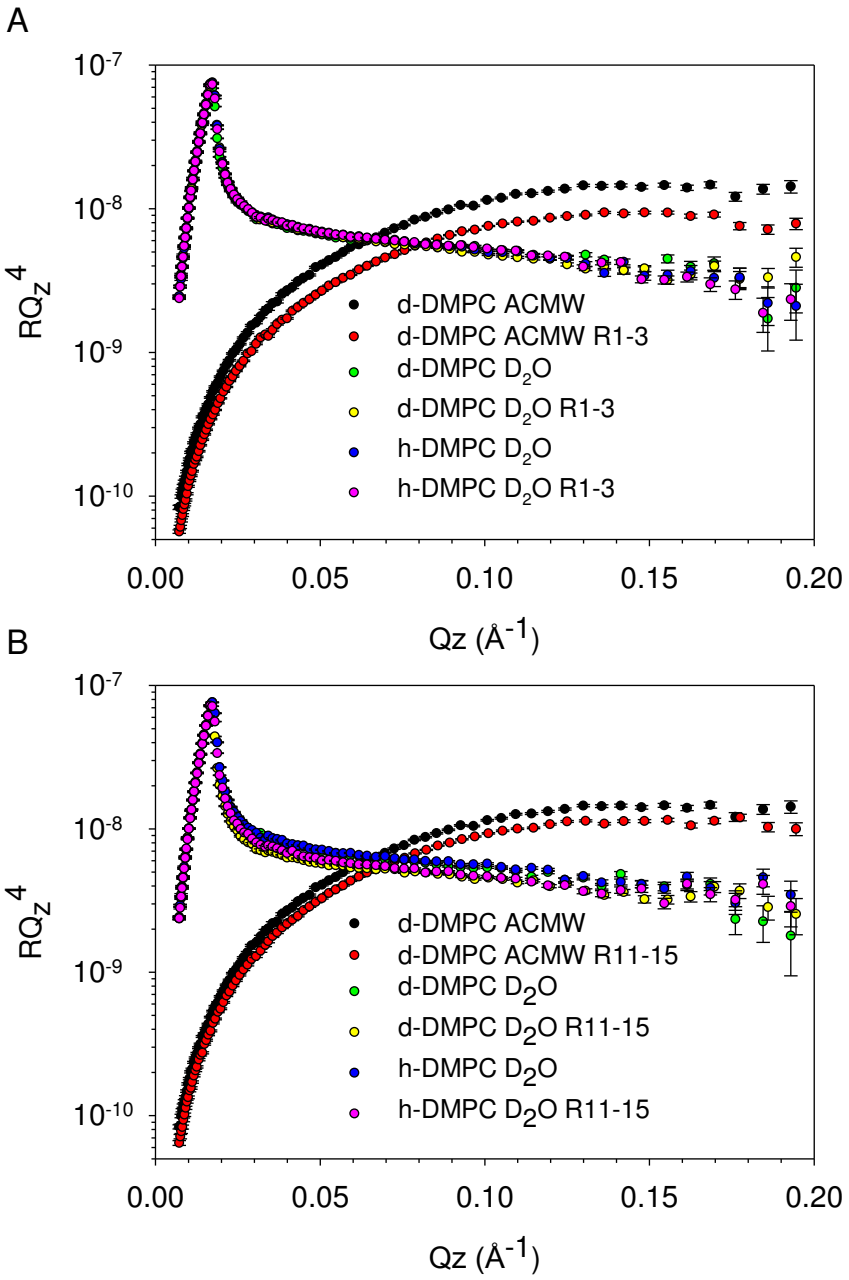


Figure 66 : Représentation en RQ_z^4 de courbes NR mesurées à l'interface air/eau d'une monocouche de h- ou d-DMPC seule ou en présence de (A) R1-3 ou (B) R11-15, aux contrastes 100% D_2O ou ACMW (se référer au code couleur des légendes).

Dans le cas de R1-3, dans les conditions de contrastes d-DMPC/D_2O et h-DMPC/D_2O, les courbes NR de la monocouche de DMPC sans protéine sont superposables à celles obtenues en présence de R1-3. *A contrario*, pour le contraste d-DMPC/ACMW, les courbes NR sont différentes (Figure 66A). Pour R11-15, aux trois contrastes étudiés, les courbes NR présentent des différences de signal (Figure 66B). Ces données sont en cours d'ajustement mais les différences, d'un point de vue qualitatif, suggèrent un mode d'interaction différent pour les deux fragments protéiques considérés. Nous avons également obtenu des données pour R1-3 en utilisant un mélange de lipides zwitterioniques et anioniques (DMPC/DMPS, 2:1). Ces données, réalisées récemment sur l'appareil REFSANS (MLZ, Garching, Allemagne), sont également en cours de traitement et d'analyse et ne sont pas présentées dans ce manuscrit.

3. Mesures NR en bicouches supportées

Dans un second temps, afin d'utiliser un modèle membranaire plus proche d'une membrane lipidique, nous avons réalisé des mesures en bicouches supportées (SLB) sur des blocs de silicium. Ces données ont été obtenues en collaboration avec Alexandros Koutsioumpas, co-responsable du réflectomètre MARIA (MLZ, Garching, Allemagne). Trois contrastes ont également été utilisés (1. d-lipides/D_2O ; 2. h-lipides/D_2O ; 3. d-lipides/SMW, SMW signifiant « silicon matched water ») et uniquement R1-3 a été étudiée pour le moment. Les bicouches utilisées sont composées de phospholipides zwitterioniques (DMPC) ou d'un mélange zwitterionique/anionique (DMPC/DMPS, 2:1), afin de conserver les mêmes phospholipides que ceux composant les bicelles. Toutes les mesures ont été faites à 30°C afin que les bicouches soient en phase fluide. Nous avons tout d'abord mesuré l'environnement échantillon sans phospholipides. L'ajustement des données permet de déterminer l'épaisseur de la couche de dioxyde de silicium (9.7 Å) ainsi que sa SLD ($3.48.10^{-6}$ $Å^{-2}$), comme précédemment rapporté (Figure 67 A) [223]. Ces valeurs seront maintenues fixées pour l'ajustement du reste des données. Nous prenons également en considération la couche d'eau (~ 3.5 Å) se trouvant entre les dioxydes de silicium et les têtes polaires des phospholipides pour le reste de l'analyse [223]. Ensuite, nous avons analysé une SLB formée de h-DMPC à deux contrastes (D_2O et SMW). D'après les ajustements de données préliminaires, l'épaisseur de la partie hydrophobe est de ~ 25 Å et celles des parties hydrophiles est de ~ 10 Å aux deux contrastes utilisés (Figure 67 A). Afin de prendre en compte les propriétés électrostatiques de la membrane, nous avons employé des mélanges phospholipidiques PC/PS. L'adsorption de phospholipides possédant une tête chargée négativement sur un support en silice peut s'avérer problématique à cause des

répulsions électrostatiques. Pour résoudre ce problème, nous avons supplémenté le tampon avec 2 mM de Ca^{2+} afin que la formation de la bicouche PC/PS puisse s'effectuer correctement. La similitude des données PC et PC/PS suggère la formation d'une bicouche homogène sur tout le support (Figure 67, A et B). Une fois des bicouches stables et homogènes obtenues, nous avons ajouté le fragment protéique R1-3. Les données obtenues sur les SLB en présence ou absence de R1-3 (Figure 67, C et D) montrent que les courbes de réflectivité sont différentes uniquement lorsque la bicouche est formée d'un mélange PC/PS.

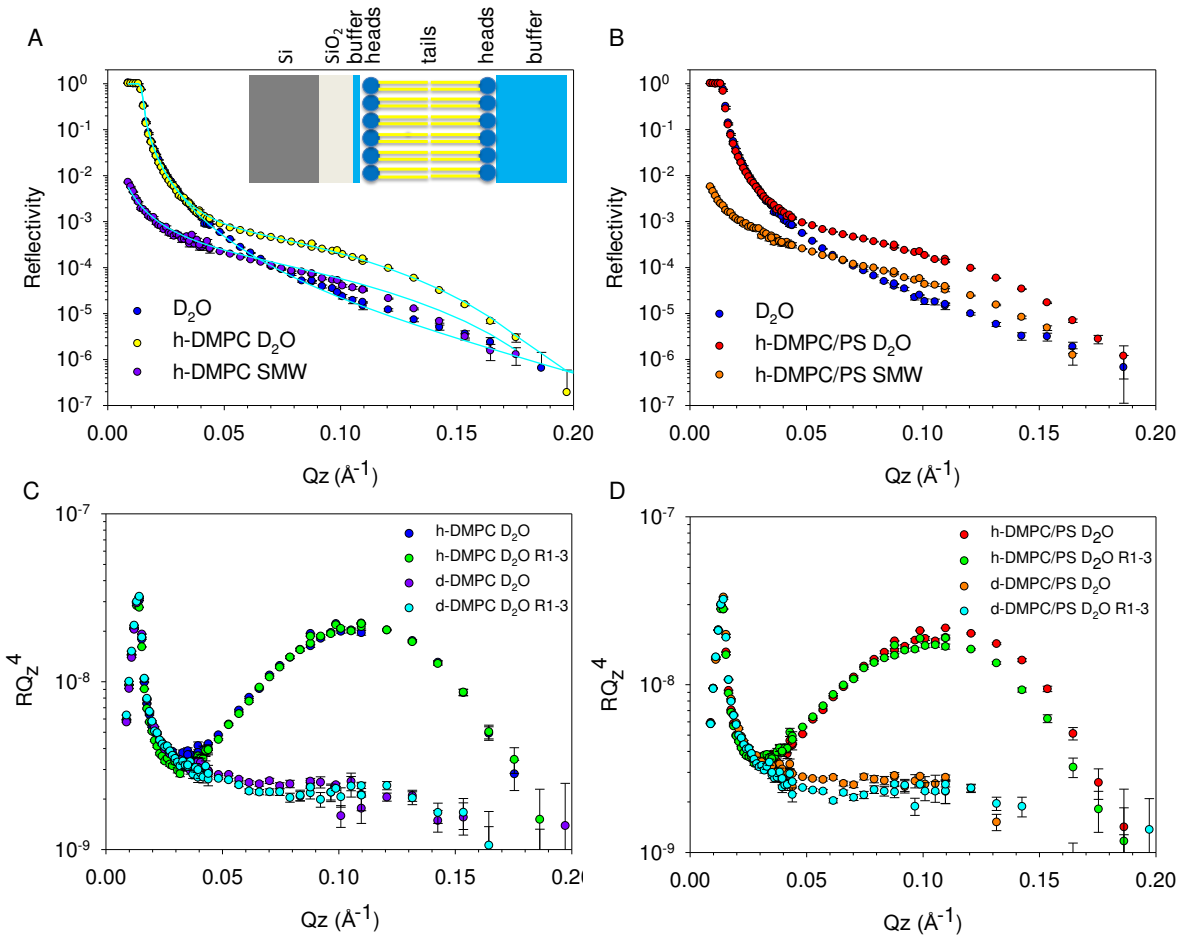


Figure 67 : Courbes NR des SLB hydrogénées formées de (A) h-DMPC et (B) h-DMPC/DMPS dans 100% D_2O (jaune et rouge, respectivement) ou SMW (violet et orange, respectivement). (C) et (D) Représentation en RQ_z^4 de SLB hydrogénées ou deutérées avant et après ajout de R1-3, uniquement dans le contraste 100% D_2O (se référer au code couleur des légendes). (A, insert) représentation schématique du système SLB.

De plus, grâce à l'analyse à différents contrastes, nous pouvons tout d'abord exclure l'insertion de R1-3 dans la partie hydrophobe de la membrane, en adéquation avec les données présentées dans l'article 2 avec les bicelles. Nous nous attelons désormais à déterminer si R1-3

interagit en s'insérant partiellement dans la partie hydrophile ou si ce fragment est uniquement apposé sur les têtes polaires. Des analyses plus approfondies sont en cours afin d'obtenir des informations plus fines et confirmer le mode d'interaction de R1-3 avec le modèle de SLB. Nous avons également pour projet de coupler les données de réflectivité avec une analyse CG-MD en collaboration avec Alexandros Koutsioumpas [223].

4. Conclusion

Dans ce chapitre, nous avons présenté les résultats préliminaires obtenus par réflectivité de neutrons. L'objectif était d'étudier une éventuelle insertion de fragments de dystrophine dans des systèmes lipidiques et, le cas échéant, le degré d'enfouissement dans ces modèles membranaires. Avant de caractériser les interactions protéine/lipide, une vérification de la qualité des modèles membranaires tout comme les environnements échantillons des grands instruments, *in situ*, était indispensable.

Un des challenges majeurs concernant ces expériences de NR consistait alors à obtenir des signaux différents entre les modèles membranaires seuls et en interaction avec les protéines d'intérêt. En effet, ce type de mesures est dans les limites de détection avec cette technique du fait que : (*i*) l'épaisseur des répétitions du domaine central de la dystrophine est seulement d'environ 2 nm ; (*ii*) le mode d'interaction des fragments de dystrophine est une adsorption ou une faible insertion dans la membrane ; et (*iii*) la protéine adsorbée ou insérée ne constitue pas une couche complète et homogène.

D'après les données obtenues avec la partie monocouche de Langmuir, nous constatons que les courbes NR de la monocouche sont différentes lorsque R1-3 est présente mais seulement à un seul contraste. Dans le cas de R11-15, les courbes sont différentes pour chaque contraste. D'après ces données préliminaires, nous pouvons en déduire que le mode d'interaction des deux fragments protéiques avec la membrane est différent.

Pour les données obtenues en SLB, nous avons été capables de former des modèles membranaires avec des phospholipides PC et également avec un mélange PC/PS, ce qui n'était pas garanti à cause des répulsions électrostatiques pouvant avoir lieu entre la tête PS et le support. Après ajout de R1-3, des différences dans les signaux NR sont observées, en particulier dans le cas de bicouches PC/PS dans le contraste 100% D_2O. Le contraste SMW permet d'obtenir uniquement le signal de la bicouche, voire même uniquement de la partie hydrophobe,

car les protéines et les têtes hydrophiles sont quasiment « éteintes » à ce contraste. Grâce à ce contraste, nous pouvons conclure des données NR que R1-3 ne pénètre pas dans la bicouche au niveau des queues hydrophobes. Ce résultat est en accord avec ceux obtenues par SANS avec le modèle « bicelle » (partie III, chapitre 2) et également avec les données précédentes obtenues par l'équipe en microscopie à force atomique [195,240]. Il nous reste à déterminer si R1-3 est adsorbée sur les têtes polaires des phospholipides ou si elle pénètre dans la partie hydrophile.

PARTIE IV : AUTRES CONTRIBUTIONS SCIENTIFIQUES

Chapitre 1 : Structure tridimensionnelle du domaine central de la dystrophine

1. Introduction

Les techniques classiques de biologie structurale, telles que la RMN et la XRC, ne peuvent pas être employées pour obtenir des informations structurales du domaine central de la dystrophine. Seule la structure de la répétition R1 a pu être résolue par XRC en ajoutant un pont disulfure artificiel afin de réduire la flexibilité du fragment protéique et permettre sa cristallisation sous forme de dimère artificiel. Au cours de cette thèse, j'ai initié des essais de cristallogenèse de R1-3 en collaboration avec Jonathan Dorival et Mirjam Czjzek de la station biologique de Roscoff. Malheureusement, les essais s'avèrent pour le moment infructueux.

Avant le début de ma thèse, l'équipe s'était tournée vers une méthode alternative combinant SAXS et modélisation moléculaire, qui a été appliquée aux fragments R11-15 et R16-17 [178]. Dans cette partie, je présente les résultats obtenus pour huit fragments natifs du domaine central de la dystrophine : R1-2, R1-3, R4-9, R11-15, R16-17, R16-19, R23 et R20-24. Le dernier fragment protéique étudié est un variant portant une mutation fréquemment retrouvée chez les patients BMD : R16-21 (Δ45-47).

Ma contribution à cet article a consisté en la réalisation et l'analyse des résultats des expériences de SEC-MALS. J'ai également participé à l'analyse des données SAXS. J'ai plus particulièrement contribué à la mise en évidence de l'état monomérique de chaque fragment, grâce entre autres à l'utilisation de la méthode développée par Rambo et Tainer, faisant appel au volume de corrélation *(V_c)* [220]. Déterminer l'état d'oligomérisation des objets est essentiel afin de ne pas se laisser entrainer dans une surinterprétation de données.

2. Publication

Les résultats sont présentés sous forme d'un article scientifique soumis à publication (actuellement en révision dans « *the Journal of Biological Chemistry* ») et intitulé:

Article 3 : Internal in-frame deletion of the DMD gene in the light of structural dissection of the dystrophin central domain

Title

Internal in-frame deletion of the *DMD* gene in the light of structural dissection of the dystrophin central domain

Running title

Structural effects of in-frame deletion of the DMD gene

Classification: Biological Sciences: Medical Genetics, Hematology and Oncology; Biophysics and Computational Biology

Authors

Delalande Olivier[1,2*], Molza Anne-Elisabeth[1,2], **Raphael Dos Santos Morais**[1,2,3,4], Angélique Chéron[1,2], Émeline Pollet[1,2], Céline Raguenes-Nicol[1,2], Christophe Tascon[1,2], Emmanuel Giudice[1,2], Marine Guilbaud[5], Aurélie Nicolas[1,2], Arnaud Bondon[1,6], France Leturcq[5,7] Nicolas Férey[8], Marc Baaden[9], Javier Perez[3], Pierre Roblin[3,10], France Piétri-Rouxel[5], Jean-François Hubert[1,2], Mirjam Czjzek[11,12] and Elisabeth Le Rumeur[1,2]

Affiliations

[1]Université de Rennes 1, 35043 Rennes, France

[2]CNRS UMR 6290, Institut de Génétique et Développement de Rennes, 35043 Rennes, France

[3]Synchrotron SOLEIL**,** 91190 Saint Aubin, France

[4]Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, Gif-sur-Yvette CEDEX F-91191, France

[5]Sorbonne Universités UPMC - INSERM - UMRS 974 - CNRS FRE 3617, Institut de Myologie, 75013 Paris, France

[6]CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, 350043 Rennes, France

[7]AP-HP, Groupe Hospitalier Cochin-Broca-Hôtel Dieu, Laboratoire de biochimie et génétique moléculaire, 75014 Paris, France

[8]CNRS UPR3251, Université Paris XI, 91403 Orsay Cedex, France

[9]CNRS UPR9080, Université Paris Diderot, Sorbonne Paris Cité, 75005 Paris, France

[10]INRA-Biopolymères, Interaction et Assemblages, 44000 Nantes, France

[11]Sorbonne Universités, UPMC Université Paris 06, 75006 Paris, France

[12]CNRS UMR 8227, Integrative Biology of Marine Models, 29688 Roscoff cedex, France

Corresponding author

Olivier Delalande

Institut de Génétique et Développement de Rennes

Faculté de Pharmacie, 35043 Rennes Cedex, France

Phone number: (+33)(0)2 23 23 30 07

olivier.delalande@univ-rennes1.fr

Short title: Structural dissection of the dystrophin central domain

OD, JFH, MC and ELR conceived and coordinated the study and wrote the paper. OD, AEM, EP, EG and AN designed, performed and analyzed the molecular simulations. OD, NF and MB designed the interactive simulation tools. AC, CRN, CT, JFH and ELR designed, performed and analyzed the protein fragment samples for SAXS experiments. OD, RDSM, JP, PR, JFH, MC and ELR designed, performed and analyzed the SAXS experiments. AB designed, performed and analyzed the NMR experiments. MG, FL and FPR designed, performed and analyzed the BMD patients cell experiments. All authors reviewed the results and approved the final version of the manuscript.

Abstract

Mutations in the *DMD* gene disrupting the reading frame prevent dystrophin production and result in the high severe Duchenne muscular dystrophy (DMD); in-frame internal deletions allow production of partly functional internally deleted dystrophin and result in the less severe Becker muscular dystrophy (BMD). Dystrophin is critical for maintaining the plasma membrane integrity during muscle contraction. It is constituted by four domains with a long central domain considered until now as a monotonous rod-shaped domain and highly involved in BMD deletions. However, in-frame deletions of the central domain do result in variable effects ranging from pauci-symptomatic to highly severe BMD. By structural studies combining small-angle X-ray scattering and molecular modeling, we showed that dystrophin central domain appears as a tortuous and complex filament which is profoundly disorganized by the most severe BMD deletion of exon 45 to 47. Despite the preservation of large parts of the neuronal nitric oxide synthase (nNOS) binding site sequence by this deletion, computational approaches revealed difficulties to recover the complex between the deleted

dystrophin and nNOS. This observation is in agreement with the total loss of nNOS immunolocalization in muscle biopsies, a parameter related to the severe phenotype of BMD patients. These findings indicate that the longest in-frame transcript produced by exon-skipping is not necessarily optimal and propose guidelines for the future design of curative therapies of DMD.

Key words: Dystrophin central domain; Becker muscular dystrophy; neuronal Nitric Oxide Synthase; Small angle X-ray scattering; multi resolution modeling

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Introduction

Mutations of the dystrophin *DMD* gene are the cause of two progressive muscle diseases of variable severity, Duchenne and Becker muscular dystrophies (DMD and BMD)(1). Dystrophin consists of a N-terminal actin-binding domain (ABD1), a central domain with 24 spectrin-like repeats interspaced by four hinge regions (2) and a cysteine-rich domain interacting with β-dystroglycan constituting the basis of the dystrophin-glycoprotein complex (DGC)(3, 4) and a C-terminal domain. According to the Monaco rule (5), DMD is mostly due to out-of-frame mutations in the *DMD* gene that result in a complete loss of the protein and a severe phenotype, while in-frame mutations of the *DMD* gene are mainly associated with BMD where modified dystrophin is produced resulting in reputed less severe phenotypes. Most BMD mutations are in-frame genomic deletions that lead to proteins lacking part of the central domain repeats (6, 7) and constitute the pattern for therapeutic strategies aiming to transform DMD patients into BMD patients (8) either by exon skipping, by injection of micro-dystrophins (9-12) or by CRSIPR/cas9 gene edition (13, 14). Indeed, the central domain has been until now considered as a monotonous rod-shaped domain which could be internally truncated without dramatic functional effects (2, 15). However, the effects of in-frame mutations involving the central domain could be more severe than expected from the Monaco rule (16, 17). Therefore, it is crucial to understand the causes of the severity of certain in-frame mutations to anticipate future therapies. This could be achieved by structural and functional studies of the dystrophin central domain and of the severe in-frame mutations.

X-ray crystallization of dystrophin repeats has been largely unsuccessful, likely due to the monomeric state of dystrophin. Repeats of two other family members — spectrin and α-actinin — have been crystallized in their biologically relevant dimer form and the only known crystal structure of a single R1 dystrophin repeat was obtained by chemically inducing dimerization (18), revealing a triple-helical coiled-coil structure typical of spectrin-like repeats. However, this structure does not allow for extrapolation of filament organization between successive repeats. Therefore, detailed structural data of larger domains are needed to acquire a precise understanding of dystrophin central domain organization and the effects of in-frame mutations on it. Due to its large size, dystrophin central domain is not amenable to current solution structural methods. Therefore, we report three-dimensional structural models of eight fragments covering the native dystrophin central domain and of the in-frame deletion of exons 45 to 47, the most frequent and severe BMD deletion, obtained by an integrative approach using small-angle X-ray scattering (SAXS) and molecular modeling. The central domain of dystrophin is far from a rod-like monotonous protein but presents numerous kinks that confer to dystrophin a tortuous and complex topology. This feature may explain how dystrophin interacts with numerous partners to form a large dystrophin-associated protein complex. The BMD in-frame mutation profoundly modified dystrophin filamentous structure and this disorganization is related to the total loss of nNOS labeling in muscle patient biopsies and an impairment of the formation of the complex associating dystrophin with nNOS. These results show that the in-frame character of the deletion does not warrant a functional protein.

Results

Native dystrophin fragments display SAXS profiles compatible with monomers in solution. We dissected dystrophin in eight native purified fragments of the central domain covering 23 over the 24 repeats of the whole central domain (Figure 1A, Table S1). Proteins were obtained as pure (Figure 1B) and α-helix folded molecules as previously observed (19, 20). The Guinier approximation analysis of the SAXS curves indicated that the fragments were non-aggregated (Figure S1). However, due to the rod-shaped objects, the Guinier plots were obtained for narrow ranges of q values (qR_g < 0.8-1.0), especially for the longest dystrophin fragments such as R4-9, R11-15, R16-19 and R20-24 which cannot be considered globular, as previously shown (21). The radii of gyration R_g, obtained from the Guinier plots range from 22.0 to 78.6 Å (Table 1). These R_g values as well as Porod volumes, both correlate linearly with the expected molecular weights (MW) (Figure S1). Moreover, the MW obtained with the methods of Rambo and Tainer (22) and using Porod volumes are in good agreement

with the MW of monomers (Table S2). These data are supported by SEC-MALS experiments performed for R1-3 and R11-15 fragments (Table S1). All these data confirm that the proteins are monomeric in solution.

The eight native fragments showed smooth SAXS scattering profiles with few structural features typical of elongated molecules (Figure 1C). The Kratky plots (Figure S2A) present a bi-modal shape with a well-defined maximum at q value of ~0.1 $Å^{-1}$ for all fragments, and a linear increase at q > 0.3 $Å^{-1}$. However, the Porod-Debye plots show a large plateau assessing the compactness of the protein fragments (Figure S2B). These features are consistent with dystrophin fragments being folded proteins, and in agreement with circular dichroism observations (19). At the same time these results are compatible with the putative inter-repeat regions variability as previously described (23). The maximum particle size (D_{MAX}) values obtained from the radial distribution function P(*r*) plots (Figure1D, Table 1) increase with the molecular mass, confirming that the fragments are elongated in solution. All R_g values obtained from P(r) functions are in line with first Guinier approximations (Table 1). Finally, we should notice that the P(r) functions tail off smoothly, especially for the largest constructs. All P(r) functions present a common peak at a short distance of ~20 Å, similar to the width of 21 Å reported for the crystallized dystrophin single repeat R1 (18), likely representing the width of the fragments. This feature is in agreement with the fact that all the fragments were in a monomeric state in solution. In all cases, except for R23, the P(r) function plots presented additional peaks, notably one peak at around 40 - 50Å which represents the distance correlation between adjacent domains (i.e. repeat length). In summary the resulting data are characteristic of rather rigid but non-linear macromolecules. These features could be attributed to specific kinks located at linker regions, and delimiting the coiled-coil structured repeats, that would disrupt a purely rod-shaped filamentous molecule.

Three-dimensional analysis evidences a tortuous and complex structure of the dystrophin central domain. We used the SASFit program (24) to obtain values for radius, contour length and Kuhn length according to three different models: long cylinder model, worm-like chain model and Kholodenko worm model (Figure S3). This analysis suggests that the Kholodenko model is the most appropriate to describe the physical characteristics of the dystrophin fragments in solution. The radius of about 10.8 Å obtained for all fragments is in accordance with the R1 monomer radius (18). Contour length higher than D_{MAX} values calculated for the Kholodenko models indicate that the dystrophin fragments are shorter than

a theoretical linear extended rod. This implies a dystrophin filament shape distinct from the rod-like structure previously assumed for the central domain of dystrophin(2, 15).

Twenty *ab initio* models, i.e. low-resolution molecular shapes, were simulated by the GASBOR method (25) for each protein fragment. All the models were superimposed using DAMSUP and analyzed by DAMAVER programs (ATSAS suite, 26). The molecular shapes converged remarkably for the R1-2, R1-3, R16-17 and R23 fragments, with normalized spatial discrepancy (NSD) values of 1 to 1.3 Å (Figure S4). For the two fragments R11-15 and R20-24, higher NSD were obtained. However, a look at the specific bulky volumes of the *ab initio* models indicates that DAMSUP placed some models head-to-tail regarding to the rest of the collection, resulting in a less efficient superimposition. Remarkably, the longest fragment R4- 9 showed a NSD of less than 2 Å, evidencing a regular molecular shape. Consistent with the SASFit analysis, the GASBOR models presented sharp kinks resulting in a highly tortuous topology.

Since most of the twenty calculated *ab initio* shapes converged to a globally homogeneous solution (Figure S4), we considered the best χ^2 *ab-initio* models – one for each fragment – as relevant guidelines to model dystrophin fragments at a higher resolution. These models were preferred to those generated by DAMAVER which do not fit the scattering curves. However, the presence of kinks precludes the simple superimposition of straight homology models obtained from spectrin templates (see Supporting Information) onto the kinked molecular shapes. Therefore, we applied the interactive flexible fitting procedure that we previously developed (27)(Supplemental Video 1) to explore the possible deformations of the spectrin-based homology models. Structure quality check (Table S3) of the final structural models indicates that the initial biological structure composed of a coiled-coil filament is properly modulated to match the SAXS-derived molecular shape (Figure 2). In the final models, kinks placed at some inter-repeat linkers disrupt the straight rod-like structure of the dystrophin fragments. Finally, experimental SAXS scattering curves and the CRYSOL theoretical curves produced from these structural models are in a range of less than 8% discrepancy for high q values that fall below 2% in low q values (Figure 2A). Consequently, these structural models suggest with confidence that the observed kinks are relevant features of the dystrophin central domain in solution (Figure 2B). All models have been successfully deposited in the SASBDB database, an international curated repository for small angle scattering data and models (28) (http://www.sasbdb.org/, see Supporting Information).

Similar to the crystal structures of spectrin and α-actinin, the B helices are generally bent in their middle (29, 30). However, the largest bends of α-helices are located at the inter-repeat linkers (Figure S5)(Supplemental Video 1). These kinks appear to be structured by several redundant features. Similar to successive spectrin repeats (29), the A/B loop of one repeat and the B'/C' loop of the following repeat interact strongly with each other and/or with the inter-repeat linker (Figure S6). In addition, the diversity in helix length in dystrophin compared to spectrin repeats (31) enables the modulation of kink amplitude and/or of the relative orientation of two successive repeats, mainly through the bending or partial unfolding of the inter-repeat linker (Figure S5). Among remarkable exceptions, we noticed the opposite behavior in the case of two of the largest observed kinks. First, the repeats R22 and R23 have respectively two particularly long A/B and B'/C' loops, which interact closely and induce a large kink at the linker (Figure S6B). The loops interaction results in a bulky extruded volume that can be clearly identified in the molecular shape (Figure 2D). Secondly, helix A of R14 and helix B' of R15 are the two shortest helices found in dystrophin repeats. *De facto*, in R14-R15, an A/B and B'/C' loop contact is not possible in a linear organization but with a large inter-repeat kink stabilized by a contact between helices B and B' (Figure S6C). Finally, kink angle values are modulated at the R1-2 and R16-17 inter-repeat linkers compared to the kinks observed in their longer overlapping fragments R1-3 and R16-19 (Table S4). This emphasizes potential dynamic behavior of the kink angles located at the inter-repeat linkers influenced by the succession of repeats.

The in-frame deletion of exons 45-47 profoundly disorganizes the structure of dystrophin at the new junction site. Building on these results, we next investigated the structural impact of the in-frame deletion of exons 45-47 in a purified mutant fragment R16-21Δ45-47 (Figure 3A and B). This mutation is the most frequently found in BMD and involves the loss of a part of the nNOS binding site (16). Guinier approximation gives a R_g value of 62.4 Å (Figure S7A). The bi-modal Kratky and Porod-Debye plots (Figure S7B and C) are consistent with an elongated and folded protein fragment. P(r) function from SAXS scattering curve of the mutant is leading to R_g and D_{MAX} values of 60.1 Å and 210 Å, respectively (Figure 3C, Figure S7D). Then, as previously done for the native fragments, the experimental SAXS data were used to constrain the molecular modeling calculations. The final high resolution model fitted into the best χ^2 *ab-initio* model showed a highly disordered topology by comparison to the filamentous structure of the corresponding native fragments

(Figure 3D). A well maintained coiled-coil structure is observed for the four repeats R16, R19, R20 and R21 conserved from the native dystrophin form (Figure 3D). However, the R17//R18 non-native repeat shows a loss of coiled-coil organization that is replaced by loosely folded loops (Figure 3D inset). All this lead to a mutant model of less quality than those obtained for the native fragments (Table S3), probably due to the poorly structured regions, including the Hinge 3 (H3). This last H3 region is composed of 40 residues between the repeats R19 and R20 and could present a loosely tendency to fold into beta-hairpins (Figure S8). In our final mutant model, the A/B loop of R16 seems to lose its interaction with the B'/C' loop of R17 as it does in the R16-17 native fragment. This feature induces a disorganization of both the R16-R17 and the R18-19 inter-repeat linkers on each side of the deletion, which drastically alters the filament organization. Interestingly, this disorganization is more dramatic than previously inferred from modeling not using any experimental data (16).

Despite the preservation of a large part of the nNOS binding site in Δ45-47 mutant, exon deletion profoundly impairs the dystrophin – nNOS interaction. Two approaches were undertaken to report for the nNOS binding in the dystrophin bearing the Δ45-47 deletion. First, a rigid docking strategy similar to the one used in our previous work (32) showed that the contact frequencies of nNOS-PDZ onto the SAXS-derived Δ45-47 model were significantly decreased compared to the wild type (Figure 4A). In particular, the two binding sites situated in the A/B helix of R16 and the linker R16-17 conserved in the Δ45-47 mutant did not show any contacts with nNOS-PDZ. Low frequency contacts on the deleted dystrophin C-terminal domain are not sufficient for its interaction with nNOS-PDZ (Figure 4B). Secondly, immunostaining experiments of dystrophin and nNOS localization were carried out on muscle sections of five Δ45-47 BMD patients and compared to normal muscle (Figure 4C). Dystrophin from normal muscle appears as an intense labeling surrounding the fibers. The nNOSμ positioning was validated by its sarcolemmal co-localization with dystrophin in normal muscle. In the muscles of the patients, dystrophin was detected at a level comparable to that of muscle control and was properly localized at the inner face of the membrane fibers. Despite the presence of dystrophin, in none of the patients labeling of nNOS could be detected. These results show that the nNOS binding is not maintained by this in-frame deletion mutant, despite the preservation of two sequences of the wild type protein involved in the binding (32, 33). This observation should be mainly related to the disordered structure induced by the Δ45-47 deletion.

Discussion

The results presented here indicate that the classical concept of the central domain of dystrophin as a rod-shaped succession of 24 monotonous repeats (2, 15) should be revisited. Evidence that the structure of this domain is complex and tortuous was obtained using experimental SAXS methodology associated with molecular modeling. We suggest that the linker regions connecting rigid repeats confer kinked and non rod-shaped features to the central domain. Such large deviations from filament linearity have not yet been observed in crystal structures of the three-repeat spectrin fragments nor of the four-repeat α-actinin, but angle values in the range of 10 - 15° were reported for inter-repeat linkers in spectrin dimers by molecular simulation (34). The large kink angles observed here in dystrophin might be due to its monomeric nature, whereas spectrin and α-actinin filaments are tetramers or dimers. Notably, they could also be 'coded' by specific inter-region linker sequences. Indeed, the dimeric - tetrameric filaments of α-actinin and spectrin are structurally maintained by inter-repeat interactions enabled by similar lengths of all repeats, whereas, in the monomeric dystrophin filament, the repeats display highly heterogeneous lengths, thereby precluding dimerization (31). The presence of specific kinks within the dystrophin filament raises the question of their biological relevance and whether they represent hinge regions with adaptable angles. Indeed, the final structural models were obtained by interactively exploring the possible angles between successive repeats without unfolding of the repeats themselves, indicating that modifications of these kink angles could occur at low energy cost. It is interesting to note that the kink angle values between successive repeats are slightly modified when appearing in two different fragments such as the R1-2 kink angle in R1-2 and R1-3 and the R16-17 kink angle in R16-17 and R16-19. However, these modifications are restricted to specific spatial directions as imposed by the structural biology restraints established for coiled-coils. Incidentally, the looseness of the linkers could preclude the crystallization of the central domain of dystrophin.

The tortuous filamentous organization of the central dystrophin domain raises important questions about the biological construction of the entire dystrophin scaffold. Large inter-repeat kink angles break the spatial alignment of the central domain by changing the direction of neighboring repeats. This feature suggests that dystrophin should no longer be considered as a succession of independent building blocks (repeats) but rather as a succession of highly interdependent multi-repeat blocks constituting structural and/or interaction modules. This

view is in line with the cooperative behavior of dystrophin repeats observed under unfolding conditions when stretched by optical tweezers (35). This compartmentalization may allow dystrophin to recruit protein partners simultaneously while bound to the plasma membrane (36). In particular, the bulkiest molecules, intermediate filaments and F-actin, are recruited by the large R11-17 domain while microtubules are recruited by R20-23 (37) and due to steric hindrance, these interactions could not occur simultaneously with a straight central domain. Remarkably, these regions overlap with *DMD* gene mutational hotspots (exons 44 to 55) (6, 7). Other modules for binding smaller protein partners emerge from the kinked filament topology of dystrophin, including the PAR-1b by repeats 8 and 9 (38, 39) and nNOS by repeats 16 and 17 (32, 33).

The emergence of multi-repeat blocks as functional interaction modules is particularly evidenced here by the consequences of the BMD in-frame 45-47 deletion. Even though a large part of the dystrophin sequence to bind nNOS is present in the Δ45-47 mutant, the disorganization of the R16-17 module by the deletion explains why we observed (i) the total loss of nNOS skeletal muscle localization in patients together with (ii) the impossibility to associate nNOS and the deleted dystrophin *in silico*. This observation could be related to the severity of this in-frame deletion (16) and could account at least partly for the heterogeneity of BMD phenotypes encountered with different deletions which may impair the nNOS binding to various degrees (40). Among these deletions, we are currently studying the deletion of exons 45 to 55 which could rescue 65% of DMD patients (13, 41).

Therefore, the maintenance of the best native-like dystrophin functional modules appears to be crucial in the design of therapeutic *micro-dystrophins* or exon-skipping/gene-editing products. Shortened *micro-dystrophins* have been designed based on BMD phenotypes by deleting some repeats that were assumed to be more dispensable than others (10, 12). Most of them were designed according to the less severe BMD deletion of exon 17 to 48 (8) and to the repeat phasing proposed by Harper (42). However, these *micro-dystrophins* were designed by considering repeats as elemental building blocks, without a true knowledge about the structural impact of disruption of the interdependent multi-repeat blocks and this remains to be investigated. *Exon-skipping* is a transcriptional therapeutic strategy expected to change DMD patients into BMD patients by subtracting additional exon(s) to restore the reading frame (9, 11). BMD deletions thus serve as patterns for DMD cure and until now they were all considered as equivalent based on the hypothesis that parts of the central domain could be

subtracted without severely affecting dystrophin structure and function. Our results indicate clearly that this is not the case and that the longest in-frame transcript produced by exon-skipping is not necessarily optimal. Therefore, BMD deletions should be studied more in detail from a protein structure to function point of view, in order to anticipate the results of therapy of DMD patients before inducing their therapeutic production.

Materials and Methods

Cloning, protein expression and purification and biochemical characterization. The cloning, expression and purification procedures for five of the eight fragments (R1-2, R1-3, R11-15, R23 and R20-24) have been described extensively in our previous papers (19) (details about sequences of fragments are shown in Table S1). R16-17 was obtained from Nick Menhart's team (20). The fragments R4-9 and R16-19 were newly overexpressed in *E. coli* and purified by similar procedures. The deletion mutant R16-21Δ45-47 was cloned, expressed and purified as previously described in detail (16). Proteins were characterized by circular dichroism (for details, see Supporting Information). The Hinge 3 region of human dystrophin was produced by Proteogenix (http://www.proteogenix.fr/) as a peptide of 47 residues with a molecular weight of 4.9 kDa and a purity of > 95%, analyzed by circular dichroism and nuclear magnetic resonance.

Small-Angle X-ray Scattering (SAXS) experiments. SAXS measurements were conducted either at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on the ID14-eh3 beamline (sample R11-15) or at the French synchrotron SOLEIL (St. Aubin, France) on the SWING beamline (all other samples). All experiments were performed at 15°C. For both set-ups, the scattering vector is defined as $q = 4\,\pi/\lambda\, sin\theta$, where 2θ is the scattering angle. Data were collected in a q-range of 0.005–0.5 $Å^{-1}$.

ID14-eh3 set-up. The data were collected on a 2D Pilatus 1M detector at a distance of 2.43 m (λ= 0.933 Å). For R11-15 a series of 5 concentrations (1.1, 2.1, 4.6, 6.4 and 10.8 mg/ml) was prepared prior to injection into the SAXS capillary. Ten frames of 1.5 seconds each were collected with alternating TNE buffer (Tris 20 mM at pH 7.5, NaCl 150 mM and EDTA 0.1 mM) and R11-15 protein samples at increasing concentrations. All frames were then averaged for each data-set after visual inspection. SAXS data were directly evaluated using the Primus software, as implemented in the ATSAS 2.3 suite (26). The experimental SAXS data for all

samples were linear in a Guinier plot of the low q region. These observations indicated that the sample did not aggregate. The radius of gyration R_g for R11-15 was derived from the Guinier approximation $I(q) = I(0) \exp(-q^2R_g^2/3)$ for $qR_g < 0.8$. Data collected at high q using high protein concentrations and at low q using low protein concentrations were then merged, with the aim to minimize the particle interference occurring at low q and to maximize the signal/noise ratio at high q.

*SWING set-up.*_All other samples data were recorded using an AVIEX170170 CCD detector at the distance of 1.807 to 1.845 m (λ= 1.033 Å). For the constructs R1-2, R1-3, R4-9, R16-17, R16-19, R16-21Δ45-47, R23 and R20-24 a stock solution of each fragment was prepared at a final concentration between 8 and 15 mg/ml. A volume between 60 to 120 µl of protein samples was injected into a size exclusion column (Bio SEC-3 300 Å, Agilent) and eluted directly into the SAXS flow-through capillary cell at a flow rate of 0.2 ml/min. The overall SEC-SAXS setup has already been described in previous work (43). The elution buffer consisted of TNE supplemented with 2% glycerol and 10 % acetonitrile for R16-19 and the BMD mutant. Two hundred fifty SAXS frames were collected continuously during the elution at a frame duration of 1.5 s and a dead time between frames of 0.5 s. One hundred frames accounting for buffer scattering were collected before the void volume. The averaged buffer scattering was then subtracted from the protein signal. SAXS curves displaying a constant R_g in a R_g versus frame number plot were averaged (Figure S1A) and were used for further characterization. Data reduction to absolute units and frame averaging and subtraction were performed in FOXTROT (7).

SAXS data analysis. All subsequent data processing and analysis were conducted with the Scatter (http://www.bioisis.net/), SASFit (44) or PRIMUS software and other programs of the ATSAS suite. The SAXS data have been successfully deposited in the SASBDB database (28) and were validated following a expert quality control procedure (http://www.sasbdb.org/, see Supporting Information).

Guinier plots assess that all samples were not aggregated (Figure S1B and S7A). The distance-distribution function P(r) and the maximum particle diameter (D_{MAX}) were calculated using the GNOM program (45). This approach offers an alternative calculation of R_g, which is based on the entire scattering spectrum (Table 1). Within a margin of error, these R_g values match those deduced from the first Guinier approximation. Scattering patterns I(q) are also

shown as Kratky plots ($q^2I(q)$ *vs* q, Figure S2A) indicating that all proteins are properly folded in solution. Typically, a globular, structured protein exhibits a pronounced maximum (bell-shaped curve), whereas a random chain (for example, an unfolded protein) will plateau (46-47). Then, Porod-Debye plots ($q^4I(q)$ vs. q^4) were generated and confirmed the compactness of all proteins (Figure S2B).

Molecular weights (MW) are generally derived from the SAXS data using the I_0 or using a standard protein, generally BSA. Both methods need an accurate determination of the protein concentration. In the SEC-SAXS experiments, the UV detector saturated and the protein concentration cannot be accurately determined. Therefore, we relied on alternative methods to control the MWs of dystrophin fragments in solution: *i.e* size-exclusion chromatography multi-angle light scattering (SEC-MALS) for two of our constructs (R1-3 and R11-15), the Porod volume-based method and that proposed by Rambo and Tainer (21). SEC-MALS experiments were performed with an HPLC system (Agilent) equipped with an UV detector coupled with light scattering (miniDAWN Treos, Wyatt) and refractive index (Viscotek, Malvern) detectors. The refractive index increment value (dn/dc) of the proteins used to determine their molecular weight was 0.185 mL/g. Data were collected using a customized Bio SEC3 column (Agilent) equilibrated with TN buffer (Tris 20 mM pH 7.4, NaCl 150 mM) with a flow rate of 0.3 $mL.min^{-1}$. Data were processed with the ASTRA software v6.1. For the Porod volume-based method, we assume that the Porod volume in nm^3 is about 1.6 times the MW in kDa (45). The method developed by Rambo and Tainer (21) was used to assess the monomeric state of all protein fragments employing the volume of correlation (Vc).

Structural models derived from the SAXS data. The overall ab initio models of the protein fragments were obtained from the SAXS experimental data using the GASBOR program (26). Following the standard procedure and using 50 harmonics, the scattering profiles were fitted to a q_{max} = 0.45 $Å^{-1}$ for all samples – except for R16-19 and R16-21Δ45-47, q_{max} = 0.3 $Å^{-1}$ (Figure S4A). Twenty independent GASBOR computations were performed for each scattering profile (each protein sample) and were superimposed using the SUPCOMB and DAMAVER programs (Figure S4B and S4C). This allowed to identify general structural features over the models of a given fragment, and thus, the result of the GASBOR computation with the smallest χ^2 was finally conserved as the unique *ab initio* model of each dystrophin fragment. This unique *ab initio* model was converted to a volume grid constraints (molecular shape) to guide the interactive flexible fitting simulations, as described in previous

work (27). Dystrophin fragments homology models (with spectrin as structural pattern) were adjusted to the molecular shape by multiple interactive simulations exploring head-to-tail initial positions of the protein fragment (see Supplemental Movie and Supporting Information for details). All the final models were refined by a standard energy-minimization (27). Evaluation of the final structural models was performed using standard quality controls (Table S3, data provided by http://services.mbi.ucla.edu/SAVES). Alpha helices curvature in the structural models (Figure S5) was measured using the measure of kink amplitude between two consecutive dystrophin repeats by defining a main axis for each coiled-coil (Table S4). Validation of these models was done by calculating their theoretical SAXS curves with the CRYSOL program (26).

Analysis of dystrophin and nNOS interaction. First, a docking strategy already described (32) was used to search for contact frequencies between the SAXS-derived structural model of Δ45-47 dystrophin mutant and the nNOS PDZ subdomain. Secondly, the nNOS immunolocalization was performed on muscle biopsies obtained from five BMD patients presenting the deletion of exons 45-47 from the French population (7). All details are described in the Supplementary experimental procedures.

Acknowledgements

This research was supported by research grants from the Association Française Contre les Myopathies, AFM–Téléthon, the Conseil Régional de Bretagne, RTR Biologie-Santé of the Université Européenne de Bretagne and the Grand Equipement National de Calcul Intensif - GENCI program (DYSIM project). AEM was recipient of a fellowship from AFM-Téléthon. AN was recipient from a Centre National de la Recherche Scientifique (CNRS) fellowship. We thank Dr Nick Menhart for providing purified R16-17 protein. We also thank the synchrotron facilities SOLEIL (St. Aubin) and ESRF (Grenoble) for allocating regular beam time and their dedicated staffs for technical help with the beamlines SWING and ID14-EH3, respectively. We thank UMS BioSit for access to the spectroscopy facility. We warmly thank Jean-Claude Kaplan for skillful discussions and critical review of the manuscript. We gratefully thank Gabriel David and Blandine Pineau from synchrotron SOLEIL (St Aubin) for support in MALS experiments.

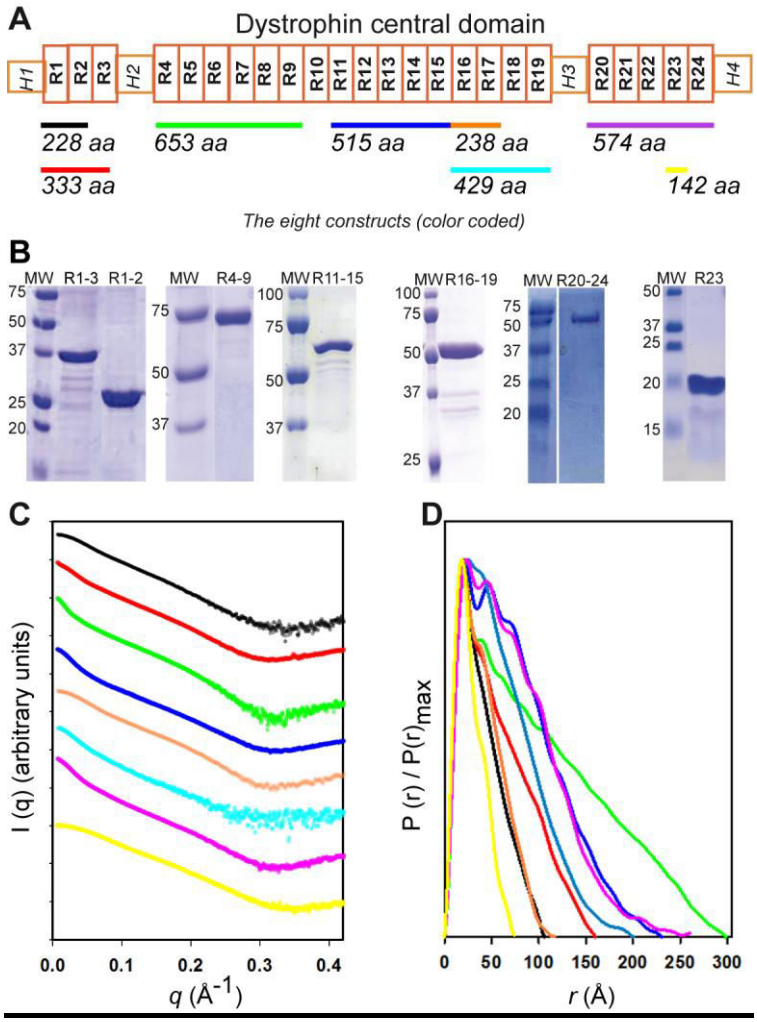
References

1. Hoffman, E. P., Brown, R. H., and Kunkel, L. M. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* 51, 919-928.
2. Koenig, M., Monaco, A. P., and Kunkel, L. M. (1988) The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* 53, 219-226.
3. Campbell, K. and Kahl, S. (1989) Association of dystrophin and an integral membrane glycoprotein. *Nature* 338, 259-262.
4. Ervasti, J. and Campbell, K. (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell* 66, 1121-1131.
5. Monaco, A., *et al.* (1988) An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* 2, 90-95.
6. Flanigan, K. M., *et al.* (2009) Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort. *Hum Mutat* 30, 1657-1666.
7. Tuffery-Giraud, S., *et al.* (2009) Genotype-phenotype analysis in 2,405 patients with a dystrophinopathy using the UMD-DMD database: a model of nationwide knowledgebase. *Hum Mutat* 30, 934-945.
8. England, S., *et al.* (1990) Very mild muscular dystrophy associated with the deletion of 46% of the dystrophin. *Nature* 343, 180-182.
9. Aartsma-Rus, A., *et al.* (2009) Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum Mutat* 30, 293-299.
10. Seto, J. T., Bengtsson, N. E., and Chamberlain, J. S. (2014) Therapy of Genetic Disorders-Novel Therapies for Duchenne Muscular Dystrophy. *Curr Pediatr Rep* 2, 102-112.
11. Ricotti, V., Muntoni, F., and Voit, T. (2015) Challenges of clinical trial design for DMD. *Neuromuscul Disord* 25, 932-935.
12. Wilton, S. D., Veedu, R. N., and Fletcher, S. (2015) The emperor's new dystrophin: finding sense in the noise. *Trends Mol Med*.
13. Young, C. S., *et al.* (2016) A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle cells. *Cell Stem Cell* 18, 1-8.
14. Wojtal, D., *et al.* (2016) Spell Checking Nature: Versatility of CRISPR/Cas9 for Developing Treatments for Inherited Disorders. *Am J Hum Genet* 98, 90-101.
15. Rahimov, F. and Kunkel, L. M. (2013) The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. *J Cell Biol* 201, 499-510.
16. Nicolas, A., *et al.* (2015) Becker muscular dystrophy severity is linked to the structure of dystrophin. *Hum Mol Genet* 24, 1267-1279.
17. Findlay, A. R., *et al.* (2015) Clinical phenotypes as predictors of the outcome of skipping around DMD exon 45. *Ann Neurol* 77, 668-674.
18. Muthu, M., Richardson, K. A., and Sutherland-Smith, A. J. (2012) The crystal structures of dystrophin and utrophin spectrin repeats: implications for domain boundaries. *PLoS One* 7, e40066.
19. Legardinier, S., *et al.* (2009) Mapping of the lipid-binding and stability properties of the central rod domain of human dystrophin. *J Mol Biol* 389, 546-558.
20. Mirza, A., Sagathevan, M., Sahni, N., Choi, L. and Menhart, N. (2010) A biophysical map of the dystrophin rod. *Biochim Biophys Acta* 1804, 1796-1809.
21. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* 40, 191-285.

22. Rambo, R.P. and Tainer, J.A.(2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496 (7446), 477-481.
23. Bernado, P. (2010) Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. *Eur Biophys J* 39, 769-780.
24. Kohlbrecher, J. (2010), ed. Villigen, P. S. I.
25. Koch, M. H. J. and Svergun, D. I. (2003) Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* 66, 1735-1782.
26. Petoukhov, M. V.*, et al.* (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* 45, 342-350.
27. Molza, A.-E.*, et al.* (2014) Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly. *Faraday Discussions* 169, 45-62.
28. Valentini, E.*, et al.* (2015) SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res* 43, D357-363.
29. Grum, V. L., Li, D., MacDonald, R. I., and Mondragon, A. (1999) Structures of two repeats of spectrin suggest models of flexibility. *Cell* 98, 523-535.
30. Ribeiro Ede, A., Jr.*, et al.* (2014) The structure and regulation of human muscle alpha-actinin. *Cell* 159, 1447-1460.
31. Winder, S. J., Gibson, T. J., and Kendrick-Jones, J. (1995) Dystrophin and utrophin: the missing links! *FEBS Letters* 369, 27-33.
32. Molza, A. E.*, et al.* (2015) Structural Basis of Neuronal Nitric-Oxide Synthase Interaction with Dystrophin Repeats 16 and 17. *J Biol Chem* 290, 29531-29541.
33. Lai, Y., Zhao, J., Yue, Y., and Duan, D. (2013) alpha2 and alpha3 helices of dystrophin R16 and R17 frame a microdomain in the alpha1 helix of dystrophin R17 for neuronal NOS binding. *Proc Natl Acad Sci U S A* 110, 525-530.
34. Mirijanian, D. T., Chu, J. W., Ayton, G. S., and Voth, G. A. (2007) Atomistic and coarse-grained analysis of double spectrin repeat units: the molecular origins of flexibility. *J Mol Biol* 365, 523-534.
35. Bhasin, N.*, et al.* (2005) Molecular extensibility of mini-dystrophins and a dystrophin rod construct. *J Mol Biol* 352, 795-806.
36. Bajanca, F.*, et al.* (2015) In vivo dynamics of skeletal muscle Dystrophin in zebrafish embryos revealed by improved FRAP analysis. *Elife* 4, 06541.
37. Belanto, J. J.*, et al.* (2014) Microtubule binding distinguishes dystrophin from utrophin. *Proc Natl Acad Sci U S A* 111, 5723-5728.
38. Yamashita, K.*, et al.* (2010) The 8th and 9th tandem spectrin-like repeats of utrophin cooperatively form a functional unit to interact with polarity-regulating kinase PAR-1b. *Biochem Biophys Res Commun* 391, 812–817.
39. Dumont, N. A.*, et al.* (2015) Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nat Med* 21, 1455-1463.
40. Gentil, C.*, et al.* (2012) Variable phenotype of del45-55 Becker patients correlated with nNOSmu mislocalization and RYR1 hypernitrosylation. *Hum Mol Genet* 21, 3449-3460.
41. Beroud, C.*, et al.* (2007) Multiexon skipping leading to an artificial DMD protein lacking amino acids from exons 45 through 55 could rescue up to 63% of patients with Duchenne muscular dystrophy. *Hum Mutat* 28, 196-202.
42. Harper, S. Q.*, et al.*(2002) Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nature Med* 8, 253-261.
43. Perez, J. and Nishino, Y. (2012) Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr Opin Struct Biol* 22, 670-678.

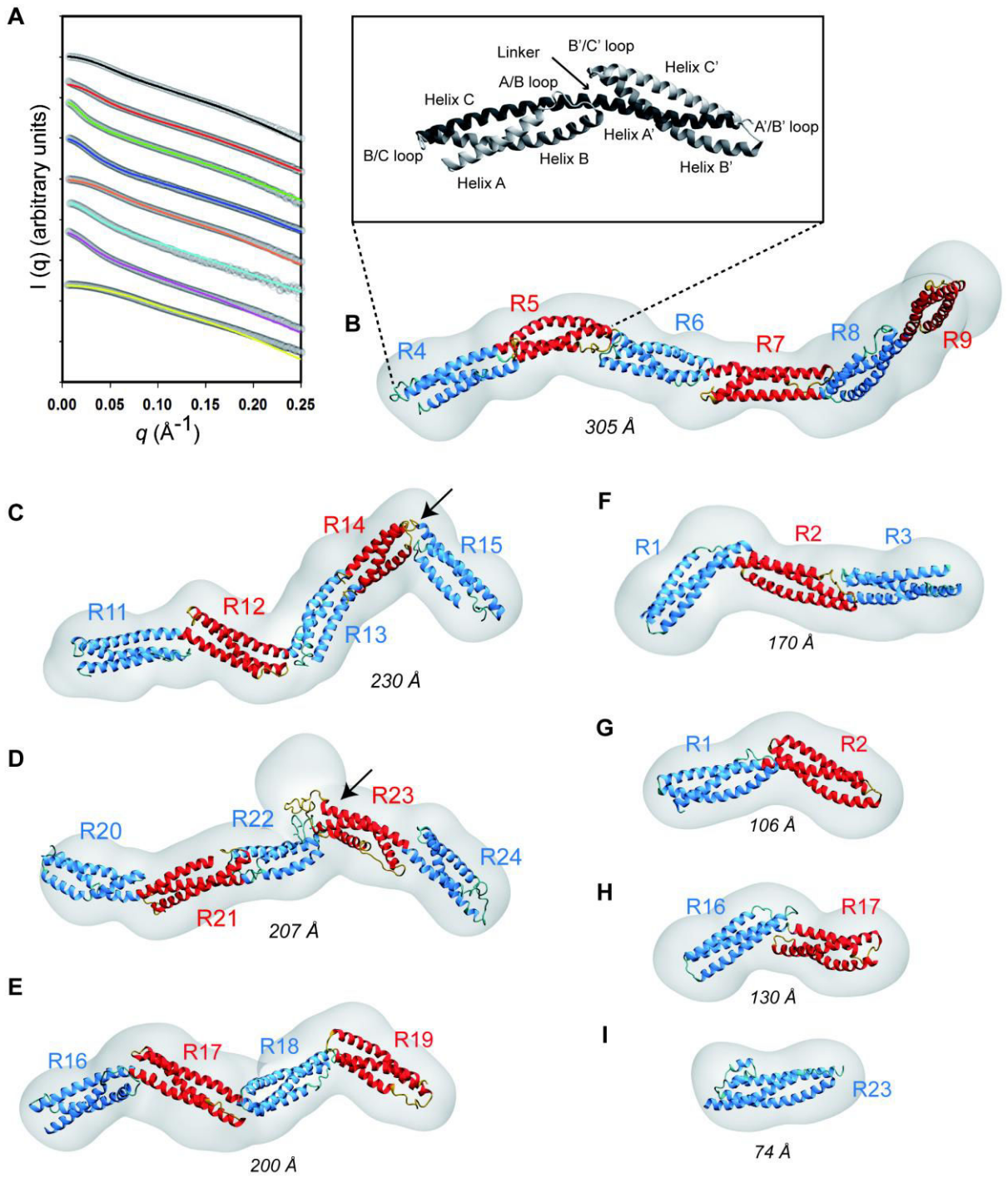
44. Bressier, I. Kohlbrecher, J. and Thünemann, A.F. (2015) *J Appl Crystallogr* 20,1587-1598.
45. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80, 2946-2953.
46. Koch, M. H. J., Vachette, P., andSvergun, D. I. (2003) Small angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys* 36**,** 147–227.
47. Skou, S., Gillilan, R. E., and Ando, N. (2014) Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat Protoc* **9,** 1727-1739.

Figure Legends



<u>Figure 1</u>. Experimental SAXS data obtained for the eight distinct native dystrophin fragments

(A) Schematic representation of the dystrophin central rod domain. It is comprised of 24 repeats (white boxes numbered R1 to R24) and four hinges (H1 to H4). The fragments studied here are highlighted by color-coded horizontal bars with indication of their length (number of amino acids). (B) SDS-PAGE of the fragments (R16-17 is not shown) indicates that they are highly pure and that they migrate at the expected molecular weight. (C) Experimental SAXS scattering profiles. The intensity values of each curve are shifted by applying a simple scale factor along the logarithmic axis to achieve clearer visualization. From top to bottom: curves R1-2, R1-3, R4-9, R11-15, R16-17, R16-19, R20-24 and R23 are displayed in the colors used in A. (D) The P(*r*) distribution functions obtained by GNOM analysis are shown for the eight fragments (color-coded as in A); the maximum *r* value (D_{MAX}) of each fragment corresponds to its maximum dimension in abscissa.



<u>Figure 2</u>. SAXS-derived three-dimensional structural models for the eight native dystrophin fragments

(A) Theoretical scattering curves obtained from the final structural models by a CRYSOL program calculation performed for each fragment. All theoretical curves are shown for q < 0.25 Å in a plain line superimposed onto its corresponding experimental grey-dotted line (color code as in Figure 1). (B to I) The molecular shape obtained from GASBOR with the best χ^2 is shown for each fragment as a grey volume. The SAXS-derived structural models of each dystrophin fragment are overlaid onto its corresponding molecular shape in a cartoon representation. The α-helices of a coiled-coil repeat are colored in red or blue for successive repeats. From the longest to the shortest fragment are shown (B) R4-9, (C) R11-15, (D) R20-

24, (E) R16-19, (F) R1-3, (G) R1-2, (H) R16-17 and (I) R23. The D_{MAX} in Å is indicated for each fragment. Arrows indicate the two more pronounced kinks observed at the linker regions on the R11-15 and R20-24 fragments. The inset shows the structural organization of a two-repeat coiled coil with the three helices A, B and C of the first repeat and the helices A', B' and C' of the second repeat. The helices are connected by the A/B or A'/B' and B/C or B'/C' loops. The linker connecting the two repeats corresponds to a eight amino acids common helix bridging C and A'.

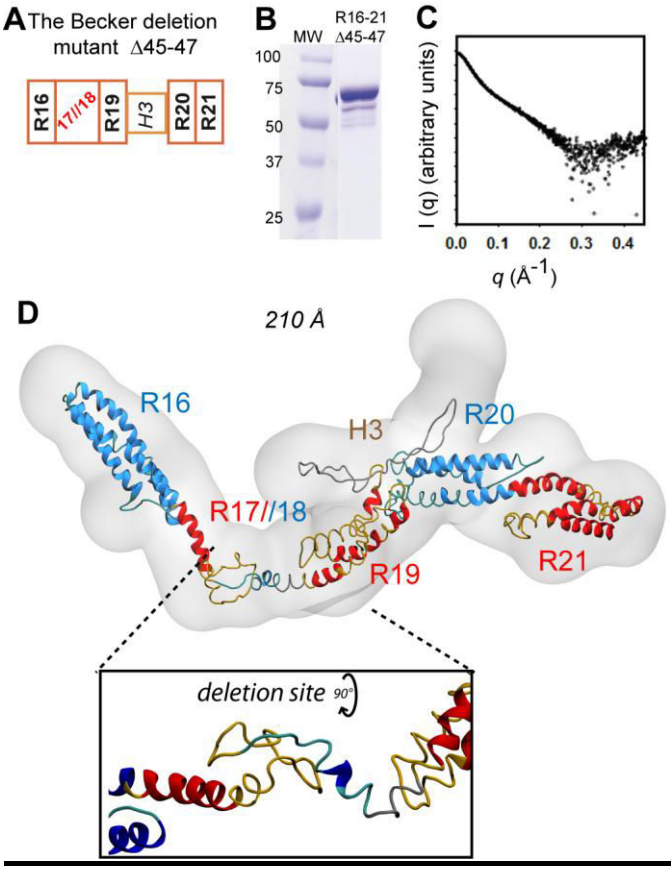


Figure 3. SAXS-based model of the R16-21Δ45-47 deletion mutant

(A) Schematic drawing of the Becker Δ45-47deletion in the R16-21 fragment of dystrophin. This mutation results in the deletion of the C-terminal part of R17 and the N-terminal part of R18, creating a non-native repeat termed R17//18. (B) SDS-PAGE of the fragment indicates that it is highly pure and migrates at the expected molecular weight. (C) The experimental SAXS scattering curve shows a lower signal-to-noise ratio for large q (>0.2 $Å^{-1}$) than in native fragments due to a lower sample concentration. (D) The structural model of the Becker Δ45-47 deletion mutant is shown overlaid on the SAXS-derived molecular shape in grey volume. The *inset* highlights filament disorganization at the newly created junction site caused by the deletion as obtained through flexible fitting simulation.

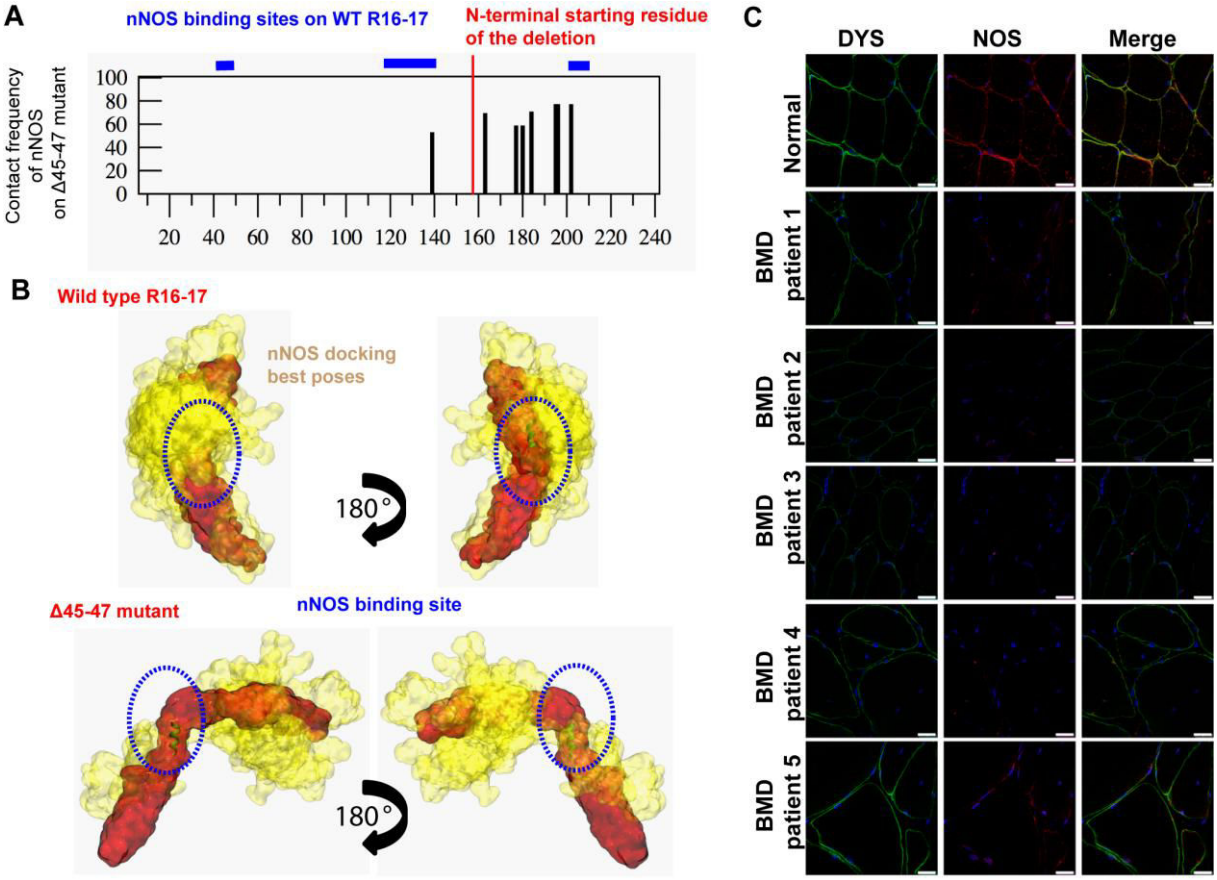


Figure 4: nNOS and dystrophin R16-17 binding in the BMD Δ45-47 deletion mutant. (A) Contact frequency of nNOS on the deletion mutant Δ45-4 compared to the three binding sites of the wild type dystrophin repeats R16-17 (in blue lines above the graph). Abscissa represents the primary sequence of dystrophin fragment where the amino acids 1 to 158 are similar for the two proteins and the amino acids from 159 to 240 are specific for each of the proteins (the deletion starts at the residue 159). (B) Interaction sites with a high contact frequency (>60%) projected onto dystrophin wild type R16-17 (top) and Δ45-47 mutant (bottom). The dystrophin proteins are shown in red and the ensemble of best nNOS poses are shown in yellow. (C) nNOS localization in Becker Δ45-47 muscles. Fixed frozen muscle sections (8 µm) from normal control (Normal) and 5 Becker patients (BMD patient 1 to 5) were immune labelled with anti-C-terminal dystrophin (green) and anti-nNOSµ (red) antibodies, nuclei with DAPI (blue) and imaged by confocal microscopy. Scale bars are 25 µm.

Supporting information (SI)

Delalande et al.

SI Materials and methods

Cloning and protein purification. The plasmid pTG11025 harboring cDNA for the Dp427m muscle isoform of human dystrophin (P11532) (National Center for Biotechnology Information Nucleotide Data Base NM-004006, provided by S. Braun Transgene, France) was used as a template for the PCR amplification of eight wild type fragments and the Becker deletion mutant (Table S1).

The cloning, expression and purification procedures for five of the eight fragments (R1-2, R1-3, R11-15, R23 and R20-24) have been described extensively in our previous papers (1, 2). Fragments were concentrated using centrifugal concentrators and the fragments were conditioned in Tris pH 7.5, NaCl 150 mM, EDTA 0.1 mM buffer with 2% glycerol for radiation protection (TNEG buffer). R16-17 was obtained from Nick Menhart (3) and transferred to TNEG buffer. The fragments R4-9 and R16-19 were newly cloned, expressed, purified and concentrated by following similar procedures. Both GST-tagged fragments were purified by affinity chromatography. The R4-9 fragment required an additional ion exchange chromatography step, and the R16-19 fragment required isoelectric focusing for purification followed by size exclusion chromatography. The latter fragment was very difficult to obtain in the amount required for SAXS acquisition without aggregation. The addition of 10 % acetonitrile to the TNEG buffer prevented aggregation after the last purification step; thus, SAXS acquisitions were made in the presence of acetonitrile for the R16-19 fragment.

The deletion mutant R16-21Δ45-47 was cloned, expressed and purified as previously described in detail (4). Briefly, the protein was recovered in *E. coli* in the form of inclusion bodies, was solubilized with N-lauryl-sarcosine (0.1%), and was purified by affinity chromatography followed by hydrophobic chromatography. It was concentrated in TNE buffer (5% glycerol). Size exclusion chromatography was performed at the end of the purification process for each protein before concentration. Purity was confirmed by SDS-PAGE; all proteins appeared as unique bands at the expected molecular weights. Protein concentration was calculated using the absorbance at 280 nm and the theoretical molar extinction coefficient obtained from the ProtParam server (http://web.expasy.org/protparam/).

Biochemical characterization

Circular dichroism was performed with a JASCO J-810 spectropolarimeter (Nantes, France) at a protein concentration of approximately 2.5 µM in TNE buffer. Spectra were acquired within the 200–250 nm range at 20°C with a path length of 0.2 cm. The percentage of α-helix was obtained using a 100% α-helix value of -36000 deg cm^2 $dmol^{-1}$ at 222 nm as previously described (2, 4). Thermal unfolding at 222 nm was followed by a temperature increase of 1 deg/min from 15 to 85°C. Because partial refolding after heating to 70°C was observed,

indicating that denaturation was reversible, the CD signal was fitted to a two-state transition as previously described (2).
Acquisitions showed that all fragments were folded into alpha helices, as expected. The CD signal ratio observed at 208/222 nm was approximately 1, indicating that the R4-9 and R16-19 fragments were folded in coiled-coils, as previously shown for the fragments R1-2, R1-3, R11-15, R20-24, R23 and the deletion mutant R16-21Δ45-47 (2, 4).

Thermal unfolding of the R1-2, R1-3, R4-9, R11-15, R16-17, R23, R20-24 and deletion mutant R16-21Δ45-47 fragments was reported previously, and the melting temperatures ranged from 50 to 67°C (1, 2, 4). The R16-17 fragment was shown to have a melting temperature of 69.2°C (3). Melting temperatures for the two new fragments R4-9 and R16-19 were determined to be 61°C and 70°C, respectively. These values are comparable to those previously reported for R1-2, R1-3, R4-9, R11-15, R16-17 R23, R20-24 and the deletion mutant R16-21Δ45-47 fragments which range from 50°C to 69.2°C (1-4).

Small-Angle X-ray Scattering (SAXS) experiments.

SAXS measurements were conducted either at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on the ID14-eh3 beamline (sample R11-15) or at the French synchrotron SOLEIL (St. Aubin, France) on the SWING beamline (all other samples). All experiments were performed at 15°C. The purified protein solutions were centrifuged for ten minutes at 10,000 rpm prior to X-ray analysis to eliminate all aggregates, and their concentrations were then measured by UV absorption at λ=280 nm on a Thermo Scientific NanoDrop 1000 Spectrophotometer. For both set-ups, the scattering vector is defined as $q = 4\pi/\lambda \sin\theta$, where 2θ is the scattering angle.

ID14-eh3 set-up. The data were collected on a 2D Pilatus 1M detector at a distance of 2.43 m (λ= 0.933 Å). For R11-15 a series of 5 concentrations (1.1, 2.1, 4.6, 6.4 and 10.8 mg/ml) was prepared and stored at 4°C prior to injection into the SAXS capillary using an automatic liquid dispensing robot (sample changer). Ten frames of 1.5 seconds each were collected for alternating TNE buffer (Tris 20 mM at pH 7.5, NaCl 150 mM and EDTA 0.1 mM) and R11-15 protein samples at increasing concentrations, and were then averaged for each data-set after visual inspection. SAXS data were directly evaluated using the Primus software, as implemented in the ATSAS 2.3 suite (5). With the exception of a slight concentration dependence arising from particle interference in solution at low q values, the scattering profiles of all concentrations superimposed well at high q values. Also, the experimental SAXS data for all samples were linear in a Guinier plot of the low q region. These observations indicated that the sample did not aggregate. The radius of gyration R_g for R11-15 was derived from the Guinier approximation $I(q) = I(0) \exp(-q^2R_G^2/3)$ for $qR_g < 0.8$. Data collected at high q using high protein concentrations and at low q using low protein concentrations were then merged, with the aim to minimize the particle interference occurring at low q and to maximize the signal/noise ratio at high q.

SWING set-up. All other samples data were recorded using an AVIEX170170 CCD detector at the energy of 12 keV. The detector was positioned at 1.807 to 1.845 m, and data were collected in a q-range of 0.005–0.5 $Å^{-1}$ (λ= 1.033 Å). For the constructs R1-2, R1-3, R4-9, R16-17, R16-19, R16-21Δ45-47, R23 and R20-24 a24a stock solution of each fragment was prepared at a final concentration between 8 and 15 mg/ml. A volume between 60 to 120 µl of the protein samples were injected into a size exclusion column (Bio SEC-3 300 Ǻ, Agilent) and eluted directly into the SAXS flow-through capillary cell at a flow rate of 0.2 ml/min. .The overall SEC-SAXS setup is described elsewhere (6, 7). The elution buffer consisted of TNE supplemented with 2% glycerol and 10 % acetonitrile for R16-19 and the Becker deletion mutant. Two hundred fifty SAXS frames were collected continuously at a frame duration of 1.5 s and a dead time between frames of 0.5 s. Selected frames corresponding to the main elution peak were averaged using FOXTROT (7). One hundred frames accounting for buffer scattering were collected before the void volume. The averaged buffer scattering was then subtracted from the protein signal. SAXS curves displaying the same R_g in a R_g versus frames plot were averaged and used for further characterization (Figure S1). Data reduction to absolute units and frame averaging and subtraction were performed in FOXTROT.

SAXS data analysis

All subsequent data processing (Guinier approximation leading to R_g, D_{max}, Porod volumes and Kratky plots) and analysis steps were conducted with the Scatter (http://www.bioisis.net/) or PRIMUS software and other programs of the ATSAS suite (5). The non-aggregated state of our fragments was also inferred by plotting the R_g value of an individual frame versus the frame number, following the chromatographic elution during data acquisition with SEC-SAXS that showed a constant R_g value throughout (Figure S1). Concerning the R11-15 fragment, SAXS was acquired using several concentrations of the sample. The final SAXS curve was obtained by merging the scattering signal for low and high values of q, since R11-15 showed a concentration dependent increase of the R_g value indicating intermolecular interactions; the ideal R_g value was thus calculated by extrapolation to zero concentration (Table S2).. The distance distribution function P(r) and the maximum particle diameter (D_{MAX}) were calculated by Fourier inversion of the scattering intensity I(q). The GNOM program (8). was used to compute the distance-distribution function P(r). This approach also features the maximum dimension of the macromolecule (D_{MAX}) and offers an alternative calculation of R_g, which is based on the entire scattering spectrum. Within a margin of error, these R_g values match those deduced from the Guinier approximation (Table S2).. D_{MAX} was chosen carefully to minimize the differences between calculated R_G results and to maximize the "total estimate," which indicates the agreement of the values of each criterion with their "ideal" values.

Scattering patterns $I(q)$ are also shown as Kratky plots, generated by plotting $q^2I(q)$ *vs.* q, to assess whether directly report all proteins are properly folded in solution (Figure S2A). Typically, a globular, structured protein exhibits a pronounced maximum (bell-shaped curve), whereas a random chain (for example, an unfolded protein) will plateau (9-11). Then, Porod-

Debye plots ($q^4I(q)$ *vs.* q^4) was generated and allowed the compactness of all proteins to be highlighted (Figure S2B). Additional EOM analysis fails to compute structural models in agreement with R_g and D_{MAX} (Figure S2C).

Determination of Construct Molecular Weights. Molecular weights (MW) are generally derived from the SAXS data using the I(0) or using a protein of known concentration and MW, generally BSA Both methods need an accurate determination of the protein concentration. In our SEC-SAXS experiments conditions, the UV detector saturated and the protein concentration cannot be accurately determined. Therefore, we relied on alternative methods to control the MWs of our constructs in solution: *i.e* size-exclusion chromatography multi-angle light scattering (SEC-MALS) for two of our constructs (R1-3 and R11-15),) and the method developed by Rambo and Tainer (12) employing the volume of correlation (Vc) from the SAXS measurements for all fragments (Table S2). SEC-MALS experiments were performed with an HPLC system (Agilent) equipped with an UV detector module, coupled with light scattering (miniDAWN TreosminiDAWNTreos, Wyatt) and refractive index (Viscotek, Malvern) detectors. Refractive index increment value (dn/dc) of the proteins used to determine their molecular weight was 0.185 mL/g(ref). Data were collected using a customized Bio SEC3BioSEC3 column (Agilent) equilibrated with TN buffer (Tris 20 mM pH 7.4, NaCl 150 mM). The flow rate was 0.3 mL/min. Data were processed with the ASTRA software v6.1.4.25. The method developed by Rambo and Tainer (12) was used to assess the monomeric state of all proteins fragments.

Structural models derived from the SAXS data. Fitting procedures for form factor determination were done with the SASfit software (Figure S3) but to obtain more accurate models, we used the dummy residue modeling method implemented in GASBOR (13). This method builds *ab-initio* protein-like models (monte-carlo distribution of one particle per amino acid) with an average distance between dummy residues of 0.38 nm, corresponding to the mean value of the inter-residue distance in proteins. For comparison, the widely used DAMMIN approach computes models without *a priori* distance between the beads. Therefore, the GASBOR program improves the resolution and reliability of low-resolution protein models compared to DAMMIN. The scattering profiles were fitted to a q_{max} = 0.45 $Å^{-1}$ for all samples(except for R16-19 and R16-21Δ45-47, q_{max} = 0.3 $Å^{-1}$). Twenty independent GASBOR searches were calculated for each scattering profile (Figure S4). The residuals between GASBOR calculations and the initial scattering curves are plotted and represent less than 4% of the signal. This approach allowed general structural features of each reconstruction to be identified and is an assessment of the consistency of *ab initio* solutions by the normalized spatial discrepancy (NSD) (13, 14).

The model with the smallest χ^2 among the twenty independent runs by GASBOR calculations was taken for each dystrophin fragment and was converted to volume grid constraints for the interactive flexible fitting procedure, as described in previous work (15). In the aim to perform efficient interactive simulations, the straight homology models obtained previously

and based on a spectrin-derived structural template (16) were converted into augmented Elastic Network Models at coarse-grain resolution. The filamentous coiled-coil structure was thus stabilized through a spring network presenting a lower density in the vicinity of inter-repeat junctions (i.e. linkers). Six independent interactive flexible fitting simulations were performed in both orientations of the N-terminal and C-terminal ends toward the extremities of the low-resolution molecular shape by following the BioSpring protocol developed previously (17). All-atom reconstructions were produced through a non-interactive BioSpring molecular dynamics simulation of 20,000 steps. Finally, the twelve all-atom models obtained for each dystrophin fragment were optimized using a standard energy-minimization protocol (15). Evaluation of the atomic models was performed using the SAVES web server (http://services.mbi.ucla.edu/SAVES). The best model for each dystrophin fragment was selected from the best quality scores provided by the ERRAT, Prove and ProCheck programs and according to the lowest residual signal between scattering curve and SAXS theoretical curve calculated with the CRYSOL program (ATSAS Suite, all default parameters, except the use of 50 harmonics). An ERRAT quality factor equal or higher than 95% was obtained for R1-2, R4-9, R11-15, R16-17 and R23, this value was about 90% for R1-3 and the two R16-19 and R20-24 have factors both at 84%. Ramachandran statistics with high percent of allowed and favored residues and low percent of outlier atoms were obtained (Table S3). The number of disallowed residues was high for several fragments but these residues are all located in loops, in correlation with the lower accuracy for positioning these highly flexible regions. The selected models were analyzed using the Bendix applet available for VMD. Alpha helix curvature in the final SAXS-derived all atom models was measured (18, Figure S5). Chimera software enabled the definition of the kink amplitude between two consecutive dystrophin repeats by defining a main axis for each coiled-coil and measuring the deviation angle between them (19, Table S4). All models have been successfully deposited in the SASBDB database, an international curated repository for small angle scattering data and models (20) (http://www.sasbdb.org/, see Supporting Information).

Analysis of Hinge 3 (H3) by circular dichroism and NMR. The Hinge 3 region of human dystrophin was produced by Proteogenix (http://www.proteogenix.fr/) as a peptide of 47 residues with a molecular weight of 4933.2 Da and a purity of >95%. The sequence, as follows, contains 6 PRO, 8 THR, 6 LEU and 5 VAL residues:

```
     10                  20                  30                  40
QPDLAPGLTT IGASPTQTVT LVTQPVVTKE TAISKLEMPS SLMLEVP
```

Circular dichroismspectra at 20°C from 185 to 260 nm were obtained at a concentration of 0.1 mg/mL in water at varying values of pH, ranging from 4.8 to 10 adjusted using HCl or NaOH. All spectra showed a negative peak at 198–200 nm, which is characteristic of a predominantly disordered secondary structure (Figure S8A). However, analysis of the CD spectra on the Dichroweb website (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml) (21, 22) using the algorithm CDSSTR and the reference set3 indicates that 30% of β-strands are present, together with 45% of unordered regions and 25% of turns. The NMR spectra display amide proton resonance in a relatively narrow region centered at approximately 8.1 ppm, suggesting

a random coil conformation for this peptide (Figure S8B). The weakness of the NOE connectivity (NOESY not shown) corroborates the absence of a well-defined structure for Hinge 3 in solution (Figure S8C).

Homology and *de novo* modelling of H3. Homology models used as starting structures for the interactive flexible fitting procedure were obtained as reported in our previous study on dystrophin fragments (16). Modelling of the H3 region was improved by following *de novo* structure prediction to avoid potential errors induced by poor sequence homology with putative patterns of the Protein Data Bank. The PEPfold program (http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=PEP-FOLD - forms::PEP-FOLD), as well as I-TASSER and Robetta, predicted a beta sheet arrangement for a portion of the H3 peptide, in agreement with our analysis of the CD spectra. Based on the structure evaluation of the predicted hinge models, we selected a H3 fold for the R16-21Δ45-47 BMD fragment bearing beta-strands.

Immunostaining of Dystrophin and nNOS on Becker muscular dystrophy skeletal muscle. Five BMD patients with the deletion of exons 45-47 and 1 patient as a control were studied for the immune-localization of dystrophin and nNOSμ. The five patients were followed in neuromuscular centers in France and they were symptomatic. The five patients were biopsied for diagnostic purposes after informed consent. The normal control muscle biopsy was obtained as surgical waste from orthopedic surgery of an individual without neuromuscular diseases. All biopsies were flash frozen in isopentane cooled in liquid nitrogen. For immunostaining muscle analysis, sections of tissues were performed at 8 µm on a cryostat (Leica CM3050S), fixed on glass slides. Slides were rehydrated in phosphate-buffered saline (PBS), fixed with paraformaldehyde 4% for 10 min, permeabilized with 0.5% Triton X-100 (Sigma-Aldrich) and blocked in PBS/4% bovine serum albumin/0.1% Triton X-100 for 1 h. Sections were incubated in PBS/2% BSA/0.1% Triton X-100 with the anti-nNOS (1:500), or with the anti-C-terminal dystrophin (1:500) overnight at room temperature, washed in PBS, incubated for 1 h with secondary antibodies, thoroughly washed in PBS, incubated with 4′,6′-diamidino-2-phenylindole for nuclear staining for 5 min and mounted in Fluoromount (Southern Biotech). Images were acquired using ×40 NA 1.4 objective lenses with Leica DM 2500 microscope the brightness/contrast was adjusted with Photoshop CS version 9.0.

SI figures and Tables

Table S1.The eight fragments of the dystrophin central rod domain and the Becker deletion mutant

Fragments (number of residues)	Starting residue	Ending residue	N-terminus sequence	C-terminus sequence
R1-2 (228)	338	563	*GS***E**VNLD…	…QD**I**LLKWQR
R1-3 (333)	338	668	*GS***E**VNLD…	…KSTAQISQ**A**
R4-9 (653)	718	1368	*GS***S**EIRKRLD**…**	**…**LLEQ**S**
R11-15 (515)	1461	1973	*GS***F**QKPAN…	…RLNFA**Q**
R16-17 (238)*	1984	2216	*GS*VMTEDMPLEI**S**YVP…	… EE**Q**KNILSEF*QPV*
R16-19 (429)	1994	2420	*GS***S**YVPS**…**	…LLQEL*(RAKQPD**L**)***
R23 (142)	2800	2939	*GS*LEAS**S**DQW…	…IDE**T**LERLQEL
R20-24 (574)	2469	3040	*GS***V**PALA…	…VRQLH**E**

Becker deletion mutant (number of residues)	First residue N-terminus sequence	Last residue C-terminus sequence	Residue in N-terminal of the deletion	Residue in C-terminal of the deletion	Sequence of the junction of the deletion	Molecular weight (kDa)
R16-21Δ45-47 (548)	1991 *GS*LEI**S**YVPS**…**	2694 …EE**T**HRLLQQF	2146	2305	...*WYLK*VSRA…	63.9

*Residues in **bold** indicate the start and end of the repeats according to the Winder alignment [36]. The GS in italics are residues added during cloning that remained after hydrolysis of the GST-tag by thrombin. *This construct was the kind gift of Nick Menhart [28]. **The sequence in parentheses was not cloned but is shown to indicate the end residue of the repeat R19.*

Table S2:SAXS parameters for the eight native dystrophin fragments and the mutant

	R1-2	R1-3	R4-9	R11-15	R16-17	R16-19	R20-24	R23	R16-21Δ45-47
qR_g range (*)	0.21-0.80 (46)	0.78-0.99 (11)	0.54-0.98 (11)	0.38-0.95 (22)	0.32-0.79 (29)	0.78- 0.99 (11)	0.61-0.90 (11)	0.34-0.80 (43)	0.60 to 0.95 (11)
R_g (Å) from Guinier**	29.9 ± 0.1	43.2 ± 0.1	78.6 ± 0.3	59.4 ± 0.4	31.0 ± 0.1	46.9 ± 0.6	58.6 ± 0.2	22.0 ± 0.1	62.4 ± 0.8
R_g (Å) from P(r)***	31.2	45.8	85.4	60.9	32.0	49.9	60.9	22.0	60.1
Porod volume (nm^3)	42	68	123	87	47	70	107	20	184
$MW_{expected}$ (kDa)	26.5	38.5	76.0	60.1	27.5	50.2	67.2	16.8	64.0
MW_{Porod}	26.3	42.5	76.9	54.4	29.4	43.8	66.9	12.5	115.0
MW_{Vc} (kDa)	21.1	29.1	49.1	46.6	22.8	42.4	53.9	12.7	89.1
$MW_{Vc\ real}$ (kDa)	20.6	28.1	47.9	45.9	22.2	41.1	52.2	12.7	87.5

* The number in parentheses indicates the number of points used
** Guinier approximation results in linear plots (FigureS1), allowing for the determination of R_g
*** Pair distribution function is shown in Figure1D

Table S3: Quality controls of the molecular all-atom models under SAXS-derived volume constraints

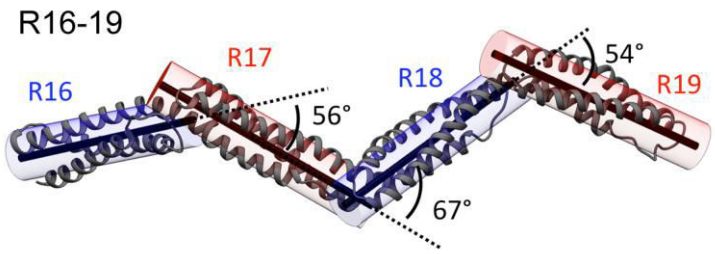
	ERRAT	PROVE	Ramachandran statistics by ProCheck			
	Quality factor	% outlier atoms	Favored	Additional allowed	Generously allowed	Disallowed
R1-2	94.8	3.9	90.8	4.9	1.9	2.4
R1-3	88.3	4.0	92.7	5.0	1.3	0.9
R4-9	96.7	3.0	87.3	8.3	1.6	2.8
R11-15	78.1	4.4	85.9	9.4	1.4	3.3
R16-17	96.5	7.4	84.2	7.7	5.0	3.2
R16-19	82.9	10.7	88.2	8.0	2.0	1.8
R20-24	84.4	7.7	74.2	14.0	6.1	5.7
R23	97.0	4.4	86.9	7.7	3.8	1.5
RΔ45-47	39.5	7.2	44.6	29.1	13.4	13.0

Table S4:Kink angles (in degree) measured between the main axes of two successive coiled-coil repeats, related to Figure 3 and 4

Fragments	R_n-R_{n+1}*	R_{n+1}-R_{n+2}	R_{n+2}-R_{n+3}	R_{n+3}-R_{n+4}	R_{n+4}-R_{n+5}
R1-2	40	-	-	-	-
R1-3	75	20	-	-	-
R4-9	40	20	35	55	20
R11-15	30	80	55	90	-
R16-17	45	-	-	-	-
R16-19	55	65	55	-	-
R20-24	40	45	65	55	-

*Coiled-coil repeat R_n is the first one of a given fragment and R_{n+1} is the following repeat.

**The angles are defined by Chimera (19). Main axis of a coiled coil is the axis of the encompassing cylinder for the three alpha-helices of a given repeat.



Inter-repeat kink angle measurements: the example of the R16-19 fragment model. Axis tool of the Chimera software allowed us to define a main axis for a set of atoms corresponding to dystrophin repeats in all the SAXS-derived models. Kink angles corresponding to the inter-repeat axis deviation were measured between the main axes of two consecutive repeats all along the filament.

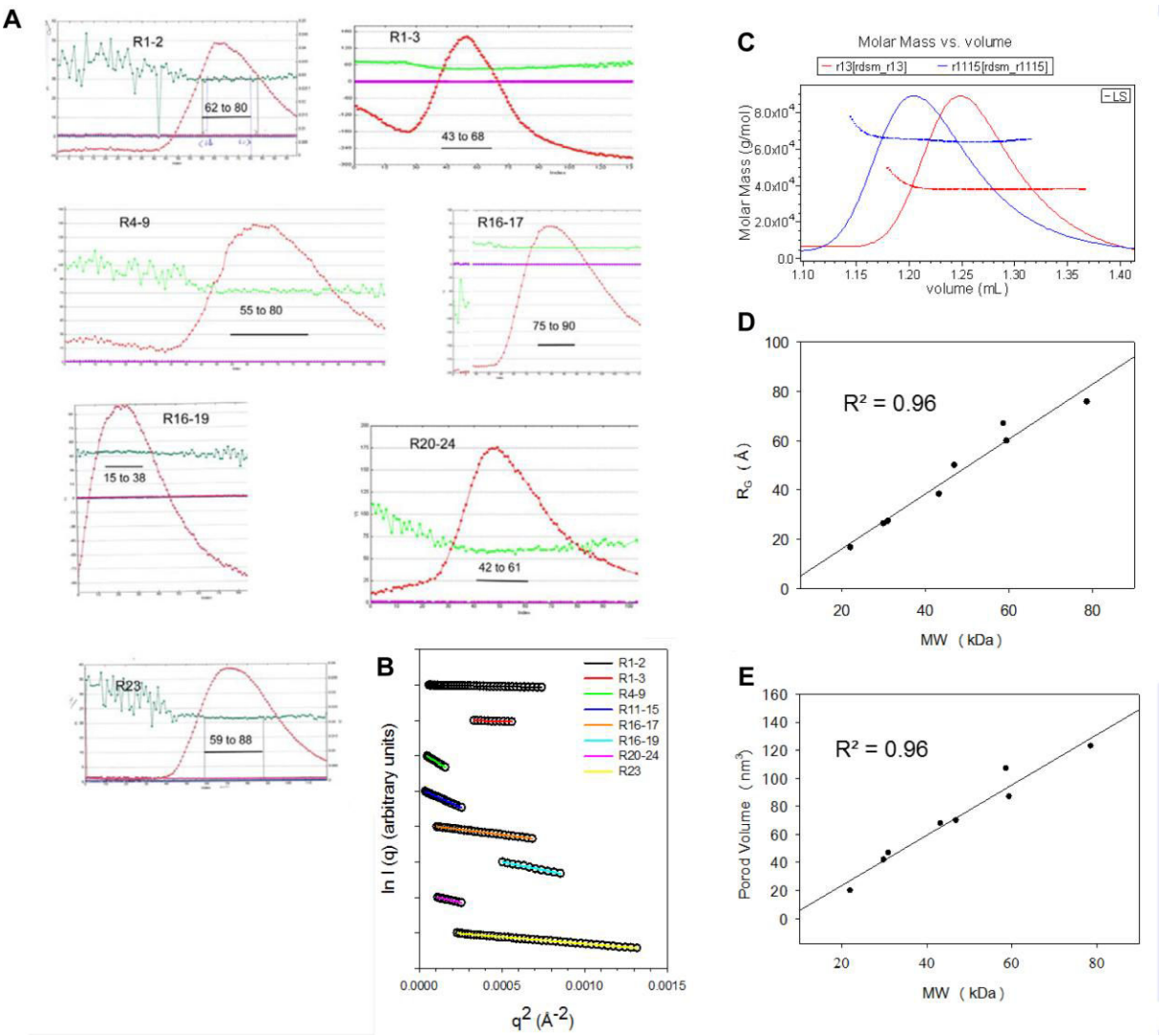
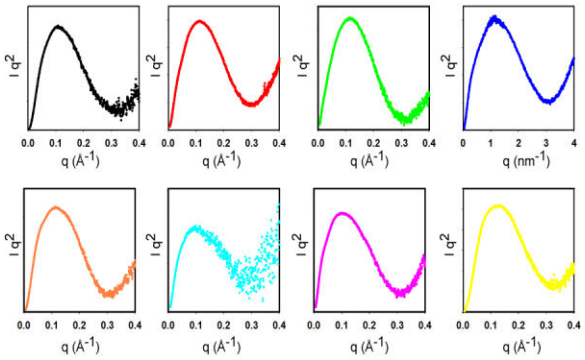
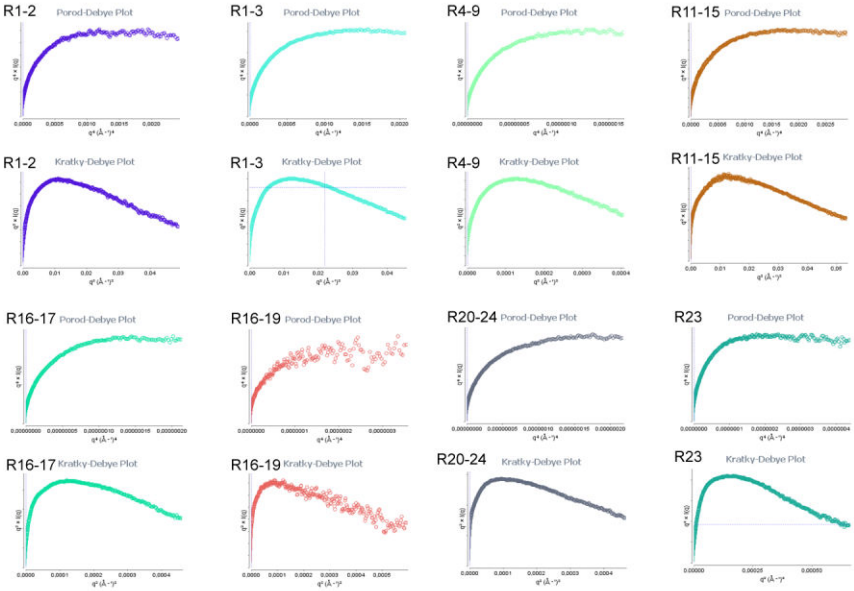


Fig. S1. Experimental SAXS data obtained for the eight distinct native dystrophin fragments. **(A)** R_g versus the frame number plots of the SEC-SAXS data. The scattering curves presented in figure 1 were obtained by averaging the indicated frames, **(B)** Guinier plots in the lowest q region of the SAXS data (the limit is defined by qRg <0.8-1.0). The plots of the eight dystrophin native fragment scattering curves are color-coded as in diagram of Figure 1A (from top to bottom: R1-2, R1-3, R4-9, R11-15, R16-17, R16-19, R20-24, and R23). (C) SEC-MALS data of R1-3 and R11-15 assessing the monomeric state of the two proteins with experimental MWs of 38.2 ± 0.1 and 65.6 ± 0.4 kDa (top of the peak), respectively. R_g (D) and Porod volumes (E) *versus* the expected MW plots for the eight protein fragments.

A. Kratky Plots



B. Porod-Debye Plots



C. EOM analysis

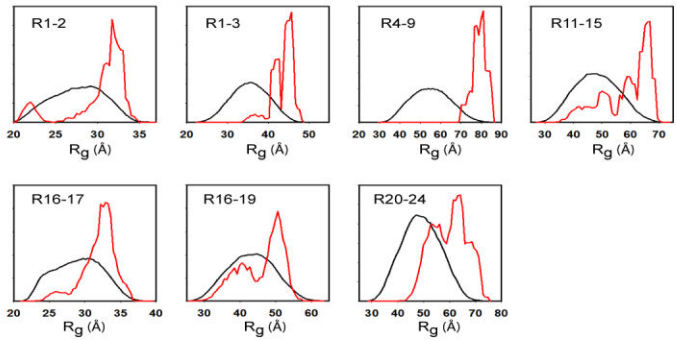


Fig. S2. (A) Kratky plots with the scattering vector S given in [Å^{-1}] for all plots except for R11-15, for which it is in [nm^{-1}]. (B) Porod-Debye and Kratky-Debye plots were obtained from the on-line server BIoIsis (http://www.bioisis.net/welcome). The plots for all fragments showed a plateau at the Porod-Debye and not at the Kratky-Debye, indicating that the fragments are more compact than flexible.(C) EOM analysis was performed through the server version of the program available online (http://www.embl-hamburg.de/biosaxs/atsas-online/). EOM is classically used to compute an ensemble of structural models in accordance to R_G and D_{MAX} obtained from analysis of SAXS data (see Table 1). This approach is particularly relevant in the case of independent subdomains separated from each other by poorly structured loop regions. In our study of dystrophin fragments, we defined as

inter-domains regions the four/eight last and four/eight first amino acids of two successive repeats. Ensembles of 10, 000 models were computed by EOM program to fit R_G and D_{MAX} parameters. Final results obtained for the seven different fragments bearing inter-repeat linkers (except the single repeat R23) were unable to propose model ensembles consistent with the SAXS-derived parameters (very low reliability chi-square values). Moreover, among the proposed models, most of them presented recurrent aberrant structural features as (i) dimeric association of two repeats and/or (ii) long range positioning of successive C-terminal amino acid of a given repeat towards the N-terminal amino acid of the following repeat. These results are in line with our proposition that repeats are modular subdomains separated by short linkers showing restricted flexibility.

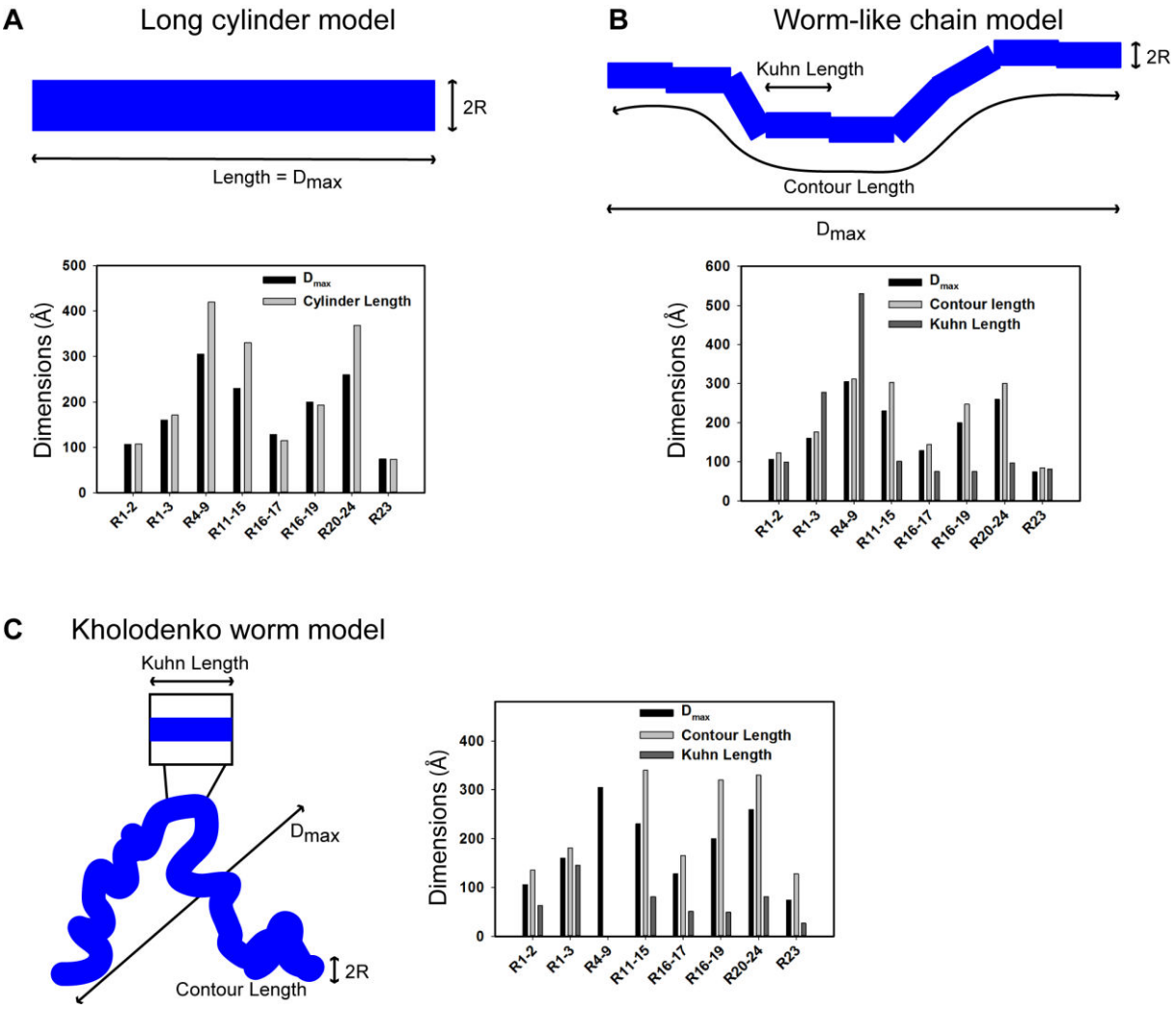
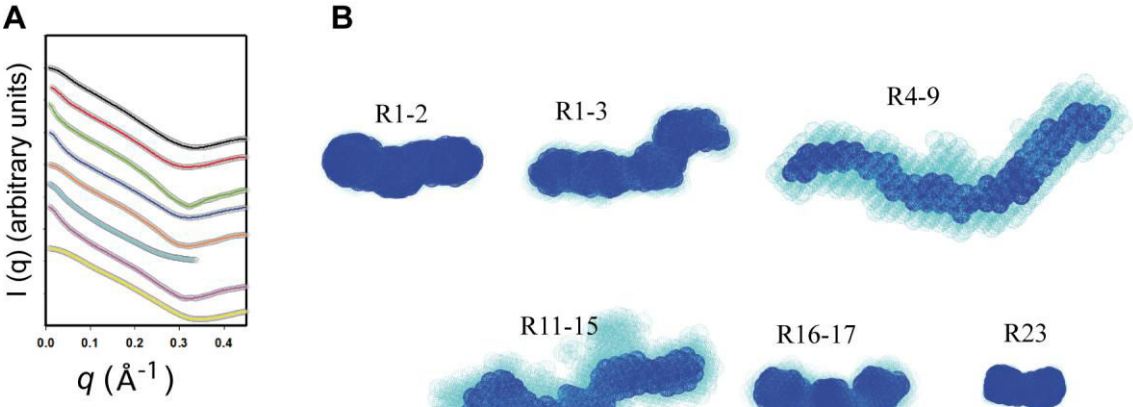


Fig. S3. **SASfit analysis of the SAXS scattering curves for the eight dystrophin fragments.**

The scattering curves were analyzed by SASFit with (A) a long cylinder model, (B) a worm-like chain model (WLCM) and (C) a Kholodenko worm model (23). The radii calculated are 12.6 ± 1 Å, 11.4 ± 0.5 Å and 10.8 ± 0.8 Å for the three models, respectively. They are all significantly different from each other ($p<0.05$). Compared to the ~10.5 Å value of the radius of the R1 crystal (24) and to the 20 Å diameter value from the P(r) distribution, the radius from the long cylinder model is ~2 Å higher and the radius from the worm-like chain model is ~1 Å higher (A and B). By contrast, the value from the Kholodenko worm model is similar to the radius of 10.5 Å calculated from the crystal structure and to the 20 Å diameter value from the P(r) (C). This agreement suggests that the Kholodenko model is more appropriate to describe the physical characteristics of our dystrophin fragments in solution as derived from their SAXS curves. The cylindrical or contour length values computed by the SASFit program for each SAXS curve were compared to the D_{MAX} obtained from the P(*r*) plot. The D_{MAX} value is actually the experimentally obtained maximum dimension of each fragment and if the fragments were pure elongated structures, there should be an adequacy between the D_{MAX} value and the cylindrical or contour lengths. The length of the long cylinder model fits with the D_{MAX} values for the shortest fragments R1-2, R1-3, R16-17, R16-19 and R23, whereas for the three other fragments, the length was longer than the D_{MAX}, showing an inadequacy of our fragments with a cylindrical model. This is confirmed by the data of the two other models, where the contour lengths are always longer than the D_{MAX}, to the exception of the computation for R4-9 fragment, which did not converge. The distortion from a cylindrical model and a contour length higher than D_{MAX} calculated by the two

other models indicate that the dystrophin fragments are shorter than a theoretical linear extended rod and, consequently, should present kinks along their length. This analysis implies a molecular shape distinct from the rod-like structure previously assumed for the central domain of dystrophin (25, 26).



C

Fragments	NSD by DAMAVER (Å)
R1-2	0.987 ± 0.029
R1-3	1.278 ± 0.056
R4-9	1.996 ± 0.196
R11-15	2.219 ± 0.143
R16-17	0.987 ± 0.017
R16-19	1.802 ± 0.076
R20-24	2.121 ± 0.102
R23	0.8 ± 0.014

Fig. S4. Analysis of a*b initio* molecular shapes from GASBOR

(A) GASBOR fit for the *ab initio* models with the lowest χ^2 of each fragment (color-coded as in Figure 1A).(B) Twenty *ab initio* molecular shapes of the eight dystrophin fragments obtained by GASBOR program and analyzed using the ATSAS Suite (27). In light blue, the DAMAVER results corresponding to the average molecular shape after superimposition fit obtained by DAMSUP; in navy blue, the DAMFILT results allowing filtering of the averaged models. (C) Normalized spatial discrepancy (NSD) values for the 20 shapes are indicated.

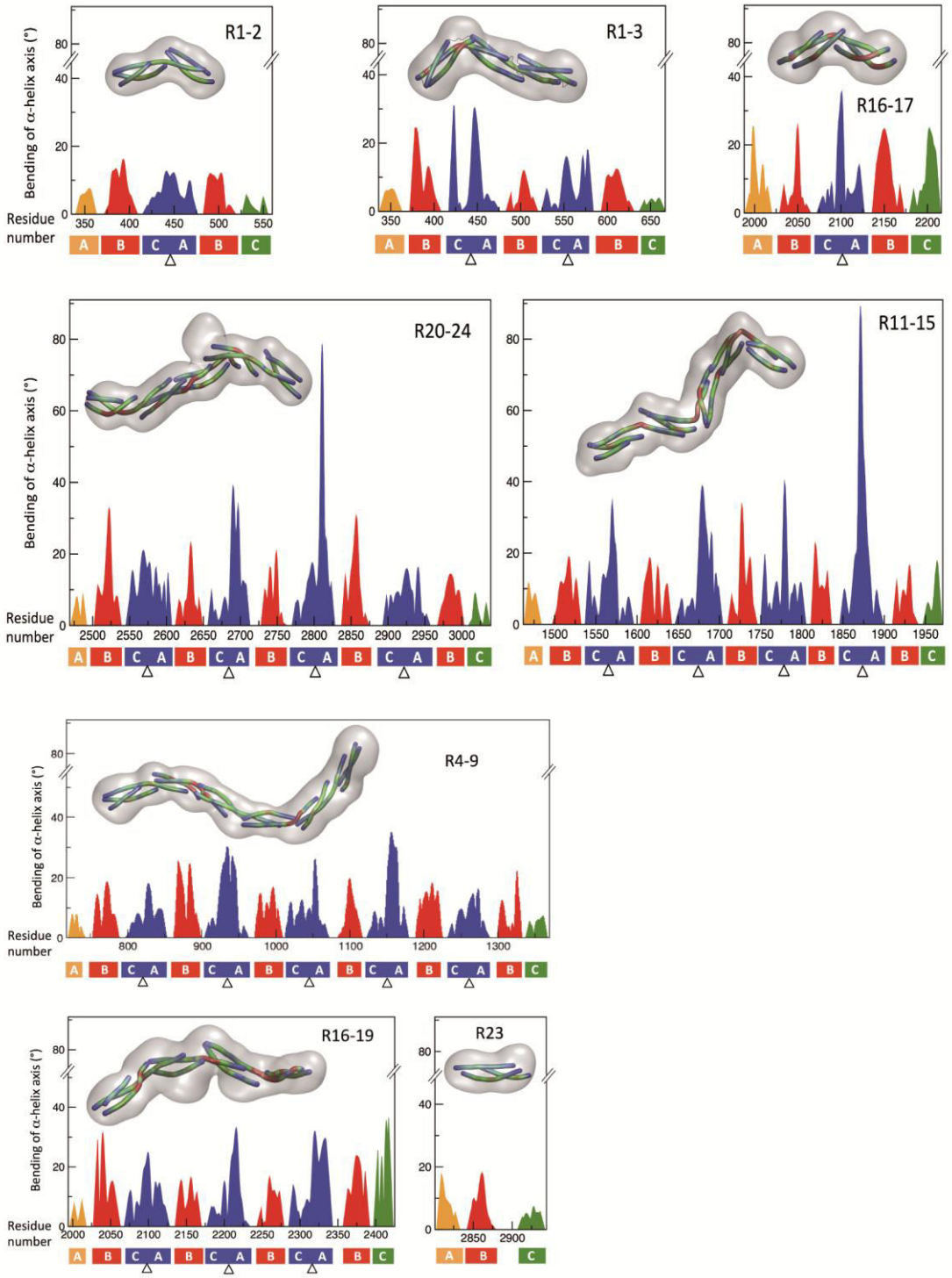


Fig. S5. Alpha helix bending analysis for all eight dystrophin fragment models
The best models obtained by SAXS-driven flexible fitting were analyzed using the Bendix (18) applet, available for VMD (28). The bending angle of alpha helices toward a straight theoretical axis was plotted (in °) according to the primary sequence of each protein. Alpha helices are indicated under the primary sequence index, and inter-repeat linkers are represented by triangles. GASBOR shapes used for the flexible fitting of dystrophin models are drawn as translucent grey volumes. Model helices are drawn as colored cylinders following their bending intensity from light (blue) to moderate (green) to strong (red) intensities.

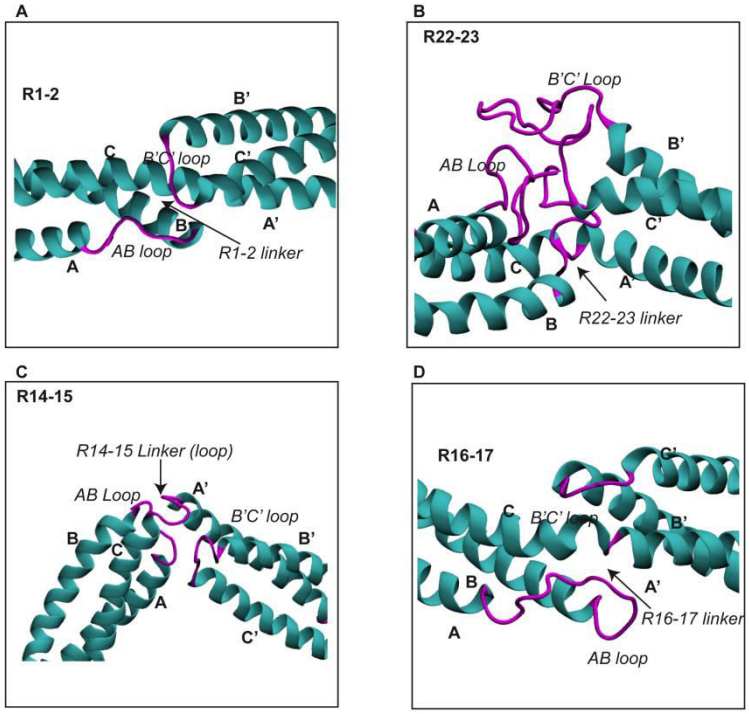


Fig. S6. Visualization of four inter-repeat linker topologies
(A) The interaction of the A/B loop of R1 with the B'/C' loop of R2 in the R1-2 model leads to a moderate kink (40°) in the linker. (B) Interaction of the long A/B and B'/C' loops of R22 and R23. (C) The linker between R14 and R15 is highly kinked and the A/B loop of R14 interacts strongly with the linker and with the B/C loop of R15 (D) A direct contact between helices B and B' is observed in the R16-17 structure that bears only a moderately kinked linker (45°), whereas no interaction between the R16 A/B loop and the R17 B'/C' loop occurs.

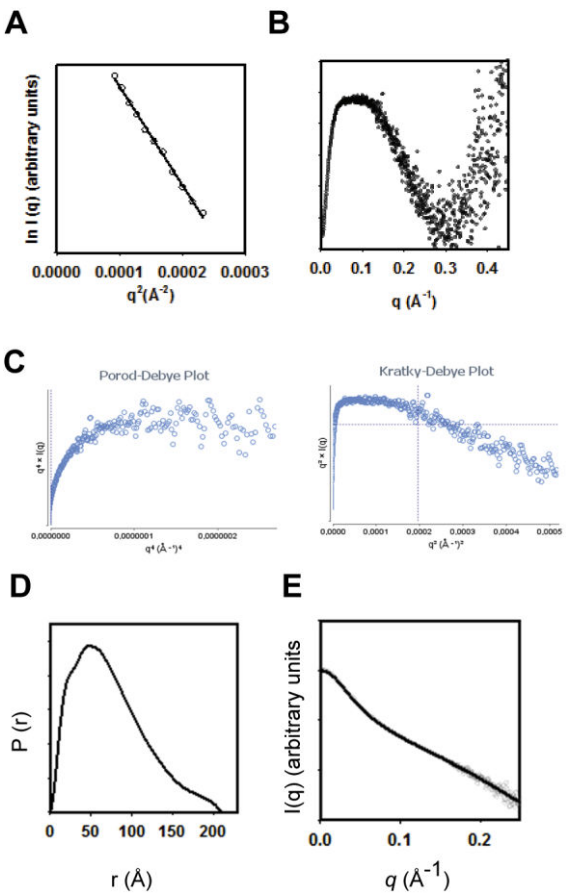


Fig. S7. SAXS-based model of the R16-21Δ45-47 deletion mutant

(A) Guinier plots in the lowest q region of the SAXS data (the limit is defined by qRg<1.0). (B) Kratky plots with the scattering vector S given in [$Å^{-1}$]. (C) Porod-Debye plot as well as Kratky-Debye plot doesn't present a clear plateau could be related to a. rather flexible protein fragment. (D) The P(*r*) distribution indicates a shoulder at a distance of 20 Å and a maximum at 48-50 Å. (E) GASBOR fit for the shape with the lowest χ^2.

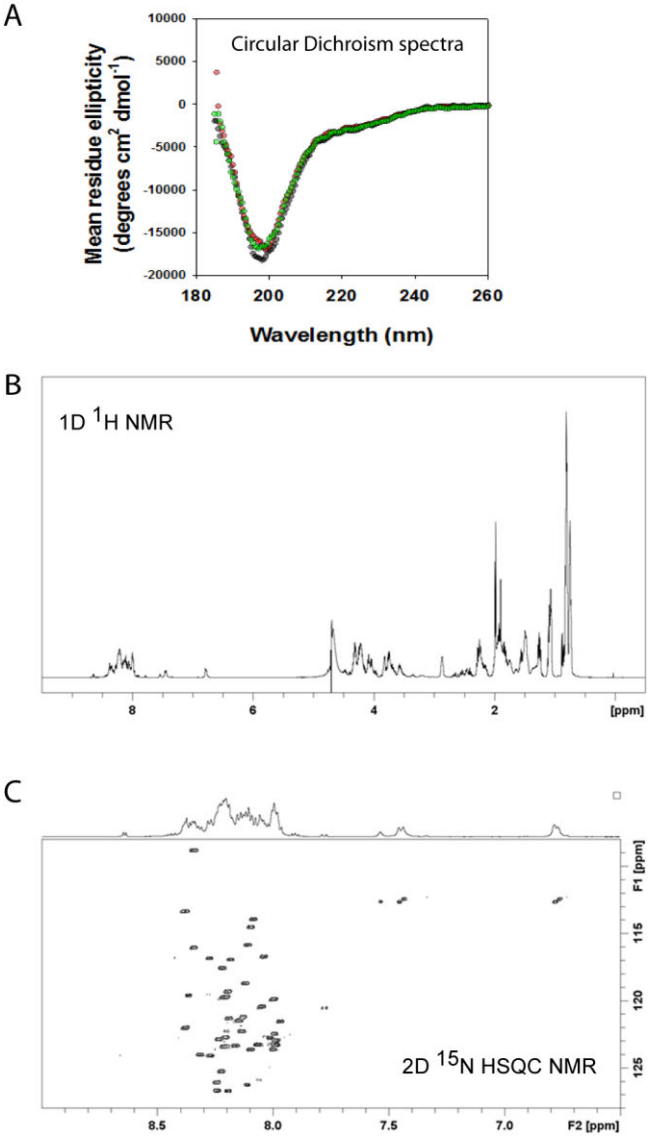
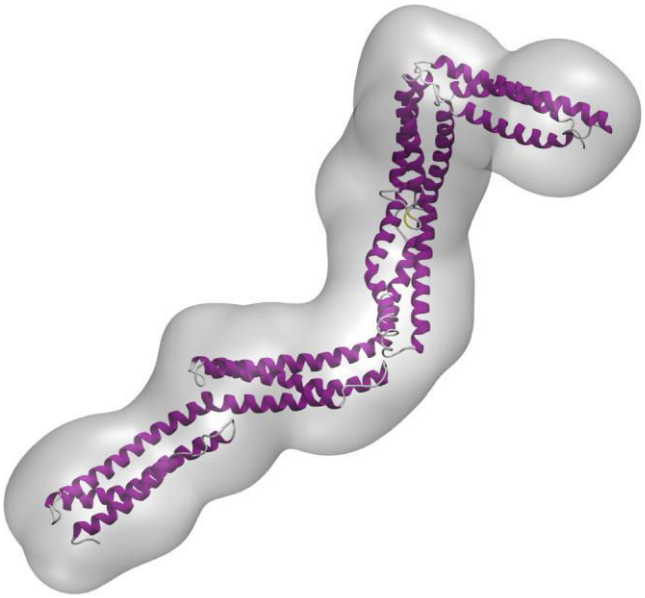


Fig. S8. Biochemical and biophysical analysis of the Hinge 3 peptide

(A) Circular Dichroism (CD) spectra of Hinge 3: the peptide concentration was 0.1 mg/mL in water and was examined at three different pH values: pH 5 (black), pH 7 (red) and pH 8.5 (green). Analysis of the spectra via the CDSSTR method using the Dichroweb server (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml)(21) showed that less than 10% of the peptide is composed of alpha helix, approximately 30% of beta-strand, 25% of turns and 35% of unordered secondary structure. (B)1H–1D NMR acquisitions: NMR spectra were acquired on a Bruker Avance 500 spectrometer equipped with a 5-mm triple-resonance cryoprobe (^{1}H, ^{13}C, ^{15}N). The concentration of unlabeled protein was 15 mM at pH 4.5. (C) 2D ^{15}N-HSQC spectra were recorded at 298 K in phase-sensitive mode using the Echo anti-Echo method as matrices of 256 (t1) × 4096 (t2) complex data points; 160 scans per t 1 increment with a repetition time of 1.2 s and spectral widths of 12 and 40 ppm for ^{1}H and ^{15}N nuclei were used. Spectra were processed with Topspin (Bruker Biospin).

Supplemental video S1



Supplemental video S1. Movie showing the reconstruction of structural models from SAXS data and molecular modelling and dynamics

References

1. Legardinier, S., *et al.*(2008) Sub-domains of the dystrophin rod domain display contrasting lipid-binding and stability properties. *Biochim Biophys Acta* 1784, 672-682.
2. Legardinier, S., *et al.* (2009) Mapping of the lipid-binding and stability properties of the central rod domain of human dystrophin. *J Mol Biol* 389, 546-558.
3. Mirza, A., *et al.* (2010) A biophysical map of the dystrophin rod. *Biochim Biophys Acta* 1804, 1796-1809.
4. Nicolas, A., *et al.* (2015) Becker muscular dystrophy severity is linked to the structure of dystrophin. *Hum Mol Genet* 24, 1267-1279.
5. Petoukhov, M. V., *et al.* (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* 45, 342-350.
6. Perez, J. & Nishino, Y. (2012) Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr Opin Struct Biol* 22, 670-678.
7. David, G. & Perez, J. (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small angle X-ray scattering experiments at the synchrotron SOLEIL SWING beamline. *J Appl Cryst* 42, 892-900.
8. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* 25, 495-503.
9. Koch, M. H. J., Vachette, P., &Svergun, D. I. (2003) Small angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* 36 147–227.
10. Skou, S., Gillilan, R. E., & Ando, N. (2014) Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat Protoc* 9, 1727-1739.
11. Svergun, D. I., Petoukhov, M. V., & Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80, 2946-2953.
12. Rambo, R. P. &Tainer, J. A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496, 477-481.
13. Kozin, M. B. &Svergun, D. I. (2000) A software system for rigid-bodymodelling of solution scattering data. *J Appl Cryst* 33, 775-777.
14. Volkov, V. V. &Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst* 36, 860-864.
15. Molza, A.-E., *et al.* (2014) Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly. *Faraday Discussions* 169, 45-62.
16. Legrand, B., *et al.* (2011) Computational Study of the Human Dystrophin Repeats: Interaction Properties and Molecular Dynamics. *Plos One* 6, e23819.
17. Tek, A., *et al.* eds. (2012) *Advances in Human-Protein Interaction - Interactive and Immersive Molecular Simulations* (InTech).
18. Dahl, A. C. E., Chavent, M., & Sansom, M. S. P. (2012) Bendix: intuitive helix geometry analysis and abstraction. *Bioinformatics* 28, 2193-2194.
19. Pettersen, E. F., *et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612.
20. Valentini, E., *et al.* (2015) SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res* 43, D357-363.
21. Whitmore, L. & Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* 32, W668-673.
22. Whitmore, L. & Wallace, B. A. (2008) Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* 89, 392-400.
23. Kohlbrecher, J. (2010), ed. Villigen, P. S. I.
24. Muthu, M., Richardson, K. A., & Sutherland-Smith, A. J. (2012) The crystal structures of dystrophin and utrophin spectrin repeats: implications for domain boundaries. *PLoS One* 7, e40066.
25. Koenig, M., Monaco, A. P., & Kunkel, L. M. (1988) The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* 53, 219-226.

26. Rahimov, F. & Kunkel, L. M. (2013) The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. *J Cell Biol* 201, 499-510.
27. Petoukhov, M. V.*, et al.* (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* 45, 342-350.
28. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J Mol Graph* 14, 33-38.

3. Résultats complémentaires

En compléments des données SAXS obtenues pour les fragments protéiques présentés ci-avant, nous présentons l'analyse des données SAXS d'autres fragments de dystrophine (Figure 68). Durant ces trois années, l'équipe a réussi à produire des fragments de dystrophine beaucoup plus grands que ceux présentés dans l'article 3, à savoir les fragments R4-15, R11-19 et R16-24. En plus de ces longs fragments, des variants protéiques issus de délétions d'exons ont également pu être produits et purifiés avec succès. C'est notamment le cas des variants R16-24 (Δ45-55) et R16-24 (Δ45-51) qui sont d'intérêt thérapeutique majeur. Tous ces fragments ont été produits en cellules d'insectes (sf9) et purifiés par chromatographie d'affinité suivi d'une chromatographie d'exclusion stérique. Le tampon utilisé pour les expériences de SAXS est PNEG (NaP 20 mM pH 7,5, NaCl 300 mM, EDTA 1 mM, Glycérol 2%). L'étiquette histidine a été conservée. La production et la purification protéique ont été réalisées par Thomas Chenuel.

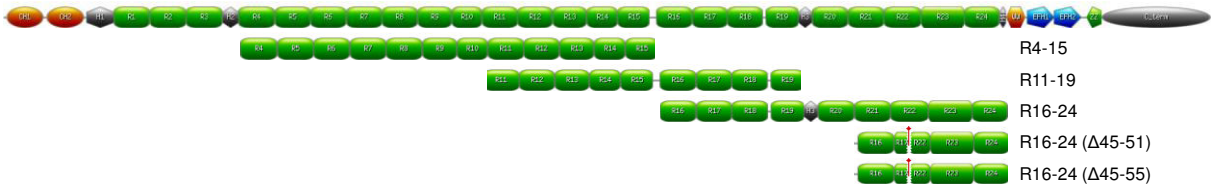


Figure 68 : Représentation schématique des longs fragments du domaine central de la dystrophine produits au laboratoire et de deux variants issus de délétions d'exons.

Deux des cinq fragments, R4-15 et R16-24, ont été analysé par SEC-MALS avant une analyse SEC-SAXS. Comme pour tous les autres fragments protéiques déjà étudiés, les résultats SEC-MALS (Figure 69) permettent de mettre en évidence la présence des fragments protéiques sous formes monomériques. Les MW déterminées sont respectivement de ~ 150 et ~ 130 kDa pour R4-15 et R16-24, résultats expérimentaux en accord avec la MW attendue selon leur structure primaire (Tableau 7). L'analyse par SEC donne également une information importante. Le volume d'élution de R4-15 est inférieur à celui de R16-24. Cette observation implique que le R_h de R4-15 est aussi inférieur à celui de R16-24. Une explication serait que R4-15 est moins étendu que R16-24, ce qui pourrait être dû au coude putatif identifié [173] entre R9 et R10 par notre équipe.

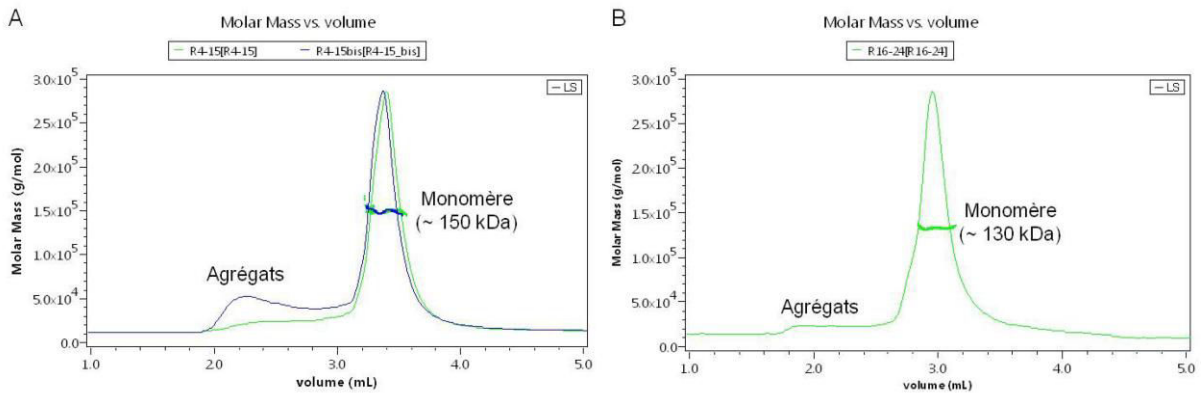


Figure 69 : Analyse SEC-MALS de R4-15 (A) et R16-24 (B) mettant en évidence la présence des fragments protéiques sous forme monomérique.

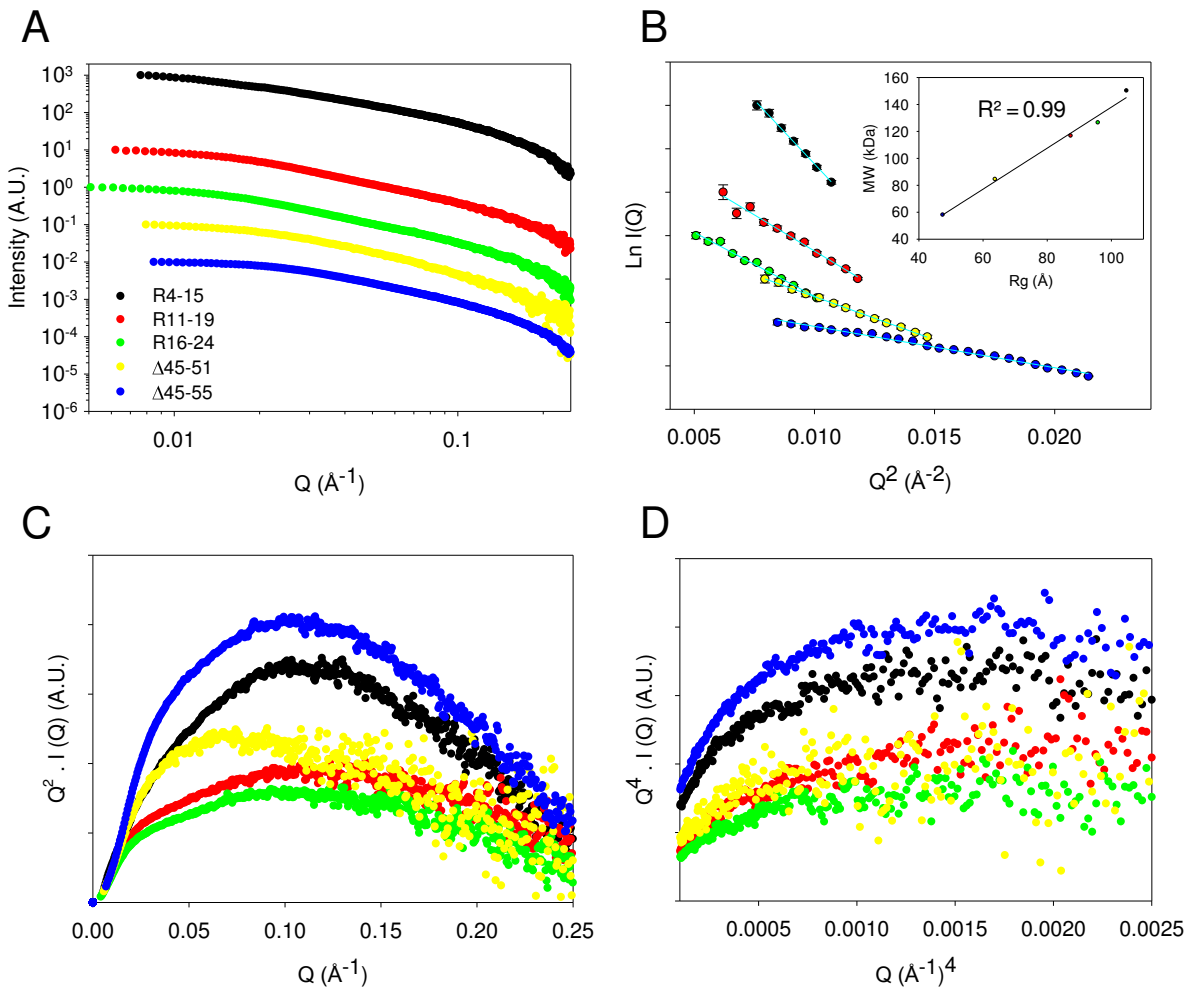


Figure 70 : (A) Courbes SAXS obtenues pour les fragments protéiques R4-15 (noir), , R11-19 (rouge), R16-24 (vert), R16-24 Δ45-51 (jaune) et R16-24 Δ45-55 (bleu). (B) Représentation de Guinier. Insert MW = f(R_g). (C) représentation de Kratky et (D) représentation de Porod-Debye (D) avec le même code couleur.

Avant une investigation et une analyse plus poussées, la qualité des données SAXS a été attestée par la détermination du R_g. Le reploiement correct des fragments protéiques a été démontré grâce aux représentations de Kratky et Porod-Debye (Figures 70 C et D). La MW a été estimée en utilisant la méthode de Rambo et Tainer [220] pour ces fragments protéiques et également pour le modèle R1 obtenu par XRC [175] sous formes monomérique et dimérique (Table 7).

Tableau 7 : Tableau récapitulatif des paramètres SAXS obtenus.

Nom	R_g	MW selon structure 1re (kDa)	MW expérimentale SAXS (V_c) (kDa)
R1 XRC monomère	18.8	13.6	14
R1 XRC dimère	21.7	27.2	28
R4-15	103.6	150	110
R11-19	90.1	116	148
R16-24	92.8	126.6	187
R16-24 (Δ45-51)	71.4	84.6	121
R16-24 (Δ45-55)	46.6	58.1	68

- ***R16-24 et variants***

D'après les données SAXS, les variants R16-24 Δ45–51 et Δ45–55 sont des fragments protéiques correctement repliés. Ces données, permettent d'émettre l'hypothèse que les jonctions entre les zones résultant de délétions d'exons sont repliées en une structure en coiled-coil, selon le concept de « hybrid repeat ». Ce type de repliement après saut d'exon étant corrélé à des formes peu sévères de BMD [180].

Le SAXS est une technique relativement rapide pouvant attester du repliement correct des protéines. Il permet un « screening » efficace de la qualité de possibles produits de thérapies géniques et donc augmente les chances de concevoir des dystrophines minimales fonctionnelles.

4. Conclusion

La communauté des scientifiques, qui s'intéressent à la structuration des cellules musculaires et aux traitements des myopathies, considère le domaine central de la dystrophine comme une structure linéaire et passive. Les résultats que nous présentons dans l'article 3

démontrent que ce concept doit être revisité. En effet, nous mettons en évidence que la structure du domaine central présente des zones coudées qui sont le signe d'une organisation tortueuse délimitant des plateformes d'interactions spécifiques et variées.

Chapitre 2 : Caractérisation des interactions dystrophine/actine

1. Introduction

Cette partie s'inscrit dans les travaux de thèse de Dominique Mias-Lucquin, actuellement en deuxième année de thèse dans l'équipe SIM (Rennes), dont l'un des objectifs est de caractériser le complexe R11-15/actine-F par cryo-EM. En parallèle du sujet principal de ma thèse, j'ai contribué à l'étude des interactions des domaines ABD1 (NT) et ABD2 (R11-15) de la dystrophine avec l'actine du cytosquelette.

2. Dystrophine/actine-F et thermophorèse

Classiquement, pour déterminer une constante de dissociation K_d entre l'actine et un partenaire, une approche par co-sédimentation est utilisée suivie d'une analyse SDS-PAGE, où l'actine est titrée par la protéine d'intérêt. Bien que relativement simple à mettre en œuvre, cette approche peut s'avérer longue et laborieuse. Nous proposons d'utiliser la technologie MST qui facilite la détermination d'un K_d quel que soit le partenaire potentiel de l'actine-F. Tout comme les bicelles, l'actine-F possède une concentration critique : environ 0,1-0,2 µM de monomères d'actine sont en équilibre avec la forme filamenteuse. Dès lors, une titration de la protéine par de l'actine-F est impossible. Une contrainte pour ce type d'études est que la concentration en actine doit être maintenue au-dessus de 5 µM pour avoir une proportion d'actine-F supérieure à 99%. Pour rappel, la polymérisation d'actine-G est déclenchée par ajout d'ATP et de Mg^{2+}, et est suivie par diffusion de lumière. L'actine-G est marquée *via* ses résidus cystéines avec une sonde maléimide fluorescente nécessaire pour les expériences MST. Marquer l'actine sur ses résidus lysines empêche la polymérisation, vraisemblablement du fait d'un encombrement stérique engendré par la sonde fluorescente. Effectivement, ces résidus lysines sont retrouvés à l'interface de polymérisation.

Beaucoup de mises au point ont été et sont encore nécessaires afin de définir les meilleures conditions de préparation des échantillons pour l'obtention systématique de données MST de haute qualité. Les résultats obtenus pour le complexe R11-15/actine-F sont très encourageants. Un K_d de l'ordre de 10 µM a été obtenu et est en adéquation avec les valeurs de la littérature obtenues par co-sédimentation [182].

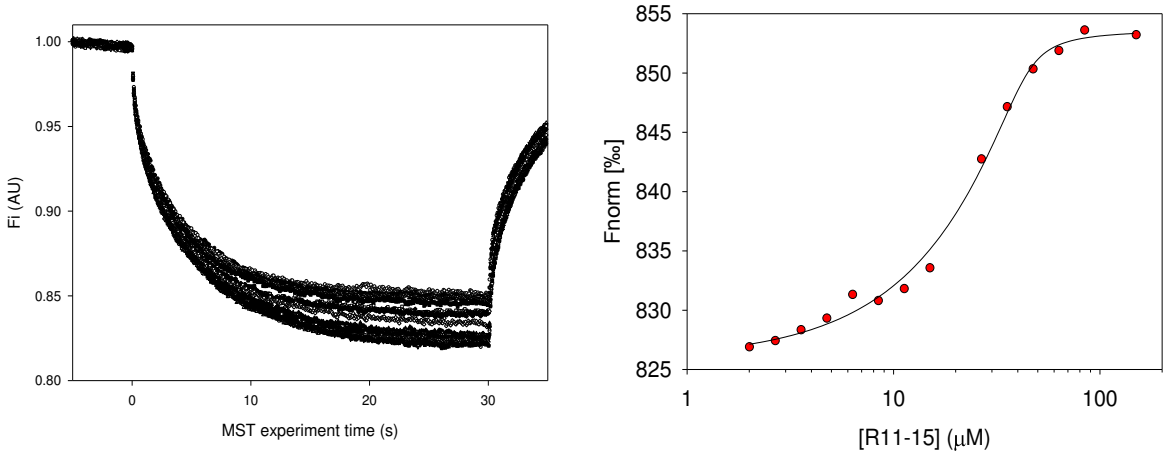


Figure 71 : Analyse MST du complexe R11-15/actine-F. Les données brutes permettent de mettre en évidence la faisabilité de la caractérisation des interactions protéines/actine-F par cette technique. La courbe de titration résultante permet de déterminer un K_d de 10 µM, résultat en accord avec la littérature [182].

3. Cartographie des interactions dystrophine/actine

Nous avons utilisé une approche de « chimie-click » couplée à la spectrométrie de masse pour cartographier les zones d'interaction dystrophine/actine. L'analyse SDS-PAGE a permis d'identifier une bande correspondant dans les deux cas à une stœchiométrie 1:1.

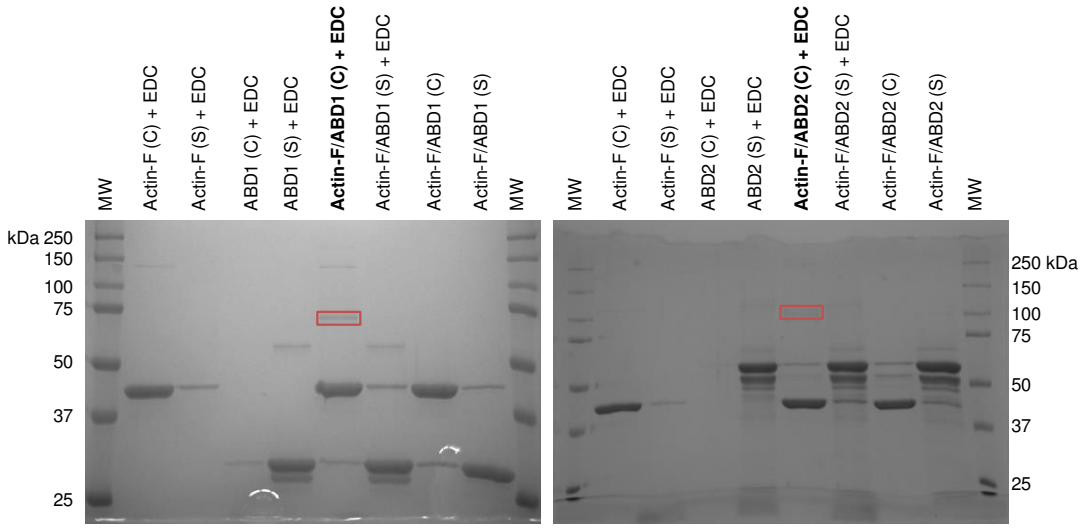


Figure 72 : Analyse SDS-PAGE des expériences de « cross-link » couplées à de la co-sédimentation. Les bandes excisées et analysées par MS sont encadrées en rouge. (C) : culot, (S) : surnageant.

Les bandes d'intérêt ont été excisées pour une analyse par spectrométrie de masse. Dans le cas du complexe ABD1/actine-F, nous avons pu identifier deux zones d'interaction entre les deux partenaires, une analyse plus approfondie des données étant en cours. Pour R11-15/actine-F, deux zones d'interactions ont également pu être mises en évidence. Notons que dans les deux cas, en plus de dipeptides intermoléculaires, des dipeptides intramoléculaires ont également été identifiés. Les données obtenues pour le complexe actine-F/ABD2 sont en adéquation avec des modèles obtenus *in silico* précédemment par A-E. Molza au cours de sa thèse [59]. L'affinage de ces modèles est en cours au laboratoire. Des études *in silico* du complexe actine-F/ABD1 sont également envisagées.

4. Conclusion

Dans ce chapitre, nous avons tout d'abord proposé une méthode alternative pour déterminer un K_d entre l'actine et ses partenaires. En effet, en utilisant la MST, nous proposons une nouvelle approche qui pourrait s'avérer bien moins fastidieuse que les expériences de co-sédimentation. Ensuite, en utilisant la chimie-click couplée à la spectrométrie de masse nous avons pu établir une cartographie précise des interactions dystrophine/actine effectuées *via* les domaines ABD1 et ABD2. Ces données serviront de contraintes expérimentales pour l'élaboration de modèles fiables des complexes ABD1 et ABD2/actine. L'ensemble de ces données permettra une meilleure compréhension des bases moléculaires impliquées dans le soutien du sarcolemme.

Tableau 8 : Tableau récapitulatif des dipeptides identifiés par MS lors des expériences de « cross-link » protéine/protéine. Les résidus impliqués dans la formation de la liaison amide sont en rouge.

Actin-F/ABD1	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To
Intermoléculaire								
	[EDVQKK]	>ABD1	15	20	[EITALAPSTmK]	>Actin	318	328
	[LLDPEDVDTTYPDKK]	>ABD1	212	226	[VAPEEHPTLLTEAPLNPK]	>Actin	98	115
	[LLDPEDVDTTYPDKK]	>ABD1	212	226	[VAPEEHPTLLTEAPLNPK]	>Actin	98	115
Intramoléculaire								
	[EKGSTRVHALNNVNKALR]	>ABD1	66	83	[KTFTKWVNAQFSKFGKQHIENLFSDLQDGR]	>ABD1	20	49
	[DIKEK]	>Actin	213	217	[SYELPDGQVITIGNERFR]	>Actin	241	258
Actin-F/ABD2	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To
Intermoléculaire								
	[KQQLEK]	>R1115	105	110	[AGFAGDDAPR]	>Actin	21	30
	[AGFAGDDAPR]	>Actin	21	30	[KKEELNAVR]	>R1115	464	472
Intramoléculaire								
	[EEIK]	>R1115	387	390	[KQQLEK]	>R1115	105	110
	[EIDR]	>R1115	456	459	[EEIKIK]	>R1115	387	392
	[EIEK]	>R1115	162	165	[KQQLEK]	>R1115	105	110
	[LQESK]	>R1115	13	17	{GSFQKPANFEQR]	>R1115	0	12

DISCUSSION GENERALE ET PERSPECTIVES

La dystrophine est une protéine indispensable au soutien du sarcolemme. Une absence totale ou un déficit en dystrophine conduit à de sévères myopathies actuellement incurables, les dystrophies de Duchenne et de Becker. Comprendre les bases moléculaires des interactions dystrophine/sarcolemme par une approche structure/fonction permet (*i*) de contribuer à une meilleure connaissance du rôle biologique de la dystrophine et (*ii*) d'apporter des informations essentielles au développement de thérapies géniques dédiées au traitement des dystrophinopathies. C'est dans ces perspectives que ce travail de thèse a été initié.

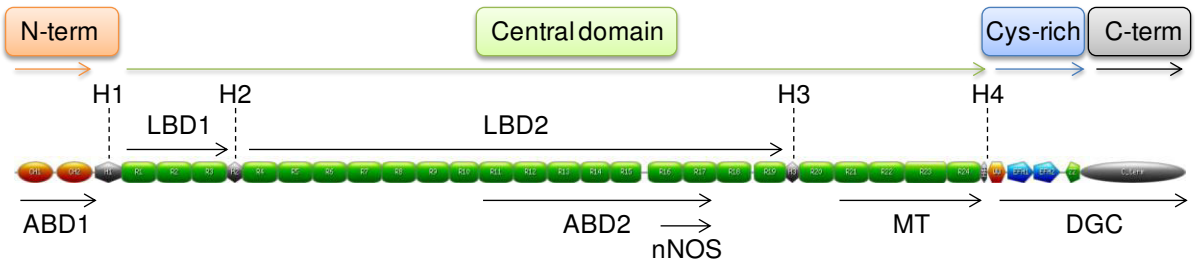


Figure 32 : Représentation schématique de l'organisation de la dystrophine et des interactions avec ses partenaires majeurs. H1-4, charnières (« hinges ») 1-4; ABD, domaine de liaison à l'actine ; LBD : domaine de liaison aux lipides. nNOS : domaine de liaison à la synthase neuronale de l'oxide nitrique MT : domaine de liaison au microtubule, DGC : domaine de liaison au complexe dystrophine et glycoprotéines associées.

Plusieurs techniques classiques pour l'analyse structurale de protéines ne sont pas adaptées à l'étude du, long et flexible, domaine central de la dystrophine. De nouvelles méthodologies ont dû être envisagées afin d'obtenir des informations fondamentales sur la structure du domaine central de la dystrophine. Ce n'est donc que très récemment que des informations structurales ont pu être obtenues. A mon arrivée dans l'équipe « Structures et Interactions Moléculaires » en septembre 2014, des avancées significatives sur la structure du domaine central de la dystrophine étaient émergentes, grâce à la combinaison du SAXS et de la modélisation moléculaire [178]. Les résultats présentés dans l'article 3 (partie IV, chapitre 1) de ce manuscrit démontrent que le long domaine central de la dystrophine ne doit plus être considéré comme un « linker » passif entre le domaine NT en amont et les domaines CR et CT en aval. Suite à ces résultats, il apparaît que la structure longtemps admise [170,241] d'un domaine en bâtonnet, parfois dénommé « rod domain », doit être reconsidérée. Ce domaine présente en effet des régions coudées, signe d'une structure plus complexe que celle rapportée jusqu'à présent et considérée comme rectiligne.

La dystrophine est une protéine membranaire périphérique qui assure un rôle de charpente au sein des cellules musculaires *via* les interactions avec les parties protéiques et lipidiques du sarcolemme. Un point à prendre en considération est que le domaine central de la dystrophine est stable dans un environnement aqueux mais possède, au moins partiellement, des propriétés d'association avec les lipides membranaires. Au début de la thèse, la structure de la dystrophine en présence de membranes demeurait inconnue. Comme pour la protéine seule en solution, il était donc nécessaire de développer des approches et des outils permettant de déterminer d'éventuels changements de conformation de la protéine lors de sa liaison aux lipides membranaires. Nous avons enrichi l'approche SAXS/modélisation moléculaire avec du SANS afin de répondre à cette problématique.

Choisir un modèle membranaire pertinent capable de mimer au mieux le sarcolemme faisait partie des prérequis nécessaires afin d'analyser la structure adoptée par la dystrophine en présence de lipides membranaires. Les bicelles présentent l'avantage de posséder une bicouche lipidique, contrairement aux micelles de détergents, et pour les analyses par modélisation moléculaire, sont de taille compatible avec des temps de calculs accessibles, ce qui n'est pas le cas pour des liposomes. Nous avons dans un premier temps mis en évidence le potentiel de ce modèle membranaire pour l'obtention d'informations structurales de PMP *via* une approche de variation de contraste en SANS (partie III, chapitre 1). Nous faisons l'hypothèse que cette approche peut être étendue à la fois à d'autres PMP et également à des IMP nécessitant un environnement membranaire plus proche des biomembranes que celui apporté par exemple par les micelles de détergents. En effet, la littérature est riche en exemples de complexes IMP/bicelle pour des études par RMN ou XRC [41,54] . Une approche par SANS peut s'avérer être un complément extrêmement utile pour vérifier par exemple le degré de similitude entre les structures cristallines et les structures en solution. Il peut s'agir également d'une alternative pour l'obtention de données structurales sur de grandes protéines non cristallisables. Nous avons montré ensuite que les bicelles sont compatibles avec un système HPLC-SEC dans le cadre d'expériences de DLS, de MALS et de SANS. De plus, les bicelles peuvent être utilisées pour caractériser les interactions protéine/lipide par chimie-click/spectrométrie de masse et MST (partie III, chapitre 2). Le modèle de bicelle était déjà connu pour être un outil polyvalent : nos travaux montrent qu'il peut désormais ajouter d'autres cordes à son arc!

Dans le chapitre 2 de la partie III, nous avons pleinement exploité le potentiel des bicelles afin de caractériser les interactions dystrophine/lipides par des approches de biochimie, de biophysique, de biologie structurale et de modélisation moléculaire. Nous avons démontré que les fragments protéiques R1-3 et R11-15 préservent leur conformation lorsqu'ils sont associés à des bicelles zwitterioniques mais – et il s'agit de la conclusion majeure de ce travail – qu'ils adoptent une structure plus étendue lors de leur liaison à des bicelles anioniques. En couplant SANS, CG-MD et MS nous proposons pour la première fois un modèle tout-atome de R1-3 en interaction avec des phospholipides membranaires. D'après ce modèle, nous proposons qu'une ouverture de la structure en faisceaux d'hélices de la répétition 1 (R1) du domaine central accompagne la liaison à un modèle membranaire à haut degré de courbure. Nous pouvons émettre l'hypothèse que cette ouverture de R1 est nécessaire à l'ancrage de la dystrophine au sarcolemme.

- ***Objectifs à court terme***

Les approches SANS et SAXS ont été menées en parallèle, avec les mêmes objectifs. Nous avons indéniablement eu le plus de succès avec l'approche SANS. Cependant, après avoir testé de nombreuses conditions expérimentales (partie III, chapitre 1), nous pensons être tout proches d'obtenir des données SAXS de bonne qualité pour des complexes entiers protéine/bicelle. Ceci grâce notamment aux différentes mises au point réalisées sur SWING et, plus récemment, grâce à l'apport des données MST. Ces données nous permettront de travailler dans des conditions expérimentales dans lesquelles uniquement des complexes protéine/bicelle sont présents et permettraient donc de nous passer *a priori* d'un système HPLC. De plus, un projet consistant en l'extension du logiciel Memprot [127] aux bicelles est actuellement en cours de développement. Ce logiciel est pour l'instant limité aux traitements de données SAXS de PDC.

Dans le but de les comparer ensuite aux données expérimentales de SAXS, nous sommes d'ores et déjà en mesure de proposer les modèles tout-atome préliminaires obtenus par CG-MD pour des complexes entiers protéine/bicelle (Figure 73). La conversion en modèle tout-atome du complexe R1-3/bicelle entier est effectuée selon des protocoles similaires à ceux décrits dans l'article 2. Ce résultat est encore considéré comme préliminaire, car il nécessiterait une simulation tout-atome afin d'obtenir une bicelle équilibrée. Cependant, l'obtention de modèles tout-atome de complexes entiers, constitués de fragments de dystrophine/bicelles en adéquation

avec les données expérimentales SAXS, est un objectif qui devrait être atteint à relativement court terme.

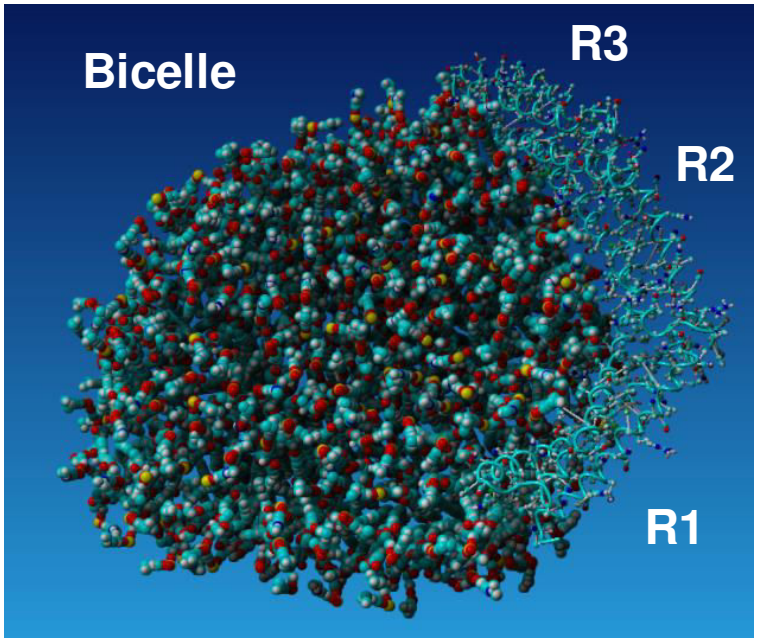


Figure 73 : Modèle tout-atome préliminaire du complexe R1-3/HZB.

En parallèle de la thématique principale de cette thèse, qui concerne la conformation de fragments de dystrophine en interaction avec des lipides membranaires, nous nous sommes intéressés au degré d'enfouissement des fragments R1-3 et R11-15 au sein de la membrane plasmique. Nous avons eu pour cela recours à la technique de réflectivité des neutrons, en utilisant comme modèles membranaires des monocouches à l'interface air/eau et des bicouches supportées. Les données sont actuellement en cours de traitement. Néanmoins, les résultats préliminaires nous permettent d'affirmer que les différences observées sur les membranes en absence ou en présence de fragments protéiques présagent de l'obtention d'importantes informations sur le mode d'interaction dystrophine/membrane. En effet, ces premiers résultats suggèrent que R1-3 et R11-15 interagissent différemment avec des monocouches de Langmuir de type PC. De plus, d'après les données obtenues en SLB sur le système R1-3, nous pouvons d'ores et déjà conclure que R1-3 ne s'insère pas dans la partie hydrophobe de la membrane, en accord avec le modèle de bicelle. Nous nous attelons désormais à déterminer le mode d'interaction de R1-3 avec les SLB de manière plus fine. Nous avons comme perspectives d'étendre l'approche NR en SLB à R11-15 et également pour projet de coupler l'ensemble de ces données avec une analyse CG-MD en collaboration avec Alexandros Koutsioumpas.

Les fragments R1-3 et R11-15 étaient les systèmes d'étude de cette thèse. Nous nous sommes plus focalisés sur R1-3 que sur R11-15, le fragment protéique le plus petit s'avérant plus facile à étudier par certaines approches. Cependant, nous avons également obtenu des modèles de complexes R11-15/HAB en CG-MD, modèles en adéquation avec les données expérimentales de chimie-click/MS (partie III, chapitre 2). Une approche CG-IMD sous contrainte des données SANS est envisagée pour ce complexe afin de proposer un modèle tout-atome de R11-15 en interaction avec les lipides membranaires. L'intérêt majeur de R11-15 réside dans le fait qu'il peut interagir simultanément avec l'actine-F du cytosquelette et les phospholipides du sarcolemme [183]. Ainsi, d'autres expériences actuellement en cours de préparation, consisteront à réaliser un double « cross-link » R11-15/bicelle (pacFA PC) et R11-15/actine-F (EDC). De cette façon, les zones d'interaction du complexe tripartite pourront être clairement définies. Ces informations contribueront fortement au design de modèles actine/R11-15/bicelle sous contraintes expérimentales.

Dans le même ordre d'idée, pour comprendre le fonctionnement des interactions dystrophine-actine-membrane à une échelle supérieure, inclure dans les projets un fragment protéique constitué de l'ABD1-hinge1-R1-3 serait d'un grand intérêt. Pendant cette thèse, nous avons réussi à obtenir ce fragment protéique, avec cependant des produits instables et des rendements de purification incompatibles avec les besoins des expériences de SAS. La révision des systèmes de production, de purification et de conservation est en cours. L'équipe est désormais capable de produire des fragments du domaine central de la dystrophine constitués d'un nombre de répétitions « coiled-coil » élevé (~10 répétitions) (partie IV, chapitre 1). Notamment, l'équipe teste la possibilité de produire ABD1-R1-15 en cellules d'insectes. La présence des deux sites natifs ABD sur la même protéine recombinante devrait permettre d'affiner les connaissances du comportement de la dystrophine vis-à-vis de son partenaire cystosquelettique [182].

- ***Objectifs à moyen/long terme***

Toutes les expériences présentées dans ce manuscrit de thèse ont été effectuées *in vitro*. Bien que pertinentes pour comprendre les bases moléculaires régissant les interactions dystrophine/sarcolemme, les conclusions de ces expériences doivent pouvoir être étendues à une échelle cellulaire. Dans l'article 2 (partie III, chapitre 2), nous avons démontré la faisabilité de l'approche chimie-click/spectrométrie de masse pour cartographier les interactions dystrophine/lipides Par conséquent, un projet en cours au sein de l'équipe consiste à établir une

cartographie des liaisons dystrophine entière/sarcolemme *in vivo* dans un modèle cellulaire. L'approche principale est l'analyse par spectrométrie de masse de produits de pontages covalents entre la dystrophine et les lipides du sarcolemme. Des expériences préliminaires d'immunofluorescence effectuées sur des cellules musculaires murines en culture montrent que la dystrophine est exprimée et localisée à la membrane (Figure 74) une dizaine de jours après différenciation cellulaire.

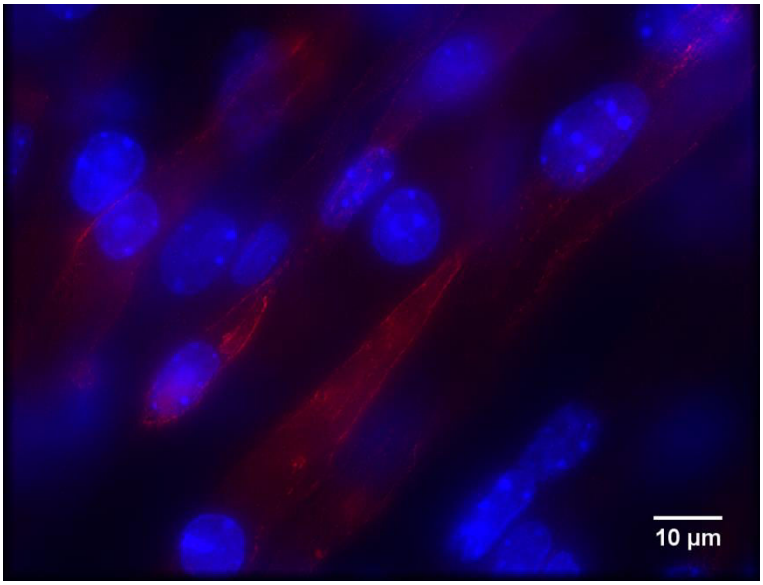


Figure 74 : Immunofluorescence de cellules musculaires murines. Les noyaux sont colorés au DAPI (bleu). Les anticorps anti-dystrophine sont représentés en rouge et attestent de la localisation de la dystrophine au sarcolemme.

Les premières expériences ont consisté en la recherche de PTM de type palmitoylation, la dystrophine possédant deux sites putatifs. Il apparaît que les peptides identifiés par spectrométrie de masse ne couvrent que 5% de la dystrophine et, parmi ceux-ci, aucun n'est porteur de ce type de PTM. La faible couverture de séquence est vraisemblablement due aux limites quantitatives de la méthode employée et à la faible abondance de la dystrophine dans les cellules (~ 0.002% [148]). Les expériences visant à utiliser des quantités suffisantes de matériel biologique pour une couverture maximale de la protéine en spectrométrie de masse sont en cours. Pour la cartographie *in vivo*, l'acide gras pacFA sera utilisé à la place de la forme phospholipidique (pacFA PC) utilisé pour les bicelles. Haberkant *et al.* [207] ont démontré en effet que le pacFA est métabolisable en phospholipides par des cellules en culture. L'identification avérée de peptides de dystrophine porteurs de pacFA représenterait une preuve expérimentale que la dystrophine interagit directement avec les lipides du sarcolemme. Ces expériences permettraient en outre l'obtention d'une cartographie précise de ces interactions.

En résumé, nous pensons que les résultats obtenus lors de ces travaux de thèse :

(i) **contribuent à une meilleure compréhension du rôle biologique de la dystrophine en vue de thérapies dédiées aux patients DMD et BMD**

(ii) **constituent un apport méthodologique significatif pour l'étude de protéines membranaires.**

REFERENCES BIBLIOGRAPHIQUES

(1) Feijó Delgado, F.; Cermak, N.; Hecht, V. C.; Son, S.; Li, Y.; Knudsen, S. M.; Olcum, S.; Higgins, J. M.; Chen, J.; Grover, W. H.; et al. Intracellular Water Exchange for Measuring the Dry Mass, Water Mass and Changes in Chemical Composition of Living Cells. *PloS One* **2013**, *8* (7), e67590.

(2) Protein Function | Learn Science at Scitable https://www.nature.com/scitable/topicpage/protein-function-14123348

(3) Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 6th ed.; Garland Science: New York, NY, 2014.

(4) Callen, J.-C.; Perasso, R.; Mounolou, J.-C. *Biologie cellulaire : Des molécules aux organismes, Cours, questions de révision et QROC*, 2e édition.; Dunod: Paris, 2005.

(5) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane Lipids: Where They Are and How They Behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (2), 112–124.

(6) Fiehn, W.; Peter, J. B.; Mead, J. F.; Gan-Elepano, M. Lipids and Fatty Acids of Sarcolemma, Sarcoplasmic Reticulum, and Mitochondria from Rat Skeletal Muscle. *J. Biol. Chem.* **1971**, *246* (18), 5617–5620.

(7) Leventis, P. A.; Grinstein, S. The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys.* **2010**, *39* (1), 407–427.

(8) Singer, S. J.; Nicolson, G. L. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* **1972**, *175* (4023), 720–731.

(9) Nicolson, G. L. The Fluid—Mosaic Model of Membrane Structure: Still Relevant to Understanding the Structure, Function and Dynamics of Biological Membranes after More than 40 Years. *Biochim. Biophys. Acta BBA - Biomembr.* **2014**, *1838* (6), 1451–1466.

(10) H. Lorent, J.; Quetin-Leclercq, J.; Mingeot-Leclercq, M.-P. The Amphiphilic Nature of Saponins and Their Effects on Artificial and Biological Membranes and Potential Consequences for Red Blood and Cancer Cells. *Org. Biomol. Chem.* **2014**, *12* (44), 8803–8822.

(11) Goñi, F. M. The Basic Structure and Dynamics of Cell Membranes: An Update of the Singer-Nicolson Model. *Biochim. Biophys. Acta* **2014**, *1838* (6), 1467–1476.

(12) Dowhan, W. Molecular Basis for Membrane Phospholipid Diversity: Why Are There so Many Lipids? *Annu. Rev. Biochem.* **1997**, *66*, 199–232.

(13) Phase Transition Temperatures for Glycerophospholipids | Avanti Polar Lipids https://avantilipids.com/tech-support/physical-properties/phase-transition-temps

(14) Luzzati, V.; Gulik-Krzywicki, T.; Tardieu, A. Polymorphism of Lecithins. *Nature* **1968**, *218* (5146), 1031–1034.

(15) Brown, M. F. Curvature Forces in Membrane Lipid–Protein Interactions. *Biochemistry (Mosc.)* **2012**, *51* (49), 9782–9795.

(16) Brown, M. F. Soft Matter in Lipid–Protein Interactions. *Annu. Rev. Biophys.* **2017**, *46* (1), 379–410.

(17) Israelachvili, J. N. *Intermolecular and Surface Forces*; Academic Press: London, 1992.

(18) Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. Theory of Self-Assembly of Lipid Bilayers and Vesicles. *Biochim. Biophys. Acta BBA - Biomembr.* **1977**, *470* (2), 185–201.

(19) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; Wiley, 1980.

(20) Akbarzadeh, A.; Rezaei-Sadabady, R.; Davaran, S.; Joo, S. W.; Zarghami, N.; Hanifehpour, Y.; Samiei, M.; Kouhi, M.; Nejati-Koshki, K. Liposome: Classification, Preparation, and Applications. *Nanoscale Res. Lett.* **2013**, *8* (1), 102.
(21) Walde, P.; Cosentino, K.; Engel, H.; Stano, P. Giant Vesicles: Preparations and Applications. *Chembiochem Eur. J. Chem. Biol.* **2010**, *11* (7), 848–865.
(22) Moussa, H. G.; Martins, A. M.; Husseini, G. A. Review on Triggered Liposomal Drug Delivery with a Focus on Ultrasound. *Curr. Cancer Drug Targets* **2015**, *15* (4), 282–313.
(23) Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbus, L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLOS ONE* **2013**, *8* (5), e62488.
(24) Seddon, A. M.; Curnow, P.; Booth, P. J. Membrane Proteins, Lipids and Detergents: Not Just a Soap Opera. *Biochim. Biophys. Acta* **2004**, *1666* (1–2), 105–117.
(25) Frey, L.; Lakomek, N.-A.; Riek, R.; Bibow, S. Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics. *Angew. Chem. Int. Ed Engl.* **2017**, *56* (1), 380–383.
(26) Tribet, C.; Audebert, R.; Popot, J. L. Amphipols: Polymers That Keep Membrane Proteins Soluble in Aqueous Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (26), 15047–15050.
(27) Zoonens, M.; Popot, J.-L. Amphipols for Each Season. *J. Membr. Biol.* **2014**, *247* (9–10), 759–796.
(28) 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.
(29) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2* (8), 853–856.
(30) Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y.-P.; Dafforn, T.; Overduin, M. Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *J. Am. Chem. Soc.* **2009**, *131* (22), 7484–7485.
(31) Bayburt, T. H.; Sligar, S. G. Membrane Protein Assembly into Nanodiscs. *FEBS Lett.* **2010**, *584* (9), 1721–1727.
(32) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J. Am. Chem. Soc.* **2004**, *126* (11), 3477–3487.
(33) Dörr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schäfer, M.; van Walree, C. A.; Killian, J. A. The Styrene-Maleic Acid Copolymer: A Versatile Tool in Membrane Research. *Eur. Biophys. J. EBJ* **2016**, *45* (1), 3–21.
(34) Eeman, M.; Deleu, M. From biological membranes to biomimetic model membranes. *BASE* **2010**.
(35) Richter, R. P.; Bérat, R.; Brisson, A. R. Formation of Solid-Supported Lipid Bilayers: An Integrated View. *Langmuir* **2006**, *22* (8), 3497–3505.
(36) Clifton, L. A.; Neylon, C.; Lakey, J. H. Examining Protein-Lipid Complexes Using Neutron Scattering. *Methods Mol. Biol. Clifton NJ* **2013**, *974*, 119–150.
(37) Lombard, J. Once upon a Time the Cell Membranes: 175 Years of Cell Boundary Research. *Biol. Direct* **2014**, *9*, 32.
(38) Raffard, G.; Steinbruckner, S.; Arnold, A.; Davis, J. H.; Dufourc, E. J. Temperature−Composition Diagram of Dimyristoylphosphatidylcholine−Dicaproylphosphatidylcholine “Bicelles” Self-

Orienting in the Magnetic Field. A Solid State 2H and 31P NMR Study. *Langmuir* **2000**, *16* (20), 7655–7662.

(39) van Dam, L.; Karlsson, G.; Edwards, K. Direct Observation and Characterization of DMPC/DHPC Aggregates under Conditions Relevant for Biological Solution NMR. *Biochim. Biophys. Acta* **2004**, *1664* (2), 241–256.

(40) Harroun, T. A.; Koslowsky, M.; Nieh, M.-P.; de Lannoy, C.-F.; Raghunathan, V. A.; Katsaras, J. Comprehensive Examination of Mesophases Formed by DMPC and DHPC Mixtures. *Langmuir* **2005**, *21* (12), 5356–5361.

(41) Dürr, U. H. N.; Gildenberg, M.; Ramamoorthy, A. The Magic of Bicelles Lights up Membrane Protein Structure. *Chem. Rev.* **2012**, *112* (11), 6054–6074.

(42) Dürr, U. H. N.; Soong, R.; Ramamoorthy, A. When Detergent Meets Bilayer: Birth and Coming of Age of Lipid Bicelles. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *69*, 1–22.

(43) Sanders, C. R.; Landis, G. C. Reconstitution of Membrane Proteins into Lipid-Rich Bilayered Mixed Micelles for NMR Studies. *Biochemistry (Mosc.)* **1995**, *34* (12), 4030–4040.

(44) Small, D. M.; Penkett, S. A.; Chapman, D. Studies on Simple and Mixed Bile Salt Micelles by Nuclear Magnetic Resonance Spectroscopy. *Biochim. Biophys. Acta* **1969**, *176* (1), 178–189.

(45) Mazer, N. A.; Benedek, G. B.; Carey, M. C. Quasielastic Light-Scattering Studies of Aqueous Biliary Lipid Systems. Mixed Micelle Formation in Bile Salt-Lecithin Solutions. *Biochemistry (Mosc.)* **1980**, *19* (4), 601–615.

(46) Sanders, C. R.; Prestegard, J. H. Magnetically Orientable Phospholipid Bilayers Containing Small Amounts of a Bile Salt Analogue, CHAPSO. *Biophys. J.* **1990**, *58* (2), 447–460.

(47) Bian, J.; Roberts, M. F. Phase Separation in Short-Chain Lecithin/Gel-State Long Chain Lecithin Aggregates. *Biochemistry (Mosc.)* **1990**, *29* (34), 7928–7935.

(48) Lin, T. L.; Liu, C. C.; Roberts, M. F.; Chen, S. H. Structure of Mixed Short-Chain Lecithin/Long-Chain Lecithin Aggregates Studied by Small-Angle Neutron Scattering. *J. Phys. Chem.* **1991**, *95* (15), 6020–6027.

(49) Sanders, C. R.; Schwonek, J. P. Characterization of Magnetically Orientable Bilayers in Mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphosphatidylcholine by Solid-State NMR. *Biochemistry (Mosc.)* **1992**, *31* (37), 8898–8905.

(50) Prosser, R. S.; Hwang, J. S.; Vold, R. R. Magnetically Aligned Phospholipid Bilayers with Positive Ordering: A New Model Membrane System. *Biophys. J.* **1998**, *74* (5), 2405–2418.

(51) Loudet, C.; Manet, S.; Gineste, S.; Oda, R.; Achard, M.-F.; Dufourc, E. J. Biphenyl Bicelle Disks Align Perpendicular to Magnetic Fields on Large Temperature Scales: A Study Combining Synthesis, Solid-State NMR, TEM, and SAXS. *Biophys. J.* **2007**, *92* (11), 3949–3959.

(52) Loudet-Courreges, C.; Nallet, F.; Dufourc, E. J.; Oda, R. Unprecedented Observation of Days-Long Remnant Orientation of Phospholipid Bicelles: A Small-Angle X-Ray Scattering and Theoretical Study. *Langmuir ACS J. Surf. Colloids* **2011**, *27* (15), 9122–9130.

(53) van Dam, L.; Karlsson, G.; Edwards, K. Morphology of Magnetically Aligning DMPC/DHPC Aggregates-Perforated Sheets, Not Disks. *Langmuir ACS J. Surf. Colloids* **2006**, *22* (7), 3280–3285.

(54) Poulos, S.; Morgan, J. L. W.; Zimmer, J.; Faham, S. Bicelles Coming of Age: An Empirical Approach to Bicelle Crystallization. *Methods Enzymol.* **2015**, *557*, 393–416.

(55) Ujwal, R.; Abramson, J. High-Throughput Crystallization of Membrane Proteins Using the Lipidic Bicelle Method. *J. Vis. Exp. JoVE* **2012**, No. 59, e3383.

(56) Ujwal, R.; Bowie, J. U. Crystallizing Membrane Proteins Using Lipidic Bicelles. *Methods San Diego Calif* **2011**, *55* (4), 337–341.
(57) Ye, W.; Lind, J.; Eriksson, J.; Mäler, L. Characterization of the Morphology of Fast-Tumbling Bicelles with Varying Composition. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (19), 5488–5496.
(58) Mineev, K. S.; Nadezhdin, K. D.; Goncharuk, S. A.; Arseniev, A. S. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir* **2016**, *32* (26), 6624–6637.
(59) Beaugrand, M.; Arnold, A. A.; Hénin, J.; Warschawski, D. E.; Williamson, P. T. F.; Marcotte, I. Lipid Concentration and Molar Ratio Boundaries for the Use of Isotropic Bicelles. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (21), 6162–6170.
(60) Jiang, Y.; Wang, H.; Kindt, J. T. Atomistic Simulations of Bicelle Mixtures. *Biophys. J.* **2010**, *98* (12), 2895–2903.
(61) Vácha, R.; Frenkel, D. Stability of Bicelles: A Simulation Study. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (15), 4229–4235.
(62) Vestergaard, M.; Kraft, J. F.; Vosegaard, T.; Thøgersen, L.; Schiøtt, B. Bicelles and Other Membrane Mimics: Comparison of Structure, Properties, and Dynamics from MD Simulations. *J. Phys. Chem. B* **2015**, *119* (52), 15831–15843.
(63) Yang, P.-W.; Lin, T.-L.; Hu, Y.; Jeng, U.-S. Small-Angle X-Ray Scattering Studies on the Structure of Mixed DPPC/DiC7PC Micelles in Aqueous Solutions. *Chin. J. Phys.* **2012**, *50* (2), 349–356.
(64) Yang, P.-W.; Lin, T.-L.; Hu, Y.; Jeng, U.-S. Formation of Divalent Ion Mediated Anionic Disc Bicelle-DNA Complexes. *Soft Matter* **2014**, *10* (13), 2313–2319.
(65) Luchette, P. A.; Vetman, T. N.; Prosser, R. S.; Hancock, R. E.; Nieh, M. P.; Glinka, C. J.; Krueger, S.; Katsaras, J. Morphology of Fast-Tumbling Bicelles: A Small Angle Neutron Scattering and NMR Study. *Biochim. Biophys. Acta* **2001**, *1513* (2), 83–94.
(66) Nieh, M.-P.; Glinka, C. J.; Krueger, S.; Prosser, R. S.; Katsaras, J. SANS Study on the Effect of Lanthanide Ions and Charged Lipids on the Morphology of Phospholipid Mixtures. Small-Angle Neutron Scattering. *Biophys. J.* **2002**, *82* (5), 2487–2498.
(67) Nieh, M.-P.; Raghunathan, V. A.; Glinka, C. J.; Harroun, T. A.; Pabst, G.; Katsaras, J. Magnetically Alignable Phase of Phospholipid "Bicelle" Mixtures Is a Chiral Nematic Made Up of Wormlike Micelles. *Langmuir* **2004**, *20* (19), 7893–7897.
(68) Li, M.; Morales, H. H.; Katsaras, J.; Kučerka, N.; Yang, Y.; Macdonald, P. M.; Nieh, M.-P. Morphological Characterization of DMPC/CHAPSO Bicellar Mixtures: A Combined SANS and NMR Study. *Langmuir ACS J. Surf. Colloids* **2013**, *29* (51), 15943–15957.
(69) Nieh, M.-P.; Raghunathan, V. A.; Pabst, G.; Harroun, T.; Nagashima, K.; Morales, H.; Katsaras, J.; Macdonald, P. Temperature Driven Annealing of Perforations in Bicellar Model Membranes. *Langmuir ACS J. Surf. Colloids* **2011**, *27* (8), 4838–4847.
(70) Minto, R. E.; Adhikari, P. R.; Lorigan, G. A. A 2H Solid-State NMR Spectroscopic Investigation of Biomimetic Bicelles Containing Cholesterol and Polyunsaturated Phosphatidylcholine. *Chem. Phys. Lipids* **2004**, *132* (1), 55–64.
(71) Struppe, J.; Whiles, J. A.; Vold, R. R. Acidic Phospholipid Bicelles: A Versatile Model Membrane System. *Biophys. J.* **2000**, *78* (1), 281–289.
(72) Yamaguchi, T.; Suzuki, T.; Yasuda, T.; Oishi, T.; Matsumori, N.; Murata, M. NMR-Based Conformational Analysis of Sphingomyelin in Bicelles. *Bioorg. Med. Chem.* **2012**, *20* (1), 270–278.
(73) Liebau, J.; Pettersson, P.; Zuber, P.; Ariöz, C.; Mäler, L. Fast-Tumbling Bicelles Constructed from Native Escherichia Coli Lipids. *Biochim. Biophys. Acta* **2016**, *1858* (9), 2097–2105.

(74) Beaugrand, M.; Arnold, A. A.; Juneau, A.; Gambaro, A. B.; Warschawski, D. E.; Williamson, P. T. F.; Marcotte, I. Magnetically Oriented Bicelles with Monoalkylphosphocholines: Versatile Membrane Mimetics for Nuclear Magnetic Resonance Applications. *Langmuir ACS J. Surf. Colloids* **2016**, *32* (49), 13244–13251.
(75) Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R. Structural Evaluation of Phospholipid Bicelles for Solution-State Studies of Membrane-Associated Biomolecules. *Biophys. J.* **2001**, *81* (4), 2163–2171.
(76) Tausk, R. J.; van Esch, J.; Karmiggelt, J.; Voordouw, G.; Overbeek, J. T. Physical Chemical Studies of Short-Chain Lecithin Homologues. II. Micellar Weights of Dihexanoyl- and Diheptanoyllecithin. *Biophys. Chem.* **1974**, *1* (3), 184–203.
(77) Draney, A. W.; Smrt, S. T.; Lorieau, J. L. Use of Isotropically Tumbling Bicelles to Measure Curvature Induced by Membrane Components. *Langmuir ACS J. Surf. Colloids* **2014**, *30* (39), 11723–11733.
(78) Critical Micelle Concentrations (CMCs) | Avanti Polar Lipids https://avantilipids.com/tech-support/physical-properties/cmcs
(79) Vold, R. R.; Prosser, R. S. Magnetically Oriented Phospholipid Bilayered Micelles for Structural Studies of Polypeptides. Does the Ideal Bicelle Exist? *J. Magn. Reson. B* **1996**, *113* (3), 267–271.
(80) Triba, M. N.; Warschawski, D. E.; Devaux, P. F. Reinvestigation by Phosphorus NMR of Lipid Distribution in Bicelles. *Biophys. J.* **2005**, *88* (3), 1887–1901.
(81) Situ, A. J.; Schmidt, T.; Mazumder, P.; Ulmer, T. S. Characterization of Membrane Protein Interactions by Isothermal Titration Calorimetry. *J. Mol. Biol.* **2014**, *426* (21), 3670–3680.
(82) Mineev, K. S.; Nadezhdin, K. D.; Goncharuk, S. A.; Arseniev, A. S. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir ACS J. Surf. Colloids* **2016**.
(83) Träuble, H.; Haynes, D. H. The Volume Change in Lipid Bilayer Lamellae at the Crystalline-Liquid Crystalline Phase Transition. *Chem. Phys. Lipids* **1971**, *7* (4), 324–335.
(84) Petrache, H. I.; Tristram-Nagle, S.; Gawrisch, K.; Harries, D.; Parsegian, V. A.; Nagle, J. F. Structure and Fluctuations of Charged Phosphatidylserine Bilayers in the Absence of Salt. *Biophys. J.* **2004**, *86* (3), 1574–1586.
(85) Lee, A. G. How Lipids Affect the Activities of Integral Membrane Proteins. *Biochim. Biophys. Acta* **2004**, *1666* (1–2), 62–87.
(86) Lee, A. G. How Lipids Interact with an Intrinsic Membrane Protein: The Case of the Calcium Pump. *Biochim. Biophys. Acta* **1998**, *1376* (3), 381–390.
(87) Pilot, J. D.; East, J. M.; Lee, A. G. Effects of Bilayer Thickness on the Activity of Diacylglycerol Kinase of Escherichia Coli. *Biochemistry (Mosc.)* **2001**, *40* (28), 8188–8195.
(88) Newton, A. C. Interaction of Proteins with Lipid Headgroups: Lessons from Protein Kinase C. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 1–25.
(89) Gibson, N. J.; Brown, M. F. Lipid Headgroup and Acyl Chain Composition Modulate the MI-MII Equilibrium of Rhodopsin in Recombinant Membranes. *Biochemistry (Mosc.)* **1993**, *32* (9), 2438–2454.
(90) Veit, M.; Serebryakova, M. V.; Kordyukova, L. V. Palmitoylation of Influenza Virus Proteins. *Biochem. Soc. Trans.* **2013**, *41* (1), 50–55.
(91) Nosjean, O.; Briolay, A.; Roux, B. Mammalian GPI Proteins: Sorting, Membrane Residence and Functions. *Biochim. Biophys. Acta* **1997**, *1331* (2), 153–186.

(92) van Genderen, H. O.; Kenis, H.; Hofstra, L.; Narula, J.; Reutelingsperger, C. P. M. Extracellular Annexin A5: Functions of Phosphatidylserine-Binding and Two-Dimensional Crystallization. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2008**, *1783* (6), 953–963.
(93) Vermes, I.; Haanen, C.; Steffens-Nakken, H.; Reutelingsperger, C. A Novel Assay for Apoptosis. Flow Cytometric Detection of Phosphatidylserine Expression on Early Apoptotic Cells Using Fluorescein Labelled Annexin V. *J. Immunol. Methods* **1995**, *184* (1), 39–51.
(94) Antonny, B. Mechanisms of Membrane Curvature Sensing. *Annu. Rev. Biochem.* **2011**, *80*, 101–123.
(95) Antonny, B.; Huber, I.; Paris, S.; Chabre, M.; Cassel, D. Activation of ADP-Ribosylation Factor 1 GTPase-Activating Protein by Phosphatidylcholine-Derived Diacylglycerols. *J. Biol. Chem.* **1997**, *272* (49), 30848–30851.
(96) Vanni, S.; Vamparys, L.; Gautier, R.; Drin, G.; Etchebest, C.; Fuchs, P. F. J.; Antonny, B. Amphipathic Lipid Packing Sensor Motifs: Probing Bilayer Defects with Hydrophobic Residues. *Biophys. J.* **2013**, *104* (3), 575–584.
(97) Snead, D.; Eliezer, D. Alpha-Synuclein Function and Dysfunction on Cellular Membranes. *Exp. Neurobiol.* **2014**, *23* (4), 292–313.
(98) Ulmer, T. S.; Bax, A.; Cole, N. B.; Nussbaum, R. L. Structure and Dynamics of Micelle-Bound Human Alpha-Synuclein. *J. Biol. Chem.* **2005**, *280* (10), 9595–9603.
(99) Jao, C. C.; Hegde, B. G.; Chen, J.; Haworth, I. S.; Langen, R. Structure of Membrane-Bound α-Synuclein from Site-Directed Spin Labeling and Computational Refinement. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (50), 19666–19671.
(100) 7.4: Proteins - Biology LibreTexts https://bio.libretexts.org/TextMaps/Map%3A_Microbiology_(OpenStax)/07%3A_Microbial_Biochemistry/7.4%3A_Proteins (accessed Aug 29, 2017).
(101) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinforma. Oxf. Engl.* **2006**, *22* (5), 623–625.
(102) Fagerberg, L.; Jonasson, K.; von Heijne, G.; Uhlén, M.; Berglund, L. Prediction of the Human Membrane Proteome. *Proteomics* **2010**, *10* (6), 1141–1149.
(103) Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A. M.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; et al. Mass-Spectrometry-Based Draft of the Human Proteome. *Nature* **2014**, *509* (7502), 582–587.
(104) Membrane Proteins of Known Structure http://blanco.biomol.uci.edu/mpstruc
(105) White, S. H. Biophysical Dissection of Membrane Proteins. *Nature* **2009**, *459* (7245), 344–346.
(106) Hendrickson, W. A. Atomic-Level Analysis of Membrane Protein Structure. *Nat. Struct. Mol. Biol.* **2016**, *23* (6), 464–467.
(107) Henderson, R.; Unwin, P. N. T. Three-Dimensional Model of Purple Membrane Obtained by Electron Microscopy. *Nature* **1975**, *257* (5521), 28–32.
(108) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. Structure of the Protein Subunits in the Photosynthetic Reaction Centre of Rhodopseudomonas Viridis at 3|[Angst]| Resolution. *Nature* **1985**, *318* (6047), 618–624.
(109) Fernandez-Leiro, R.; Scheres, S. H. W. Unravelling Biological Macromolecules with Cryo-Electron Microscopy. *Nature* **2016**, *537* (7620), 339–346.
(110) Wisedchaisri, G.; Reichow, S. L.; Gonen, T. Advances in Structural and Functional Analysis of Membrane Proteins by Electron Crystallography. *Struct. Lond. Engl. 1993* **2011**, *19* (10), 1381–1393.
(111) Smith, S. M. Strategies for the Purification of Membrane Proteins. *Methods Mol. Biol. Clifton NJ* **2011**, *681*, 485–496.

(112) Shi, D.; Nannenga, B. L.; Iadanza, M. G.; Gonen, T. Three-Dimensional Electron Crystallography of Protein Microcrystals. *eLife* **2013**, *2*, e01345.
(113) Neutze, R.; Brändén, G.; Schertler, G. F. X. Membrane Protein Structural Biology Using X-Ray Free Electron Lasers. *Curr. Opin. Struct. Biol.* **2015**, *33*, 115–125.
(114) Kaptein, R.; Wagner, G. NMR Studies of Membrane Proteins. *J. Biomol. NMR* **2015**, *61* (0), 181–184.
(115) Grassucci, R. A.; Taylor, D. J.; Frank, J. Preparation of Macromolecular Complexes for Cryo-Electron Microscopy. *Nat. Protoc.* **2007**, *2* (12), 3239–3246.
(116) Merk, A.; Bartesaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M. I.; Pragani, R.; Boxer, M. B.; Earl, L. A.; Milne, J. L. S.; et al. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **2016**, *165* (7), 1698–1707.
(117) Chavent, M.; Duncan, A. L.; Sansom, M. S. Molecular Dynamics Simulations of Membrane Proteins and Their Interactions: From Nanoscale to Mesoscale. *Curr. Opin. Struct. Biol.* **2016**, *40*, 8–16.
(118) MemProtMD http://memprotmd.bioch.ox.ac.uk
(119) Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, S.; Sansom, M. S. P. MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Struct. Lond. Engl. 1993* **2015**, *23* (7), 1350–1361.
(120) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. A.; Mosberg, H. I. Positioning of Proteins in Membranes: A Computational Approach. *Protein Sci. Publ. Protein Soc.* **2006**, *15* (6), 1318–1333.
(121) Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. Anisotropic Solvent Model of the Lipid Bilayer. 2. Energetics of Insertion of Small Molecules, Peptides, and Proteins in Membranes. *J. Chem. Inf. Model.* **2011**, *51* (4), 930–946.
(122) Petoukhov, M. V.; Svergun, D. I. Analysis of X-Ray and Neutron Scattering from Biomacromolecular Solutions. *Curr. Opin. Struct. Biol.* **2007**, *17* (5), 562–571.
(123) Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. A Brief Survey of State-of-the-Art BioSAXS. *Protein Pept. Lett.* **2016**, *23* (3), 217–231.
(124) Svergun, D. I.; Nierhaus, K. H. A Map of Protein-RRNA Distribution in the 70 S Escherichia Coli Ribosome. *J. Biol. Chem.* **2000**, *275* (19), 14432–14439.
(125) Valentini, E.; Kikhney, A. G.; Previtali, G.; Jeffries, C. M.; Svergun, D. I. SASBDB, a Repository for Biological Small-Angle Scattering Data. *Nucleic Acids Res.* **2015**, *43* (Database issue), D357-363.
(126) Berthaud, A.; Manzi, J.; Pérez, J.; Mangenot, S. Modeling Detergent Organization around Aquaporin-0 Using Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2012**, *134* (24), 10080–10088.
(127) Pérez, J.; Koutsioubas, A. Memprot: A Program to Model the Detergent Corona around a Membrane Protein Based on SEC–SAXS Data. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71* (1), 86–93.
(128) Skar-Gislinge, N.; Kynde, S. A. R.; Denisov, I. G.; Ye, X.; Lenov, I.; Sligar, S. G.; Arleth, L. Small-Angle Scattering Determination of the Shape and Localization of Human Cytochrome P450 Embedded in a Phospholipid Nanodisc Environment. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71* (Pt 12), 2412–2421.
(129) Heller, W. T. Small-Angle Neutron Scattering and Contrast Variation: A Powerful Combination for Studying Biological Structures. *Acta Crystallogr. Sect. D* **2010**, *66* (11), 1213–1217.
(130) Breyton, C.; Gabel, F.; Lethier, M.; Flayhan, A.; Durand, G.; Jault, J.-M.; Juillan-Binard, C.; Imbert, L.; Moulin, M.; Ravaud, S.; et al. Small Angle Neutron Scattering for the Study of Solubilised Membrane Proteins. *Eur. Phys. J. E Soft Matter* **2013**, *36* (7), 71.

(131) Hunt, J. F.; McCrea, P. D.; Zaccaï, G.; Engelman, D. M. Assessment of the Aggregation State of Integral Membrane Proteins in Reconstituted Phospholipid Vesicles Using Small Angle Neutron Scattering. *J. Mol. Biol.* **1997**, *273* (5), 1004–1019.
(132) Gabel, F.; Lensink, M. F.; Clantin, B.; Jacob-Dubuisson, F.; Villeret, V.; Ebel, C. Probing the Conformation of FhaC with Small-Angle Neutron Scattering and Molecular Modeling. *Biophys. J.* **2014**, *107* (1), 185–196.
(133) Breyton, C.; Flayhan, A.; Gabel, F.; Lethier, M.; Durand, G.; Boulanger, P.; Chami, M.; Ebel, C. Assessing the Conformational Changes of Pb5, the Receptor-Binding Protein of Phage T5, upon Binding to Its *Escherichia Coli* Receptor FhuA. *J. Biol. Chem.* **2013**, *288* (42), 30763–30772.
(134) Warschawski, D. E.; Arnold, A. A.; Beaugrand, M.; Gravel, A.; Chartrand, É.; Marcotte, I. Choosing Membrane Mimetics for NMR Structural Studies of Transmembrane Proteins. *Biochim. Biophys. Acta* **2011**, *1808* (8), 1957–1974.
(135) Maric, S.; Skar-Gislinge, N.; Midtgaard, S.; Thygesen, M. B.; Schiller, J.; Frielinghaus, H.; Moulin, M.; Haertlein, M.; Forsyth, V. T.; Pomorski, T. G.; et al. Stealth Carriers for Low-Resolution Structure Determination of Membrane Proteins in Solution. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70* (Pt 2), 317–328.
(136) Wang, X.; Mu, Z.; Li, Y.; Bi, Y.; Wang, Y. Smaller Nanodiscs Are Suitable for Studying Protein Lipid Interactions by Solution NMR. *Protein J.* **2015**, *34* (3), 205–211.
(137) Grinkova, Y. V.; Denisov, I. G.; Sligar, S. G. Engineering Extended Membrane Scaffold Proteins for Self-Assembly of Soluble Nanoscale Lipid Bilayers. *Protein Eng. Des. Sel. PEDS* **2010**, *23* (11), 843–848.
(138) Dürr, U. H. N.; Yamamoto, K.; Im, S.-C.; Waskell, L.; Ramamoorthy, A. Solid-State NMR Reveals Structural and Dynamical Properties of a Membrane-Anchored Electron-Carrier Protein, Cytochrome B5. *J. Am. Chem. Soc.* **2007**, *129* (21), 6670–6671.
(139) Dev, J.; Park, D.; Fu, Q.; Chen, J.; Ha, H. J.; Ghantous, F.; Herrmann, T.; Chang, W.; Liu, Z.; Frey, G.; et al. Structural Basis for Membrane Anchoring of HIV-1 Envelope Spike. *Science* **2016**, *353* (6295), 172–175.
(140) Caffrey, M. A Comprehensive Review of the Lipid Cubic Phase or in Meso Method for Crystallizing Membrane and Soluble Proteins and Complexes. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2015**, *71* (Pt 1), 3–18.
(141) Faham, S.; Bowie, J. U. Bicelle Crystallization: A New Method for Crystallizing Membrane Proteins Yields a Monomeric Bacteriorhodopsin Structure11Edited by D. Rees. *J. Mol. Biol.* **2002**, *316* (1), 1–6.
(142) Hosaka, T.; Okazaki, M.; Kimura-Someya, T.; Ishizuka-Katsura, Y.; Ito, K.; Yokoyama, S.; Dodo, K.; Sodeoka, M.; Shirouzu, M. Crystal Structural Characterization Reveals Novel Oligomeric Interactions of Human Voltage-Dependent Anion Channel 1. *Protein Sci. Publ. Protein Soc.* **2017**.
(143) Nakane, T.; Hanashima, S.; Suzuki, M.; Saiki, H.; Hayashi, T.; Kakinouchi, K.; Sugiyama, S.; Kawatake, S.; Matsuoka, S.; Matsumori, N.; et al. Membrane Protein Structure Determination by SAD, SIR, or SIRAS Phasing in Serial Femtosecond Crystallography Using an Iododetergent. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (46), 13039–13044.
(144) Efremov, R. G.; Gatsogiannis, C.; Raunser, S. Lipid Nanodiscs as a Tool for High-Resolution Structure Determination of Membrane Proteins by Single-Particle Cryo-EM. *Methods Enzymol.* **2017**, *594*, 1–30.
(145) The Editors of Encyclopædia Britannica. skeletal muscle | anatomy https://www.britannica.com/science/skeletal-muscle.

(146) Davies, R. E. .; Newsom-Davis, J. M. muscle https://www.britannica.com/science/muscle/Tetrapod-musculature.
(147) Byers, T. J.; Kunkel, L. M.; Watkins, S. C. The Subcellular Distribution of Dystrophin in Mouse Skeletal, Cardiac, and Smooth Muscle. *J. Cell Biol.* **1991**, *115* (2), 411–421.
(148) Hoffman, E. P.; Brown, R. H.; Kunkel, L. M. Dystrophin: The Protein Product of the Duchenne Muscular Dystrophy Locus. *Cell* **1987**, *51* (6), 919–928.
(149) Flanigan, K. M.; Dunn, D.; von Niederhausern, A.; Soltanzadeh, P.; Gappmaier, E.; Howard, M. T.; Sampson, J.; Mendell, J.; Wall, C.; King, W.; et al. Mutational Spectrum of DMD Mutations in Dystrophinopathy Patients: Application of Modern Diagnostic Techniques to a Large Cohort. *Hum. Mutat.* **2009**, *30* (12), 1657–1666.
(150) Stark, A. E. Determinants of the Incidence of Duchenne Muscular Dystrophy. *Ann. Transl. Med.* **2015**, *3* (19).
(151) Nance, M. E.; Hakim, C. H.; Yang, N. N.; Duan, D. Nanotherapy for Duchenne Muscular Dystrophy. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2017**.
(152) Becker, P. E.; Kiener, F. [A new x-chromosomal muscular dystrophy]. *Arch. Psychiatr. Nervenkrankh. Ver. Mit Z. Gesamte Neurol. Psychiatr.* **1955**, *193* (4), 427–448.
(153) Monaco, A. P.; Neve, R. L.; Colletti-Feener, C.; Bertelson, C. J.; Kurnit, D. M.; Kunkel, L. M. Isolation of Candidate CDNAs for Portions of the Duchenne Muscular Dystrophy Gene. *Nature* **1986**, *323* (6089), 646–650.
(154) McGreevy, J. W.; Hakim, C. H.; McIntosh, M. A.; Duan, D. Animal Models of Duchenne Muscular Dystrophy: From Basic Mechanisms to Gene Therapy. *Dis. Model. Mech.* **2015**, *8* (3), 195–213.
(155) Lim, K. R. Q.; Maruyama, R.; Yokota, T. Eteplirsen in the Treatment of Duchenne Muscular Dystrophy. *Drug Des. Devel. Ther.* **2017**, *11*, 533–545.
(156) England, S. B.; Nicholson, L. V. B.; Johnson, M. A.; Forrest, S. M.; Love, D. R.; Zubrzycka-Gaarn, E. E.; Bulman, D. E.; Harris, J. B.; Davies, K. E. Very Mild Muscular Dystrophy Associated with the Deletion of 46% of Dystrophin. *Nature* **1990**, *343* (6254), 180–182.
(157) Chamberlain, J. R.; Chamberlain, J. S. Progress toward Gene Therapy for Duchenne Muscular Dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther.* **2017**, *25* (5), 1125–1131.
(158) Harper, S. Q.; Hauser, M. A.; DelloRusso, C.; Duan, D.; Crawford, R. W.; Phelps, S. F.; Harper, H. A.; Robinson, A. S.; Engelhardt, J. F.; Brooks, S. V.; et al. Modular Flexibility of Dystrophin: Implications for Gene Therapy of Duchenne Muscular Dystrophy. *Nat. Med.* **2002**, *8* (3), 253–261.
(159) Long, C.; Amoasii, L.; Mireault, A. A.; McAnally, J. R.; Li, H.; Sanchez-Ortiz, E.; Bhattacharyya, S.; Shelton, J. M.; Bassel-Duby, R.; Olson, E. N. Postnatal Genome Editing Partially Restores Dystrophin Expression in a Mouse Model of Muscular Dystrophy. *Science* **2016**, *351* (6271), 400–403.
(160) Nelson, C. E.; Hakim, C. H.; Ousterout, D. G.; Thakore, P. I.; Moreb, E. A.; Castellanos Rivera, R. M.; Madhavan, S.; Pan, X.; Ran, F. A.; Yan, W. X.; et al. In Vivo Genome Editing Improves Muscle Function in a Mouse Model of Duchenne Muscular Dystrophy. *Science* **2016**, *351* (6271), 403–407.
(161) Tabebordbar, M.; Zhu, K.; Cheng, J. K. W.; Chew, W. L.; Widrick, J. J.; Yan, W. X.; Maesner, C.; Wu, E. Y.; Xiao, R.; Ran, F. A.; et al. In Vivo Gene Editing in Dystrophic Mouse Muscle and Muscle Stem Cells. *Science* **2016**, *351* (6271), 407–411.
(162) Counsell, J. R.; Asgarian, Z.; Meng, J.; Ferrer, V.; Vink, C. A.; Howe, S. J.; Waddington, S. N.; Thrasher, A. J.; Muntoni, F.; Morgan, J. E.; et al. Lentiviral Vectors Can Be Used for Full-Length Dystrophin Gene Therapy. *Sci. Rep.* **2017**, *7*, 44775.
(163) Tomás, H. A.; Rodrigues, A. F.; Alves, P. M.; Coroadinha, A. S. Lentiviral Gene Therapy Vectors: Challenges and Future Directions. **2013**.

(164) Sienkiewicz, D.; Kulak, W.; Okurowska-Zawada, B.; Paszko-Patej, G.; Kawnik, K. Duchenne Muscular Dystrophy: Current Cell Therapies. *Ther. Adv. Neurol. Disord.* **2015**, *8* (4), 166–177.
(165) Sharma, A.; Sane, H.; Badhe, P.; Gokulchandran, N.; Kulkarni, P.; Lohiya, M.; Biju, H.; Jacob, V. C. A Clinical Study Shows Safety and Efficacy of Autologous Bone Marrow Mononuclear Cell Therapy to Improve Quality of Life in Muscular Dystrophy Patients. *Cell Transplant.* **2013**, *22 Suppl 1*, S127-138.
(166) Koenig, M.; Hoffman, E. P.; Bertelson, C. J.; Monaco, A. P.; Feener, C.; Kunkel, L. M. Complete Cloning of the Duchenne Muscular Dystrophy (DMD) CDNA and Preliminary Genomic Organization of the DMD Gene in Normal and Affected Individuals. *Cell* **1987**, *50* (3), 509–517.
(167) Muntoni, F.; Torelli, S.; Ferlini, A. Dystrophin and Mutations: One Gene, Several Proteins, Multiple Phenotypes. *Lancet Neurol.* **2003**, *2* (12), 731–740.
(168) Sadoulet-Puccio, H. M.; Kunkel, L. M. Dystrophin and Its Isoforms. *Brain Pathol. Zurich Switz.* **1996**, *6* (1), 25–35.
(169) Nicolas, A.; Lucchetti-Miganeh, C.; Yaou, R. B.; Kaplan, J.-C.; Chelly, J.; Leturcq, F.; Barloy-Hubler, F.; Le Rumeur, E. Assessment of the Structural and Functional Impact of In-Frame Mutations of the DMD Gene, Using the Tools Included in the EDystrophin Online Database. *Orphanet J. Rare Dis.* **2012**, *7*, 45.
(170) Koenig, M.; Monaco, A. P.; Kunkel, L. M. The Complete Sequence of Dystrophin Predicts a Rod-Shaped Cytoskeletal Protein. *Cell* **1988**, *53* (2), 219–228.
(171) Le Rumeur, E.; Winder, S. J.; Hubert, J.-F. Dystrophin: More than Just the Sum of Its Parts. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2010**, *1804* (9), 1713–1722.
(172) Kusunoki, H.; Minasov, G.; Macdonald, R. I.; Mondragón, A. Independent Movement, Dimerization and Stability of Tandem Repeats of Chicken Brain Alpha-Spectrin. *J. Mol. Biol.* **2004**, *344* (2), 495–511.
(173) Legrand, B.; Giudice, E.; Nicolas, A.; Delalande, O.; Le Rumeur, E. Computational Study of the Human Dystrophin Repeats: Interaction Properties and Molecular Dynamics. *PloS One* **2011**, *6* (8), e23819.
(174) Norwood, F. L.; Sutherland-Smith, A. J.; Keep, N. H.; Kendrick-Jones, J. The Structure of the N-Terminal Actin-Binding Domain of Human Dystrophin and How Mutations in This Domain May Cause Duchenne or Becker Muscular Dystrophy. *Struct. Lond. Engl. 1993* **2000**, *8* (5), 481–491.
(175) Muthu, M.; Richardson, K. A.; Sutherland-Smith, A. J. The Crystal Structures of Dystrophin and Utrophin Spectrin Repeats: Implications for Domain Boundaries. *PloS One* **2012**, *7* (7), e40066.
(176) Huang, X.; Poy, F.; Zhang, R.; Joachimiak, A.; Sudol, M.; Eck, M. J. Structure of a WW Domain Containing Fragment of Dystrophin in Complex with β-Dystroglycan. *Nat. Struct. Mol. Biol.* **2000**, *7* (8), 634–638.
(177) Yang, J.; Zhang, Y. Protein Structure and Function Prediction Using I-TASSER. *Curr. Protoc. Bioinforma. Ed. Board Andreas Baxevanis Al* **2015**, *52*, 5.8.1-5.815.
(178) Molza, A.-E.; Férey, N.; Czjzek, M.; Rumeur, E. L.; Hubert, J.-F.; Tek, A.; Laurent, B.; Baaden, M.; Delalande, O. Innovative Interactive Flexible Docking Method for Multi-Scale Reconstruction Elucidates Dystrophin Molecular Assembly. *Faraday Discuss.* **2014**.
(179) Tek, A.; Chavent, M.; Baaden, M.; Delalande, O.; Bourdot, P.; Ferey, N. Advances in Human-Protein Interaction - Interactive and Immersive Molecular Simulations. In *Protein-Protein Interactions - Computational and Experimental Tools*; Cai, W., Ed.; InTech, 2012.

(180) Nicolas, A.; Raguénès-Nicol, C.; Ben Yaou, R.; Ameziane-Le Hir, S.; Chéron, A.; Vié, V.; Claustres, M.; Leturcq, F.; Delalande, O.; Hubert, J.-F.; et al. Becker Muscular Dystrophy Severity Is Linked to the Structure of Dystrophin. *Hum. Mol. Genet.* **2015**, *24* (5), 1267–1279.
(181) Allen, D. G.; Whitehead, N. P.; Froehner, S. C. Absence of Dystrophin Disrupts Skeletal Muscle Signaling: Roles of Ca2+, Reactive Oxygen Species, and Nitric Oxide in the Development of Muscular Dystrophy. *Physiol. Rev.* **2016**, *96* (1), 253–305.
(182) Rybakova, I. N.; Humston, J. L.; Sonnemann, K. J.; Ervasti, J. M. Dystrophin and Utrophin Bind Actin through Distinct Modes of Contact. *J. Biol. Chem.* **2006**, *281* (15), 9996–10001.
(183) Sarkis, J.; Vié, V.; Winder, S. J.; Renault, A.; Le Rumeur, E.; Hubert, J.-F. Resisting Sarcolemmal Rupture: Dystrophin Repeats Increase Membrane-Actin Stiffness. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2013**, *27* (1), 359–367.
(184) Campbell, M. G.; Smith, B. C.; Potter, C. S.; Carragher, B.; Marletta, M. A. Molecular Architecture of Mammalian Nitric Oxide Synthases. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (35), E3614-3623.
(185) Lai, Y.; Thomas, G. D.; Yue, Y.; Yang, H. T.; Li, D.; Long, C.; Judge, L.; Bostick, B.; Chamberlain, J. S.; Terjung, R. L.; et al. Dystrophins Carrying Spectrin-like Repeats 16 and 17 Anchor NNOS to the Sarcolemma and Enhance Exercise Performance in a Mouse Model of Muscular Dystrophy. *J. Clin. Invest.* **2009**, *119* (3), 624–635.
(186) Harper, S. Q. Molecular Dissection of Dystrophin Identifies the Docking Site for NNOS. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (2), 387–388.
(187) Stamler, J. S.; Meissner, G. Physiology of Nitric Oxide in Skeletal Muscle. *Physiol. Rev.* **2001**, *81* (1), 209–237.
(188) Belanto, J. J.; Mader, T. L.; Eckhoff, M. D.; Strandjord, D. M.; Banks, G. B.; Gardner, M. K.; Lowe, D. A.; Ervasti, J. M. Microtubule Binding Distinguishes Dystrophin from Utrophin. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (15), 5723–5728.
(189) Ervasti, J. M. *Structure and Function of the Dystrophin-Glycoprotein Complex*; Landes Bioscience, 2013.
(190) DeWolf, C.; McCauley, P.; Sikorski, A. F.; Winlove, C. P.; Bailey, A. I.; Kahana, E.; Pinder, J. C.; Gratzer, W. B. Interaction of Dystrophin Fragments with Model Membranes. *Biophys. J.* **1997**, *72* (6), 2599–2604.
(191) Le Rumeur, E.; Fichou, Y.; Pottier, S.; Gaboriau, F.; Rondeau-Mouro, C.; Vincent, M.; Gallay, J.; Bondon, A. Interaction of Dystrophin Rod Domain with Membrane Phospholipids. Evidence of a Close Proximity between Tryptophan Residues and Lipids. *J. Biol. Chem.* **2003**, *278* (8), 5993–6001.
(192) Le Rumeur, E.; Pottier, S.; Da Costa, G.; Metzinger, L.; Mouret, L.; Rocher, C.; Fourage, M.; Rondeau-Mouro, C.; Bondon, A. Binding of the Dystrophin Second Repeat to Membrane Di-Oleyl Phospholipids Is Dependent upon Lipid Packing. *Biochim. Biophys. Acta* **2007**, *1768* (3), 648–654.
(193) Legardinier, S.; Hubert, J.-F.; Le Bihan, O.; Tascon, C.; Rocher, C.; Raguénès-Nicol, C.; Bondon, A.; Hardy, S.; Le Rumeur, E. Sub-Domains of the Dystrophin Rod Domain Display Contrasting Lipid-Binding and Stability Properties. *Biochim. Biophys. Acta* **2008**, *1784* (4), 672–682.
(194) Legardinier, S.; Raguénès-Nicol, C.; Tascon, C.; Rocher, C.; Hardy, S.; Hubert, J.-F.; Le Rumeur, E. Mapping of the Lipid-Binding and Stability Properties of the Central Rod Domain of Human Dystrophin. *J. Mol. Biol.* **2009**, *389* (3), 546–558.
(195) Sarkis, J.; Hubert, J.-F.; Legrand, B.; Robert, E.; Chéron, A.; Jardin, J.; Hitti, E.; Le Rumeur, E.; Vié, V. Spectrin-like Repeats 11-15 of Human Dystrophin Show Adaptations to a Lipidic Environment. *J. Biol. Chem.* **2011**, *286* (35), 30481–30491.

(196) Ameziane-Le Hir, S.; Raguénès-Nicol, C.; Paboeuf, G.; Nicolas, A.; Le Rumeur, E.; Vié, V. Cholesterol Favors the Anchorage of Human Dystrophin Repeats 16 to 21 in Membrane at Physiological Surface Pressure. *Biochim. Biophys. Acta* **2014**, *1838* (5), 1266–1273.
(197) Fantini, J.; Barrantes, F. J. How Cholesterol Interacts with Membrane Proteins: An Exploration of Cholesterol-Binding Sites Including CRAC, CARC, and Tilted Domains. *Front. Physiol.* **2013**, *4*.
(198) Ameziane-Le Hir, S.; Paboeuf, G.; Tascon, C.; Hubert, J.-F.; Le Rumeur, E.; Vié, V.; Raguénès-Nicol, C. Dystrophin Hot-Spot Mutants Leading to Becker Muscular Dystrophy Insert More Deeply into Membrane Models than the Native Protein. *Biochemistry (Mosc.)* **2016**, *55* (29), 4018–4026.
(199) Zhao, J.; Kodippili, K.; Yue, Y.; Hakim, C. H.; Wasala, L.; Pan, X.; Zhang, K.; Yang, N. N.; Duan, D.; Lai, Y. Dystrophin Contains Multiple Independent Membrane-Binding Domains. *Hum. Mol. Genet.* **2016**, *25* (17), 3647–3653.
(200) Winder, S. J.; Gibson, T. J.; Kendrick-Jones, J. Dystrophin and Utrophin: The Missing Links! *FEBS Lett.* **1995**, *369* (1), 27–33.
(201) Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227* (5259), 680–685.
(202) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2006**, *1* (6), 2876–2890.
(203) Mehboob, S.; Luo, B. H.; Patel, B. M.; Fung, L. W. Alpha Beta Spectrin Coiled Coil Association at the Tetramerization Site. *Biochemistry (Mosc.)* **2001**, *40* (41), 12457–12464.
(204) Duhr, S.; Braun, D. Why Molecules Move along a Temperature Gradient. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (52), 19678–19682.
(205) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction Analysis and Beyond. *J. Mol. Struct.* **2014**, *1077*, 101–113.
(204) Carbodiimide Crosslinker Chemistry | Thermo Fisher Scientific https://www.thermofisher.com
(207) Haberkant, P.; Raijmakers, R.; Wildwater, M.; Sachsenheimer, T.; Brügger, B.; Maeda, K.; Houweling, M.; Gavin, A.-C.; Schultz, C.; van Meer, G.; et al. In Vivo Profiling and Visualization of Cellular Protein-Lipid Interactions Using Bifunctional Fatty Acids. *Angew. Chem. Int. Ed Engl.* **2013**, *52* (14), 4033–4038.
(208) Götze, M.; Pettelkau, J.; Schaks, S.; Bosse, K.; Ihling, C. H.; Krauth, F.; Fritzsche, R.; Kühn, U.; Sinz, A. StavroX--a Software for Analyzing Crosslinked Products in Protein Interaction Studies. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (1), 76–87.
(209) Gabel, F. Small-Angle Neutron Scattering for Structural Biology of Protein-RNA Complexes. *Methods Enzymol.* **2015**, *558*, 391–415.
(210) Cotton, J. P. Variations on Contrast in SANS: Determination of Self and Distinct Correlation Functions. *Adv. Colloid Interface Sci.* **1996**, *69* (1), 1–29.
(211) Hammouda, B. SANS from Polymers—Review of the Recent Literature. *Polym. Rev.* **2010**, *50* (1), 14–39.
(212) Svergun, D. I.; Koch, M. H. J. Small-Angle Scattering Studies of Biological Macromolecules in Solution. *Rep. Prog. Phys.* **2003**, *66* (10), 1735.
(213) Franke, D.; Petoukhov, M. V.; Konarev, P. V.; Panjkovich, A.; Tuukkanen, A.; Mertens, H. D. T.; Kikhney, A. G.; Hajizadeh, N. R.; Franklin, J. M.; Jeffries, C. M.; et al. ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions. *J. Appl. Crystallogr.* **2017**, *50* (4), 1212–1225.
(214) Scatter http://www.bioisis.net/tutorial/9

(215) Jeffries, C. M.; Graewert, M. A.; Blanchet, C. E.; Langley, D. B.; Whitten, A. E.; Svergun, D. I. Preparing Monodisperse Macromolecular Samples for Successful Biological Small-Angle X-Ray and Neutron-Scattering Experiments. *Nat. Protoc.* **2016**, *11* (11), 2122–2153.
(216) David, G.; Pérez, J. Combined Sampler Robot and High-Performance Liquid Chromatography: A Fully Automated System for Biological Small-Angle X-Ray Scattering Experiments at the Synchrotron SOLEIL SWING Beamline. *J. Appl. Crystallogr.* **2009**, *42* (5), 892–900.
(217) Jordan, A.; Jacques, M.; Merrick, C.; Devos, J.; Forsyth, V. T.; Porcar, L.; Martel, A. SEC-SANS: Size Exclusion Chromatography Combined in Situ with Small-Angle Neutron Scattering. *J. Appl. Crystallogr.* **2016**, *49* (6), 2015–2020.
(218) Glatter, O.; Kratky, O. *Small Angle X-Ray Scattering*; Academic Press, 1982.
(217) Biomolecular Scattering Length Density Calculator http://psldc.isis.rl.ac.uk/Psldc
(220) Rambo, R. P.; Tainer, J. A. Accurate Assessment of Mass, Models and Resolution by Small-Angle Scattering. *Nature* **2013**, *496* (7446), 477–481.
(221) Kline, S. R. Reduction and Analysis of SANS and USANS Data Using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39* (6), 895–900.
(222) Nelson, A. Co-Refinement of Multiple-Contrast Neutron/X-Ray Reflectivity Data Using *MOTOFIT*. *J. Appl. Crystallogr.* **2006**, *39* (2), 273–276.
(223) Koutsioubas, A. Combined Coarse-Grained Molecular Dynamics and Neutron Reflectivity Characterization of Supported Lipid Membranes. *J. Phys. Chem. B* **2016**, *120* (44), 11474–11483.
(224) Richter, R. P.; Brisson, A. R. Following the Formation of Supported Lipid Bilayers on Mica: A Study Combining AFM, QCM-D, and Ellipsometry. *Biophys. J.* **2005**, *88* (5), 3422–3433.
(225) Small, D. M. *The Physical Chemistry of Lipids: From Alkanes to Phospholipids*; Plenum Press, 1986.
(224) Neutron Activation Calculator https://www.ncnr.nist.gov/resources/activation
(227) Bordner, A. J. Force Fields for Homology Modeling. *Methods Mol. Biol. Clifton NJ* **2012**, *857*, 83–106.
(228) Notman, R.; Anwar, J. Breaching the Skin Barrier--Insights from Molecular Simulation of Model Membranes. *Adv. Drug Deliv. Rev.* **2013**, *65* (2), 237–250.
(229) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824.
(230) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4* (5), 819–834.
(231) Dos Santos Morais, R.; Delalande, O.; Pérez, J.; Mouret, L.; Bondon, A.; Martel, A.; Appavou, M.-S.; Le Rumeur, E.; Hubert, J.-F.; Combet, S. Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS. *Langmuir* **2017**, *33* (26), 6572–6580.
(232) Lin, T. L.; Chen, S. H.; Gabriel, N. E.; Roberts, M. F. The Use of Small-Angle Neutron Scattering to Determine the Structure and Interaction of Dihexanoylphosphatidylcholine Micelles. *J. Am. Chem. Soc.* **1986**, *108* (12), 3499–3507.
(233) Lipfert, J.; Columbus, L.; Chu, V. B.; Lesley, S. A.; Doniach, S. Size and Shape of Detergent Micelles Determined by Small-Angle X-Ray Scattering. *J. Phys. Chem. B* **2007**, *111* (43), 12427–12438.

(234) Kiselev, M. A.; Lombardo, D. Structural Characterization in Mixed Lipid Membrane Systems by Neutron and X-Ray Scattering. *Biochim. Biophys. Acta* **2016**.
(235) Gabriel, N. E.; Roberts, M. F. Short-Chain Lecithin/Long-Chain Phospholipid Unilamellar Vesicles: Asymmetry, Dynamics, and Enzymatic Hydrolysis of the Short-Chain Component. *Biochemistry (Mosc.)* **1987**, *26* (9), 2432–2440.
(236) Tristram-Nagle, S.; Liu, Y.; Legleiter, J.; Nagle, J. F. Structure of Gel Phase DMPC Determined by X-Ray Diffraction. *Biophys. J.* **2002**, *83* (6), 3324–3335.
(237) Kučerka, N.; Liu, Y.; Chu, N.; Petrache, H. I.; Tristram-Nagle, S.; Nagle, J. F. Structure of Fully Hydrated Fluid Phase DMPC and DLPC Lipid Bilayers Using X-Ray Scattering from Oriented Multilamellar Arrays and from Unilamellar Vesicles. *Biophys. J.* **2005**, *88* (4), 2626–2637.
(238) Gimpl, K.; Klement, J.; Keller, S. Characterising Protein/Detergent Complexes by Triple-Detection Size-Exclusion Chromatography. *Biol. Proced. Online* **2016**, *18*.
(239) Kusunoki, H.; Minasov, G.; Macdonald, R. I.; Mondragón, A. Independent Movement, Dimerization and Stability of Tandem Repeats of Chicken Brain Alpha-Spectrin. *J. Mol. Biol.* **2004**, *344* (2), 495–511.
(240) Vié, V.; Legardinier, S.; Chieze, L.; Le Bihan, O.; Qin, Y.; Sarkis, J.; Hubert, J.-F.; Renault, A.; Desbat, B.; Le Rumeur, E. Specific Anchoring Modes of Two Distinct Dystrophin Rod Sub-Domains Interacting in Phospholipid Langmuir Films Studied by Atomic Force Microscopy and PM-IRRAS. *Biochim. Biophys. Acta* **2010**, *1798* (8), 1503–1511.
(241) Rahimov, F.; Kunkel, L. M. The Cell Biology of Disease: Cellular and Molecular Mechanisms Underlying Muscular Dystrophy. *J. Cell Biol.* **2013**, *201* (4), 499–510.

Résumé

La dystrophine est une grande protéine membranaire périphérique qui assure un rôle de soutien du sarcolemme permettant aux cellules musculaires de résister aux stress mécaniques engendrés lors des processus de contraction/élongation. Des mutations génétiques conduisent à sa production sous forme tronquée voire à un déficit total en protéine engendrant de sévères myopathies actuellement incurables. Concevoir des thérapies adaptées passe par une meilleure compréhension du rôle biologique de la dystrophine. Par une approche structure/fonction, notre objectif est de déterminer les bases moléculaires impliquées dans les interactions de la dystrophine avec les lipides membranaires du sarcolemme. Grâce à une approche de diffusion aux petits angles (SAXS et SANS) combinée à de la modélisation moléculaire, nous montrons dans un premier temps que les bicelles constituent un modèle expérimental particulièrement adapté aux analyses de structures de protéines qui y sont associées. Ce développement méthodologique original a été exploité dans un deuxième temps pour caractériser les modifications structurales subies par la dystrophine lorsqu'elle interagit avec les lipides. Nous montrons particulièrement que la liaison aux lipides induit l'ouverture significative de la structure en triple hélice « coiled-coil » de la répétiton 1 du domaine central, et proposons en conclusion un modèle tout atome de la protéine en présence de bicelles. Ces travaux de thèse (i) constituent un apport méthodologique significatif pour l'étude de protéines membranaires, (ii) contribuent à une meilleure compréhension du rôle biologique de la dystrophine en vue de thérapies dédiées aux patients atteints de myopathies.

Abstract

Dystrophin is a large peripheral membrane protein that provides a supporting role for sarcolemma allowing muscle cells to withstand the mechanical stresses generated during contraction / elongation processes. Genetic mutations lead to dystrophin production in truncated form or even to a total deficit in the protein leading to severe myopathies currently incurable. Designing adapted therapies requires a huge knowledge of the biological role of dystrophin. Using a structure / function approach, our aim is to determine the molecular bases involved in the interactions of dystrophin with the membrane lipids of the sarcolemma. Using a small-angle scattering approach (SAXS and SANS) combined with molecular modeling, we show that bicelles constitute a versatile membrane mimic that is particularly adapted to analyze the structure of membrane proteins. This original methodological development was exploited to characterize the structural changes undergone by dystrophin upon lipid binding. We highlight in particular that the lipid binding induces a significant opening of the coiled-coil structure of the repeat 1 of the central domain and, in conclusion, we propose an all-atom model of the protein bound to a bicelle. These thesis works (i) constitute a significant methodological contribution for the study of membrane proteins, (ii) contribute to a better understanding of the biological role of dystrophin for therapies dedicated to patients with myopathies.