



**HAL**  
open science

# Dynamique et mécanique de complexes dystrophine-actine-lipides membranaires

Dominique Mias-Lucquin

► **To cite this version:**

Dominique Mias-Lucquin. Dynamique et mécanique de complexes dystrophine-actine-lipides membranaires. Biochimie, Biologie Moléculaire. Université de Rennes, 2018. Français. NNT : 2018REN1B024 . tel-01998542

**HAL Id: tel-01998542**

**<https://theses.hal.science/tel-01998542>**

Submitted on 29 Jan 2019

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

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

# Thèse de doctorat de

L'UNIVERSITE DE RENNES 1

Comue Université Bretagne Loire

ÉCOLE DOCTORALE N° 605

*Biologie Santé*

Spécialité : *Biologie Moléculaire et Structurale, Biochimie*

Par

**Dominique Mias-Lucquin**

## **Dynamique et mécanique de complexes Dystrophine-Actine-Lipides membranaires**

Thèse présentée et soutenue à Rennes, le 24 Septembre 2018

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)

### **Rapporteurs avant soutenance :**

Catherine ETCHEBEST

Raphael GUEROIS

Professeur

Chercheur CEA

Université Paris Diderot

I2BC (Institut de biologie intégrative de la cellule)

### **Composition du Jury :**

Examineurs :

Catherine ETCHEBEST

Raphael GUEROIS

France PIETRI-ROUXEL

Christian DELAMARCHE

Professeur

Chercheur CEA

Directrice de recherche CNRS

Professeur

Université Paris Diderot

I2BC

Institut de Myologie

Université Rennes 1

Dir. de thèse :

Jean-Francois HUBERT

Professeur

Université Rennes 1

Co-dir. de thèse :

Olivier DELALANDE

Maitre de conférences

Université Rennes 1



"— [...] Ensuite il y a les mages qui sont dessus, qui doivent tous être des hydrophobes chevronnés...

— Tu veux dire qu'ils détestent l'eau ?

— Non, ça ne marcherait pas. La haine, c'est une force d'attraction, tout comme l'amour. Ils l'ont vraiment en horreur, l'idée même de l'eau les révolte. Un très bon hydrophobe doit subir un entraînement à l'eau déshydratée dès sa naissance. Je veux dire, ça coûte une fortune rien qu'en magie. Mais ils font de grands mages climatiques. Les nuages de pluie n'insistent pas et vont voir ailleurs.

Terry Pratchett, Les annales du disque monde, tome 1 : « La Huitième Couleur »





## Remerciements

Je remercie l'Association Française contre les Myopathies (AFM) d'avoir financé ma thèse, et pour son soutien financier à l'équipe qui a permis de réaliser une partie des expériences présentées dans cette thèse.

Ces travaux ont été réalisés au sein de l'Institut de Génétique et de Développement de Rennes (IGDR, UMR CNRS 6290) de l'Université Rennes 1, dans l'équipe DysSim (ex SIM). Tout cela n'aurait pas été possible sans le Professeur Jean-François Hubert et le Docteur Olivier Delalande, qui m'ont supervisé dans ce travail, et bien plus encore. Un grand merci à Elisabeth Le Rumeur, initiatrice du projet, et sans laquelle je n'aurais pas eu ce financement.

Merci à Angélique Chéron, à Thomas Chénuel, à Raphael Dos-Santos Morais (Docteur depuis le 27 Octobre 2017 !), à Agnès Allaire, à Anne-Elisabeth Molza, à Céline Raguénès-Nicol, à Noémie Buisard, membres ou anciens membres de l'équipe que j'ai eu la chance de côtoyer.

Un grand merci aux personnes du bâtiment 13 : Denis Chrétien, Laurence Duchesne et Claire Heichette, avec qui j'ai eu la chance de travailler ; et aussi à ceux que je ne faisais que croiser. Et merci aussi à Agnès Burel, responsable de la plateforme microscopie électronique Biosit.

Je remercie chaleureusement Louis Marage et Sean Guegan, pour nos pauses jeux. Je tiens également à remercier toutes les personnes qui partagent mon quotidien au sous-sol du bâtiment 5, notamment du chat qui a squatté notre faux plafond quelques temps ; il était probablement iso-9001. Merci à Lucie, grâce à qui je suis venue à Rennes il y a maintenant 5 ans, et à Marion, pour son apparition dans mon court métrage.

Merci à Claire Rome, Nicolas Coquery, et tous ceux dont les conseils m'ont permis d'en arriver là où j'en suis aujourd'hui.

Merci à toute ma famille, ma maman, mon frère, ma mamie, mon oncle, mes tantes et mes cousins pour leur soutien moral. Merci à Lisa, pour presque tout, et à Yvette, pour nous avoir maintenus en vie dans les moments difficiles. Merci à mon chaton, qui a appris à rapporter la balle durant la rédaction de cette thèse.

Afin de n'oublier personne, merci à ....., pour son ..... et sa ..... . Sans cet individu, rien n'aurait été possible.



## Table des matières

Table des abréviations .....	9
I. Introduction .....	11
A. Dystrophine .....	11
1. Gène.....	11
2. Isoformes .....	11
3. Structure de dp427-M .....	12
4. Superfamille de la spectrine et utrophine.....	15
B. Partenaires de la dystrophine .....	18
1. Lipides .....	18
2. Protéines.....	23
C. Myopathie de Duchenne .....	35
1. Historique .....	35
2. Muscle.....	35
3. Pathologie.....	38
D. Thérapies .....	42
1. Thérapies cellulaires .....	42
2. Thérapies géniques.....	42
E. Objectifs de thèse.....	48
II. Méthodes .....	51
A. Champs de force.....	51
B. Minimisation d'énergie .....	53
C. Simulation de dynamique moléculaire .....	54
1. Approches haute-résolution tout-atome .....	55
2. Approches basse-résolution gros-grain.....	56
3. Construction d'un système et simulation de dynamique moléculaire .....	58
4. Dynamique moléculaire sous contrainte et interactives .....	60
5. Amarrage moléculaire .....	61
D. Techniques expérimentales de guidage des simulations.....	65
1. Cartographie des interactions protéine/protéine et protéine/lipide .....	65
2. Diffusion aux petits angles.....	67
3. Cryo-tomo-électromicroscopie .....	69

III.	Etude des interactions hydrophobes du faisceau d'hélice R1-3.....	73
A.	Introduction.....	73
1.	Spectrine et dystrophine .....	73
2.	Interactions hydrophobes .....	76
3.	Importance biologique du fragment R1-3 de la dystrophine.....	77
B.	Article 1 : Fine mapping of hydrophobic contacts reassesses the organisation of the first three dystrophin coiled-coil repeats.....	80
C.	Discussion .....	109
1.	Compléments à l'analyse d'une répétition de dystrophine .....	109
2.	Fragment R1-3 et interaction avec un modèle de bicelle .....	111
D.	Perspectives.....	122
IV.	Modélisation de complexes de la dystrophine (actine, microtubules, lipides)...	127
A.	Introduction.....	127
B.	Microtubules .....	128
1.	Introduction .....	128
2.	Analyses des travaux de l'équipe du Pr.Ervasti .....	128
3.	Reproduction des résultats .....	131
C.	Actine.....	134
1.	Introduction .....	134
2.	Co-sédimentation .....	135
3.	Acquisitions CET.....	137
D.	Article 2 : How the central domain of dystrophin bridges F-actin to sarcolemma lipids	143
E.	Discussion .....	182
F.	Perspectives.....	188
V.	Discussion générale.....	189
	Références bibliographiques .....	193
	Annexe : Dos Santos Morais <i>et al</i> , 2018 .....	217

## Table des abréviations

+TIPs	Plus-end tracking protein
AA	All-Atom (Tout atome)
ABD	Actin Binding Domain (domaine de liaison à l'actine)
ADN	Acide DéoxyriboNucléique
ALPS	<i>ArfGAP1/Amphipathic Lipid-Packing Sensor</i>
ARN	Acide RiboNucléique
ARNm / ARNg	ARN messenger / ARN guide
Actine-F / -G	Actine Filamenteuse ou Globulaire
AON / ASO	Antisense oligonucleotides (oligonucleotides antisens)
CEM / CET	Cryo-microscopie électronique / tomo-microscopie électronique
CRISPR / Cas9	Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR associated protein 9
DGC	Dystrophin glycoprotein complex
DHPC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
DMD / BMD	Duchenne muscular dystrophy / Becker muscular dystrophy
DMPC	1,2-Dimyristoyl-sn-glycero-3-phosphocholine
DMPS	1,2-Dimyristoyl-sn-glycero-3-phosphoserine
DOPC	1,2-Dioléoyl-sn-glycero-3-phosphocholine
DOPE	1,2-Dioléoyl-sn-glycero-3-phosphoethanolamine
DOPS	1,2-Dioléoyl-sn-glycero-3-phosphoserine
EDC	1-éthyl-3-(3-diméthylaminopropyl)carbodiimide hydrochloride
FRET	Fluorescence Resonance Energy Transfer (Transfert d'énergie entre molécules fluorescentes)
H1 – H4	Hinge/charnière 1 – Hinge/charnière 4
IMD	Interactive molecular dynamic (Dynamique moléculaire interactive)
kDa	Kilo-Daltons
LBD	Lipid Binding Domain (domaine de liaison aux lipides)
MAP	Microtubules Associated Proteins (protéine associé aux microtubules)
MHP	Molecular Hydrophobic Potential (potentiel hydrophobe moléculaire)
nNOS	Neuronal Nitric Oxide Synthase (oxyde nitrique synthase neuronale)
pacFA-PC	1-Palmitoyl-2-(9-(3-pent-4-ynyl-3-H-diazirine-3-yl)-nonanoyl)-sn-glycero-3-phosphocholine
R1 - R24	Répétition 1 - répétition 24
SAS	Small Angle Scattering (diffusion aux petits angles)
SANS	Small Angle Neutron Scattering (diffusion des neutrons aux petits angles)
SAXS	Small Angle X-Ray Scattering (diffusion des rayons Xaux petits angles)
SMD	Steered molecular dynamic (dynamique moléculaire sous contraintes)
SUV / LUV	Small unilamellar vesicles / large unilamellar vesicles
μs	micro-seconde (10 <sup>-6</sup> s)
ns	nano-seconde (10 <sup>-9</sup> s)
ps	micro-seconde (10 <sup>-12</sup> s)
fs	micro-seconde (10 <sup>-15</sup> s)



## I. Introduction

### A. Dystrophine

#### 1. Gène

Le gène *DMD* qui code pour la dystrophine est situé sur le chromosome X (locus Xp21.2). Sa taille de 2,3 Mb fait de lui le plus long gène humain, et représente environ 1/1000 de la totalité du génome humain (Koenig et al., 1987; Monaco et al., 1986). Il est constitué de 7 promoteurs, 79 exons, et 78 introns (Figure 1). Ces introns ont la particularité d'avoir une taille variant de 107b à 360kb, et constituant 90% de la longueur totale du gène (Pons et al., 1993).

#### 2. Isoformes

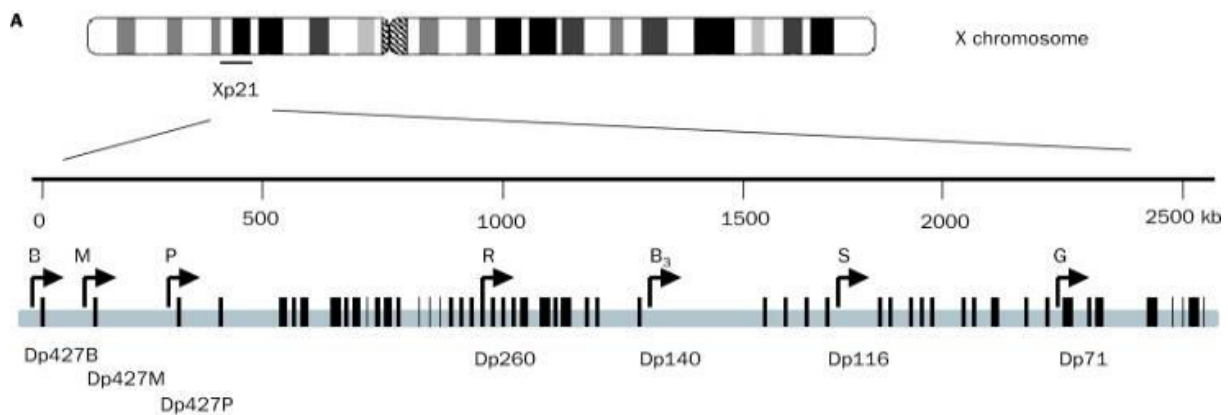


Figure 1 : Position du gène *DMD* sur le chromosome X ; position des promoteurs (Flèches) et des exons (traits noirs). (D'après Muntoni et al, 2003.)

La régulation des différents promoteurs de transcription et de l'épissage alternatif est finement contrôlée en fonction des tissus d'expression. Sept isoformes spécifiques à différents organes sont ainsi produites (Sadoulet-Puccio and Kunkel, 1996).

Trois promoteurs permettent la production de protéines complètes (*Full length*, 427kDa, ARN messenger 14kb) ; ils sont notés B, M et P, et sont respectivement exprimés dans le cerveau (Nudel et al., 1989), dans le muscle (Monaco et al., 1986), et dans les cellules de Purkinje (Górecki et al., 1992). Ils initient la traduction de l'ARN messenger (ARNm) à partir d'un premier exon qui leur est spécifique, mais dont la séquence ne diffère des autres que de 9 à 36 nucléotides. Les 78 exons suivants sont communs aux trois transcrits. L'isoforme dp427-M est la forme majoritairement exprimée dans les muscles. C'est elle qui est altérée dans les myopathies de Becker et de Duchenne et fait l'objet de cette thèse.



Le gène *DMD* contient également quatre promoteurs internes codant pour des protéines plus courtes que les formes *full length* – de la plus longue à la plus courte - (Dp260 (D'Souza et al., 1995), Dp140 (Lidov et al., 1995), Dp116 (Byers et al., 1993), et Dp71 (Lederfein et al., 1992)).

### 3. Structure de dp427-M

L'isoforme dp427-M de la dystrophine possède quatre domaines fonctionnels (un domaine ABD1, *Actin Binding Domain 1*, un domaine central, un domaine Cys-riche, et un domaine Carboxy-terminale) (Gao and McNally, 2011) (Figure 2).



Figure 2: Organisation de l'isoforme dp427-M de la dystrophine. De gauche à droite : (orange) domaine ABD1 ; (vert) domaine central ; (bleu) domaine Cys-Riche ; (gris) domaine C-terminal

A l'extrémité N-terminale se situe le domaine ABD1 (*Actin Binding Domain 1*), dont la structure a été résolue dès 2000 (PDB ID : 1DXX (Norwood et al., 2000)). Il est constitué de 246 acides aminés, et permet de lier la dystrophine au réseau d'actine cytosquelettique (Way et al., 1992). Il se compose de deux sous-domaines homologues à la calponine (*calponin homology domain* CH1 et CH2) séparés par une hélice  $\alpha$  (Norwood et al., 2000). Ces domaines font partie d'une famille de domaines structuraux qui sont retrouvés dans de nombreuses protéines se liant à l'actine (utrophine (Singh et al., 2017),  $\alpha$ -actinine (Shams et al., 2016)).

Le second domaine, ou domaine central, représente 75 % des acides aminés (2798 acides aminés) de la protéine. Il est principalement structuré en hélices  $\alpha$ , qui s'arrangent en une succession de 24 faisceaux à trois hélices (identifiées A, B et C), ou *coiled-coil*, chacun ayant une longueur d'environ 5 nm et un diamètre de 2 nm.

De manière générale, les structures en faisceaux d'hélices sont composées de plusieurs hélices  $\alpha$  (de 2 à 4) qui s'enroulent les unes autour des autres. Les hélices  $\alpha$  sont elles-mêmes formées par l'enroulement régulier de motifs répétés de 7 acides aminés tous les deux tours d'hélice (Figure 3A). Ces sept acides aminés forment un motif en heptade, dont les positions sont identifiées de *a* à *g* (Crick, 1953) (Figure 3B). En particulier, les résidus retrouvés aux positions *a* et *d* sont principalement des acides aminés présentant des propriétés hydrophobes (Akey et al., 2001). Ces résidus sont donc distants de 3,5 positions en moyenne, soit tous les tours d'hélice. Ils forment de fait sur l'une des faces de l'hélice un sillon

hydrophobe qui guide son association avec d'autres hélices présentant la même organisation (Figure 3C). Les chaînes latérales d'une hélice se placent en quinconce entre celles de l'hélice faisant face, dans un arrangement appelé *knobs-into-holes* (Walshaw and Woolfson, 2003) (Figure 3D). La substitution d'acides aminés hydrophobes par des acides aminés polaires à certaines positions *a* et *d* est associée à des modulations de structure du faisceau d'hélice, comme son diamètre ou l'état de polymérisation (Akey et al., 2001).

Dans le domaine central de la dystrophine, les faisceaux d'hélices sont constitués de trois hélices, et ce motif est répété 24 fois (identifié de R1 à R24). L'enchaînement d'une répétition avec la suivante se fait par la connexion directe de la troisième hélice d'une répétition (l'hélice C) avec la première hélice de la répétition suivante (l'hélice A), sans perte de l'hélicité. Au niveau de cette région appelée *linker*, la seconde hélice (l'hélice B) d'une répétition entre en contact avec l'hélice A de la répétition suivante et forme localement un faisceau à deux hélices (Figure 3A).

Malgré des structures similaires, les scores d'homologie entre les séquences de ces répétitions sont plutôt faibles (au maximum de 25%) (Nicolas et al., 2014a). Les répétitions sont en moyenne composées de 110 résidus et les hélices d'environ 30 acides aminés sont séparées par des boucles d'environ 10 acides aminés.

L'enchaînement des répétitions est interrompu par quatre charnières (*hinges*). Ces domaines non structurés en hélices et riches en proline segmentent et flexibilisent le domaine central (avant R1, entre R3 et R4, entre R19 et R20, et suivant R24 (Figure 2)) (Koenig and Kunkel, 1990).

À cause de sa similarité avec la spectrine, le domaine central a longtemps été considéré comme une tige rigide (Koenig et al., 1988). À cause de sa similarité avec la spectrine, le domaine central a longtemps été considéré comme une tige rigide (Koenig et al., 1988). Des travaux récents de notre équipe tendent à montrer au contraire que la présence de coudes entre certaines répétitions au niveau des régions *linker* crée une structure non-linéaire (Delalande et al., 2018). Ces coudes sont aussi potentiellement des régions dynamiques, conférant au domaine central une certaine flexibilité et délimitant des domaines fonctionnels (Legrand et al., 2011). Cette souplesse pourrait expliquer la difficulté d'obtention de structures cristallographiques recouvrant cette région (à l'exception de la répétition R1 (PDB ID 3UUN (Muthu et al., 2012), structure obtenue sous forme d'un dimère forcé).

Le domaine central de la dystrophine est impliqué dans des interactions avec de nombreux partenaires qui seront détaillés dans la partie B de ce chapitre. Par exemple, le

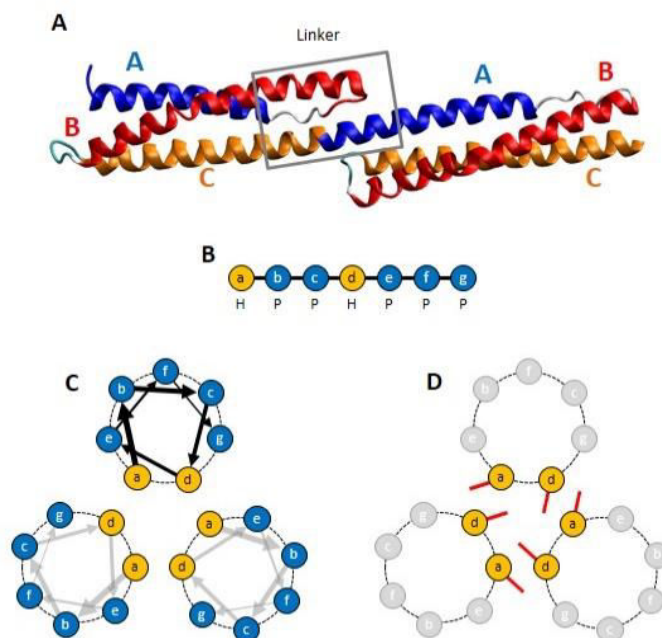


Figure 3 : Structure canonique du faisceau d'hélice à trois hélices. A) Répétition de deux faisceaux d'hélices successifs ; (bleu) Hélices A ; (rouge) Hélice B ; (orange) Hélice C ; Cadre gris : Région Linker. B) Motif en heptade (jaune) positions **a** et **d** Hydrophobes (H) ; (bleu) positions **b, c, e, f, g** polaires (P). C) Organisation des résidus de l'heptade dans les hélices  $\alpha$ . D) Principe de l'arrangement knobs-into-holes

fragment s'étendant des répétitions R1 à R3 à la capacité de se lier aux lipides composants les membranes, tout comme la région s'étendant de R11 à R15 (Sarkis et al., 2011). Cette dernière abrite également un second domaine de liaison à l'actine (ABD2) (Rybakova et al., 1996). Les répétitions R16-17 interagissent avec la protéine nNOS (Molza et al., 2015), et le domaine R20-24 apparaît nécessaire pour l'organisation du réseau de microtubules des cellules musculaires (Belanto et al., 2016).

Le troisième domaine de la dystrophine, situé à la suite de la quatrième charnière (H4), est un domaine riche en cystéine (Cys-riche), constitué de 307 acides aminés. Le domaine cys-riche contient un motif "doigt de zinc" (ou ZZ) qui se structure en présence d'ions divalents (principalement le  $Zn^{2+}$  (Ponting et al., 1996)). Ce motif constitue le site de liaison au  $\beta$ -dystroglycane (Yoshida and Ozawa, 1990) et à la calmoduline (Anderson et al., 1996). D'autres partenaires de ce domaine ont été décrits, comme l'ankyrine-B (Ayalon et al., 2008) et la synamine (Bhosle et al., 2006). À l'extrémité N-terminale de ce domaine cys-riche, un domaine "WW" permet de se lier au  $\beta$ -dystroglycane (Bork and Sudol, 1994; Huang et al., 2000). Il est suivi par deux domaines "EF" (EF1 et EF2), structurés en « hélice-boucle-hélice » (Huang et al., 2000; Rentschler et al., 1999). Ces motifs structuraux, retrouvés dans plus de 3200 protéines (PROSITE entry PS00018, 1990), sont impliqués dans la chélation d'ions divalents ( $Ca^{2+}$  ou  $Mg^{2+}$ ) (Tufty and Kretsinger, 1975). Comme le domaine WW qui les précède, les deux domaines EF permettent l'ancrage de la dystrophine à la membrane par leur liaison à l'extrémité C-terminale du  $\beta$ -dystroglycane (Rentschler et al., 1999).

À l'extrémité C-terminale se situe le quatrième domaine de la dystrophine. Il est composé de 333 résidus (Blake et al., 1995) qui ont la particularité de former deux hélices qui s'organisent en faisceau, sur le même modèle que ceux retrouvés dans le domaine central (Blake et al., 1995). Ce domaine a été décrit comme interagissant avec la syntrophine (Newey et al., 2000) et la dystrobrevine (Sadoulet-Puccio et al., 1997).

#### 4. Superfamille de la spectrine et utrophine

La dystrophine fait partie de la superfamille de la spectrine, qui regroupe les spectrines  $\alpha$  et  $\beta$ , l' $\alpha$ -actinine, ou l'utrophine. Les analyses phylogénétiques suggèrent que des réarrangements d'un ancêtre commun similaire à l' $\alpha$ -actinine (Byers et al., 1992; Dubreuil,

1991) ont donné d'une part les spectrines (Pascual et al., 1997), et d'autre part l'utrophine et la dystrophine (Dubreuil, 1991) (Figure 4).

La dystrophine, les spectrines et l'utrophine partagent la même organisation globale (un domaine ABD N-terminal, un long domaine central en faisceau d'hélices à trois hélices, deux domaines C-terminaux). Les états d'oligomérisations sont cependant différents (voir chapitre III et Figure 28). La comparaison de la structure de domaines homologues résolues expérimentalement montre une très faible divergence entre elles (par exemple similarité d'ABD1 N-terminal dystrophine avec la spectrine (Norwood et al., 2000)).

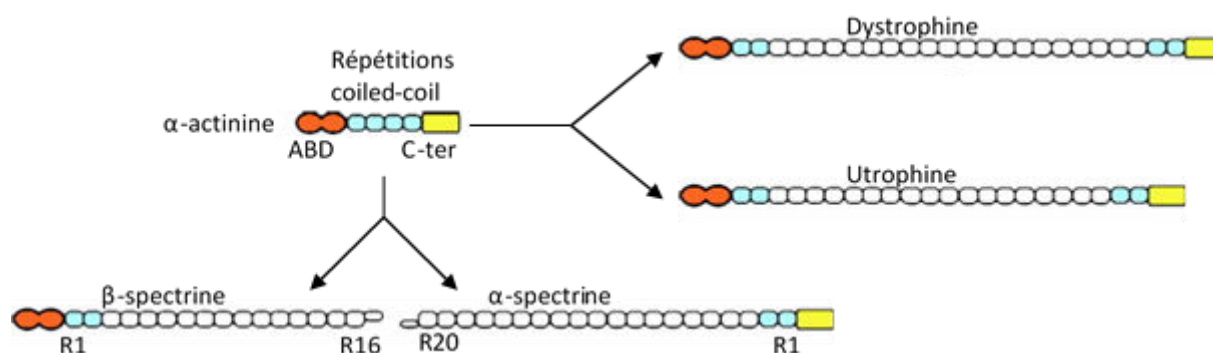


Figure 4 : Phylogénie des protéines de la famille de la spectrine. D'après Broderick et Winder, 2005

L'utrophine est une protéine de structure et de taille proche de celle de la dystrophine (respectivement 395 kDa et 427 kDa) qui a en commun certaines de ses fonctions avec celle-ci (Blake et al., 2002). L'utrophine est exprimée au niveau des jonctions neuro-musculaires des muscles matures. Elle est également retrouvée au niveau du sarcolemme dans les muscles en développement, avant d'être remplacée par la dystrophine à la naissance (Khurana and Davies, 2003). Une approche envisagée pour traiter les maladies liées à un défaut du gène codant la dystrophine est de sur-exprimer l'utrophine dans les muscles de patients atteints de myopathie de Duchenne, dans le but de restaurer une partie des fonctions normalement assurées par la dystrophine (Krag et al., 2004). C'est dans ce cadre que de nombreuses études s'attachent à comparer les fonctions de la dystrophine et de l'utrophine.

Peu de structures cristallographiques de ces domaines sont disponibles, notamment concernant les régions centrales (Delalande et al., 2017). En particulier, seules les répétitions R1 de l'utrophine et de la dystrophine ont été obtenues, sous une forme de dimère forcé par un pont disulfure. Pour les spectrines  $\alpha$  et  $\beta$ , on peut remarquer que les structures résolues

de répétitions de leurs domaines centraux se concentrent au niveau ou à proximité du domaine de tétramérisation (Tableau 1) (voir chapitre III et Figure 28).

<b>Protéine et source</b>	<b>Domaine</b>	<b>PDB ID</b>
Dystrophin (H. sapiens)	Répétition 1	3UUN
Utrophin (R. norvegicus)	Répétition 1	3UUL
Spectrin $\alpha$ (D. melanogaster)	Répétition 14	2SPC
Spectrin $\alpha$ (G. gallus)	Répétitions 15, 16, 17	1U4Q
Spectrin $\alpha$ (G. gallus)	Répétitions 15, 16	1U5P
Spectrin $\alpha$ (G. gallus)	Répétition 16	1AJ3
Spectrin $\alpha$ (G. gallus)	Répétitions 16, 17	1CUN
Spectrin $\alpha$ (G. gallus)	Site de tétramérisation	3F31
Spectrin $\alpha$ (H. sapiens)	Site de tétramérisation	1OWA
Spectrin $\alpha\beta$ (H. sapiens)	Site de tétramérisation	3LBX
Spectrin $\beta$ (H. sapiens)	Répétitions 8, 9	1S35
Spectrin $\beta$ (H. sapiens)	Répétitions 13, 14, 15	3KBT
Spectrin $\beta$ (H. sapiens)	Répétitions 14, 15	3F57, 3EDU
Spectrin $\beta$ (H. sapiens)	Répétitions 14, 15, 16	3EDV

Tableau 1 : Structures de répétitions en faisceau d'hélice de la famille de la spectrine disponibles dans la PDB ([www.rcsb.org](http://www.rcsb.org)).

## B. Partenaires de la dystrophine

### 1. Lipides

#### *a. Membranes et lipides membranaires*

Les membranes forment la limite entre l'intérieur et l'extérieur de la cellule. Elles constituent également un compartiment cellulaire à part entière qui assure plusieurs fonctions biologiques (communication avec l'environnement, transport actif et passif, fonctions métaboliques). Les membranes biologiques et les protéines qui leurs sont associées sont ainsi impliqués dans de nombreux processus biologiques, comme la neuro-transmission (Dore et al., 2017) ou la phagocytose (Rougerie et al., 2013).

Les membranes sont composées à 50% de lipides, mais également de protéines (glycoprotéines, récepteurs, canaux ioniques, transporteurs) et de glucides. Certaines protéines, dites intrinsèques ou intégrales, sont enchâssées dans la bicouche lipidique où elles sont capables d'établir des interactions hydrophiles/hydrophobes avec leur contrepartie lipidique. D'autres protéines, dites périphérique ou extrinsèques, peuvent aussi interagir de manière permanente ou transitoire avec les surfaces membranaires sans les traverser, sans posséder à la fois de régions extracellulaire et intracellulaire.

Parmi les lipides membranaires, les phospholipides représentent la classe la plus abondante au sein des membranes. Les phospholipides possèdent des propriétés amphiphiles (ou amphipathiques), c'est-à-dire qu'elles possèdent à la fois un groupement chimique hydrophobe et un groupement hydrophile. En milieu aqueux, ces lipides s'organisent spontanément en bicouches de manière à minimiser les surfaces de contact entre l'eau et leur région hydrophobe. De fait, les régions hydrophobes sont enfouies au centre de la bicouche et les groupes hydrophiles sont exposés à l'eau. Chez les eucaryotes, les phospholipides les plus retrouvés sont des phosphoglycérides (Figure 5A), comprenant entre autre la phosphatidylcholine (PC) et la phosphatidylserine (PS) (Figure 5B). Leur région hydrophobe est composée d'une ou plusieurs chaînes d'acides gras saturés (alcane linéaires  $-(CH_2)_n-CH_3$ ) (Figure 5C) ou cis-insaturés (présentant des doubles liaisons C=C) de tailles variables.

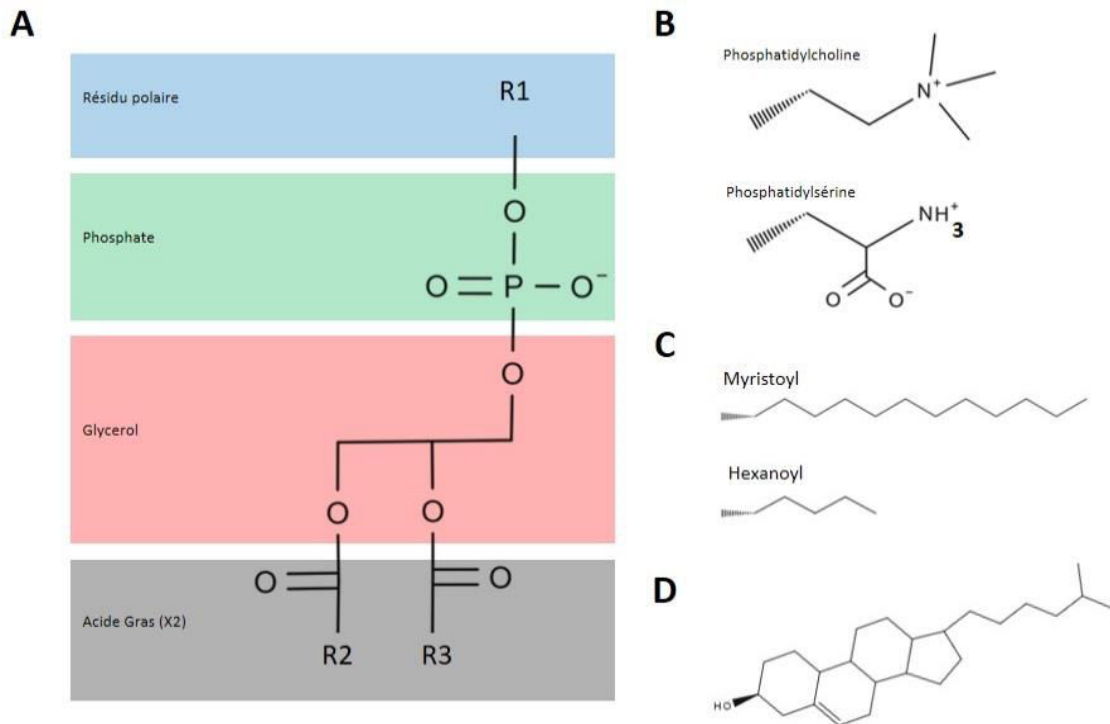


Figure 5 : A) Structure des phosphoglycérides ; B) Exemples de résidus polaires retrouvés dans les phosphoglycérides ; (Haut) Phosphatidylcholine (PC) ; (Bas) Phosphatidylsérine (PS). C) Exemples d'acides gras retrouvés dans les phosphoglycérides ; (Haut) Myristoyl ; (Bas) Hexanoyl.

Le cholestérol est également très présent au sein des membranes (Harayama and Riezman, 2018). Le cholestérol et ses dérivés sont impliqués dans des processus cellulaires en interagissant avec des protéines ou d'autres lipides membranaires. La structure du cholestérol (Figure 5D) lui confère des propriétés particulières. Sa présence entraîne une compaction des lipides avoisinants, augmentant localement la rigidité de la membrane (Griffiths et al., 2017; Ikonen, 2008).

Différents modèles de membranes lipidiques ont été développés pour décrire la structure et la dynamique de celles-ci. En 1972, Singer et Nicolson ont proposé le modèle de « mosaïque fluide » (Singer and Nicolson, 1972). À la température corporelle (~37°C), les lipides membranaires ont une consistance visqueuse, proche de celle des huiles végétales (tournesol, olive). Aussi, selon ce modèle, la membrane constitue un fluide en 2D, dans lequel les différents composants membranaires sont libres de se mouvoir. Plus tard, Kusumi et al. (Kusumi et al., 1993) présentent le modèle de « diffusion confinée ». Celui-ci suggère que la membrane fluide est compartimentée par des « barrières » composées par le réseau d'actine sous-membranaire. Ces deux composants (membrane et actine) sont liés par des « piquets »



de protéines trans-membranaires. Cette organisation de la membrane limiterait ainsi la diffusion de ses composants.

*b. Sarcolemme*

Le sarcolemme est la membrane plasmique spécialisée qui entoure les fibres musculaires. Il joue un rôle essentiel dans la structure et la fonction des muscles (Clark et al., 2002; Colognato and Yurchenco, 2000), en étant par exemple directement impliqué dans la transmission synaptique (Nishimune and Shigemoto, 2018), la propagation du potentiel d'action (Song et al., 2005), et le couplage excitation/contraction (Barrientos et al., 2017). Il est également impliqué dans le transport du glucose via les transporteurs GLUT (*Glucose Transporter*) (Klip et al., 1996) et régule de fait les apports énergétiques aux muscles. L'interaction simultanée du sarcolemme avec la dystrophine et l'actine du cytosquelette lui confère des propriétés mécaniques spécifiques, lui permettant de résister aux forces mécaniques induites lors de la contraction musculaire.

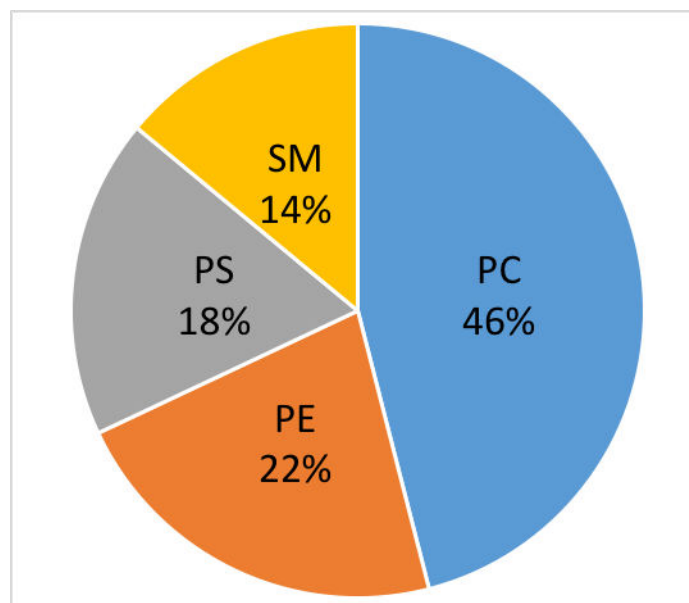


Figure 6 : : Composition globale en phospholipides du sarcolemme, d'après Fiehn et al, 1971. Étude réalisée à partir de muscle strié de rat. PC : Phosphatidylcholine ; PS : Phosphatidylsérine ; PE : Phosphatidyléthanolamine ; SM : Sphingomyéline;

D'après les mesures de compositions lipidiques de muscles de rats de Fiehn et al. (Fiehn et al., 1971), les différents types de phospholipides représentent 48% des lipides du sarcolemme (Figure 6). Le type PC (Phosphatidylcholine) représente environ 46%, les PE (Phosphatidyléthanolamine) 22% et ceux de types PS (Phosphatidylsérine) 18% des

phospholipides du sarcolemme. Les PS, chargés négativement, sont majoritairement présents au niveau du feuillet interne du sarcolemme (Leventis and Grinstein, 2010).

*c. Modèle bicelle*

Les bicelles sont un type de modèle membranaire utilisé pour l'étude des interactions protéines/lipides. Ce sont des structures discoïdales, composées de lipides à chaînes longues formant une bicouche plane, et de lipides à chaînes courtes formant le bord du disque (Figure 7). La taille des bicelles est définie par le ratio entre les lipides à chaînes longues et courtes, et leurs propriétés de surfaces sont conditionnées par les groupes formant les têtes polaires.

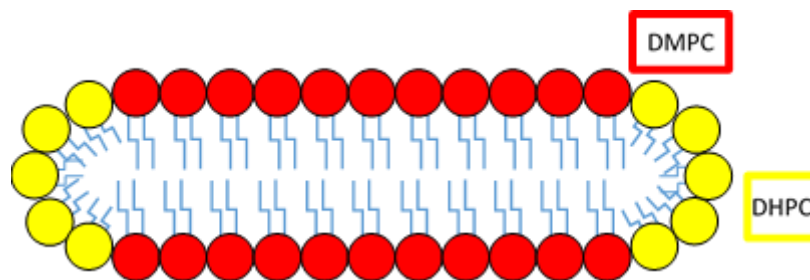


Figure 7 : Représentation schématique d'une bicelle. (rouge) Lipides à chaînes longues (DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine) ; (jaune) Lipides à chaînes courtes (DHPC, 1,2-dihexanoyl-sn-glycero-3-phosphocholine)

Par rapport à d'autres formations lipidiques (micelles, ...), les bicelles ont l'avantage de pouvoir imiter les propriétés physico-chimiques des membranes physiologiques, et ainsi de permettre l'étude des protéines membranaires (intégrales ou périphériques) dans des conditions proches de la physiologie, en particulier vis-à-vis de leur orientation par rapport à la membrane (Lau et al., 2008). Grâce à leur structure particulière, les bicelles permettent d'étudier l'interaction avec des protéines dans différents contextes de courbure et de compacité (*packing*), c'est-à-dire différents niveaux d'accessibilité du cœur hydrophobe de la membrane du fait des modulations de l'espacement entre les têtes polaires. En effet, certaines protéines interagissant avec les membranes sont sensibles à la courbure (Peter et al., 2004) tandis que d'autres peuvent l'influencer (Martens et al., 2007). Des variations locales de la composition lipidique peuvent ainsi modifier l'interaction protéine/membrane (Leikin et al., 1996). La courbure des membranes est donc un élément important à prendre en compte lors de l'étude de leurs interactions avec des protéines. Dans les cellules musculaires, les forces mécaniques que subissent les membranes varient au cours de la contraction et affectent la courbure et le *packing* des lipides.

*d. Interaction avec la dystrophine*

En se basant sur son homologie avec la spectrine, laquelle est présente à la membrane des érythrocytes (Maksymiw et al., 1987; Marchesi and Steers, 1968), il a été montré que la répétition R2 de la dystrophine pouvait interagir avec les phospholipides membranaires (DeWolf et al., 1997). Les auteurs suggèrent cependant que d'autres parties du domaine central pourraient également être impliquées dans des interactions avec la membrane. En 2003, les travaux de notre équipe d'accueil ont mis en évidence la proximité de résidus tryptophanes de la répétition R2 avec les têtes des lipides de vésicules phospholipidiques (DOPC (1,2-Dioléoyl-sn-glycéro-3-phosphocholine), DOPS (1,2-Dioléoyl-sn-glycéro-3-phosphoserine), DOPE (1,2-Dioléoyl-sn-glycéro-3-phosphoéthanolamine)) anioniques et zwitterioniques par des méthodes de spectrofluométrie (Le Rumeur et al., 2003).

La dystrophine est aussi suspectée d'être sensible à la courbure des membranes, un motif senseur de courbure prédit par des méthodes bioinformatiques à été localisé dans l'hélice B de la répétition R1 (Drin et al., 2007) (Voir III-Etude des interactions hydrophobes du faisceau d'hélice R1-3).

Notre équipe s'est attachée à caractériser l'interaction R2-lipides (Le Rumeur et al., 2007) et à évaluer la sensibilité de cette interaction à la courbure et au *packing* des lipides. Cette liaison R2-lipides est identifiée comme étant d'origine électrostatique (entre des résidus lysines ou arginines chargés, et les têtes chargées de la phosphatidylsérine) et hydrophobe (entre les groupes apolaires de certains acides aminés comme le tryptophane et les queues hydrophobes des lipides). Cette approche a été étendue aux fragments contenant les répétitions R1 à R3 (R1-3) et R20 à R24 (R20-24), tous deux étant assimilables à des sous-domaines indépendants car étant bordés par les charnières H1/H2 et H3/H4 respectivement (Legardinier et al., 2008). Il est montré que le domaine R1-3 interagit fortement avec des lipides anioniques, à l'opposé de R20-24 qui ne semble pas avoir d'affinité pour les membranes. Un second article de notre équipe, publié en 2009, met en évidence l'affinité globale de la région R4-19 envers les phospholipides anioniques et propose la définition de deux domaines de liaisons aux lipides (*lipid binding domains*) LBD1 et LBD2, constitués respectivement des fragments R1-3 et R4-19 (Legardinier et al., 2009). En 2010, des travaux montrent le caractère amphiphile de R1-3 et de R20-24 tout en confirmant l'affinité de R1-3 pour les lipides, contrairement à R20-24 (Vié et al., 2010).

Les travaux effectués au sein de notre équipe ont également montré le caractère amphiphile de R11-15 (comme R1-3 et R2) (Sarkis et al., 2011). De plus, la liaison de R11-15 aux lipides dépend de la courbure membranaire et/ou du *packing* des lipides (liaison à des SUV (*small unilamellar vesicles*) anioniques de 30 nm, mais pas aux LUV (*large unilamellar vesicles*) de 100 nm ou plus) et du *packing*. Ainsi, le fragment R11-15 montre une affinité supérieure pour des membranes zwitterioniques présentant des défauts de compaction composées de DOPC/POPE que pour des membranes anioniques DOPC/POPS plus compactes. De plus, alors que la région R1-19 du domaine central est capable de lier les membranes anioniques, R11-15 est le seul fragment qui peut également interagir avec des membranes zwitterioniques.

En 2013, il a été montré au sein de l'équipe, par des expériences de cisaillement en cuve de Langmuir, que le fragment R11-15 de la dystrophine est capable de former un pont entre l'actine-F (filamenteuse) et les lipides membranaires, augmentant drastiquement la résistance des membranes aux forces de cisaillement (Sarkis et al., 2013). *In vivo*, ce lien pourrait contribuer à la protection du sarcolemme contre les dommages mécaniques. L'équipe a également montré que R16-21 s'enfonce dans des membranes phospholipidiques contenant du cholestérol lorsque la tension de la surface lipidique est proche de la valeur physiologique (30 miliNewtons/m) (Ameziane-Le Hir et al., 2014).



Figure 8 : Représentation des zones d'interactions dystrophine / Lipides. (vert) Interactions avec des membranes anioniques ; (bleu) Interactions avec des membranes zwitterioniques.

## 2. Protéines

### *α. nNOS*

Les Oxyde nitrique synthases (NOS, Nitric oxide synthase) forment une famille d'enzymes (Expasy Enzyme ID EC 1.14.13.39) dont la fonction est de produire du monoxyde d'azote à partir d'arginine (Marletta, 1993). Le monoxyde d'azote est la plus petite molécule impliquée dans des processus de signalisation cellulaire. Elle est impliquée dans de nombreux processus cellulaires comme la réponse immunitaire (Green et al., 1990) ou la neurotransmission (Vincent, 2010).

Les NOS neuronales (nNOS) sont exprimées au niveau du système nerveux central et périphérique, dont l'isoforme nNOS $\mu$  qui est retrouvée dans les muscles. Le monoxyde d'azote est un messager permettant la vasodilatation des vaisseaux sanguins durant l'effort musculaire (Endo et al., 1994).

Dans la cellule musculaire, nNOS $\mu$  est retrouvée à la membrane. Elle co-localise avec la dystrophine (Gentil et al., 2012) au niveau des costamères (bandes Z) (Figure 11). Il a été montré que l'interaction de ces deux protéines s'effectue via le domaine PDZ de nNOS, et les répétitions R16 et R17 de la dystrophine. Lai et al. ont par ailleurs identifié les sous-domaines de R16 et R17 nécessaires à la formation du complexe (Lai et al., 2009, 2013). En se basant sur cette cartographie, notre équipe a publié en 2013 une analyse théorique du fragment R16-17 permettant d'expliquer le mode d'interaction entre les partenaires (Giudice et al., 2013). Enfin, dans une étude de 2015, des expériences d'*alanine scanning* ont identifié les résidus impliqués dans l'interaction entre R16-17 et nNOS, et ont permis la construction d'un modèle de complexe basé sur ces données (Molza et al., 2015).

#### *b. Syntrophine*

Les syntrophines ( $\alpha$ ,  $\beta$  et  $\delta$ ) sont les membres d'une famille de protéines associées au sarcolemme (Piluso et al., 2000). Ces protéines jouent un rôle important dans la localisation de nNOS à la membrane (Adams et al., 2018), mais aussi dans la transduction de signaux entre l'extérieur et l'intérieur de la cellule (Bhat et al., 2013). Récemment, Adams et al. ont caractérisé biochimiquement l'interaction entre le domaine R16-17 de la dystrophine, la syntrophine  $\alpha$  et la nNOS (Adams et al., 2018).

#### *c. Complexe dystrophine-Glycoprotéine*

Le complexe dystrophine-glycoprotéine (DGC, *dystrophin glycoprotein complex*) est un complexe protéique situé au niveau des costamères des fibres musculaires. Ce complexe a simultanément un rôle dans la stabilisation de la membrane, dans la transduction de voies de signalisations cellulaires et dans la transmission des forces mécaniques entre l'intérieur et l'extérieur de la fibre musculaire.

Le DGC est composé de nombreuses protéines (Ervasti, 2013), parmi lesquelles la dystrophine, la protéine nNOS, les sarcoglycanes ( $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\gamma$ ,  $\epsilon$  et  $\zeta$ ), les dystrobrevines ( $\alpha$ ,  $\beta$ ), les dystroglycane ( $\alpha$ ,  $\beta$ ), syntrophines ( $\alpha$ ,  $\beta$ ,  $\delta$ ), ... (Lapidos et al., 2004).

Les composants du DGC assurent des fonctions très diverses. Par exemple, le dystroglycane connecte le cytosquelette à la matrice extra-cellulaire (Henry and Campbell, 1999). De la même manière, les six membres des sarcoglycanes, qui sont des protéines transmembranaires, semblent impliqués dans l'imperméabilisation du sarcolemme (Hack et al., 1999).

#### *d. Actine*

L'actine est un composant majeur du cytosquelette très conservé dans l'évolution. On retrouve ainsi des homologues chez les bactéries, qui sont également capables de former des filaments (Carballido-López, 2006). Chez les mammifères, elle est impliquée dans de nombreuses fonctions cellulaires, comme la motilité cellulaire (Pollard and Borisy, 2003; Pollard and Cooper, 2009), le maintien de la forme (Pollard and Cooper, 2009) et de la polarité des cellules (Dalous et al., 2008) ou encore la régulation de la transcription (Miralles and Visa, 2006).

L'actine existe sous deux états : le monomère d'actine G (globulaire) et le filament (actine-F), qui se forme lors de la polymérisation de l'actine-G (Holmes et al., 1990). L'actine-G est une protéine de 375 acides aminés (42 kDa), qui se replie en deux domaines  $\alpha/\beta$ , eux-mêmes composés des sous-domaines 1 et 2 ( $\alpha$ ) et 3 et 4 ( $\beta$ ) (Otterbein et al., 2001) (Figure 9). Ces deux domaines établissent peu de contacts entre eux, et la chaîne peptidique les reliant forme un axe autour duquel ils peuvent légèrement pivoter.

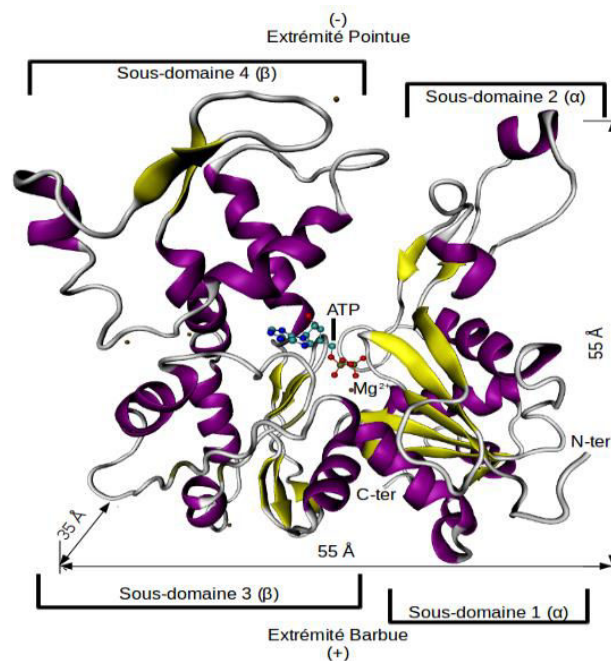


Figure 9 : Structure du monomère d'actine-G, d'après de Otterbein et al, 2001.

On trouve entre les sous-domaines 1 et 2 d'une part, et les sous-domaines 3 et 4 d'autre part, une crevasse délimitant un site d'hydrolyse de l'ATP, ainsi qu'un site de liaison à un ion divalent ( $Mg^{2+}$  dans le contexte cellulaire) nécessaire à l'hydrolyse du nucléotide (Dominguez and Holmes, 2011). Les sous-domaines 1 et 3 sont structurellement proches, et proviennent probablement d'un événement de duplication génique, alors que les sous-domaines 2 et 4

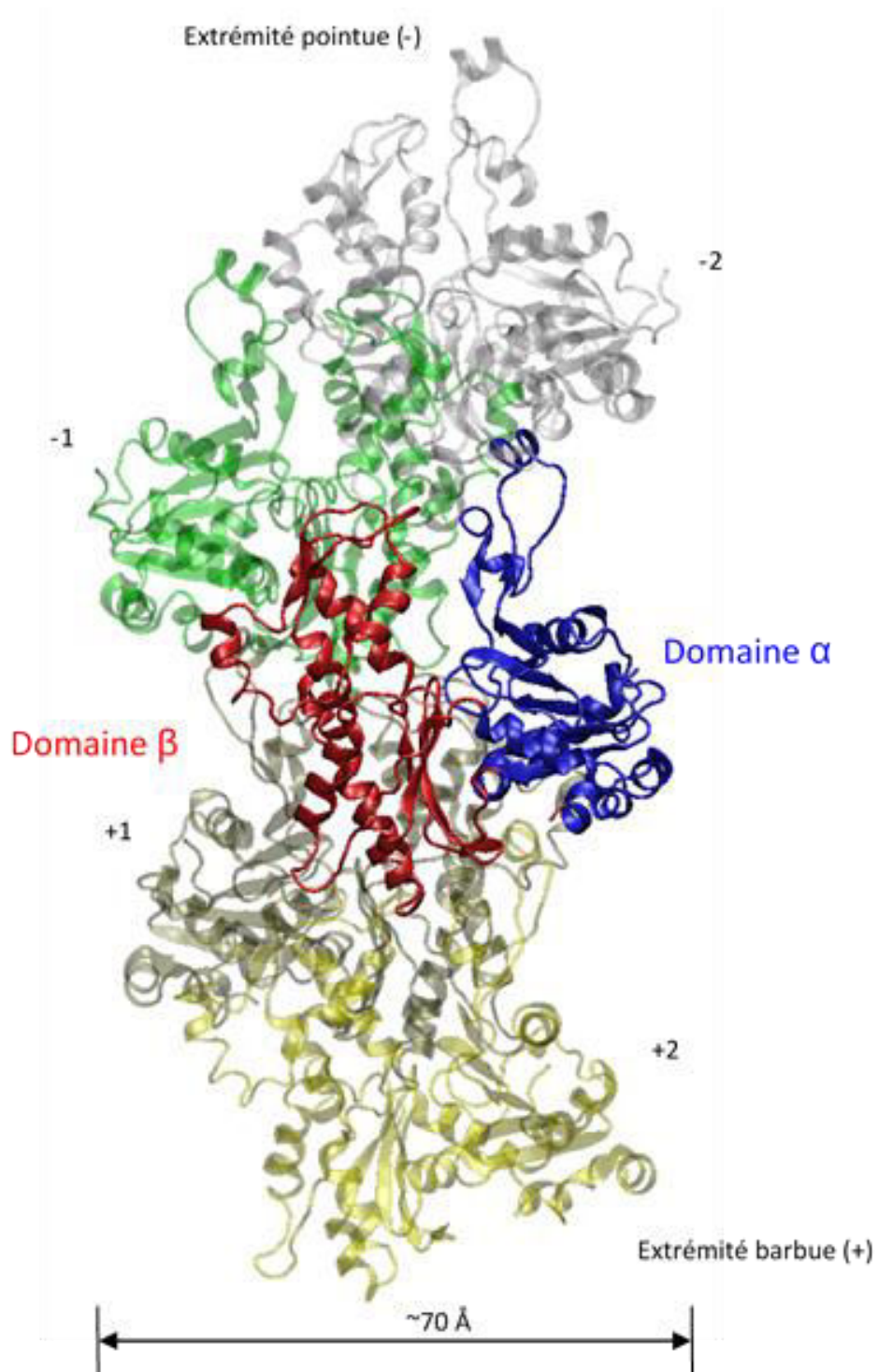


Figure 10 : Structure d'un filament d'actine (PDBID 3G37) composé de cinq monomères ; (bleu) domaine  $\alpha$ , composé des sous-domaines 1 et 2 ; (rouge) domaine  $\beta$ , composé des sous-domaines 3 et 4.

peuvent être perçus comme de larges insertions dans les domaines 1 et 3 (respectivement). Le monomère d'actine peut être considéré comme un pavé droit de dimension 55 Å x 55 Å x 35 Å. Le site d'hydrolyse de l'ATP de l'actine-G n'est pas fonctionnel, mais le devient lors de la polymérisation (Figure 9).

L'observation par microscopie électronique de fibres d'actine colorées négativement montre une structure en double-hélice droite (Cohen and Hanson, 1956; Hanson and Lowy, 1964). En réalité, les monomères s'associent en quinconce (étendant tour à tour chacune des hélices droites) avec une rotation de  $166,6 \pm 0,6^\circ$  et une translation de 27,6 Å, formant une simple hélice gauche, avec approximativement 13 molécules tous les 6 tours (Figure 10).

L'actine-F est un polymère polarisé, dont l'une des extrémités (-) correspond aux domaines 2 et 4 de l'actine G, et l'autre extrémité (+) aux domaines 1 et 3. Cette polarisation est associée à différences cinétiques d'incorporation des monomères aux deux extrémités : l'extrémité + (ou "barbue") s'allonge cinq à dix fois plus vite que l'extrémité - (ou "pointue"). Il en résulte qu'à une concentration d'actine-G intermédiaire l'extrémité + s'allonge tandis que l'extrémité - dépolymérise. Ces deux phénomènes se déroulant à la même vitesse, le filament conserve une taille constante, les sous-unités « voyageant » de l'extrémité barbue à l'extrémité pointue par un mécanisme de tapis roulant.

#### i. Fonctions

La localisation ubiquitaire de l'actine est à mettre en regard de son rôle essentiel dans un grand nombre de processus cellulaires vitaux. Dans le système nerveux, l'actine dynamique est un élément nécessaire à l'activation de boutons pré-synaptiques silencieux (Shen et al., 2006; Yao et al., 2006). Du côté post-synaptique, l'actine est essentielle au trafic et au positionnement des récepteurs aux neurotransmetteurs (Kirsch and Betz, 1995; Kuriu et al., 2006).

Dans le muscle, on retrouve deux types d'actine-F, aux fonctions différentes. L'actine « architecturale », commune à toutes les cellules de l'organisme, a pour fonction d'organiser le cytoplasme. Le second type d'actine-F, « contractile », est spécifique au muscle, et a pour fonction de produire du mouvement, en association avec la myosine (Figure 11) (Richfield, 2014). Le complexe actine-F/myosine forme des filaments, eux-mêmes organisés en sarcomères, qui constituent l'unité de contraction du muscle. Les sarcomères mesurent



environ 2,4  $\mu\text{m}$  de long, et sont délimités à chaque extrémité par des disques Z. L'extrémité + des filaments d'actine est fixée aux disques Z et leur extrémité – est orientée vers le centre du sarcomère. Au centre des sarcomères, on trouve la bande M, qui organise la myosine en vis-à-vis des filaments d'actine.

Pour assurer l'ensemble de ses fonctions, la dynamique de polymérisation/dépolymérisation de l'actine nécessite une régulation très fine, à la fois sur le plan spatial et temporel. Le recrutement de centres nucléateurs de l'actine ciblant certains compartiments subcellulaires permet de générer des structures d'actines spécifiques.

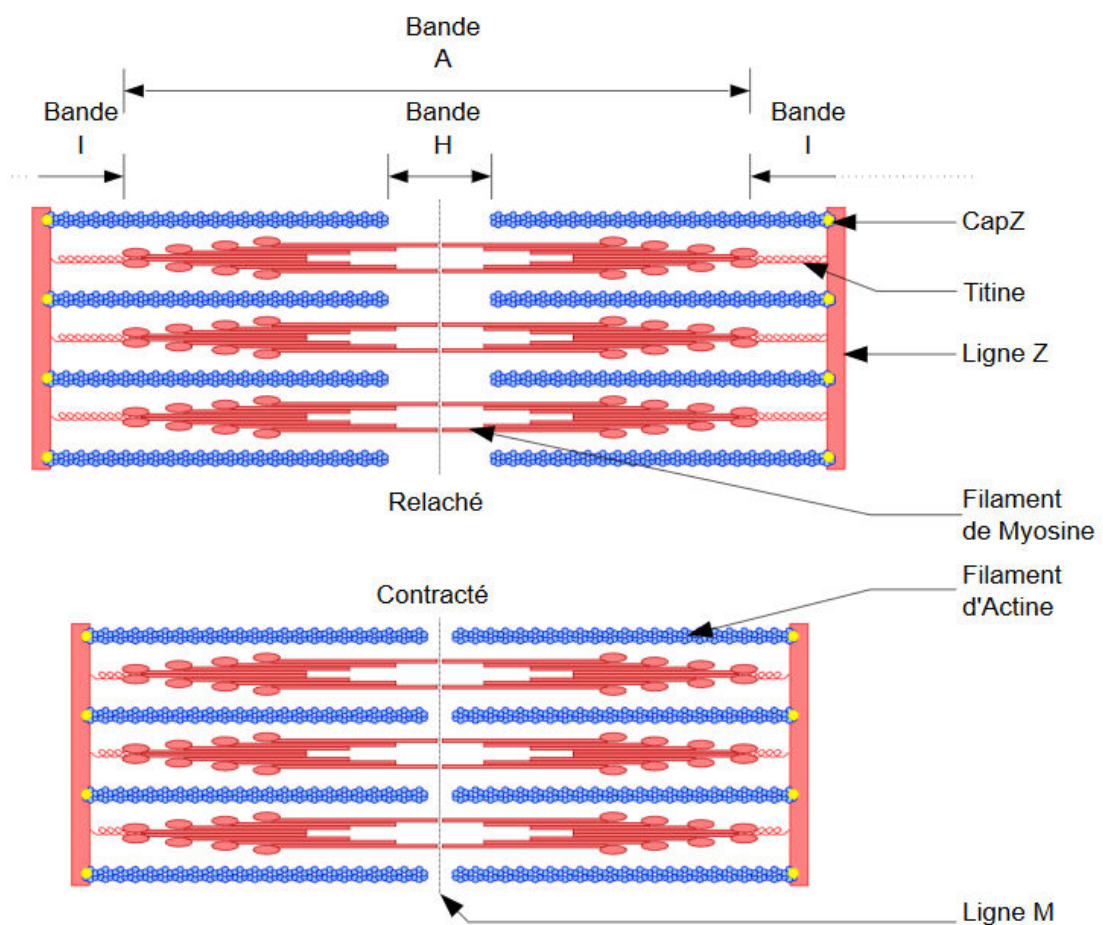


Figure 11 : Structure d'un sarcomère dans différents états de contraction musculaire. Tiré de Richfield, 2014.

## ii. Isoformes et modifications post-traductionnelles

La diversité des fonctions biologiques impliquant l'actine est possible grâce à la présence chez les mammifères de six gènes, chacun codant une isoforme. Quatre isoformes sont exprimées dans les muscles squelettiques, cardiaques et lisses ( $\alpha$ -squelettique,  $\alpha$ -cardiaque,  $\alpha$ -lisse, et  $\gamma$ -lisse), et les deux autres sont exprimées de manière ubiquitaire dans l'organisme

( $\beta$ -cytoplasmique et  $\gamma$ -cytoplasmique). Toutes partagent au moins 93 % d'identité entre elles (Rubenstein, 1990), leurs principales différences résidant au niveau de l'extrémité N-terminale (Herman, 1993).

Les différentes isoformes ont des fonctions distinctes. À titre d'exemple, dans le muscle, les isoformes  $\alpha$ -cardiaque et  $\alpha$ -squelettique sont retrouvées dans les filaments fins, responsables de la contraction musculaire, alors que l'actine- $\gamma$ -cytosolique forme le cytosquelette cortical au niveau de la membrane plasmique (Rybakova et al., 2000).

L'actine subit diverses modifications post-traductionnelles, notamment des méthylations stabilisant le filament (Nyman et al., 2002; Yao et al., 1999, 2002), des acétylations facilitant l'interaction actine/myosine dans le muscle (Abe et al., 2000), ou des phosphorylations de plus de 35 sites facilitant l'élongation (Ohta et al., 1987; Prat et al., 1993).

### iii. Actine et Dystrophine

La dystrophine contient deux domaines de liaison à l'actine (ABD1 et ABD2). ABD1 est situé dans le domaine N-terminal de la protéine et ABD2 a été localisé au milieu du domaine central, d'abord entre les répétitions R11 et R17 (Amann et al., 1999). En 2013, des travaux menés dans notre équipe montrent que la présence de R11-15 est suffisante pour se lier à l'actine (Sarkis et al., 2013). En se liant à la fois à l'actine et aux lipides, la région ABD2 augmente la résistance des membranes aux cisaillements. La dystrophine se lie au réseau d'actine costamérique, et pas contractile (Rybakova et al., 2000), avec une constante d'affinité inférieure au micro-molaire (Winder et al., 1995) (voir chapitre *IV-Modélisation de complexes de la dystrophine (actine, microtubules, lipides)*).

## e. Microtubules

### i. Introduction

Les microtubules constituent un des composants majeurs du cytosquelette eucaryote, et sont impliqués dans nombre de processus essentiels (division cellulaire, transport intracellulaire, etc).

Les microtubules sont des tubes creux de 25 nm de diamètre, composés de 13 protofilaments. Chaque protofilament est lui-même composé de l'assemblage polarisé d'hétérodimères de tubuline  $\alpha$  et  $\beta$ , deux isoformes de tubulines proches, qui fixent le GTP. Cette organisation définit la polarité des microtubules. Ceux-ci sont en effet composés d'une

extrémité « moins » (-), où sont exposées des sous-unités  $\alpha$ , et d'une extrémité « plus » (+), où sont exposées des sous-unités  $\beta$  (Figure 12). L'orientation des microtubules qui en résulte permet à des moteurs moléculaires de se déplacer le long du filament vers l'extrémité + (transport antérograde, p.ex. kinésines) ou - (transport rétrograde, p.ex. dynéines).

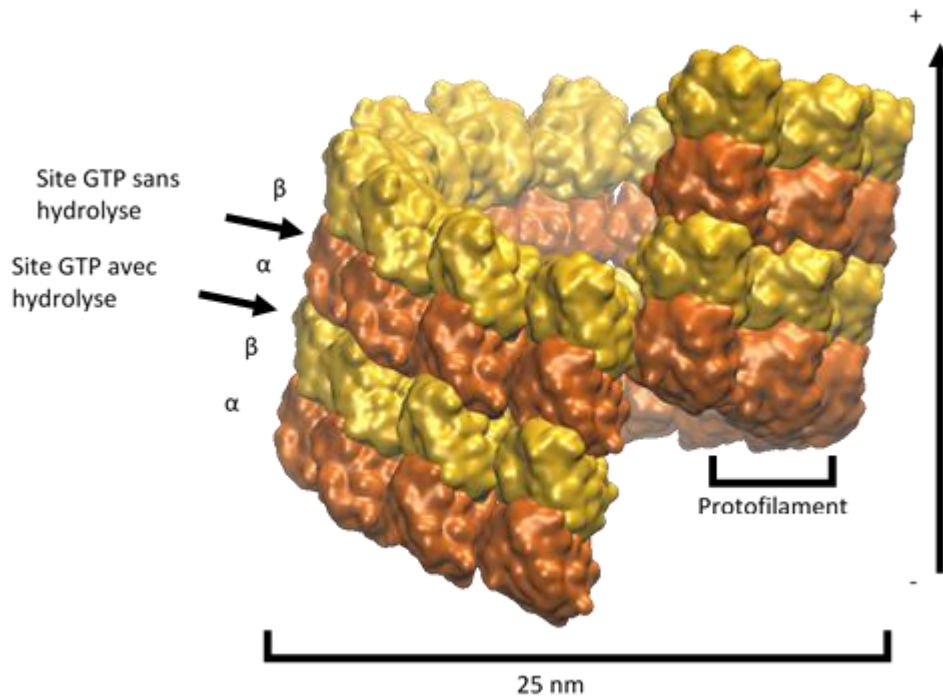


Figure 12 : Organisation des monomères de tubulines dans les microtubules. (Orange) Isoforme  $\alpha$  (Jaune) Isoforme  $\beta$ .

*In vitro*, les microtubules peuvent se former spontanément, à partir d'une certaine concentration en hétérodimères  $\alpha/\beta$ . Une des particularités de l'extrémité + des microtubules est leur propriété d'instabilité dynamique, décrite et nommée par Mitchison et al. (Mitchison and Kirschner, 1984) : en présence de GTP, les phases de croissances sont interrompues par des événements de dépolymérisation rapides, appelés catastrophes, elles-mêmes interrompues par des sauvetages permettant la reprise de la polymérisation (Wade, 2009). Ainsi, il a été montré que les propriétés dynamiques et la taille moyenne d'une population de microtubules pouvaient être décrites via quatre paramètres (Mitchison and Kirschner, 1984) : les vitesses de croissance et de dépolymérisation et les fréquences des catastrophes et des sauvetages. De plus, l'extrémité - étant beaucoup plus stable, la polarisation des microtubules influence aussi la différence de dynamique aux deux extrémités.

A partir d'un microtubule préformé, la croissance se fait par l'élongation de plusieurs protofilaments adjacents formant un feuillet courbé vers l'extérieur. A partir de six protofilaments, le feuillet se redresse en commençant à former la paroi du tube, qui se fermera avec l'élongation du treizième filament.

Dans des conditions physiologiques, la polymérisation de la tubuline repose principalement sur l'existence préalable d'un microtubule, l'initiation de microtubules *de novo* étant un évènement très rare. Dans la cellule, on retrouve des complexes, les centrosomes, qui servent de centre de nucléation aux microtubules (Petry et al., 2013). Il est à noter que ces centres de nucléation sont enrichis en isoforme de tubuline  $\gamma$ .

Le GTP est incorporé à l'interface entre toutes les sous unités  $\alpha$  et  $\beta$ . Cependant, alors que le GTP situé entre les deux sous-unités au sein d'un hétérodimère  $\alpha/\beta$  n'est ni échangé ni hydrolysé, l'hydrolyse du nucléotide portée par la sous-unité  $\beta$  localisé entre deux dimères est

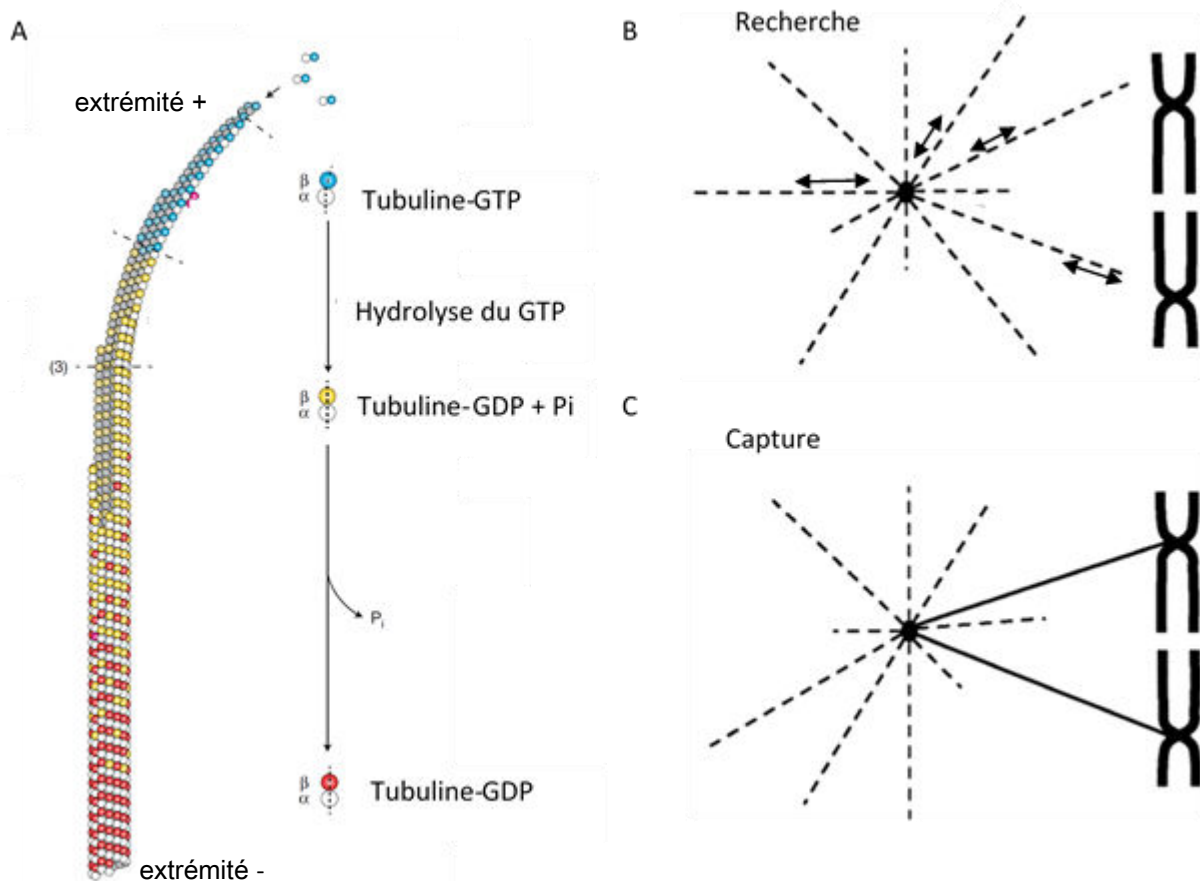


Figure 13 : (A) Dynamique d'incorporation et d'hydrolyse du GTP dans les microtubules, adapté de Guesdon et al., 2015. :le GTP intégré dans les dimères intégrés à l'extrémité + est progressivement hydrolysé en GDP lors de l'élongation du microtubule (B et C)Principe du mécanisme « recherche et capture » (search and capture). (B) Recherche : La dynamique de polymérisation/dépolymérisation des microtubules permet d'explorer l'espacejusqu'à (C) Capture : les microtubules ont été « capturés » par des +TIPs (Plus-end Tracking protein) associé à certaines structures cellulaires (centromères, membranes, ...). Adapté de Gaoglio et al. ; 1997.

essentielle à la dynamique des microtubules. *In vitro*, dans un milieu contenant du GMCPP (guanylyl-( $\alpha$ ,  $\beta$ )-methylene-diphosphonate), un analogue du GTP à la cinétique d'hydrolyse ralentie, les microtubules présentent une stabilité accrue.

L'hydrolyse du GTP par la sous-unité  $\beta$  survient après l'incorporation du dimère de tubuline à l'extrémité + en croissance (Figure 13). Ainsi, le « corps » du microtubule est composé de tubuline-GDP, et est protégé par une « coiffe » de tubuline-GTP à l'extrémité +. L'hydrolyse du GTP fragilise l'assemblage avec les dimères voisins, et lorsqu'elle se produit plus vite que l'incorporation de tubuline-GTP elle est à l'origine de catastrophes (Alushin et al., 2014).

Pour comprendre l'intérêt de ce phénomène d'instabilité dynamique, qui consomme de l'énergie « à vide », il a été proposé un mécanisme de recherche-et-capture (*search-and-capture*) (Gaglio et al., 1997; Mimori-Kiyosue and Tsukita, 2003). Dans ce modèle, le cycle polymérisation/dépolymérisation permet d'explorer l'environnement cellulaire. Les microtubules sont ensuite fixés aux composants cellulaires grâce à des +TIPs (*plus-end tracking protein*) qui leurs sont associés, comme les centromères de chromosomes.

Le besoin de spécialisation des microtubules est assuré d'une part par l'interaction avec une grande gamme de protéines associées (MAP, *Microtubules Associated Proteins*), stabilisant, déstabilisant et connectant les microtubules ou générant des forces (moteurs). D'autre part, l'expression différentielle des différents gènes codant les tubulines associé au vaste choix de modifications post-traductionnelles permet de contrôler finement l'organisation du cytosquelette de microtubules, appelé « code tubuline » (Gadadhar et al., 2017).

#### ii. Gènes, Isoformes et isotypes de la tubuline-G

Il existe au moins 7 isoformes de tubuline chez les eucaryotes ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  et  $\eta$ ) (Chang and Stearns, 2000; Dutcher, 2001; Dutcher and Trabuco, 1998; Oakley and Oakley, 1989; Ruiz et al., 2000; Vaughan et al., 2000; Wise et al., 2000), mais seulement cinq ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ) ont été identifiées chez l'Homme. Chaque isoforme est codée par plusieurs gènes, menant à la production de plusieurs protéines, appelées isotypes. Chez les mammifères, il a été ainsi identifié 9 isotypes de tubuline- $\alpha$  et 9 de tubuline- $\beta$ , 2 isotypes de tubuline  $\gamma$ , et une seule pour  $\delta$  et  $\epsilon$  (HUGO ID : TUB).

Les microtubules sont principalement composés par l'association des isoformes  $\alpha$  et  $\beta$  de la tubuline. Dans une cellule donnée, les microtubules sont constitués d'une mosaïque d'isotypes, et on ne peut pas associer à un isotype une fonction spécifique, mais certains microtubules aux fonctions spécialisés sont enrichies en certaines isotypes de l'isoforme  $\beta$  (axonèmes (Raff et al., 2008), microtubules neuronaux (Joshi and Cleveland, 1989), plaquettes (Lecine et al., 2000; Schwer et al., 2001)).

### iii. Modifications post-traductionnelles

Le second composant du « code tubuline » est représenté par les modifications post-traductionnelles. Parmi celles-ci, la (dé)tyrosination, la (poly)glutamylolation et la poly(glycylation) ont été d'abord décrite comme spécifique au microtubule, et à l'heure actuelle, cela reste vrai pour la tyrosination de la tubuline- $\alpha$  via la TTL (*Tubuline Tyrosine Ligase*) (Prota et al., 2013). Les deux autres ont été plus tard retrouvées sur d'autres substrats (van Dijk et al., 2008). D'autres modifications communes (méthylation, phosphorylation, etc.) sont aussi retrouvées sur les microtubules, influant sur la dynamique de polymérisation (Liu et al., 2015; Park et al., 2016).

### iv. Fonctions

Les microtubules ont une fonction dans un grand nombre de processus cellulaires essentiels, dont nous ne présenterons que quelques exemples dans ce chapitre. Durant la migration cellulaire, comme celle qui a lieu durant l'embryogenèse, le mouvement est induit par le complexe actine/myosine. La migration est supportée par le réseau de microtubules (Ganguly et al., 2012; Liao et al., 1995; Vasiliev et al., 1970) qui coordonne les différents aspects du processus (Kaverina and Straube, 2011; Watanabe et al., 2004; Yoo et al., 2012).

Pendant la mitose, les microtubules s'amarrent aux centromères des chromosomes via un ensemble de protéines, associées aux microtubules ou aux chromosomes, en formant le fuseau mitotique. Plus tard, et de manière synchronisée, les microtubules dépolymérisent, séparant les deux chromosomes homologues (Waters et al., 1996).

Enfin, dans les cellules musculaires, en plus de leur rôle d'échafaudage, les microtubules permettent le transport de myosine et l'assemblage correct des sarcomères (Pizon et al., 2005). Ils ont également un rôle majeur dans le positionnement des noyaux dans les myotubes (Folker and Baylies, 2013).

#### v. Dystrophine et Microtubules

Les connaissances scientifiques actuelles concernant le partenariat dystrophine-microtubules sont très limitées. Seulement deux publications traitant de l'interaction directe entre ces protéines étaient disponibles au début de cette thèse (Belanto et al., 2014; Percival et al., 2007), une troisième étant parue courant 2016 (Belanto et al., 2016). Il en sera question en détail dans le chapitre IV.

L'existence d'un partenariat directe entre la dystrophine et les microtubules est encore débattue. En particulier, des analyses protéomiques du réseau d'interaction protéine-protéine de la dystrophine montrent que cette interaction est indirecte (Murphy and Ohlendieck, 2016; Murphy et al., 2018)

## C. Myopathie de Duchenne

### 1. Historique

Les premiers cas de dystrophie musculaire ont été rapportés dès 1852 par Édouard Meryon (Meryon, 1852) chez quatre jeunes garçons présentant des faiblesses musculaires, mais sans aucune altération du système nerveux central. Meryon suspecte alors un désordre héréditaire porté par la mère mais qui n'affecterait que les garçons. Des observations similaires ont été faites à la même époque par William John Little (Little et al., 1853) pour deux frères atteints de la même pathologie.

A partir de 1861, les observations de Guillaume Duchenne de Boulogne permirent de définir l'origine musculaire de ces troubles (Duchenne, 1868). Il rapporte également un remplacement du tissu musculaire par du tissu fibreux ou conjonctif. En parallèle, il présente les premières approches thérapeutiques (massages, hydrothérapie, stimulation électrique) qui permettent d'améliorer l'état du patient dans les phases précoces de la maladie. Son nom sera associé à la forme la plus sévère de dystrophie musculaire, la myopathie de Duchenne (*Duchenne Muscular Dystrophy*, DMD).

En 1955, le professeur Becker décrit une forme moins sévère de myopathie mais d'origine identique à la DMD : la myopathie de Becker (*Becker Muscular Dystrophy*, BMD) (Becker and Kiener, 1955).

L'origine héréditaire de ces pathologies est suspectée par Meryon depuis le milieu du 19ème siècle. Le gène responsable de ces maladies a été identifié en 1986 (Koenig et al., 1987; Monaco et al., 1986) et la protéine décrite l'année suivante (Hoffman et al., 1987a).

### 2. Muscle

#### a. Rôle et Organisation

Les muscles constituent un tissu dynamique et plastique du corps humain. Ils assurent des fonctions essentielles, comme la respiration, la circulation sanguine, la motricité, le maintien de la posture ou les mouvements oculaires. Il existe différents types de muscles : les muscles cardiaques qui permettent le fonctionnement du cœur, les muscles lisses qui sont retrouvés autour d'organes (vessie, vaisseaux sanguins, etc) et qui sont autonomes, et les muscles squelettiques qui répondent au contrôle volontaire du système nerveux central.



Les muscles squelettiques représentent entre 20 et 40% de la masse corporelle selon les individus, et contiennent 50 à 75% de toutes les protéines du corps.

Histologiquement, les muscles sont entièrement entourés et séparés du reste de l'organisme par une membrane, l'épimysium. Le périmysium regroupe un faisceau de fibres musculaires, chacune enveloppée par l'endomysium (Figure 14).

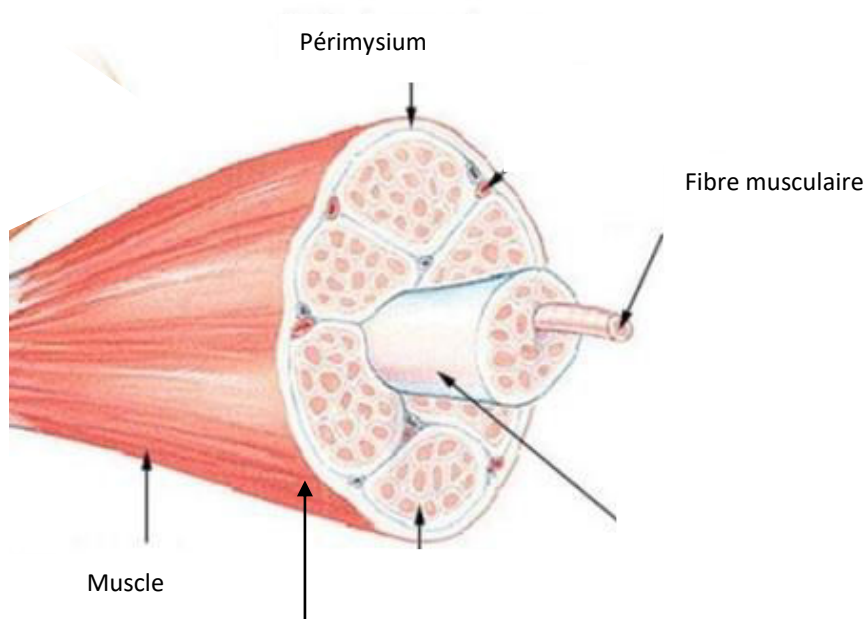


Figure 14 : Organisation structurale du muscle strié squelettique. (Adapté de [ressources.unisciel.fr/physiologie/co/2a\\_1.html](http://ressources.unisciel.fr/physiologie/co/2a_1.html))

Les fibres musculaires, ou myocytes, sont les unités fonctionnelles du muscle. Elles peuvent mesurer jusqu'à 10 mm de long pour 0,1 mm de diamètre. Ces fibres ont la particularité d'être un syncytium poly-nucléé, issu de la fusion de cellules individuelles, où chaque noyau contrôle une région spécifique de la fibre.

La membrane plasmatique du myocyte est appelée sarcolemme. On y retrouve des complexes protéiques qui permettent de connecter physiquement entre eux la matrice extracellulaire, la membrane et le cytosquelette du myocyte. Des dysfonctionnements au niveau de ces complexes, comme ceux retrouvés chez les patients DMD et BMD, provoquent des ruptures du sarcolemme qui devient perméable aux protéines et aux ions.

Le sarcolemme pénètre dans la cellule en formant les tubules transverses (ou tubules T), une structure très spécialisée qui contient un grand nombre de canaux ioniques (Figure 15).

Chaque tubule T est entouré de deux citernes terminales, qui sont des extensions du réticulum sarcoplasmique. L'une des fonctions principales du réticulum sarcoplasmique est de séquestrer le calcium du myocyte.

Les fibres musculaires sont séparées les unes des autres par une couche de matrice extracellulaire, qui contient aussi un petit réservoir de cellules souches satellites. Ce sont des cellules souches multipotentes, qui peuvent au besoin fournir de nouvelles fibres musculaires, en cas de blessure par exemple.

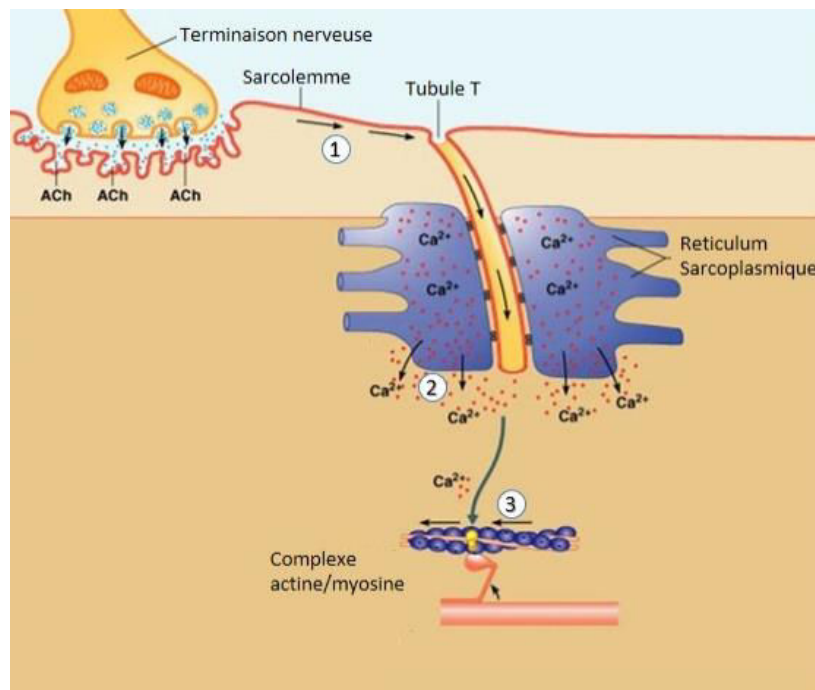


Figure 15 : Principe du couplage Excitation/Contraction. 1) Conduction du potentiel d'action le long du sarcolemme et dans les tubules T jusqu'à proximité des citernes terminales du réticulum sarcoplasmique. 2) Libération d'ion calcium du réticulum sarcoplasmique dans le sarcoplasme. 3) Contraction des sarcomères.

### b. Fonctionnement et physiologie

Lorsqu'un potentiel d'action (PA) est déclenché au niveau de la fibre musculaire par la libération d'acétylcholine d'un motoneurone (Figure 15), le PA est conduit à l'intérieur de la fibre via les tubules T, à proximité des citernes terminales du réticulum sarcoplasmique qui stockent le calcium. Des protéines sensibles au voltage (récepteur à la dihydropyridine) sur le tubule T sont activées par le PA (Ursu et al., 2001). Une fois activées, elles déclenchent une

entrée d'ions calcium extracellulaire dans le cytoplasme. Ce courant calcique déclenche les récepteurs à la ryanodine (Seo et al., 2012) de la citerne terminale et libère une grande quantité de calcium du réticulum sarcoplasmique dans le sarcoplasme. Ce calcium se lie à la troponine C, une protéine régulatrice située sur l'actine musculaire, et déclenche une série d'évènements moléculaires, dont le déplacement de la tropomyosine qui bloquait les sites de liaisons de la myosine sur l'actine (Ebashi and Ebashi, 1964). La myosine est alors en état pour hydrolyser l'ATP, nécessaire au détachement de la myosine de l'actine, à son déplacement le long du filament et à son rattachement à l'actine (Heissler and Sellers, 2016). Ce processus, appelé « couplage Excitation-contraction » (Marty and Fauré, 2016), permet le glissement de la myosine le long de l'actine, conduisant à la contraction de la fibre musculaire, et *in fine* du muscle.

### 3. Pathologie

#### a. Diagnostic/ Phénotype « macroscopique »

A la naissance, les personnes atteintes de myopathie de Duchenne (DMD) sont asymptomatiques (Blake et al., 2002). Cependant les dosages sériques de créatine kinase musculaire révèlent déjà des valeurs supérieures à la normale. Les premiers symptômes apparaissent entre l'âge de 2 et 5 ans, sous la forme d'un retard du développement moteur (marche, stature debout, difficulté à courir). Avec l'âge et la progression de la maladie, le déficit moteur (parésie) des membres inférieurs oblige l'enfant à utiliser ses membres supérieurs pour se lever et atteindre la position debout (signe de Gowers). Vers 10 ans, l'enfant commence à dépendre d'une chaise roulante. Les faiblesses musculaires atteignent ensuite les membres supérieurs puis les muscles vitaux (muscle cardiaque, diaphragme), le décès précoce survient généralement autour de 20 ans.

Les myopathies de Becker (BMD) se distinguent par une variabilité de la gravité des symptômes selon l'individu, de formes aussi grave qu'une myopathie type Duchenne à des formes avec des symptômes plus légers, apparaissant autour de 40 ou 60 ans, et ayant un impact limité sur la qualité de vie.

*b. Phénotype cellulaire*

Les études histologiques de muscles de personnes atteintes de DMD révèlent une rupture (Mokri and Engel, 1975) et une perméabilisation (Clarke et al., 1993) de la membrane des fibres musculaires.

A cause de cette fragilisation du sarcolemme, les cellules musculaires subissent des dommages lors des contractions musculaires répétées. Ces dommages sont associés à de nombreux dysfonctionnements (Péréon et al., 2015), comme l'exfiltration de protéines musculaires (créatine kinases) (Zatz et al., 1991), la dérégulation de l'homéostasie calcique (Bodensteiner and Engel, 1978; Franco and Lansman, 1990; Mongini et al., 1988; Tutdibi et al., 1999) ou la génération de stress oxydant (Ragusa et al., 1997), entraînant *in fine* la mort cellulaire par nécrose ou apoptose (Serdaroglu et al., 2002). Ces symptômes sont amplifiés par le déclenchement de processus inflammatoires (Cruz-Guzmán et al., 2015; Grounds and Torrisi, 2004; Spencer et al., 2001). Le potentiel de remplacement des fibres musculaires nécrosées par de nouvelles fibres est limité par l'incapacité des cellules souches satellites à générer suffisamment de fibres musculaires pour compenser leurs disparitions. Les cellules satellites ayant un nombre de divisions limité, leur activité accrue épuise ce potentiel de mitose à un rythme bien plus élevé que chez un individu non atteint (Renault et al., 2000). À terme, les muscles des patients DMD sont incapables de créer de nouvelles fibres. Le muscle est alors remplacé par du tissu adipeux (Ogura et al., 2014) et conjonctif, avec une accumulation excessive de matrice extracellulaire (Kharraz et al., 2014).

Au niveau génétique, les dystrophies musculaires ont pour origine des mutations du gène *DMD*. D'après Monaco et al. (Monaco et al., 1988), la différence entre BMD et DMD dépend de l'altération du cadre de lecture de la mutation sur le gène *DMD* (Figure 16). Une mutation décalant le cadre de lecture, comme celles retrouvées dans les pathologies de type DMD, mène à une absence de protéine, associée à un phénotype sévère. Dans le cas des patients BMD, la mutation n'altère pas le cadre de lecture, mais le gène code pour une protéine raccourcie et partiellement fonctionnelle, associée à un phénotype moins sévère que celui des patients DMD.

### c. *Modèles animaux*

Depuis la description du gène *DMD* (Koenig et al., 1987; Monaco et al., 1986), différents modèles animaux de myopathies ont été développés. Les deux modèles principaux sont les souris *mdx* et les chiens cDMD (*canine Duchenne muscular dystrophy*)

Les souris *mdx* sont issues de souris de souche C57BL/10ScSn (Bulfield et al., 1984), présentant une mutation ponctuelle non-sens (transition d'une cytosine en thymine avec création d'un codon « STOP ») au niveau de l'exon 23. Cette mutation empêche l'expression d'une dystrophine complète et fonctionnelle (Sicinski et al., 1989). Bien que certains symptômes puissent différer entre les souris *mdx* et les patients humains, comme l'espérance de vie réduite de 25 % contre 75 % chez l'Homme (**Erreur ! Source du renvoi introuvable.**), ce modèle murin présente de nombreuses similarités phénotypiques avec l'Homme, comme un niveau élevé en créatine kinase (Bulfield et al., 1984), des dysfonctions musculaires et une

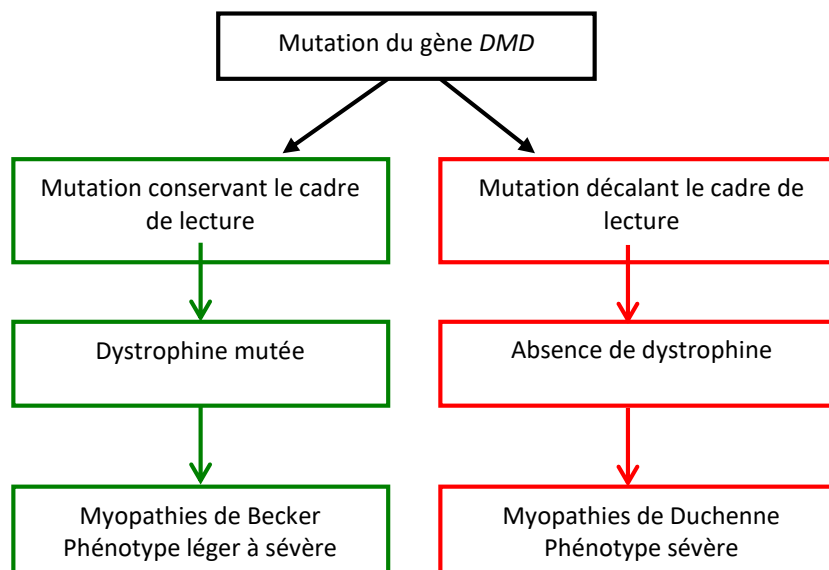


Figure 16 : Règles de Monaco, d'après (Monaco et al., 1988)

accumulation de tissu adipeux (Pastoret and Sebillé, 1995). De nombreux autres modèles murins ont depuis été développés (Yucel et al., 2018). En particulier, des doubles KO dystrophine/utrophine ont été générés (Deconinck et al., 1997), entre autre pour étudier l'importance de la compensation de l'utrophine en absence de dystrophine. Les modèles murins sont principalement utilisés pour la preuve de concept de thérapie et l'étude approfondie des différents aspects induits par l'absence de dystrophine fonctionnelle (Belanto et al., 2016).

Bien que des déficiences en dystrophine aient été identifiées chez une vingtaine de races de chiens (McGreevy et al., 2015), la souche canine GRMD (*golden retriever model of Duchenne muscular dystrophy*) est la plus utilisée comme modèle cDMD, ces chiens présentent un phénotype plus sévère que celui des souris *mdx*, et plus proche celui de de l'Homme (**Erreur ! Source du renvoi introuvable.**). En particulier, la progression de la maladie est très similaire à ce qui est observé chez les jeunes patients (McGreevy et al., 2015). Les modèles canins sont généralement utilisés à des stades plus avancés de recherche thérapeutique que les modèles murins, par exemple pour étudier leur capacité de distribution systémique (McCloy et al., 2006; Yokota et al., 2009).

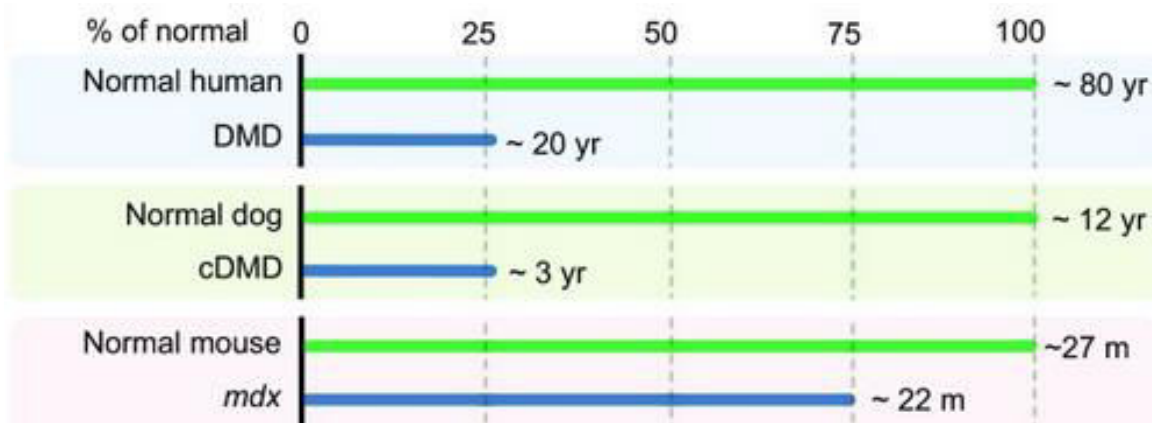


Figure 17 : Espérance de vie comparée entre des individus atteints de DMD ou de modèles animaux. Extrait de McGreevy et al, 2015.

## D. Thérapies

### 1. Thérapies cellulaires

Les thérapies cellulaires à partir de cellules souches sont des approches prometteuses dans le traitement des dystrophies musculaires. Les cellules souches présentent une capacité de mitose prolongée et de différenciation en différents types cellulaires (Weissman, 2000). Les cellules satellites constituent un type de cellules souches multipotentes spécifiques du muscle qui permettent de régénérer le muscle tout au long de la vie.

Dans le cadre des DMD et BMD, l'approche consiste à cultiver des cellules satellites portant un gène codant la dystrophine complète et fonctionnelle (Collins et al., 2005). Une fois implantées dans les muscles de malades, elles pourront se maintenir par mitose tout en produisant de nouvelles fibres musculaires exprimant une dystrophine fonctionnelle. Ces cellules peuvent provenir d'un donneur sain (Partridge, 2004), ou de cellules du patient modifiées *in vitro* pour restaurer l'expression de la dystrophine (Mendell and Clark, 2006). Bien que les différents essais sur des modèles animaux montrent des résultats encourageants, plusieurs difficultés persistent. En effet, les cellules satellites ne sont pas mobiles, et doivent être injectées en de nombreux sites d'administration. Dans le cas d'une source allogénique (donneur différent du receveur), la réponse immune doit être prise en compte pour maximiser la tolérance du traitement.

De nombreux essais ont été, et sont toujours, conduits chez l'animal (souris (Gussoni et al., 1999), chien (Sampaolesi et al., 2006)) et chez l'Homme (Huard et al., 1992; Miller et al., 1997). Cependant, bien que potentiellement efficace, l'application pratique de cette approche est compliquée par la difficulté de cultiver suffisamment de cellules.

### 2. Thérapies géniques

Cette dénomination regroupe toutes les approches qui ont pour objectif de restaurer la production d'une protéine fonctionnelle par les cellules des muscles malades.

#### *a. Gene delivery / remplacement de gène*

Ces stratégies visent à remplacer le gène *DMD* défectueux par un nouveau gène, codant une dystrophine fonctionnelle (Belanto et al., 2014).

Ces approches présentent de nombreux défis, comme la prise en compte de la grande taille du gène (1,4 Mb) et de celle de l'ADN codant (14kb), et la nécessité de faire parvenir le traitement à tous les muscles du corps. Plusieurs vecteurs ont été testés. L'un des plus prometteurs est dérivé des virus adéno-associé (AAV).

Les AAV sont des petits virus à ADN, initialement isolés de préparations d'adénovirus, (Hastie and Samulski, 2015; Rose et al., 1966). Ils sont composés d'une capsidie protéique, qui leur permet de s'amarrer aux membranes cellulaires et de pénétrer dans les cellules par endocytose, et d'un ADN mono-brin de 4,8kb, capable de s'intégrer dans le génome.

Sous une forme recombinante (rAAV), ces vecteurs peuvent être utilisés pour délivrer un gène au noyau de toutes les cellules de l'organisme. En fonction de la séquence conçue, leur ADN peut s'intégrer dans le génome et persister au cours des divisions cellulaires, ou rester hors des chromosomes (ADN episomal) mais être perdu avec les divisions cellulaires.

L'utilisation thérapeutique des vecteurs AAV, permettant une distribution systémique du gène, nécessite l'administration d'une grande quantité de vecteurs, avec l'incorporation d'une cassette d'expression spécifique aux muscles, tout en limitant la réponse immunitaire. Pour minimiser la dose injectée tout en maintenant l'efficacité de ce système, une approche consiste à fonctionnaliser les vecteurs afin de cibler un tissu spécifique (Körbelin et al., 2016; Li et al., 2008).

Les vecteurs AAV sont capables d'assurer le transport d'acides nucléiques, et ont déjà été évalués dans le cadre de la myopathie de Duchenne (clinicaltrials NCT01396239 et NCT01540409 (Mendell et al., 2016)) sans apparition des effets indésirables sévères retrouvés avec d'autres vecteurs (Cotter and Muruve, 2005). Les données concernant la sécurité d'utilisation sont rassurantes car les AAV sont incapables de se répliquer par eux-mêmes en absence d'une co-infection de l'hôte par un adénovirus (Cheung et al., 1980). Toutefois, cette approche est notamment limitée par la faible capacité d'encapsulation des AAV (5kb), nécessitant *a priori* l'utilisation de séquences *dmd* raccourcies (Dong et al., 1996). Plusieurs essais cliniques sont actuellement en préparation pour évaluer l'efficacité de telles approches dans la DMD (clinicaltrials ID NCT03375164, NCT03333590, NCT03368742).

Une alternative aux AAV consiste en l'utilisation de virus dérivés des lentivirus, issus du HIV-1 (*human immunodeficiency virus*). Ils sont déjà utilisés en thérapie génique pour



transporter des gènes dans des cellules (en division ou non) (Yáñez-Muñoz et al., 2006). Leur capacité d'encapsulation est limitée en théorie à un transgène de 10kb, et donc nécessite toujours la conception d'un gène raccourci (Shunchang et al., 2008). Des travaux récents ont cependant montré *in vitro* qu'il est possible de construire des vecteurs capables de prendre en charge la globalité de la séquence codante de dystrophine (Counsell et al., 2017; Kimura, 2010). L'efficacité de ce type de vecteur n'a été montrée qu'*in vitro*, sur des myoblastes DMD murins ou humains. Les limites de ce type de vecteur sont les mêmes que pour les AAV (charge virale nécessaire, ciblage (Frecha et al., 2008)). Cependant, ces lentivirus n'ont pas fait l'objet de développements thérapeutiques importants.

*b. Exon skipping / Saut d'exon*

Il a été montré que des dystrophines codées par des gènes présentant des délétions internes pouvaient rester partiellement fonctionnelles, générant des BMD au phénotype modérément sévère (England et al., 1990). Cette observation est à la base de la stratégie de « saut d'exon », pour transformer une DMD en une BMD la moins grave possible en interférant sur la phase d'épissage de l'ARN pré-messager, en excisant un ou plusieurs exons contenant l'erreur responsable du dysfonctionnement, afin de produire une protéine raccourcie mais fonctionnelle (Figure 18).

Selon l'erreur à l'origine de la DMD, le saut d'exon est réalisé avec différents objectifs :

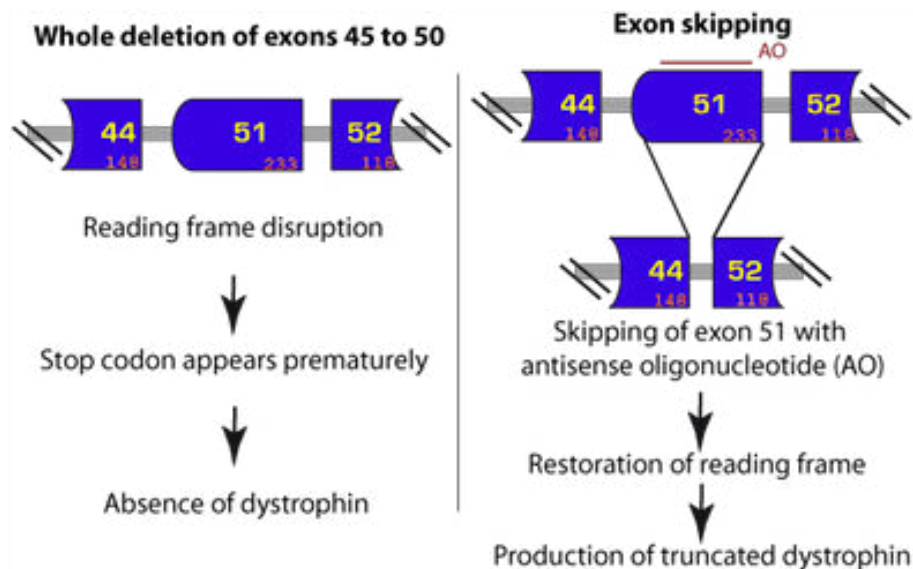


Figure 18 : Principe du saut d'exon ; Gauche) Exemple de délétion des exons 45 à 50 dans le gène DMD de patient DMD ; créant un décalage du cadre de lecture au niveau de la jonction entre 44 et 51 ; Droite) Principe du saut d'exon 51, pour joindre les exons 44 et 52, maintenant le cadre de lecture. Extrait de Nicolas et al, 2012.

- Si la maladie est causée par une mutation non-sens, le saut d'exon permet d'exciser l'exon contenant la mutation (Yokota et al., 2012)
- Si la maladie est causée par la perte du cadre de lecture (perte/duplication d'un exon ou insertion/délétion d'un nucléotide) le saut d'exon est ciblé pour connecter deux exons non-consécutifs, mais présentant le même cadre de lecture (Figure 18).

Plus de 7000 mutations ont été recensées chez les patients atteints de DMD (Bladen et al., 2015). Cependant, 13% d'entre elles impliquent l'exon 51 (répétition 20), et plus de 43% la région comprise entre les exons 45 et 55, codant les répétitions R17 à R22 (Aartsma-Rus et al., 2009). Ainsi, malgré la diversité des causes de la maladie, la mise en place d'un petit nombre de stratégies de saut d'exon différentes pourrait permettre la prise en charge d'une grande partie des patients.

Les recherches actuelles portent principalement sur l'utilisation d'oligonucléotides antisens (AON ou ASO, *antisense oligonucleotides*) permettant de moduler l'épissage de l'ARN pré-messager. Les AON sont des poly nucléotidiques, modifiées chimiquement pour augmenter leur stabilité métabolique. Les AON peuvent se lier à un ARN simple brin cible par complémentarité de bases (Watson and Crick, 1953), formant ainsi localement un ARN double brin. Cette dimérisation active différents processus, en fonction de la conception de l'AON, comme la dégradation de l'ARN cible par l'ARNase H1 (Vickers and Crooke, 2015). Dans les approches de saut d'exon, l'oligonucléotide est conçu de manière à cibler un site d'épissage de l'ARN pré-messager d'un exon **n** pour le bloquer, et ainsi forcer la cellule à connecter directement l'exon **n-1** à l'exon **n+1** (Havens and Hastings, 2016).

Les approches thérapeutiques par saut d'exon sont à l'étude depuis de nombreuses années pour le traitement des myopathies (Milligan et al., 1993), avec des travaux *in vitro* (Aartsma-Rus et al., 2004; McClorey et al., 2006) et *in vivo* sur des modèles murins (Mann et al., 2001) et canins (Miskew Nichols et al., 2016; Yokota et al., 2009). Chez l'Homme, des essais cliniques sont également en cours (Havens and Hastings, 2016).

L'Eteplirsen® est un AON de 30 nucléotides, qui cible l'exon 51 du gène *DMD*. Ce traitement est actuellement à l'étude dans plusieurs essais cliniques (ClinicalTrials.gov Identifier: NCT00159250 et NCT02420379) et a reçu en 2016 l'autorisation de mise sur le marché nord-américain par la *Food and Drug Administration* (NDA 206488, sous le nom

EXONDYS 51™). Son efficacité clinique est encore en cours d'évaluation (ClinicalTrials.gov Identifier: NCT02286947), mais ce traitement ralentirait la progression des symptômes, sans pour autant repousser l'âge du décès (Lim et al., 2017).

Le Kyndrisa® est un AON de 20 nucléotides, avec une séquence similaire à celle de l'Eteplirsén, mais raccourcie. Il provoque aussi le saut de l'exon 51. Les essais cliniques ont mis en évidence l'effet bénéfique de ce médicament durant les 25 premières semaines de traitement. Cependant, après 48 semaines, cet effet bénéfique n'est plus observé (Voit et al., 2014).

Globalement, les oligonucléotides anti-sens ne montrent pas d'effets toxiques, même à fortes doses (Relizani et al., 2017). Cependant, bien qu'ils soient modifiés afin d'être plus stables, ils sont graduellement dégradés dans l'organisme. Il est ainsi nécessaire de procéder à des injections régulières potentiellement tout au long de la vie du patient. Aussi, il est à noter que le coût de ces traitements est très élevé (300 000 \$ / an pour l'Eteplirsén®).

### *c. Correction de gène*

Initialement découvert chez les bactéries (Barrangou et al., 2007) le système CRISPR (clustered regularly interspaced short palindromic repeats) /Cas9 (CRISPR associated protein 9) constitue une alternative intéressante aux stratégies de sauts d'exons. Depuis les premières publications de sa caractérisation biochimique en 2012 (Jinek et al., 2012) et de son utilisation dans des cellules eucaryotes l'année suivante (Cong et al., 2013), il s'est imposé comme un outil polyvalent d'ingénierie des génomes eucaryotes, principalement grâce à sa facilité d'utilisation. Alors que les méthodes antérieures nécessitaient de concevoir une protéine *de novo* pour effectuer des changements sur des séquences spécifiques (TALEN (Miller et al., 2011), ZFN (Kim et al., 1996)), le système CRISPR/cas9 se compose d'un complexe protéique « ubiquitaire », dont le ciblage sur le génome est apporté par un court ARN guide (ARNg de 28 b à 37 b) (Jinek et al., 2012), plus simple à synthétiser qu'une protéine complète.

La protéine Cas9 est une nucléase utilisée en association avec un ARN guide qui permet de cibler la zone à éditer (Hsu et al., 2014). L'ARNg est composée d'une région constante de liaison à Cas9, ainsi que d'une région variable, complémentaire à la séquence d'ADN ciblée (Munshi, 2016). Cas9 se lie à une région constante de l'ARNg, permettant une hydrolyse ciblée et localisée de l'ADN. En réaction, la cellule met en œuvre différents mécanismes de

réparation de l'ADN, lesquels peuvent introduire des modifications au niveau de la coupure, comme des mutations ponctuelles (mutation conservant le cadre de lecture ou insertion/délétion le modifiant). Si un ADN contenant à ses extrémités les séquences présentes de part et d'autre du site de coupure est fourni en plus de la nucléase, celui-ci peut être introduit à ce niveau lors de la réparation. Enfin, deux ARNg ciblant deux cibles différentes sur le même chromosome permettent de supprimer du génome la séquence située entre les deux sites de coupure.

Le complexe protéine/ARNg peut être introduit dans les cellules sous différentes formes (Kouranova et al., 2016). Il peut être introduit sous forme de complexe ribonucleo-protéique dans les cellules à traiter *In vitro* (Kim et al., 2014) ou par injection locale (Stahl et al., 2017). Une distribution systémique large peut être obtenue en utilisant un vecteur viral (AAV, limité à 5 kb) encapsulant un ADN ou un ARN codant à la fois la protéine Cas9 (3,2 kb) et l'ARNg (~100bp) (Long et al., 2015). La simplicité de conception d'un ARNg en fait un outil flexible et polyvalent de premier ordre pour la mise en place de nouvelles thérapies pour la DMD et BMD.

Ce système a été testé avec succès *in vitro* sur des cellules souches musculaires déficientes en dystrophine (Li et al., 2015; Ousterout et al., 2015). Début 2016, trois articles ont été publiés dans le même numéro de la revue *Science*. Ils décrivent pour la première fois la possibilité d'utiliser le système CRISPR/Cas9 *in vivo* dans le but de restaurer l'expression d'une dystrophine fonctionnelle chez des souris *mdx* en excisant l'exon 23 (Long et al., 2015; Nelson et al., 2015; Tabebordbar et al., 2015).

A ce jour, aucun essai clinique n'a été mis en place pour évaluer le traitement de la DMD par CRISPR-Cas9 chez l'Homme.

## E. Objectifs de thèse

La dystrophine est une protéine essentielle dans la résistance du sarcolemme aux contraintes mécaniques induites durant les contractions musculaires. Cette fonction de charpente de la membrane est assurée par ses nombreuses interactions avec des partenaires protéiques ou lipidiques.

L'absence de dystrophine, causée par des mutations dans le gène *DMD*, est responsable de la myopathie de Duchenne, causant le décès prématuré des patients atteints vers l'âge de 20 ans. D'autres mutations du gène *DMD* sont responsables de la production d'une forme tronquée et partiellement fonctionnelle de dystrophine, causant la myopathie de Becker. Dans cette seconde forme, le phénotype varie selon la région absente de la dystrophine, allant de formes similaires aux myopathies de Duchenne à des formes très légères, avec une apparition des premiers symptômes qui peut être très tardive.

La dystrophine est une très grande protéine, codée par le plus long gène humain. Une partie des stratégies thérapeutiques actuellement en développement envisage la production d'une protéine raccourcie mais fonctionnelle. Ce travail nécessite une caractérisation préalable de la structure et de la fonction des différentes régions de la dystrophine, afin d'évaluer le potentiel thérapeutique dans la conception de mini-dystrophines.

Longtemps, l'absence d'études portant sur la structure du domaine central de la dystrophine, qui représente pourtant 80% de sa longueur, a limité la compréhension des mécanismes par lesquels cette protéine conditionne la résistance du sarcolemme aux contraintes mécaniques induites lors des contractions musculaires. Des travaux récents de l'équipe associant diffusion des rayons X aux petits angles (SAXS, Small Angle X-ray scattering) et modélisation moléculaire (Delalande et al., 2018) ont permis la caractérisation à l'échelle atomique de la région centrale de la dystrophine.

En se basant sur ces travaux, cette thèse a pour but de caractériser les complexes macromoléculaires qui sont formés par le domaine central de la dystrophine avec d'autres composants cellulaires. Dans un premier temps, ce travail a porté sur la caractérisation fine des interactions hydrophobes maintenant le faisceau d'hélice des répétitions R1 à R3, en lien avec l'implication de ce fragment dans l'interaction avec les lipides membranaires. Dans une

seconde partie, il a été entrepris de décrire l'interaction tripartite entre la région R11-15 de la dystrophine, l'actine-F et les microtubules du cytosquelette, et les lipides du sarcolemme.



## II. Méthodes

### A. Champs de force

Un champ de force est un ensemble d'équations et de paramètres permettant de décrire l'état énergétique d'un système moléculaire en fonction de la position de ses particules. Dans cette partie nous décrirons essentiellement les champs de force atomiques (décrivant les modèles moléculaires à l'échelle de l'atome). Les principes et équations décrits ici sont souvent transposables en partie à d'autres champs de force particuliers (notamment basse-résolution, voir partie C.2-*Approches basse-résolution gros-grain*). Pour les champs de force atomiques, les liaisons covalentes reliant des atomes sont par exemple définies par des relations harmoniques. Il existe de nombreux champs de force dont leur expression générale est de la forme :

$$E_{totale} = E_{liant} + E_{non-liant}$$

Équation 1 : Forme générale des équations de champ de force.

L'énergie totale  $E_{totale}$  correspond à la somme des énergies associées au terme liant  $E_{liant}$ , entre particules d'une même molécule qui sont séparées par 1, 2 ou 3 liaisons covalentes, et des énergies associées au terme non-liant  $E_{non-liant}$ , entre particules de deux molécules distinctes ou d'une même molécule, mais qui sont éloignées de plus de deux liaisons.

$$\begin{aligned} E_{liant} &= E_{liaison} && + E_{angle} && + E_{dièdre} \\ &= \sum_{liaison} k_b(l - l_0)^2 + \sum_{angle} k_a(\theta - \theta_0)^2 + \sum_{dièdre} \sum_n A_n [1 + \cos(n\tau - \varphi)] \end{aligned}$$

Équation 2 : calcul du terme liant du champ de force ;  $k_b$ ,  $k_a$  : constantes de force ( $kcal.mol^{-1}.\text{Å}^{-2}$ ) ;  $l$  : longueur de liaison ( $\text{Å}$ ) ;  $l_0$  : longueur de liaison à l'équilibre ( $\text{Å}$ ) ;  $\theta$  : angle ( $^\circ$ ) ;  $\theta_0$  : angle à l'équilibre ( $^\circ$ ) ;  $A_n$  : valeur de la barrière d'énergie de rotation d'ordre  $n$  ( $kcal.mol^{-1}$ ) ;  $\tau$  : valeur d'angle dièdre ( $^\circ$ ) ;  $\varphi$  : facteur de phase.

L'expression des énergies associées au terme liant  $E_{liant}$  est décomposée en trois termes (Équation 2 et Figure 19). Le premier terme, l'énergie de liaison ( $E_{liaison}$ ), modélise la contribution énergétique de l'ensemble des liaisons covalentes dans les champs de force atomiques, en l'approximant par le calcul d'une relation harmonique des atomes évoluant



autour d'une distance optimale. Le second terme ( $E_{angle}$ ) correspond au calcul de l'énergie associée à l'angle formé par trois atomes liés et se base également sur un potentiel harmonique. Enfin, le troisième terme ( $E_{dièdre}$ ) est l'énergie associée à l'angle de torsion (ou angle dièdre) formé par quatre atomes consécutivement liés par des liaisons covalentes.

$$E_{non-liant} = E_{vdW} + E_{électrostatique}$$

$$= \sum_j^{N-1} \sum_{i=j+1}^N f_{ij} \left\{ \epsilon_{ij} \left[ \left( \frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{0ij}}{r_{ij}} \right)^6 \right] - \left( \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \right\}$$

Équation 3 : calcul du terme non-liant du champ de force ; forces de Van der Waals entre les atomes  $i$  et  $j$  ;  $\epsilon_{ij}$  : énergie à l'équilibre à la distance  $r_{0ij}$  (Å) ;  $r_{ij}$  : distance entre les atomes  $i$  et  $j$  (Å) ; les termes à la puissance 12 et 6 représente les forces de répulsion et d'attraction, respectivement. Interactions électrostatiques :  $q_i, q_j$  : charges partielles ;  $\epsilon_0$  : constante diélectrique ou permittivité du vide,  $\epsilon_0 \approx 8,854 \text{ F.m}^{-1}$ .

Les énergies associées aux interactions non-liantes ( $E_{non-liant}$ ) sont approximées par deux termes. D'une part, le terme stérique est issu des interactions de Van der Waals ( $E_{vdW}$ ), émergeant de l'équilibre entre des forces attractives et répulsives entre volumes atomiques. Le second terme rend compte des interactions électrostatiques ( $E_{électrostatique}$ ) entre particules (intra ou intermoléculaires), qui peuvent être attractives ou répulsives selon la charge des deux atomes impliqués.

En ce qui concerne les interactions de Van der Waals, la composante attractive provient de dipôles induits par une asymétrie temporaire et spontanée du nuage électronique entourant un atome. D'autre part, le chevauchement des orbitales atomiques de deux atomes suffisamment proches entraîne leur éloignement l'un de l'autre. Le potentiel de Lennard-Jones 12-6 est souvent utilisé pour représenter ces interactions, les valeurs 12 et 6 correspondant aux puissances auxquelles sont élevés les différents termes (répulsif et attractif,

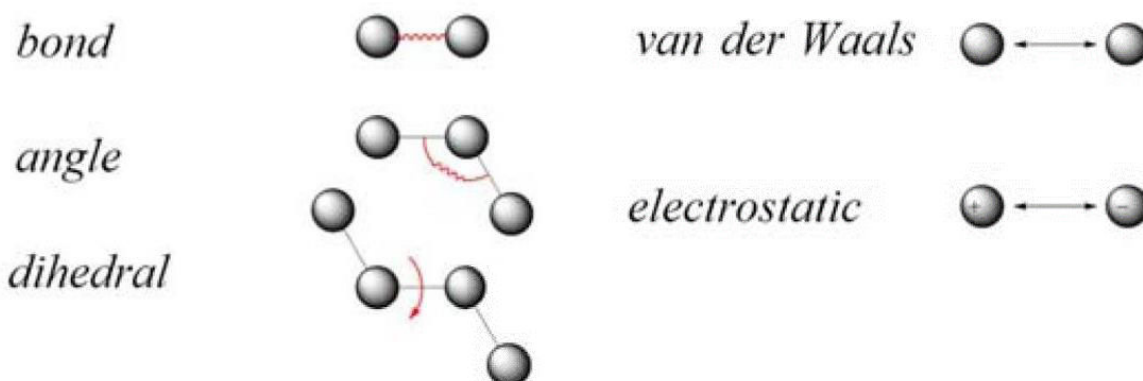


Figure 19 : Représentations schématiques des composantes des équations du champ de force. D'après Chang et al., 2016.

respectivement) de cette fonction. D'autres valeurs de puissance peuvent être employées, souvent dans l'objectif d'amoindrir l'effet répulsif, comme il est proposé dans les potentiels 9-6 (Sun, 1998) et 14-7 (Halgren, 1996)). Pour finir, le dernier terme du calcul des termes non-liants ( $E_{\text{électrostatique}}$ ) décrit les interactions électrostatiques entre particules (intra ou intermoléculaires). Celui-ci peut être attractif ou répulsif selon la charge  $q$  des deux atomes impliqués.

## B. Minimisation d'énergie

La géométrie d'une molécule peut être optimisée au cours d'un calcul consistant à en minimiser l'énergie interne, celle-ci étant calculée à partir d'un champ de force choisi. L'objectif de cette étape d'optimisation – aussi appelée « minimisation » – n'est pas de trouver la conformation du système correspondant à l'énergie globale du système la plus basse, mais plutôt la géométrie associée au minimal local d'énergie le plus proche (Figure 20). Cette étape de « minimisation » est aussi effectuée avant une simulation de dynamique moléculaire (MD, *molecular dynamic*, voir partie C-Simulation de dynamique moléculaire), pour placer le système de simulation dans un état stable. Dans ce cadre, cette optimisation réalisée avant une simulation de dynamique moléculaire permet d'éviter que des atomes soient soumis à des forces trop importantes, ce qui ferait « exploser » le système, empêchant le logiciel de poursuivre la simulation. Ces forces excessives sont généralement causées par des problèmes de recouvrement stérique introduits durant la préparation du système à simuler.

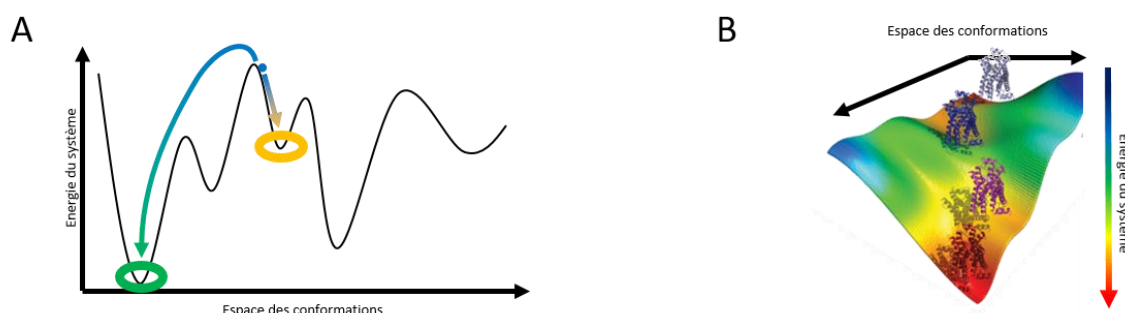


Figure 20 : A) Principe de la minimisation d'énergie. Bleu : Position initial du système dans l'espace des conformations ; Vert : Minima global d'énergie ; Jaune : Minima local accessible par le système simulé. Les algorithmes de minimisations ne permettent pas d'atteindre le minima global, ou tout autre minima local de meilleure énergie pour lequel un état de plus grande énergie doit être atteint au préalable. B) Exemple d'espace des conformations du récepteur adrenergique B2, d'après Provasi et al., 2011.

Les méthodes de minimisation sont basées sur l'algorithme de descente de gradient d'énergie, c'est-à-dire le calcul de la dérivée partielle de la fonction d'énergie du champ de force suivant les trois coordonnées spatiales  $x$ ,  $y$  et  $z$ . La position des atomes est ensuite modifiée dans le sens opposé au gradient pour diminuer l'énergie du système. Le logiciel NAMD (Phillips et al., 2005) utilise la méthode du gradient conjugué (Payne et al., 1992) qui consiste à pondérer le déplacement des atomes à une étape de calcul par ceux des étapes précédentes. Ainsi, on conserve une information sur la direction globale du gradient, pour atteindre plus rapidement l'état d'équilibre.

### C. Simulation de dynamique moléculaire

Les méthodes de simulation de dynamique moléculaire regroupent les approches numériques basées sur la simulation physique de la trajectoire des particules (atomes ou pseudo-atomes) en tenant compte de leur environnement physico-chimique. Ce sont des outils puissants, qui permettent d'explorer les propriétés physiques et dynamiques des biomolécules (protéines, membranes, acides nucléiques, ...) à l'échelle atomique (Karplus and McCammon, 2002).

Le principe général consiste à intégrer les équations du mouvement de Newton (Équation 4) suivant la méthode d'Euler. Celle-ci consiste à déterminer la vitesse et la direction de déplacement des atomes à un instant  $t$  à partir de l'équation du champ de force, puis à les appliquer durant un pas de temps défini  $\Delta t$  :

$$F_i = m_i \cdot a_i = m_i \frac{dv_i}{dt} = m \frac{d^2x_i}{dt^2} \rightarrow a_i = \frac{F_i}{m_i}$$

*Équation 4 : Equation du mouvement de Newton.  $F$  : Forces extérieures exercées sur l'atome  $i$  ;  $m$  : masse de l'atome  $i$  ;  $a$  : accélération de l'atome  $i$ .*

Cette dernière équation est intégrée successivement deux fois, permettant d'obtenir l'expression des vitesses puis des positions (Equation 5).

$$v_{t+\Delta t} = a_i * \Delta t + v_i$$

$$x_{t+\Delta t} = a_t * \Delta t^2 + v_t * \Delta t + x_t$$

*Équation 5 : Expressions de la vitesse  $v$  et de la position  $x$  d'une particule à un instant  $t+\Delta t$  à partir de l'accélération  $a$ , la vitesse  $v$  et la positions  $x$  d'une particule à un instant  $t$ .*

Dans ces équations du mouvement, on peut noter l'absence de terme relatif à la température du système. Il existe cependant une relation entre la vitesse des atomes et la température du système :

$$T \equiv \frac{\frac{1}{2} \sum_i v_i^2}{3N}$$

*Équation 6 : Relation entre la température du système et la vitesse des particules le composant. T : Température du système en Kelvin ;  $v_i$  : vitesse de la particule  $i$  ; N : nombre de particules du système.*

Différentes méthodes permettent de corriger la vitesse des particules pour que la quantité totale de mouvement corresponde à la valeur attendue pour la température du système, comme le thermostat de Berendsen ou le thermostat de Langevin.

La durée du pas d'intégration est choisie afin d'être environ un ordre de grandeur plus court que le plus petit mouvement que l'on doit simuler. Dans le cas de simulations tout-atome, l'évènement le plus rapide à modéliser est la vibration des liaisons carbone-hydrogène, avec une période de l'ordre de  $10^{-14}$ s. Ainsi, le pas d'intégration est généralement fixé à 1 ou  $2 \cdot 10^{-15}$ s (1 ou 2 fs).

#### 1. Approches haute-résolution tout-atome

Dans les approches tout-atome, tous les atomes sont considérés et simulés. Parmi les principaux champs de forces tout-atome, nous avons utilisé CHARMM (Mackerell et al., 1998, 2000) avec NAMD comme moteur de simulation (Phillips et al., 2005), et AMBER (Salomon-Ferrer et al.) sous sa variante incluse dans YASARA (Krieger et al., 2002), (YAMBER3 (Krieger et al., 2004)). Ces champs de forces sont globalement similaires, et ils peuvent être utilisés sans discrimination pour la description de système moléculaire protéique (Martín-García et al., 2015). Pour des raisons de cohérence avec les études précédentes (Giudice et al., 2013; Legrand et al., 2011; Molza et al., 2014, 2015; Nicolas et al., 2014a), nous avons utilisé de préférence CHARMM dans le cadre d'étude par simulation de dynamique moléculaire. De plus, le moteur de simulation associé (NAMD) a été développé pour des infrastructures de calculs hautement parallélisées, tel que des centres de calculs nationaux sur lesquels nous avons déployé nos calculs (IDRIS, CINES, ...), et permettant notamment de lancer des simulations sur jusqu'à 2048 processeurs en parallèle, sans pertes de performances. Dans de telles configurations, les calculs peuvent être effectués 10 à 100 fois plus rapidement que sur

un ordinateur de bureau (3 h/ns simulée contre 30 h/ns simulée pour un système de 500 000 atomes sur 512 processeurs de Turing au CINES). Le champ de force AMBER a généralement été utilisé lors des étapes suivant les reconstructions tout-atome dans les expériences d'ajustement interactif (voir partie 4-*Dynamique moléculaire sous contrainte et interactives*). En effet, la représentation tout-atome du logiciel BioSpring est basée sur le champ de force AMBER. La transition entre l'utilisation d'AMBER vers CHARMM est effectuée par deux étapes de minimisations successives, la première avec le champ de force d'origine, AMBER, la seconde basée sur CHARMM.

Comme présenté précédemment, le pas d'intégration dans les approches tout-atome est de l'ordre de la femtoseconde. Alors que ces simulations tout-atome permettent une description fine de la dynamique des molécules (mouvements de chaînes latérales et de sous-domaines protéiques, etc), la puissance de calcul nécessaire pour représenter tous les atomes d'un grand système moléculaire (solvant inclus) limite souvent les durées de simulation à quelques centaines de nanosecondes ( $100 \times 10^{-9}$ s). Bien entendu, il existe des exceptions comme Anton-2, un supercalculateur développé spécifiquement pour le calcul de simulations de dynamique moléculaire, permet de calculer de environ  $25\mu\text{s}$  ( $10^{-6}$ s) en 24h (soit 4 secondes / ns) pour un système de 300 000 atomes (Shaw et al., 2014). Les simulations sur du matériel plus « classique » (cluster ou station de calcul, tous deux étant plus abordables) peuvent ainsi de fait limiter l'intérêt des simulations de dynamique moléculaire tout-atome pour l'étude des phénomènes biologiques de plus grande ampleur, notamment de ceux qui se déroulent sur des temps de l'ordre de la micro- ou milliseconde (évolution de structure secondaire, diffusion, association inter-moléculaire, etc).

## 2. Approches basse-résolution gros-grain

Les limites de temps de simulation discutées dans le paragraphe précédent peuvent être parfois contournées en utilisant des approches dites gros-grain. Dans celles-ci, le champ de force est complété par des règles de réduction décrivant comment plusieurs atomes, géométriquement proches et/ou faisant partie d'un même groupe chimique (cycle

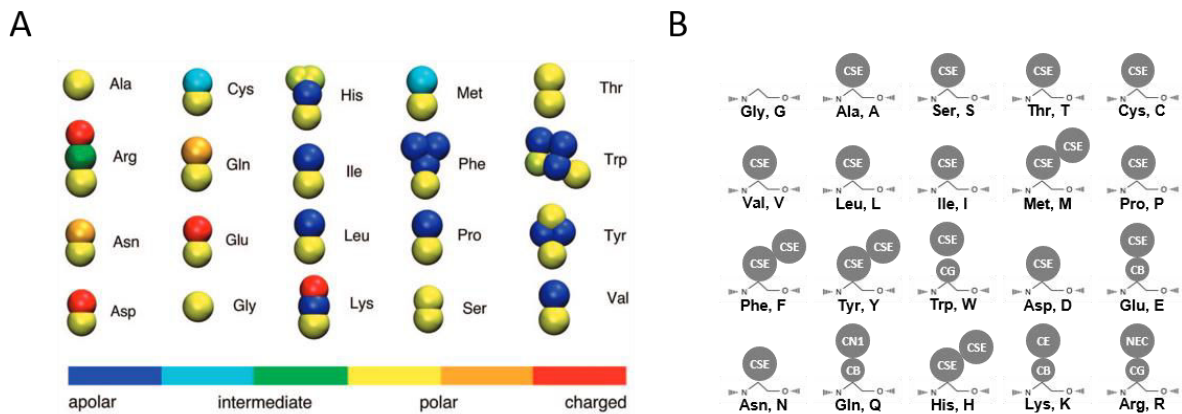


Figure 21 : A) Modèle de réduction du champ de force MARTINI, extrait de Monticelli et al., 2008. B) Modèle de réduction du champ de force ZACHARIAS.

aromatique, alcool, ...) sont regroupés en pseudo-atomes. Ce regroupement permet de réduire le nombre de particules simulées. De plus, en intégrant les atomes proches en un même grain, on diminue la fréquence des vibrations inter-particulaires. On peut alors augmenter la durée du pas d'intégration entre deux états consécutifs. Les simulations sont ainsi accélérées sur plusieurs fronts: d'une part, en diminuant le nombre de calculs paire à paire à effectuer, et d'autre part en augmentant le temps entre deux états calculés. De plus, à système équivalent, l'empreinte mémoire est réduite, ce qui permet de simuler des systèmes plus complexes, comme de grands complexes protéiques, en des temps (humains) raisonnables. Les gains de temps que permettent ces approches gros grains dépendent du type de réduction effectuée, mais peuvent être 100 à 1000 fois plus rapides, permettant d'atteindre des vitesses de simulation de l'ordre de 7 minutes/ns simulée (Abraham et al., 2015).

Au cours de cette thèse, des simulations dynamiques moléculaires gros-grain ont été effectuées entre des fragments de dystrophine et des modèles membranaires (voir Chapitres III et IV), en utilisant le champ de force MARTINI (Marrink et al., 2007; Periole and Marrink, 2013). Ce champ de force a été initialement développé pour simuler des mélanges de détergents ou de phospholipides, avant d'être étendu aux protéines (Monticelli et al., 2008) puis aux acides nucléiques (Uusitalo et al., 2015). Il est ainsi particulièrement adapté à l'étude de l'interaction entre les lipides et les protéines, tel que présenté par Monticelli et al. dans des simulations d'hélices transmembranaires (Monticelli et al., 2008), ou par Barreto-Ojeda et al. dans des simulations de la P-glycoprotéine insérée dans une bicouche lipidique (Barreto-Ojeda et al., 2018). La réduction d'une structure tout-atome en modèle gros-grain MARTINI

regroupe en moyenne 4 atomes lourds (atomes hors hydrogènes) par grain, à l'exception des cycles qui regroupent moins d'atomes. Le champ de force MARTINI considère quatre types de grains : polaires, non-polaires, apolaires, et chargés. Dans chaque type, des sous-types sont distingués selon leur capacité à créer des liaisons hydrogène (donneur et/ou accepteur) ou selon leur polarité (Figure 21A).

Ces champ de force gros-grain ne sont pas seulement applicables à des approches de simulation de dynamique moléculaire. Ainsi les expériences d'amarrage moléculaire rigide de cette thèse ont été basées sur le champ de force gros-grain ATTRACT2 (Chéron et al., 2017; Zacharias, 2005). Celui-ci a été développé spécifiquement dans le but de déterminer le mode d'association entre grandes protéines (Zacharias, 2005), puis étendu aux acides nucléiques (Setny et al., 2012). Dans le modèle de réduction associé à ATTRACT2, chaque résidu est représenté par 2 à 4 particules : deux d'entre elles permettent de décrire le squelette, et de zéro à deux particules supplémentaires sont dédiés à la chaîne latérale (Figure 21B).

### 3. Construction d'un système et simulation de dynamique moléculaire

La première étape d'un protocole standard de simulation (quelle que soit la résolution, tout-atome ou gros-grain) consiste à définir l'espace qui sera calculé (cellule ou boîte de simulation) autour du système d'intérêt. La cellule est ensuite remplie d'eau, puis certaines de ces molécules de solvant sont remplacées par des ions (généralement  $\text{Na}^+$  et  $\text{Cl}^-$ ) de manière à atteindre une force ionique choisie. En effet, certaines protéines sont sensibles à la force ionique du solvant, et si ce n'est pas le cas, l'addition d'ions  $\text{Na}^+$  ou  $\text{Cl}^-$  permet de neutraliser la charge du système. Le système est ensuite minimisé durant quelques dizaines de milliers de pas (voir B-*Minimisation d'énergie*).

Suite à cela, différentes étapes de simulations de dynamique moléculaire sont effectuées. Toutes se placent dans des conditions périodiques, pour éviter les effets de bords

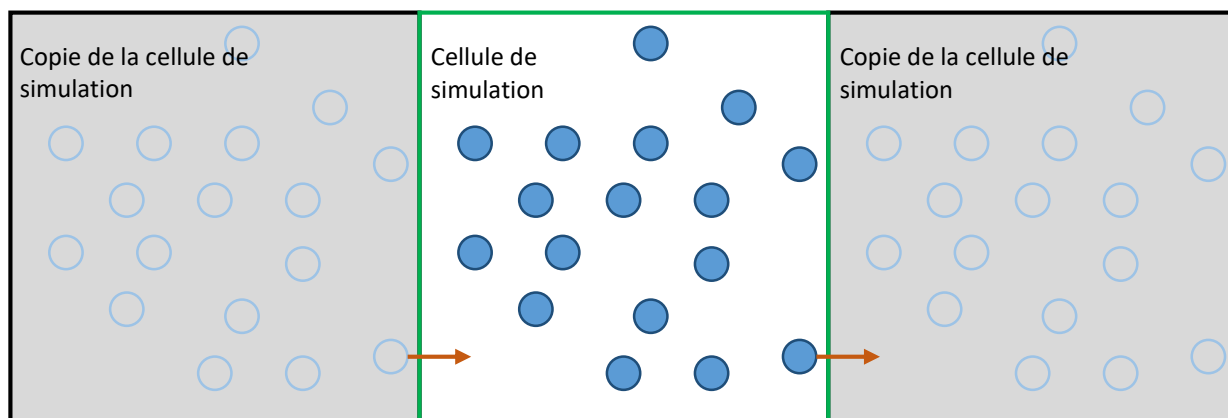


Figure 22 : Principe de conditions périodiques aux limites. Cellule Verte : cellule de simulation ; cellules noires : copies de la cellule de simulation. Flèches orange : mouvement apparent de la particule franchissant les limites de la cellule.

à proximité des limites de la cellule de simulation : l'espace est pavé par des copies de la cellule, de telle manière à ce que si un objet est contraint de franchir une de ses faces, il réapparaît au niveau de la face opposée (Figure 22). On simule ainsi virtuellement un environnement infini. Dans de telles conditions, il est ainsi nécessaire de définir une boîte de simulation suffisamment grande pour éviter l'interaction d'un objet moléculaire avec lui-même, en particulier au travers des forces électrostatiques (Équation 3). Ces forces entraînent des interactions entre des particules pouvant être distantes de plusieurs nanomètres. Une marge d'au moins 5 nm est donc prévue lors de la création de la cellule de simulation entre le soluté et le bord de la boîte.

Les forces électrostatiques sont appliquées en utilisant la méthode du maillage particulier d'Ewald (*particular mesh Ewald*, PME) (Darden et al., 1993). Brièvement, celle-ci consiste à sommer les forces électrostatiques qui s'exercent sur une particule dans l'espace réel pour les interactions proches, distantes de moins de 20 Å, et dans l'espace de Fourier pour les interactions à longue portée.

La première étape de simulation de dynamique moléculaire est la thermalisation : suite à la minimisation, le système est initialement dans un état qui correspond à une température de zéro Kelvin (0K). Le moteur de simulation augmente par paliers la température jusqu'à atteindre la température voulue (en général, entre 293K/20°C et 303K/30°C), le volume de la cellule de simulation étant conservé (ensemble thermodynamique **NVT** : **N**ombre de particules constant ; **V**olume constant ; **T**empérature constante). Une technique dérivée est le recuit simulé, dans lequel la température est augmentée par paliers jusqu'à des valeurs



extrêmes (800K) pour augmenter l'énergie du système et atteindre des états conformationnels autrement inaccessibles, avant que le système ne soit minimisé.

Suite à la thermalisation, la température est ensuite maintenue fixe pendant la durée souhaitée dans une seconde étape de simulation, de même que la pression (ensemble thermodynamique **NPT** : **N**ombre de particules constant ; **P**ression constante ; **T**empérature constante) : la taille de la cellule est ajustée tout au long de la simulation pour se stabiliser autour d'une valeur définie, généralement 1,013 bar. On peut distinguer deux périodes dans cette seconde étape de simulation. Tout d'abord la période d'équilibration, durant laquelle la distribution des vitesses des particules se stabilise autour des valeurs attendues pour un système à une température donnée. Dans les simulations tout-atome, cette étape s'étend des quelques premières nanosecondes aux premières dizaines de nanosecondes de simulation. Cette étape est suivie par celle de production, période de la simulation fournissant les informations utiles et pertinentes pour l'étude.

#### 4. Dynamique moléculaire sous contrainte et interactives

Les simulations de dynamique moléculaire permettent d'observer des changements de conformations prenant place dans des gammes de temps simulé allant de la dizaine ou de la centaine de nanosecondes (tout-atome) à la dizaine ou à la centaine de microsecondes (gros-grain). Des événements de grande amplitude peuvent nécessiter des temps de simulation supérieurs (millisecondes ou secondes), non accessibles avec les moyens matériels et logiciels actuellement disponibles. Enfin, certaines simulations de dynamiques cherchent à déterminer l'influence de l'application de forces externes sur une protéine, par exemple pour notre objet d'étude, lors de la transmission de forces mécaniques le long de la dystrophine.

Le principe des approches de dynamique moléculaire sous contraintes (SMD, *Steered MD*) est d'effectuer une dynamique moléculaire en appliquant une force extérieure à un ou plusieurs atomes du système dans une direction définie. Elles ont été par exemple utilisées pour étudier le passage de molécules d'eau par des aquaporines (Cui and Bastien, 2011) ou la dynamique de dépliement d'une répétition en triple hélice de la spectrine (Altmann et al., 2002).

Dans certaines situations, l'application d'une force dans plusieurs directions successives peut être nécessaire afin d'exercer un mouvement complexe sur l'objet étudié. Cependant,

de telles approches nécessitent d'ajuster au cours de la simulation la direction de la contrainte appliquée, afin de s'adapter à l'évolution constante du système moléculaire simulé. Ainsi, il est parfois nécessaire d'employer des méthodes de dynamique moléculaire interactives (IMD, *Interactive MD*), dans lesquelles un utilisateur guide une SMD en temps réel, à l'aide, par exemple, d'un périphérique 3D ou bien d'un dispositif haptique (retour tactile). Cependant, pour que l'interactivité reste possible et efficace, il est nécessaire que la simulation se déroule à une vitesse suffisamment élevée pour que l'utilisateur puisse observer en direct l'effet des perturbations qu'il provoque, soit au moins 20 à 25 pas de simulations par seconde. Les simulations de dynamique moléculaire peuvent difficilement être calculées à cette vitesse sur un ordinateur de bureau, elles doivent donc être adaptées pour permettre l'interactivité. Différentes stratégies peuvent alors être mises en place pour accélérer la simulation. Les techniques de simulations gros-grain basées sur le champ de force MARTINI (voir 2-*Approches basse-résolution gros-grain*) permettent ainsi de diminuer le nombre de particules du système et de rendre l'approche IMD réalisable. Une seconde approche peut être l'utilisation d'un moteur spécifique de simulation de dynamique moléculaire comme BioSpring (Ferey et al., 2012). Ce dernier, conçu pour des IMD sans solvant, combine d'une part l'utilisation de grilles précalculées pour la simulation de certains potentiels (électrostatiques ou hydrophobes) et simplifie d'autre part les interactions entre particules (atomes ou autre) par un réseau de ressort, régit par des contraintes harmoniques. Le système moléculaire peut encore être simplifié en utilisant un modèle gros-grain (Attract pour BioSpring).

#### 5. Amarrage moléculaire

Les interactions protéines-protéines sont nécessaires à de nombreuses fonctions moléculaires comme par exemple la co-localisation de deux partenaires ou la modification d'un substrat par une enzyme (phosphorylation, méthylation, etc). La résolution expérimentale de complexes protéiques peut s'avérer compliquée à obtenir. La prédiction *in silico* est ainsi un outil complémentaire aux approches expérimentales qui permet de prédire, à partir de deux molécules de structure connue, un modèle théorique du complexe qu'elles pourraient former en s'associant. Plus généralement, ces approches permettent de prédire les modes d'interaction entre molécules, et de modéliser par exemple la zone d'interaction d'une petite molécule avec un transporteur membranaire. On peut identifier deux catégories de méthodes d'amarrage moléculaire: les méthodes exhaustives et les méthodes guidées. Les

premières regroupent les méthodes qui explorent les solutions d'amarrages sans *a-priori*, alors que les approches guidées se basent sur des informations préalables pour réduire l'espace de recherche. Ces informations peuvent provenir de multiples approches expérimentales ou théoriques, comme par exemple d'analyses de complexes par spectrométrie de masse (voir chapitre *IV-Modélisation de complexes de la dystrophine (actine, microtubules, lipides)*) ou par l'analyse préalable des propriétés de surface et de détection de poches hydrophobes dans le cas de la recherche de complexes avec des petits ligands. Il est à noter que ces approches ne sont pas incompatibles, des logiciels d'amarrages proposant même les deux approches (Deplazes et al., 2016; Geng et al., 2017; Trott and Olson, 2010). Dans cette thèse, l'amarrage moléculaire *in silico* a notamment été utilisé pour déterminer la position et la conformation de fragments de dystrophine en complexe avec l'actine-F. Dans la suite de ce chapitre, le terme « récepteur » désignera la plus grande des deux protéines et le « ligand » la plus petite, convention qui découle des approches de *docking* des petites molécules actives.

#### *a. Amarrage moléculaire rigide*

Dans les approches d'amarrage rigide, la géométrie des deux protéines est fixée. L'espace de recherche et le temps de calcul nécessaire sont réduits en ne considérant que six paramètres (trois dimensions de translations X, Y, Z et trois degrés de rotations (angles d'Euler) Phi ( $\phi$ ), Thêta ( $\theta$ ), Psi ( $\psi$ )) pour décrire la position du ligand sur le récepteur. Nous avons utilisé le logiciel Ptools (Saladin et al., 2009), qui repose sur le champs de force gros-grain ATTRACT2 (Chéron et al., 2017; Zacharias, 2005) (Voir 2-Approches basse-résolution gros-grain). Cette réduction en gros-grain permet de diminuer encore le temps de calcul, permettant d'obtenir un très grand nombre de solutions d'amarrages (entre 10 000 et 100 000) en quelques heures à quelques jours, selon la taille du système (nombre de particules pour décrire les macromolécules, ligand et récepteur). L'approche rigide dans un *docking* gros-grain permet de simplifier la reconstruction en tout-atome, en appliquant les transformations (translations + rotations) calculées en gros grains au modèle tout-atome. De plus, plusieurs avantages émergent de cette description des solutions par leur transformation dans l'espace à partir d'un point d'origine. D'une part, l'espace de stockage nécessaire est bien plus limité que dans un système enregistrant les positions de chaque atome de chaque solution. D'autre

part, la première étape d'analyse consiste généralement au regroupement (*clustering*) des solutions proches en un modèle de référence. Cette description simplifiée de la position des solutions dans l'espace permet ainsi de limiter la comparaison des modèles à la comparaison de leur transformation associée.

#### *b. Ajustement flexible*

Les expériences d'ajustements flexibles, dérivées des approches de simulations interactives (Voir C.4-Dynamique moléculaire sous contrainte et interactives), ont pour objectif de permettre la modélisation de modèles tout-atome à haute résolution à partir de données expérimentales SAXS basses résolutions (Molza et al., 2014). Elles sont ici assimilables à un amarrage d'une molécule à sa propre "forme moléculaire" obtenue expérimentalement, et souvent à une résolution inférieure. D'autres méthodes existent pour ajuster des modèles dans des cartes de densités expérimentales (cartes de densité électroniques ou enveloppes de densité issue de SAXS). Le logiciel MDFF (Trabuco et al., 2008) modifie un protocole de simulation de dynamique moléculaire classique dans lequel les données expérimentales sont intégrées comme un champ de potentiel supplémentaire, guidant le système. La méthode ICFF (*Internal Coordinate Flexible Fitting* (Flores, 2014)) se base sur la définition de domaines structuraux rigides, et une simulation de dynamique moléculaire simplifiée pour ajuster un modèle à une carte de densité. Cependant, ces méthodes nécessitent respectivement de grosses infrastructures de calculs, ou une connaissance approfondie et *a priori* de la dynamique de la structure étudiée pour la définition des domaines rigides. L'utilisation d'un réseau de ressorts augmenté permet de contourner ces problématiques, d'une part grâce à une simulation de dynamique moléculaire simplifiée et un système moléculaire réduit en gros-grain (champ de force Attract), et d'autre part en ne nécessitant qu'une simple connaissance géométrique de la structure de la protéine (et non pas dynamique ou topologique, comme la connaissance préalable de l'emplacement de sous-domaines ou de charnières).

La dystrophine ayant une structure suffisamment caractérisée expérimentalement, en faisceaux d'hélices à trois hélices (voir chapitre I-Introduction), le réseau de ressorts est créé en plusieurs étapes, pour contraindre les hélices et les répétitions, tout en laissant les boucles plus flexibles. Les ressorts sont dans un premier temps créés entre les grains correspondant aux carbones C $\alpha$  des hélices  $\alpha$  distants de moins de 9 Å, créant un réseau dense conservant la

structure secondaire de la protéine. Un second réseau est ensuite créé, pour relier les autres pseudo-atomes distants de moins de 5 Å. Il est finalement fusionné au premier réseau en joignant les particules distantes de moins de 5 Å.

Pour chaque fragment étudié par SAXS, entre 20 et 100 enveloppes basses-résolutions sont générées par le programme GASBOR (Franke et al., 2017), puis les trois meilleures enveloppes moléculaires sont sélectionnées en fonction de leur concordance aux données expérimentales (variable  $\chi^2$  la plus basse). Lors de la simulation d'IMD, ces enveloppes seront utilisées pour appliquer une contrainte supplémentaire sur les particules du système pour aider l'ajustement. L'expérimentateur procède ensuite à vingt ajustements par enveloppe, en guidant interactivement le modèle moléculaire dans l'enveloppe SAXS. Au cours de cette thèse, un nouveau protocole a été mis en place, pour permettre de limiter un biais trop important quant à la position initiale des modèles dans les enveloppes, et ainsi rendre l'étape d'ajustement flexible plus exhaustive et plus reproductible, et donc plus fiable (voir chapitre *IV-Modélisation de complexes de la dystrophine (actine, microtubules, lipides)*).

Pour la reconstruction d'un modèle gros-grain ajusté dans une enveloppe en un modèle tout-atome correspondant, la seule correspondance exacte entre pseudo-atomes du champ de force Attract et atomes d'un champ de force tout-atomes correspond aux atomes de carbone C $\alpha$ . Ainsi, un réseau de ressorts est créé à partir du modèle tout-atome original, avec des distances de liaison de 9 Å entre les carbones C $\alpha$  des hélices, et 3 Å entre les autres paires d'atomes lourds (sans hydrogène). De nouvelles simulations sont alors effectuées, en fixant les atomes de carbone C $\alpha$  à la position des pseudo-atomes gros-grain équivalents après ajustement dans l'enveloppe SAXS. La simulation du réseau de ressorts tout-atome par BioSpring va alors repositionner tous les atomes (hors atomes de carbone C $\alpha$ ) de manière à obtenir un modèle haute résolution du modèle gros-grain ajusté précédemment par IMD. Plusieurs étapes de minimisation courtes (entre 2000 et 5000 pas) et une étape de recuit simulé (16 000 pas) effectuée par NAMD permettent ensuite d'optimiser la position des atomes du système. Les modèles tout-atome sont ensuite évalués et comparés suivant plusieurs paramètres : l'énergie du système après minimisation (NAMD, (Phillips et al., 2005)), la qualité stéréochimique (diagramme de Ramachandran, PROCHECK (Laskowski et al., 1993)), la fraction de résidus impliqués dans des hélices (DSSP, (Kabsch and Sander, 1983)), et la correspondance aux données SAXS (CRY SOL (Franke et al., 2017)). Les meilleurs modèles sont

définis comme ceux ayant le plus de ces paramètres au-dessus de la médiane des autres résultats.

### c. Amarrage semi-flexible interactif

Les approches de *docking* semi-flexible considèrent un récepteur rigide sur lequel sera amarré un ligand flexible. Ces stratégies permettent de prendre en compte les changements de conformations que peuvent subir les partenaires lors de la formation du complexe. Elles sont généralement implémentées dans des logiciels pour le *docking* de petits ligands (AutoDock/Vina (Trott and Olson, 2010), Haddock (Dominguez et al., 2003; van Zundert et al., 2016)). Dans le travail présenté dans cette thèse, nous avons utilisé des approches d'IMD (voir *C.4-Dynamique moléculaire sous contrainte et interactives*) de modèles tout-atome basées sur le logiciel BioSpring et le champ de force tout-atome AMBER (voir *C.1-Approches haute-résolution tout-atome*).

Dans celles-ci, l'utilisateur interagit dans un système composé d'une part d'un récepteur rigide et fixé dans l'espace, et d'un ligand flexible, décrit par un réseau de ressort. L'expérimentateur va ainsi agir sur le ligand, afin de l'amener dans une conformation souhaitée, et guidée par des données expérimentales (voir *C.4-Dynamique moléculaire sous contrainte et interactives*). Le réseau de ressorts du ligand est créé par trois sous-réseaux : un réseau reliant les atomes de carbone C $\alpha$  distants de moins de 9 Å, un second reliant les autres atomes à moins de 5 Å, et le dernier joignant les particules distantes de moins de 5 Å des deux réseaux précédents. Après simulation, les modèles amarrés sont évalués de manière similaire à ceux obtenus lors des ajustements flexibles à l'exception du critère de correspondance aux données SAXS qui peut être remplacé par tout autre critère de correspondance à des résultats expérimentaux, comme des expériences de mutagénèse dirigée ou de spectroscopie de masse après *cross-link* chimique (voir *D.1-Cartographie des interactions protéine/protéine et protéine/lipide*).

## D. Techniques expérimentales de guidage des simulations

### 1. Cartographie des interactions protéine/protéine et protéine/lipide

Les approches de réticulations (*cross-link*) consistent à lier de manière covalente deux molécules au travers d'un groupement chimique réactif. Les complexes ainsi formés sont

ensuite étudiés par spectrométrie de masse pour déterminer les régions d'interactions entre les partenaires moléculaires.

*a. Cross-link protéine/protéine*

Le fragment de dystrophine R11-15 et de l'actine-F sont mis en présence d'EDC (1-éthyl-3-(3-diméthylaminopropyl)carbodiimide hydrochloride) à 5 mM pendant une heure à 20°C. L'EDC est un agent de pontage dit « zéro distance », liant les groupements carboxylates d'une protéine (aspartate et glutamate) aux amines de la seconde (lysine). Les atomes de carbone C $\alpha$  des acides aminés correspondant sont alors séparés d'environ 12Å en moyenne (Shi et al., 2014) (Figure 23).

Les échantillons sont ensuite ultracentrifugés (100 000g, 30 minutes, 4°C) de façon à ce que l'actine-F et les protéines qui lui sont liées sédimentent. Le culot est analysé par SDS-PAGE. Le gel est découpé pour isoler les bandes de migrations correspondantes aux complexes réticulés, avant analyse par spectrométrie de masse (voir *b*-Cartographie des interactions protéiques par Spectrométrie de masse).

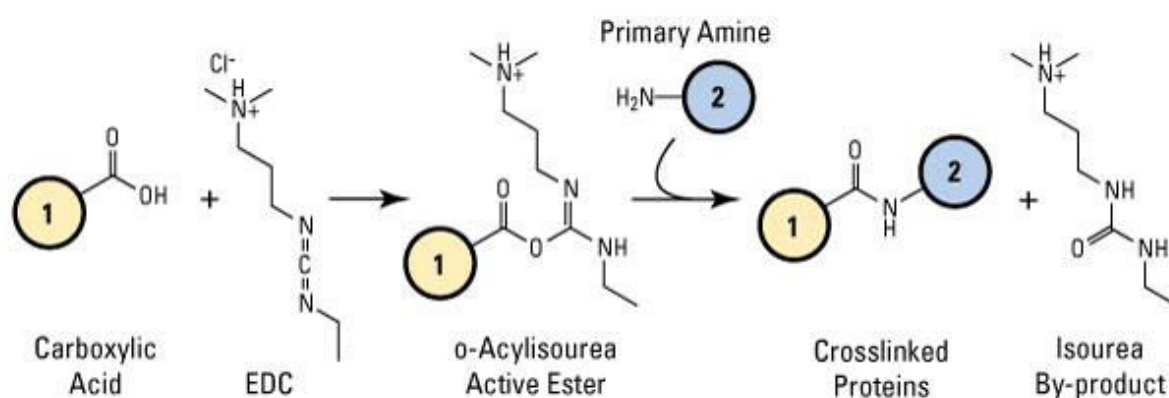


Figure 23 : Représentation schématique du cross-link « zéro distance » par l'EDC entre un groupement Jaune) carboxylate et (bleu) amine. D'après Carbodiimide Crosslinker Chemistry, [www.thermofisher.com](http://www.thermofisher.com).

*b. Cross-link protéine/lipides*

Pour étudier l'interaction entre une protéine et des modèles de membranes bicelles (voir chapitre III et IV), les bicelles ont été complétées par un lipide « cliquable », le pacFA-PC ( 1-palmitoyl-2- ( 9- ( 3-pent-4-ynyl-3-H-diazirine-3-yl ) -nonanoyl ) -sn-glycéro-3-phosphocholine, Avanti Polar Lipids). Il s'agit d'un acide gras qui contient un groupe photo-réactif diazirine. Lors d'une exposition aux UVA ce groupement chimique devient réactif et

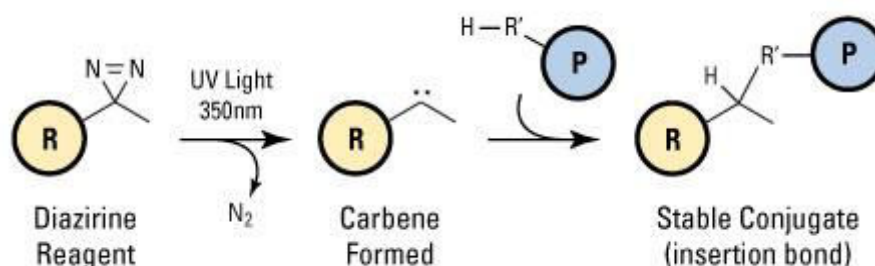


Figure 24 : Réaction de cross-link entre un groupe diazirine, porté par le pacFA-PC (R) et la chaîne latérale d'une protéine (P). D'après Diazirine reaction chemistry, [www.thermofisher.com](http://www.thermofisher.com).

permet la création d'une liaison covalente avec les résidus protéiques situés à proximité (<12Å) (Figure 24).

La protéine est incubée avec les bicelles pacFA-PC, puis le complexe est exposé cinq minutes aux UVA (5 x 15 W Stratalinker, GE Healthcare). L'échantillon est analysé par SDS-PAGE. Le gel est découpé pour isoler les bandes de migrations correspondant aux complexes réticulés, avant analyse par spectrométrie de masse.

### c. Cartographie des interactions protéiques par Spectrométrie de masse

L'analyse par spectrométrie de masse des éléments isolés précédemment permet d'identifier, selon l'expérience de *cross-link*, les fragments de la dystrophine ont été réticulés avec l'actine-F, ou les fragments protéiques qui ont été liés aux lipides cliquables. Le protocole détaillé peut être trouvé dans Dos-Santos Morais et al., 2018 (voir Annexe : Dos Santos Morais et al, 2018). Il est ainsi possible d'obtenir une cartographie « 1D », le long de la séquence protéique, des régions en interaction avec une autre protéine ou avec les lipides. Dans le cas de protéine dont on possède un modèle, il est également possible de reporter cette information 1D en 3D et de l'utiliser pour filtrer les données issues des expériences d'amarrage moléculaire rigide ou de diriger les expériences interactives d'amarrage semi-flexible (voir C.5-Amarrage moléculaire).

## 2. Diffusion aux petits angles

La diffusion aux petits angles (SAS, *Small Angle Scattering*) rassemble les techniques basées sur la déviation (la diffusion) d'un faisceau colimaté de rayons X (SAXS, *Small Angle Xray Scattering*) ou de neutrons (SANS, *Small Angle Neutron Scattering*) par un échantillon biologique en solution. La variation d'intensité du rayonnement en fonction de l'angle de



diffusion permet d'obtenir des informations sur la taille et la forme globale de l'échantillon (globulaire ou filamenteuse) (Figure 25).

Les travaux de cette thèse sont basés sur l'étude de fragments de dystrophine dont les structures sont obtenues à partir de données de SAXS (Delalande et al., 2018) et de SANS (Dos Santos Morais et al., 2017). Dans les expériences de SAXS, les rayons X incidents sont déviés par le nuage électronique entourant les atomes et permettent d'étudier la conformation de protéines en solution. L'approche par SANS repose sur l'interaction des neutrons avec le noyau des atomes de la protéine. Elle a été utilisée pour étudier les éventuels changements de conformation de la protéine au cours de l'interaction protéine/bicelle. Les atomes d'hydrogène sont les contributeurs majoritaires du signal SANS, et celui-ci peut être modulé en introduisant du deutérium dans l'échantillon. Ainsi par exemple, une protéine hydrogénée est invisible dans une eau deutérée à 42%, et une bicelle deutérée est invisible dans une eau deutérée à 100%. Il est ainsi possible d'obtenir le signal des deux partenaires en complexe de manière indépendante ainsi que le signal du complexe total.

Les données issues des expériences de SAXS ou de SANS sont globalement similaires, et elles peuvent être utilisées à différents niveaux, soit dans l'étude de modèles structuraux, soit dans la génération de modèles sous contraintes expérimentales.

Au cours des études de SAS, il est possible d'isoler certains paramètres de l'échantillon étudié et de les comparer directement à ceux obtenus à partir de modèles moléculaires, comme le rayon de giration de la protéine ou du complexe (distance pondérée par la masse de la distance moyenne des atomes d'un objet à son barycentre), ou alors la  $P(r)$ , c'est-à-dire la distribution de la distance des atomes pair-à-pair. Enfin, à partir d'un modèle moléculaire,

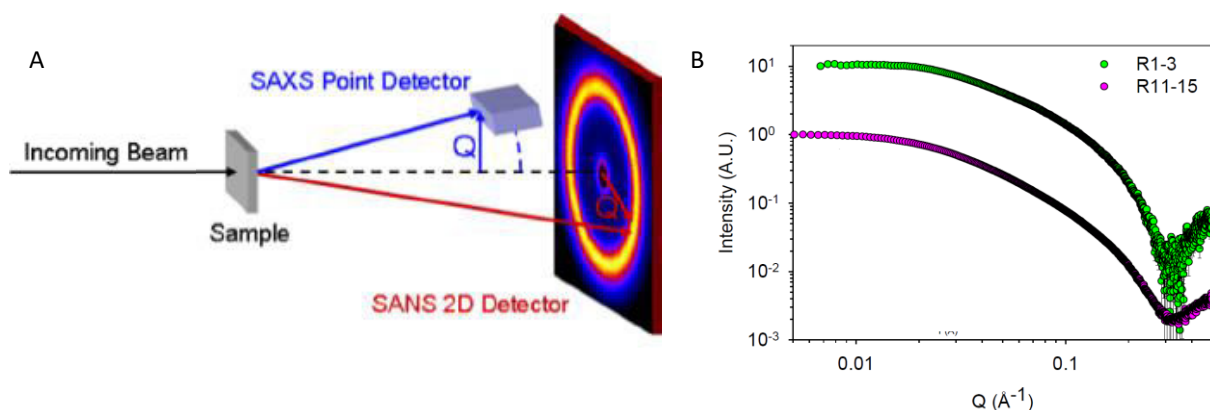


Figure 25 : A) Principe des expériences SAS et des mesures (bleu) SAXS et (rouge) SANS ; B) Exemple de courbes SAXS obtenues à partir des fragments (vert) R1-3 et (violet) R11-15.

il est également possible de produire une courbe SAS théorique (avec les programmes CRY SOL pour le SAXS et CRYSON pour le SANS, tous deux disponibles dans la suite de logiciel ATSAS (Franke et al., 2017)) et de comparer celle-ci avec la courbe expérimentale.

Les approches SAS peuvent aussi permettre d'obtenir des informations basses résolutions sur la structure d'une protéine. Des outils tels que DAMMIN/DAMMIF ou GASBOR (Franke et al., 2017) permettent de générer des volumes moléculaires basses résolutions à partir des courbes expérimentales. Ces volumes servent de contrainte expérimentale lors des simulations d'ajustements flexibles (voir *C.5.b-Ajustement flexible*). Deux stratégies existent pour la création des volumes moléculaires. Pour les logiciels DAMMIN et DAMMIF, l'espace est représenté par une organisation de sphères densément organisées, tels les atomes dans un cristal. Ces sphères sont ensuite définies comme appartenant au solvant ou à la protéine pour concorder au mieux avec les données expérimentales. Bien que cette méthode soit plus rapide que GASBOR, la représentation « cristalline » de l'espace limite la création de modèles présentant des irrégularités prononcées, tels que des coudes de grandes amplitudes. Dans le programme GASBOR, la protéine est initialement représentée par des grains (un grain par acide aminé) placés dans l'espace sous la forme d'une grande sphère dont la taille est définie par le rayon de giration expérimental. Les positions des grains sont ensuite successivement modifiées suivant un algorithme de type Monte-Carlo pour concorder au mieux avec les données expérimentales. Cette méthode, plus lente que DAMMIN/DAMMIF, permet cependant de mieux représenter une protéine en solution car la distance entre deux grains est calibrée de manière à correspondre à la distance moyenne observée entre deux atomes de carbone C $\alpha$  dans les structures protéiques.

### 3. Cryo-tomo-électromicroscopie

Dans les approches par cryo-microscopie électronique (CEM, Cryo-electro-microscopie) ou de cryo-tomo-microscopie électronique (CET, Cryo-electron Tomography), l'échantillon est déposé sur une grille de microscopie électronique puis l'excès de liquide est éliminé par application d'un papier absorbant pour créer un film fin d'environ 0,1  $\mu\text{m}$  d'épaisseur. Pour un même échantillon, cette opération est effectuée sur trois grilles différentes, avec des durées d'application du papier de 1s, 1,2s et 1,4s. Les grilles sont ensuite rapidement plongées dans de l'éthane liquide maintenu à -193°C par de l'azote liquide. Cette étape permet de figer

l'échantillon dans une glace amorphe, une glace cristalline diffracterait les électrons et empêcherait l'observation des objets qu'elle contient.

Les grilles d'observations sont recouvertes d'un film de carbone aux propriétés hydrophobes. Préalablement aux dépôts, il est donc nécessaire d'ioniser les grilles afin d'augmenter l'hydrophilie du carbone pour permettre à l'échantillon de s'étaler sur toute la surface de la grille. Cette étape est effectuée grâce à un appareil de *Glow Discharge*, dans lequel la grille est placée dans une chambre à vide partiel, et un courant d'environ 500 Volts est appliqué entre la grille et une anode. Cette différence de potentiel ionise négativement le gaz, qui se dépose ensuite sur la grille, donnant à la surface de carbone un caractère hydrophile. Les grilles sont ensuite conservées dans l'azote liquide jusqu'à observation.

Sur des échantillons suffisamment concentrés, des approches semi-automatisées peuvent être employées pour isoler les objets d'intérêt sur les micrographies (Macé et al., 2018) et reconstruire des volumes moléculaires 3D à partir d'images 2D.

Dans le cas de filaments comme l'actine ou les microtubules, les préparations doivent avoir une concentration relativement faible ( $\sim 10\mu\text{M}$  de monomères) pour éviter un chevauchement des fibres qui empêcherait de les isoler pour les étudier. D'un autre côté, en limitant la quantité d'actine présente dans l'échantillon, on limite la capacité à acquérir suffisamment de micrographies en 2D pour les reconstructions 3D. Pour obtenir une information volumique, nous avons choisi d'utiliser une technique de tomographie, ou CET (*Cryo-electron Tomography*). Celle-ci consiste à la prise de vue de la même région de

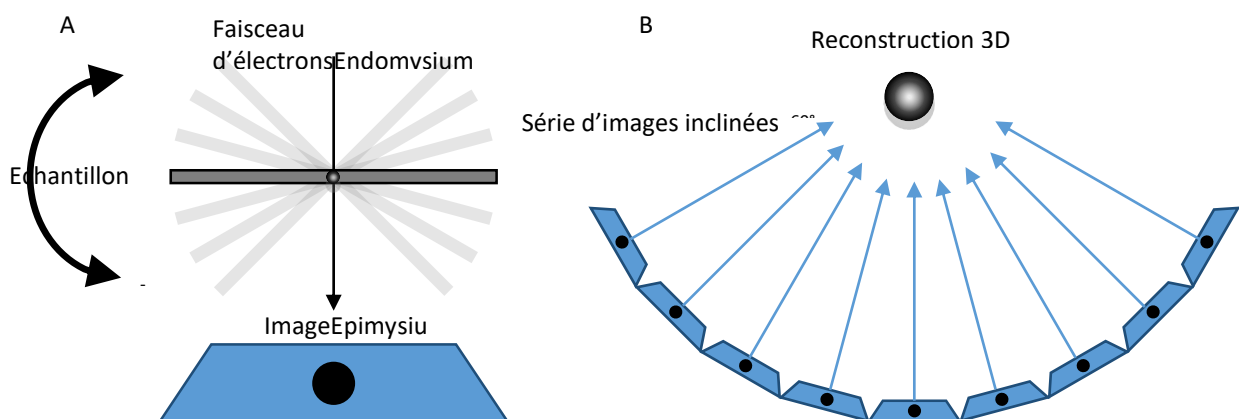


Figure 26 : Principe de la CET : A) acquisition et de B) reconstruction.

l'échantillon sous différents angles (en général de  $-60^\circ$  à  $+60^\circ$ ), permettant après traitement informatique de reconstruire une information 3D (Figure 26).



### III. Etude des interactions hydrophobes du faisceau d'hélice R1-3

#### A. Introduction

##### 1. Spectrine et dystrophine

Nous avons vu dans l'introduction de cette thèse que les faisceaux d'hélices sont des éléments structuraux retrouvés dans de nombreuses protéines (Lupas and Bassler, 2017; Truebestein and Leonard, 2016), comme la spectrine localisée à la membrane des érythrocytes (Zhang et al., 2013), les laminines qui structurent la matrice extra-cellulaire (Domogatskaya et al., 2012) ou bien les complexes SNARE qui permettent le trafic de vésicules intra et inter-cellulaires (Hong and Lev, 2014).

Le domaine central de la dystrophine est souvent décrit comme ayant une structure similaire à celle de la spectrine, évoquant une baguette droite (*rod domain*) et avec une flexibilité limitée (Grum et al., 1999; Koenig et al., 1988). Cette similitude structurale a été proposée d'une part à cause de la proximité phylogénétique entre les deux protéines (Byers et al., 1992; Dubreuil, 1991), et d'autre part du fait de la faible différence de structure entre leurs sous-domaines N-terminaux ABD1 (Norwood et al., 2000). De plus, dans le cas de la spectrine, la linéarité et la rigidité de son domaine central sont en partie associés à l'état de multimérisation des protéines de spectrine- $\alpha$  et de spectrine- $\beta$  qui forment un hétérotétramère  $(\alpha\text{-}\beta)_2$  (Figure 27).

Cette régularité est retrouvée dans d'autres paramètres de la spectrine. Ainsi, l'étude de la séquence des répétitions de la spectrine (Conway and Parry, 1991; Parry et al., 1992) indique une grande régularité dans la conservation et dans la répétition des heptades, celle-ci divergeant très peu de la forme canonique. Ainsi, environ 70% des positions *a* et *d* sont occupées par des résidus apolaires (leucine, valine, isoleucine, méthionine, phénylalanine, tyrosine, alanine, lysine) (Conway and Parry, 1991). De plus, la longueur des hélices varie très peu d'une répétition à une autre. En effet, sur ce dernier point, on observe que les hélices A des spectrines  $\alpha$  et  $\beta$  sont systématiquement composées de 27 acides aminés, à l'exception de l'hélice A de la répétition R10 de la spectrine  $\alpha$ , et R15 de la spectrine  $\beta$ , qui n'en possèdent que 26 (Nicolas et al., 2014a).

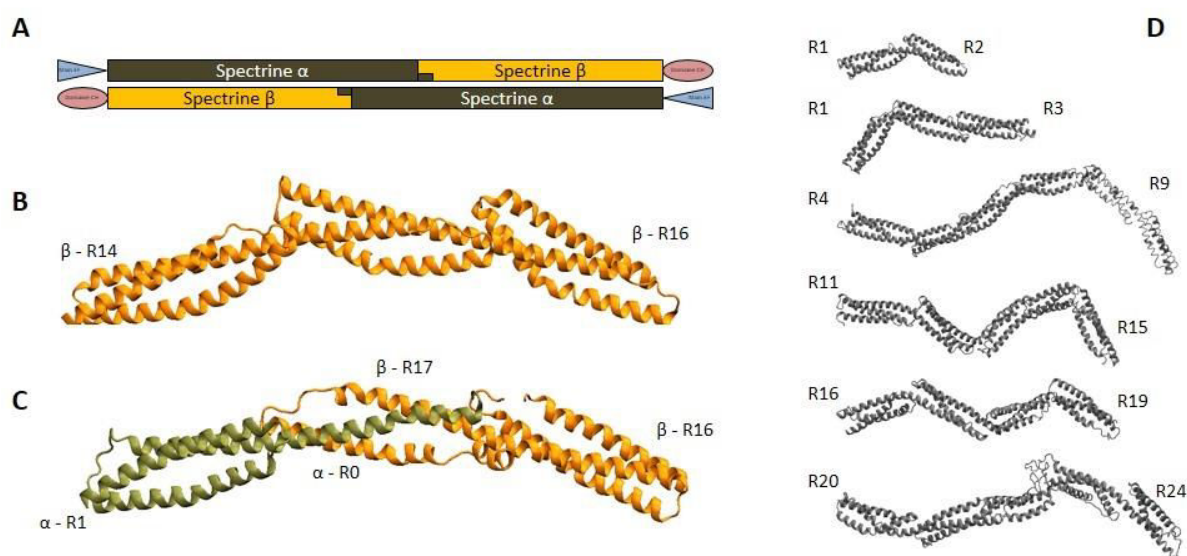


Figure 27: Comparaison de structures de faisceaux d'hélices de la spectrine et de modèles SAXS de faisceaux d'hélices de la dystrophine ; A) Représentation schématique du tétramère de spectrine  $\alpha/\beta$ . En vert : spectrine  $\alpha$ , en jaune : spectrine  $\beta$  ; B) Structure cristallographique (PDBID : 3EDV, Homo sapiens) des répétitions R14 à R16 du domaine central de la spectrine  $\beta$  C) Structure cristallographique (PDBID : 3LBX, Homo sapiens) du site de tétramérisation du domaine central de la spectrine ; En vert, répétitions R0 et R1 de la spectrine  $\alpha$ , en jaune : répétitions R16 et R17 de la spectrine  $\beta$  D) Modèles structuraux obtenus sous contraintes SAXS de fragments du domaine central de la dystrophine humaine. On observe que les fragments de la spectrine tendent à être linéaires, alors que des coudes importants sont observés entre certaines répétitions du domaine central de la dystrophine.

Les axes principaux des trois hélices ne sont pas rectilignes, les hélices s'enroulent les unes autour des autres pour former un faisceau ou *coil* (Yan et al., 1993). Il a été également noté que des acides aminés brisant ou perturbant les hélices  $\alpha$ , comme les prolines ou les glycines, sont systématiquement retrouvés aux extrémités C-terminales des hélices A et B (Nicolas et al., 2014a), induisant la présence d'une boucle A/B entre les hélices A et B et d'une boucle B/C entre les hélices B et C. Cette même observation n'a pas été faite en ce qui concerne la fin des hélices C: la région connectant l'hélice C d'une répétition à l'hélice A de la répétition suivante, aussi appelée *linker*, a ainsi été prédite comme étant hélicale dans un premier temps, et cela a été vérifié expérimentalement (Grum et al., 1999). Cependant, il n'est toujours pas tranché dans la littérature si cette région *linker* est plutôt rigide (Bloch and Pumplin, 1992; Speicher and Marchesi, 1984) ou bien flexible (Nicolas et al., 1998). Cependant, comme cette dernière étude porte sur l'étude de fragments monomériques, les variations d'angles observées pourraient être artefactuels.

C'est en se basant sur les structures de fragment de spectrine disponibles dans la PDB (R15-16-17 de la spectrine  $\alpha$  (*G. gallus* PDB ID 1U4Q), répétitions R8-R9 (*H. sapiens* PDB ID 1S35) et R14-15-16 de la spectrine  $\beta$  (*H. sapiens* PDB ID 3EDV)) qu'ont été établis les premiers

modèles par homologie de répétitions de la dystrophine (Nicolas et al., 2014a). Cependant, la dystrophine est fonctionnelle sous la forme d'un monomère, alors que deux des modèles de spectrine qui ont servi comme modèles se situent à proximité du site de tétramérisation.

Contrairement à la spectrine, les faisceaux d'hélices de la dystrophine sont légèrement modulés par rapport au modèle canonique, avec par exemple une grande hétérogénéité de la longueur des hélices de la dystrophine (par exemple, de 31 à 37 résidus pour les hélices A), et aussi des limites de la définition des hélices B et C qui varient d'une répétition à une autre (Nicolas et al., 2014a). De plus, on observe une divergence de composition en acides aminés du motif en heptade, aux niveaux des positions *a/d* classiquement hydrophobes (Winder et al., 1995). Ces irrégularités dans la séquence et la structure de la dystrophine semblent associées à l'absence de prédisposition de celle-ci à adopter une forme multimérique.

Récemment, l'analyse de données structurales acquises par SAXS a permis l'obtention des premiers modèles de répétitions successives sous contraintes expérimentales de la dystrophine (Delalande et al., 2018). Ceux-ci mettent clairement en évidence la non-linéarité du domaine central de la dystrophine: des coudes ( $>30^\circ$ ) aux régions *linker* sont notamment observés entre les répétitions R1-R2, R4-5, R6 à R8, entre toutes les répétitions de R11 à R15, R16 à R18, et R21 à R23. De plus, un certain degré de flexibilité peut être attendu entre certaines répétitions suivant le contexte moléculaire. Ainsi, dans les deux modèles SAXS contenant les répétitions R1 et R2 (R1-2 et R1-3), l'angle du coude de la région *linker* varie entre  $40^\circ$  et  $75^\circ$ , respectivement.

Notre équipe a précédemment émis l'hypothèse que ces coudes, observés auparavant *in silico*, pourraient avoir une fonction essentielle en délimitant des domaines d'interactions sur le domaine central de la dystrophine (Legrand et al. 2011). En effet, la diversité de partenaires avec lesquels interagit, potentiellement de manière simultanée, cette grande région ne peut pas s'expliquer par la répétition monotone d'un même motif structural. Les coudes observés au niveau des régions *linker* et les charnières (*hinges*, voir *I-Introduction*) pourraient ainsi permettre de délimiter les sous-domaines individualisés dictant la diversité et la spécificité des interactions de la dystrophine avec ses partenaires. Les différences observées dans les répétitions *coiled-coil* de la spectrine et de la dystrophine suggèrent la nécessité d'une nouvelle caractérisation du faisceau d'hélice de la dystrophine, notamment pour souligner ses propres spécificités vis-à-vis des répétitions de type spectrine. De plus, une analyse



approfondie de la structuration des régions *linker* permettrait de mieux comprendre les propriétés dynamiques du domaine central de la dystrophine.

## 2. Interactions hydrophobes

Dans une publication de 1805, Thomas Young définit l'hydrophobie d'un liquide en fonction de la forme d'une goutte de ce composé déposé sur une surface de verre (Young, 1805) (Figure 28A). Berthelot a ensuite émis en 1872 le principe du coefficient de partage entre une phase hydrophile et une phase hydrophobe d'un composé (Berthelot and Jungfleisch, 1872). En 1971, Leo et ses collaborateurs formuleront l'expression moderne de ce coefficient de partage, noté  $\log P$ , entre l'eau (phase hydrophile) et l'octanol (phase hydrophobe) (Leo et al., 1971). Le  $\log P$  correspond au logarithme de base 10 du ratio des concentrations du composé solubilisé dans l'1-octanol et l'eau ( $\log P = \log \frac{C_{\text{octanol}}}{C_{\text{eau}}}$ ). Une valeur positive marque la plus grande affinité du produit étudié avec la phase lipophile, le composé étant ainsi défini comme hydrophobe. À l'échelle atomique, l'hydrophobie correspond à la propriété physique d'un atome à ne pas modifier son nuage électronique dans un environnement polaire et qui ne forment pas de liaisons hydrogène avec l'eau ou un groupement polaire. Une interaction hydrophobe correspond à l'induction de dipôles faibles

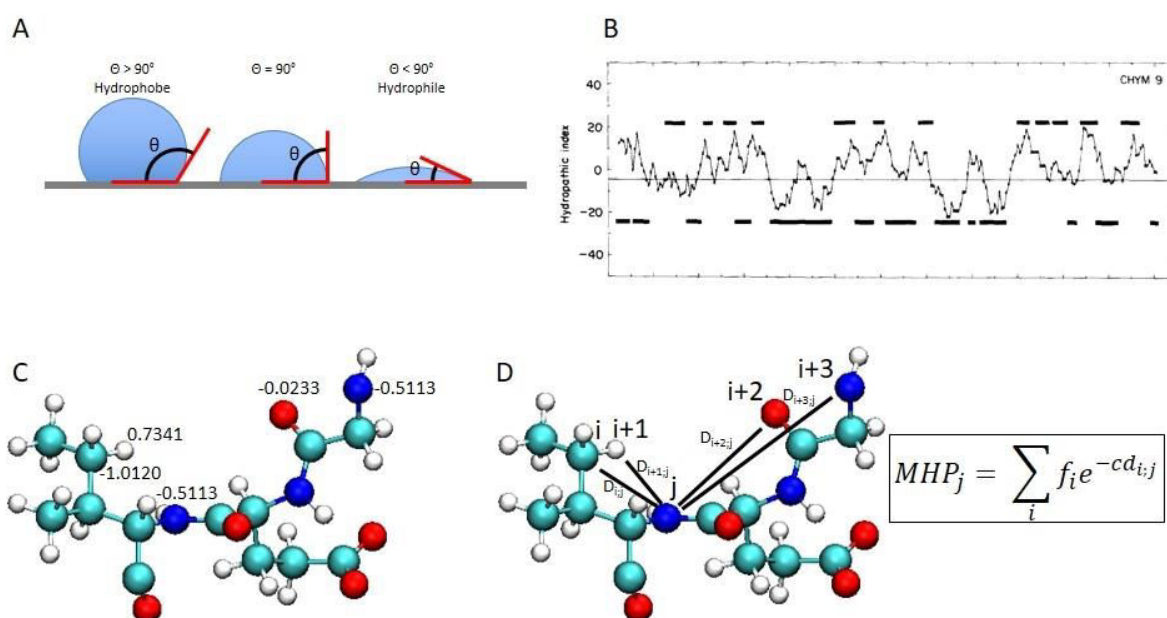


Figure 28 : A) Principe selon Young de la détermination de l'hydrophobie d'un composé liquide (bleu) déposé sur une lamelle de verre hydrophile. B) Exemple d'analyse de l'hydrophatie le long de la séquence du chymotrypsinogène proposé par Kyte & Doolittle (Kyte and Doolittle, 1982). C) Exemple d'assignation de quelques valeurs de potentiels atomiques d'hydrophobie selon les paramètres de Ghose (Ghose and Crippen, 1987) pour le peptide (Gly-Glu-Ile). D) Gauche : schéma simplifié du calcul du MHP d'un atome  $j$ . Les auteurs de PLATINUM suggèrent de fixer le seuil au-dessus duquel on considère un atome hydrophobe à 0.03. Droite : Haut : Formule du potentiel hydrophobe moléculaire (MHP) d'un atome  $j$ .

entre deux atomes hydrophobes. Dans une protéine, ces interactions se forment généralement entre des atomes de groupements aliphatiques ou aromatiques.

D'autres méthodes ont été mises en œuvre pour étudier l'hydrophobie dans les protéines. On notera particulièrement les travaux de Kyte et Doolittle permettant d'évaluer l'hydrophobie le long de la séquence d'une protéine (Kyte and Doolittle, 1982). Cette approche est utilisée pour identifier ou caractériser des sous-domaines de la protéine (Figure 28B).

En parallèle, dans une publication de 1986, Taylor décrit les propriétés générales des acides aminés en les organisant dans un diagramme de Venn. Bien qu'encore très utilisée de nos jours, cette classification simpliste ne propose qu'une définition qualitative de l'hydrophobie (Taylor, 1986).

Par la suite, des méthodes informatiques ont été développées dans le but de prédire la valeur d'hydrophobie d'une molécule (CLOGP (BioLoom 5.0), ALOGP (Ghose and Crippen, 1987)) en estimant la participation au logP de chacun des atomes qui la composent. Ces méthodes reposent sur le calcul préalable de tables attribuant une valeur d'hydrophobie (Figure 28C) aux atomes dans différents contextes physico-chimiques (Ghose et al., 1998). Dans le cas de grandes molécules telles que les protéines, ce calcul n'est pas seulement lié à la nature de l'atome étudié mais il nécessite également de prendre en compte l'environnement distant. Les méthodes calculant le potentiel hydrophobe moléculaire (MHP, *Molecular Hydrophobic Potential*) pondèrent la valeur d'hydrophobie d'un atome par la valeur d'hydrophobie des atomes voisins en fonction de leur distance (Figure 28D). PLATINUM (Pyrkov et al., 2009) est un programme accessible à partir d'un serveur en ligne qui permet de calculer rapidement le MHP d'une molécule.

### 3. Importance biologique du fragment R1-3 de la dystrophine

Le fragment composé par les répétitions R1 à R3 de la dystrophine (R1-3) correspond à l'extrémité N-terminale de son domaine central. Ce sous-domaine délimité par deux *hinges* (H1 et H2) a été précédemment caractérisé comme étant une région LBD (*lipid binding domain*) se liant aux membranes lipidiques *in vitro* (Legardinier et al., 2008) et *in vivo* (Zhao et al., 2016).

Une cartographie biochimique et biophysique exhaustive du domaine central de la dystrophine a montré la grande variabilité des propriétés mesurées pour chaque répétition

de la protéine (Mirza et al., 2010). Les auteurs de cette étude ont aussi montré que certaines répétitions en tandem forment un ensemble plus stable que les répétitions les composant individuellement. Ce comportement a notamment été souligné pour les répétitions R2 et R3 et la répétition en tandem correspondante R2-3. Ainsi, dans la région formée par les répétitions R1, R2 et R3, les fractions repliées en hélice sont de 66%, 98% et 51%, respectivement alors que dans les répétitions en tandems R1-2 et R2-3, ces valeurs sont de 63% et 91%. Les auteurs de cette étude observent ainsi une nette stabilisation de la structure en hélice de la répétition R3 par la répétition R2, mais pas de celle de la répétition R1 par la répétition R2. Des résultats similaires ont été obtenus pour d'autres paramètres tels que l'enthalpie apparente de formation  $\Delta H$  qui correspond à une valeur thermodynamique associée à la stabilité du fragment protéique. Mirza et ses collaborateurs ont pu ainsi avancer l'hypothèse que le fragment R1-3 pouvait être considéré comme constitué de deux sous-domaines: l'un étant composé de la répétition R1 et l'autre du tandem R2-3, ce second sous-domaine étant plus stable que chacune des répétitions R2 et R3 prises séparément. Ce concept de sous-domaines fonctionnels est retrouvé dans l'efficacité des différentes mini- et micro-dystrophines à visée thérapeutique, dans lesquelles des constructions contenant uniquement R1-2 injectées dans des souris *mdx* ne restaurent pas autant de capacité de contraction que des constructions contenant R1-3 (Meng et al., 2016; Sakamoto et al., 2002).

Il a été prédit un motif senseur de courbure des membranes au niveau de la moitié C-terminale de l'hélice B de la répétition R1 (Drin et al., 2007). Ce motif, nommé ALPS (*ArfGAP1/Amphipathic Lipid-Packing Sensor*), est composé d'une hélice  $\alpha$  amphiphile présentant une face polaire et une face apolaire (Figure 29A) s'adsorbant aux membranes présentant une courbure positive ou un défaut de compacité des lipides. Dans la dystrophine, les résidus identifiés comme faisant partie du motif ALPS prédit sont partiellement enfouis dans le faisceau d'hélice (Figure 29B) car ils sont situés aux positions *g/a* ou *d/e*, fortement impliqués dans les contacts inter-hélices (Val 453, Ile 456, Gly 460, Leu 463, voir publication de ce chapitre).

Ainsi, si la fonction de liaison aux lipides du domaine ALPS de la dystrophine était avérée, elle serait associée à l'ouverture du faisceau d'hélice de la répétition R1 qui a été observée expérimentalement par notre équipe grâce à une approche couplée de SANS, de spectroscopie de masse (MS) et de simulation de dynamique moléculaire (Figure 29C).

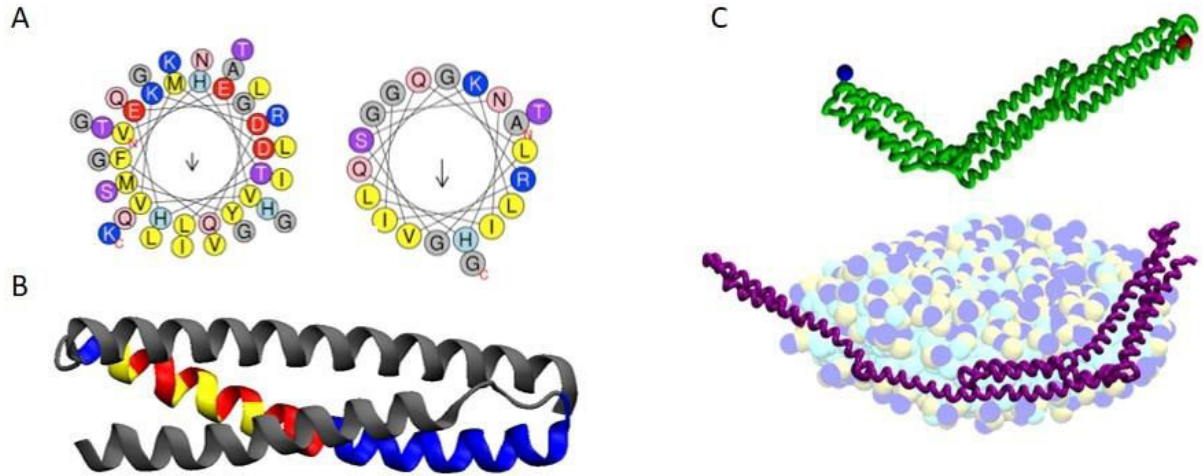


Figure 29 : A) Projection axiale des résidus (gauche) de l'hélice B de la répétition R1 (droite) de la région ALPS de l'hélice B de la répétition R1; jaune : Résidus hydrophobes. B) Structure cristallographique de la répétition R1 (PDB ID : 3UUN) ; bleu : hélice B, rouge : motif ALPS, jaune : Résidus hydrophobes de l'ALPS. C) Haut : Modèle SANS du fragment R1-3 de la dystrophine en solution ;Bas : Modèle SANS du fragment R1-3 de la dystrophine liée à une bicelle anionique

(DMPC / DMPS / DHPC)

B. Article 1 : Fine mapping of hydrophobic contacts reassesses the organisation of the first three dystrophin coiled-coil repeats

L'étude des interactions au sein des répétitions R1, R2 et R3 de la dystrophine réalisée au cours de cette thèse a fait l'objet d'une publication scientifique [Mias-Lucquin et al., 2018 ; Article en révision dans la revue *Protein Science*]. Nous y présentons une approche basée sur le calcul du potentiel hydrophobe de cartographie des interactions hydrophobes à l'échelle atomique, qui nous a permis de redéfinir les spécificités du faisceau d'hélices d'une répétition de dystrophine et de caractériser l'organisation structurale de la région « *linker* » entre deux répétitions. Cette méthode, implémentée dans le logiciel Pynte, a permis d'explorer efficacement les interactions inter-hélices au sein d'un *coiled-coil* et de prédire les bases physico-chimiques qui sous-tendraient la flexibilité du domaine central de la dystrophine.

Ce travail est complémentaire à celui que nous avons publié récemment (Voir Annexe : Dos Santos Morais et al, 2018), et qui porte sur l'étude de la structure et de la dynamique du fragment R1-3 de dystrophine en interaction avec des bicelles zwitterioniques (DMPC/DHPC 0,57:0,43) ou anioniques (DMPC/DMPS/DHPC 0,38:0,19:0,43).

Cette étude a notamment montré que pour l'interaction dystrophine-bicelle anionique,

- i) le fragment R1-3 interagit avec le pourtour des bicelles, lieu présentant un défaut de compacité des assemblages de lipides;
- ii) le fragment R1-3 aborde un large changement conformationnel correspondant à une ouverture du faisceau d'hélice de la répétition R1;
- iii) le complexe dystrophine-bicelle est stabilisé par le biais d'interactions hydrophobes impliquant notamment des acides aminés du motif en heptade et du motif ALPS prédit.

# **Fine mapping of hydrophobic contacts reassesses the organisation of the first three dystrophin coiled-coil repeats**

*Dominique Mias-Lucquin,<sup>1</sup> Angélique Chéron,<sup>1</sup> Elisabeth Le Rumeur,<sup>1</sup> Jean-François Hubert,<sup>1</sup> and Olivier Delalande<sup>\*, 1</sup>*

1 Univ Rennes, CNRS, IGDR - UMR6290, F-35000 Rennes, France

\* Corresponding author

Olivier Delalande

Institut de Génétique et Développement de Rennes

Faculté de Médecine, 35043 Rennes Cedex, France

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

Fax number: (+33)(0)2 23 23 46 06

RUNNING TITLE: Plasticity of first dystrophin coiled-coil repeats

MANUSCRIPT INFORMATION: 25 pages; 4 supplementary material pages; no table;  
4 figures

SUPPLEMENTARY MATERIAL CONTENT: Table ST 1 and Figures SM 1 to SM 3 in  
a file named MiasLucquin\_et\_al\_SI.pdf

**ABSTRACT**

Coiled-coil domain is a structural motif found in proteins crucial for achievement of central biological processes, such as cellular cohesion or neuro-transmission. The coiled-coil fold consists of alpha-helices bundle that can be repeated to form larger filament. Hydrophobic residues, distributed following a regular seven-residues pattern, named heptad pattern, are commonly admitted to be essential for the formation and the stability of canonical coiled-coil repeats. Here we investigated the three first coiled-coil repeats of the central domain of dystrophin, a scaffolding protein in muscle cells whose deficiency leads to Duchenne and Becker Muscular Dystrophies. By an automated approach based on the Molecular Hydrophobic Potential (MHP), we highlighted (i) that coiled-coil filament conformational changes are associated to specific patterns of inter-helices hydrophobic contacts, (ii) that inter-repeat hydrophobic interactions determine the behaviour of linker regions including filament kinks, and (iii) that a non-strict conservation of the heptad patterns is leading to a relative plasticity of the dystrophin coiled-coil repeats. These structural properties could play an essential role in folding/unfolding processes related to their biological function. By considering an atomic description of the hydrophobic interactions in biomolecular structures, we propose clues to understand the relation between conformational modulation and mechanical properties of filamentous proteins made of coiled-coil repeats. More specifically, our structural description of the three first dystrophin repeats could allow great improvement in the protein design related to gene therapies for the cure of muscular dystrophies.

**KEYWORDS:** dystrophin; coiled-coil fold; molecular hydrophobicity potential

**BROADER AUDIENCE ABSTRACT**

Coiled-coil fold consists of alpha-helices bundle formed and stabilized through hydrophobic contacts. By the observation of their modulation, we characterized the plasticity of the dystrophin coiled-coil filament, a scaffolding protein whose deficiency leads to muscular dystrophies. The understanding of these structural properties that could play an essential role in its biological function should allow a great improvement in the protein design related to gene therapies for the cure of dystrophin-related pathologies.



## INTRODUCTION

The coiled-coil structural fold is found in numerous proteins involved in central biological processes like transcription<sup>1</sup>, neurotransmission<sup>2</sup>, membrane trafficking<sup>3</sup> or cellular cohesion<sup>4</sup>, and it also serve as a basis for the engineering of synthetic molecules<sup>5</sup>. The coiled-coil fold consists of a bundle of multiple alpha-helices wrapped together to form a super-helix<sup>6</sup>. It is well established that alpha helices involved in coiled-coil include a regular winding of seven residues called heptad pattern forming two helix turns<sup>7,8</sup>. The positions of these seven residues are noted from *a* to *g*, with positions *a* and *d* being canonically occupied by hydrophobic residues. The alpha-helix unwinding from 3.6 to 3.5 residues by turn<sup>8</sup> leads the *a* and *d* residues to create hydrophobic stripes which associate into a knobs-into-holes packing<sup>9</sup> between different alpha-helices. Coiled-coils are found in the spectrin superfamily as three consecutive alpha-helices (A, B and C) of approximately 33 residues wrapped into a bundle forming a unit called “repeat”<sup>10</sup>. In multirepeat filamentous proteins, the third helix of a repeat is directly extended to the first helix of its following repeat by a short region called linker and therefore common to both successive repeats.

Within the context of spectrin superfamily, dystrophin is a protein bearing a central domain composed of 24 coiled-coil repeats (R1 to R24) that is involved in severe muscular pathologies like Duchenne and Becker Muscular Dystrophies and that presents some deviations from the standard properties of a canonical coiled-coil fold<sup>11-13</sup>. First, while the heptad pattern is highly conserved in spectrin<sup>14</sup>, a probable structural requirement for its efficient dimerization, this is less strictly the case for dystrophin as proposed from the sequence alignment by Winder *et al*<sup>11</sup>. As this observation could be related to the monomeric nature of dystrophin<sup>10</sup>, we should nevertheless mitigate this observation since it has been appropriately highlighted in some spectrin repeats that charged residues with long nonpolar chains could recover hydrophobic pattern<sup>15</sup>. Secondly, some differences are observed in the length of spectrin and dystrophin

repeats or helices even if a structural comparison between single coiled-coil repeat of each of these proteins allows to establish a high degree of similarities in their organisation<sup>12</sup>. For instance, the conservation of an hydrophobic residue to start a stammer (*abc* triad before heptad) at the beginning of helix B was previously noticed in spectrin<sup>15</sup> and can also be found in dystrophin from sequence alignment. On the contrary, polar residues promoting salt-bridges and/or H-bonding found at *e/g* positions in spectrin<sup>16</sup> are not as well conserved in dystrophin<sup>11</sup>. Finally, the bending of helices observed in coiled-coil repeats of spectrin<sup>16</sup> or alpha-actinin<sup>17</sup> are also not systematically observed in the same way in a dystrophin single repeat structure<sup>18</sup>. In dystrophin first coiled-coil repeat, only helix B is bent in its middle<sup>18</sup> as anticipated by our studies of secondary structure prediction and homology modelling<sup>19</sup>. This region of helix B is corresponding to a *d* position that is associated to a very low conservation of hydrophobic residue in the dystrophin sequence<sup>11</sup>. It has been previously noticed that this spot of weakened secondary structure is close to exon phasing position<sup>12</sup>.

The linker region was first predicted to be helical in spectrin<sup>16</sup> and this was confirmed later from experimental determination of the structure of a tandem coiled-coil repeat of spectrin<sup>15</sup>. Nevertheless, several studies presented contradictory clues concerning the spectrin linker rigidity<sup>20,21</sup> or flexibility<sup>22,23</sup>. The dystrophin central domain was described as a nested structure<sup>24</sup>, a property that originates from the contacts made by the linker region with flanking loops or helices as anticipated from the C-terminal part of the structure of a single spectrin repeat<sup>16</sup>. Indeed, contacts made by loop AB (located between helix A and helix B) with the 2-helices coiled-coil part of a repeat (involving helices B and C) is described in spectrin<sup>16</sup> as well as in dystrophin structures<sup>18</sup>. Once more, the tandem repeat structure has brought precise insights relative to linker region interactions (loop AB of repeat N and loop BC of repeat N+1) describing the flexible behaviour of this region presented as a ball-and-socket hinge with flexibility constraints in some directions<sup>15</sup>. All these structural features were confirmed as key

aspects by theoretical studies of molecular dynamics of spectrin tandem repeats<sup>25,26</sup> As a whole, it was recognised that in a monomeric filament some spectrin-like tandem repeats exhibit flexibility at their linker regions. However, atomic information that could help to interpret and to predict linker bending is lacking.

The differences inherent to the dimeric and monomeric nature of spectrin and dystrophin, respectively, raised the question about the relevance of a comparison between their dynamical properties. A recent study of multi-repeat fragments of dystrophin in solution by Small Angle X-ray Scattering<sup>27</sup>, revealed the existence of high degrees of bending at some linker regions. Modulations in the amplitude and directionality of linker bending have been observed, notably for the bending angle of the filament axis between repeats R1 and R2 that is 40° in the R1-2 fragment and 75° in the R1-3 fragment. By looking at these results, it appears necessary to revisit the coiled-coil organisation and linker structural behaviour to better describe the flexibility properties of the first three repeats of dystrophin central domain, repeats which are frequently included in chimera protein designed for therapy strategies<sup>28-30</sup>. To this aim, we applied to crystal structure of repeat R1 and to SAXS-based models of R1-2 and R1-3 fragments a new approach named Pynte and based on Molecular Hydrophobic Potential (MHP)<sup>31-33</sup> which allows the measure and the graphical rendering of the hydrophobic interactions in coiled-coil folds. Thanks to this approach we described precisely the variety of structural changes accessible to coiled-coil dystrophin first repeats and the related effects on the dynamical and mechanical properties of its central domain. The description of this structural plasticity of the first three dystrophin repeats could be of interest for the improvement of protein design used in gene therapies for the cure of Duchenne and Becker Muscular Dystrophies.

## RESULTS



the loop AB and should also contact the helix A of the next repeat, playing a role in the organisation of the linker region. By comparison between theoretical CCB model of a three-helice coiled-coil and the crystallographic structure of dystrophin repeat R1, we can observe in the latest an increased packing of the helix bundle (Figure 1C and 1D). This appears to be mainly caused by more bent and supercoiled helices in the experimental structure of repeat R1, modulating the burying of *a/d* and *e/g* positions.

***Atomic mapping of hydrophobicity reveals dense and asymmetric inter-helices contacts in the dystrophin coiled-coil.*** For a long time, hydrophobicity has been only defined as a qualitative global parameter relative to the nature of the amino acid sidechain<sup>34</sup>. Thus, hydrophobic contacts were simply defined as the closeness of two residues recognised as "hydrophobic". As a finer alternative, Tina et al. proposed an approach based on the interaction between a predefined list of apolar groups in hydrophobic sidechain<sup>35</sup>. To go even deeper, we developed an atomic-level approach based on MHP to allow a precise mapping of hydrophobic contacts, a contact being defined from a cut-off distance between two atoms with a positive MHP. While an arginine residue is classically considered as not hydrophobic, our method allows to highlight the involvement of aliphatic part of arginines in some hydrophobic contacts stabilizing the dystrophin coiled-coil repeat R1 (Figure 2A). Indeed, an interaction such as the one involving Arg395 of helix B with Val351 of helix A would have remained undetected using a residue-level approach to define hydrophobic contacts. Even if such contacts have been described promptly in spectrin coiled-coil structures<sup>15, 16</sup>, it remains difficult to accurately analyse them for a large set of structures. Overall, we detected 20% more hydrophobic contacts in dystrophin repeat R1 by comparison to an approach using a fixed definition of hydrophobic residues (Figure 2B and 2C).

From a heptad pattern point of view, atomic-level definition of hydrophobic contacts allows to characterize more precisely the interaction network stabilizing the dystrophin coiled-coil repeat.

Unlike what could be expected from a classical coiled-coil organisation, we observed that in the dystrophin repeat R1, many hydrophobic contacts involve other residues than those in *a* or

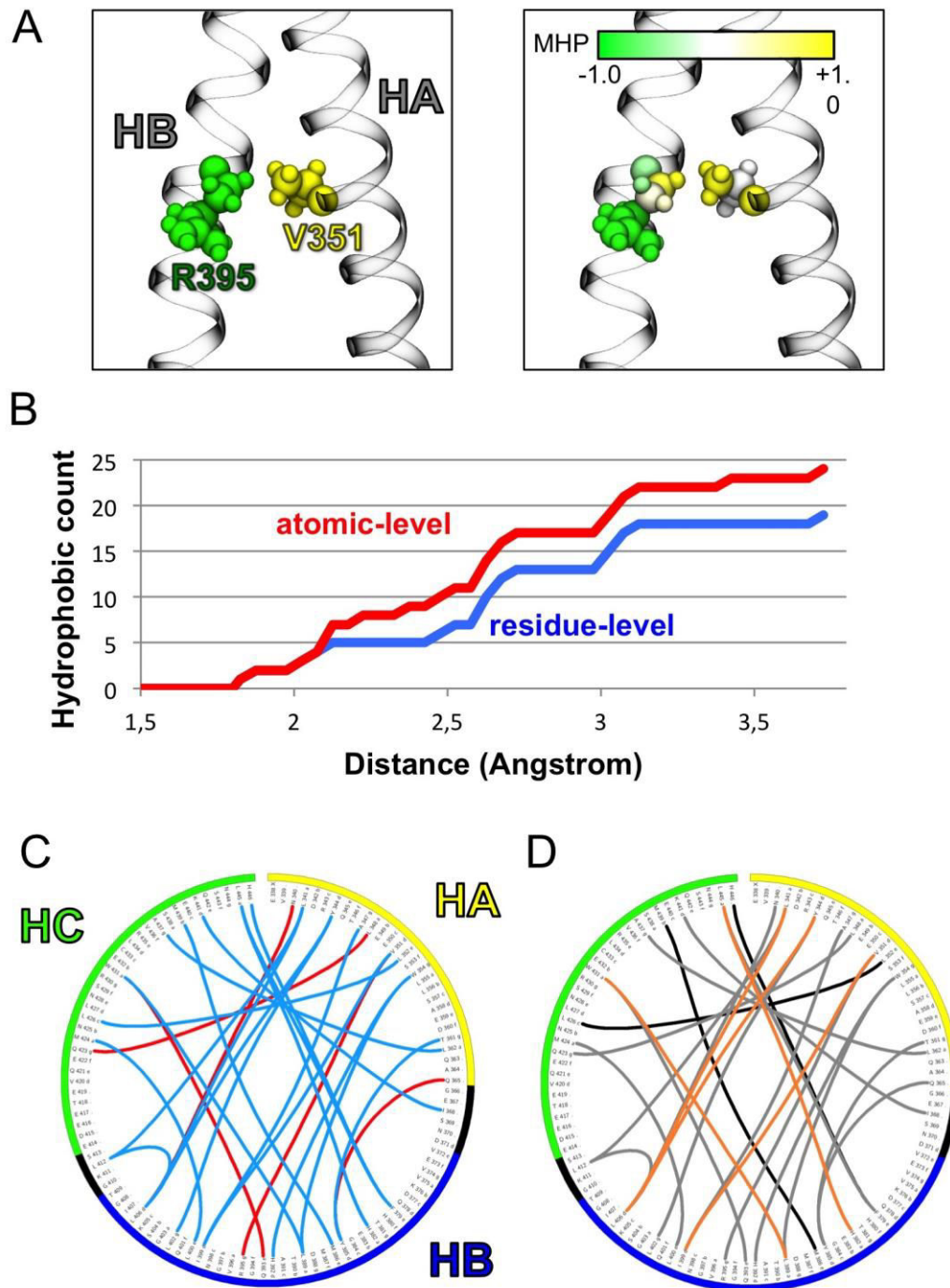


Fig. 2 : A. Inter-helices contact found in the dystrophin repeat 1 (PDB 3UUN) between Val351 and Arg395 and (on the left) following the classical residue-level qualitative nomenclature of amino acids properties (hydrophobic in yellow, hydrophilic in green) or (on the right) following the quantitative atomic-level Molecular Hydrophobicity Potential (MHP) ; B. Cumulative count of hydrophobic inter-helices interactions, in blue for residue-level definition and in red for atomic-level definition ; C and D. Circular projection of hydrophobic interactions found between helices forming the coiled-coil fold in dystrophin R1 according to (C) the residue-level definition (blue) or to the atomic-level definition (blue and red), or (D) the position of the residue in the heptad pattern : *a/d* to *a/d* contact in orange, *a/d* contact with another position in grey, and contact involving no *a/d* position in black.

*d* positions, forging a very dense interaction network between the three alpha-helices (Figure 2D). Moreover, only a minority of hydrophobic contacts are involving two residues located at *a* or *d* positions but are mainly involving a residue at *a* or *d* position with a residue at another position. The circular projection of atomic-level hydrophobic contacts highlights the asymmetry in the interaction of helices A or C with helix B, primarily due to the bending of its axis. While helix C over its entire length is contacting helix B through sparse but stronger interactions at its N-terminal end, helix A mainly interacts with the C-terminal end of helix B following a more regular pattern (Figure 2D and Figure SM1B). Only few interactions are achieved between residues at *a* or *d* positions in helices and hydrophobic residues located in AB or BC loops.

***Hydrophobic contacts between successive repeats are organized through the linker region.***

By investigating the inter-repeat contacts surrounding or involving the linker region of the dystrophin fragments constituting our structural data set, we have noticed that hydrophobic interactions are largely preponderant in comparison to the H-bonds or to the salt bridges (Figure SM2). In the light of the atomic-level hydrophobicity mapping, in all structure and models, the N-terminal end of helix B of repeat R1 is interacting systematically with the N-terminal end of helix A of repeat R2 (Figure 3A-C and Figure SM2).

Same behaviour is observed for helix B of repeat R2 and helix A of repeat R3 in the R1-3 SAXS-based model. Helix B is thus the main partner of the linker region, bridging a repeat to the following one by the formation of a short two-helices coiled-coil track stabilized through hydrophobic interactions. We also noticed that the various amplitude kinks observed in the filament axis between the repeats R1 and R2 do not strongly modify this feature.

Moreover, the projection of these hydrophobic interactions along the main filament axis of dystrophin fragment revealed that as inter-repeats interactions between repeats R2 and R3 interweave with inter-helices hydrophobic contacts stabilizing the R3 coiled-coil, while interactions between repeats R1 and R2 do not (Figure SM3). This can be observed either in the R1-2 fragment or in the R1-3 fragment, even if they present some differences in their kink between R1 and R2 axis. The overlap of inter- and intra-repeat hydrophobic interactions may lead to a stiffer structure at the linker region between the repeats achieved by sharing a track of helix A between the short two-helices and the dystrophin three-helices coiled-coils. On the other hand, a disruption in this bridging interaction network may be associated with an increase in the linker bending. As we finally observed only few contacts of any type (hydrophobic, H-bonds and salt-bridges) of the AB and BC loops with the linker regions (Figure 2, SM2 and SM3), we aimed to characterize the short two-helices coiled-coil defining the linker region by

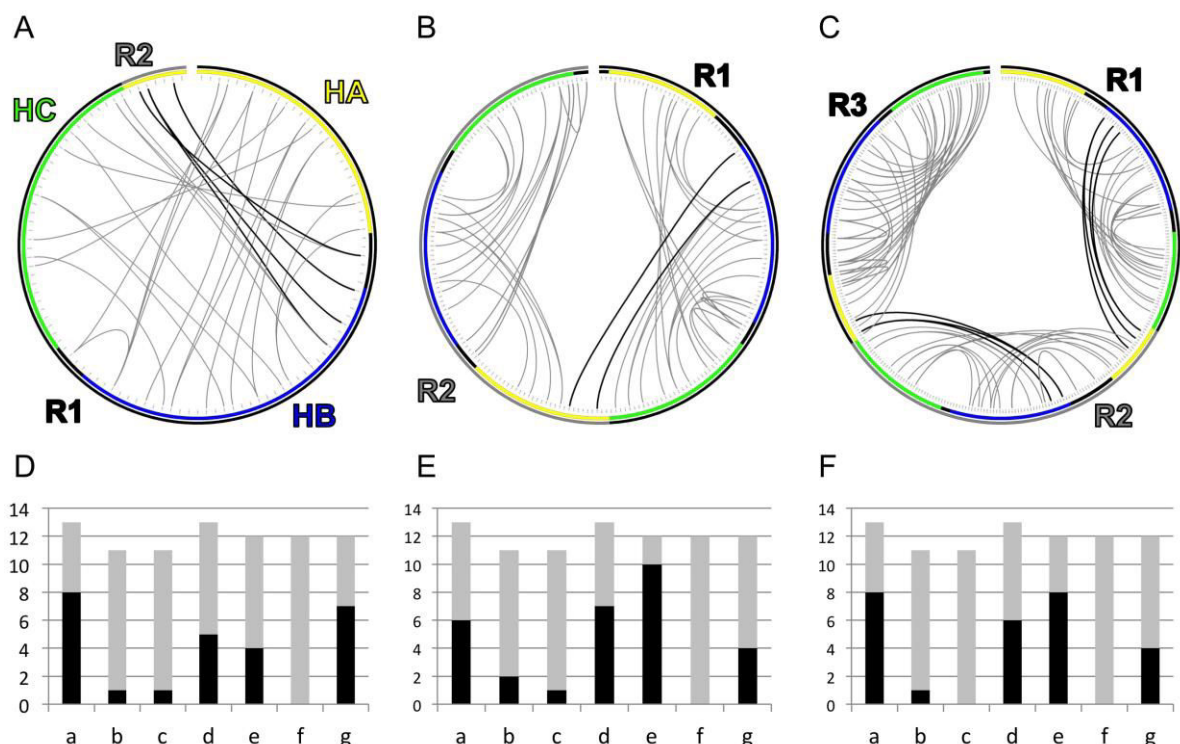


Fig. 3 : A, B and C. Circular projection of hydrophobic interactions along sequence of dystrophin fragments including the R1 repeat: dystrophin R1 crystallographic structure, holding the N-terminal start of R2 repeat (A) and dystrophin R1-2 (B) and R1-3 (C) structural models refined under SAXS restraints (Delalande et al., 2018b; Molza et al., 2014). Black lines indicate inter-helices hydrophobic interactions involving residues of two different repeats. (D to F) Heptad pattern positions involved in helix-to-helix hydrophobic contact (black) over the total number of all these positions available in repeat R1 according to the helices definition (grey) in the dystrophin crystallographic structure (D) and in the R1-2 (E) and R1-3 (F) SAXS-based models.



measuring the frequencies of the heptad positions involved in the inter-helices interactions (Figure 3D to 3F). It appears that these interactions involve alternative *e* and *g* positions as often as canonical *a* and *d* positions. At the three-helices coiled-coil level, this structural feature could allow some flexibility of the filament towards longitudinal stretching forces. In the case of the short two-helices coiled-coil track, and by considering the position *a* starting the/an *abc* stammer before the first heptad of helix B, only 50% of hydrophobic interactions are involving at least one canonical position in this region. Moreover, 67% of these interactions in the linker region are involving one or two alternative *e* and *g* positions (Figure 3A to 3C). Indeed, the frequency of MHP-positive residues at these alternative positions at the N-terminal end of helix B, at the C-terminal end of helix C and at the N-terminal end of helix A (of the following repeat) could be associated with linker structure and dynamics.

***Hydrophobic network achieved by a specific helix across different coiled-coil repeats reveals the diversity of coiled-coil in dystrophin central domain.*** Analysis related to heptad positions usage in hydrophobic contacts has been achieved for a given helix across the different coiled-coil repeats (Figure 4). By merging the interactions provided by the helix A of the repeat R1 of all dystrophin structure and SAXS-based models (Figure 4A), we noticed a regular heptad pattern usage, with some rescue from alternative *e* and *g* positions. However, for the helix A of repeats R2 and R3, we observed a track of hydrophobic interactions at its N-terminal end with the involvement of contiguous residues including rarely featured residues (*b*, *c* and *f* positions). This hydrophobic interaction track is stopped at a highly involved *a* position corresponding to the centre of helix A. Indeed, this seems to confirm the linker region to be principally stabilized through contacts between the helix A of one coiled-coil repeat with the helix B from the previous coiled-coil repeat. In the case of helix B, in any structure or models and by comparison to helices A and C, we could notice a more regular pattern across the different repeats (Figure 4B), resulting in higher frequencies at canonical (*a* and *d*) or alternative (*e* and *g*) positions.

Nevertheless, the N-terminal end of helix B of all repeats shows a slight deficiency in the high-frequency interacting residues, even if all of those that are involved in the two-helices coiled-coil of the linker region are conserved.

This relative lack of interaction is stopped with a high-frequency peak of interaction made by phenylalanine at a non-canonical *e* position in repeat R1. A similar high-frequency interaction peak is observed at the C-terminal end of the helix B of repeat R1 ; both seem to define a structural lock for the coiled-coil fold. Finally, by considering helices C of all repeats of our

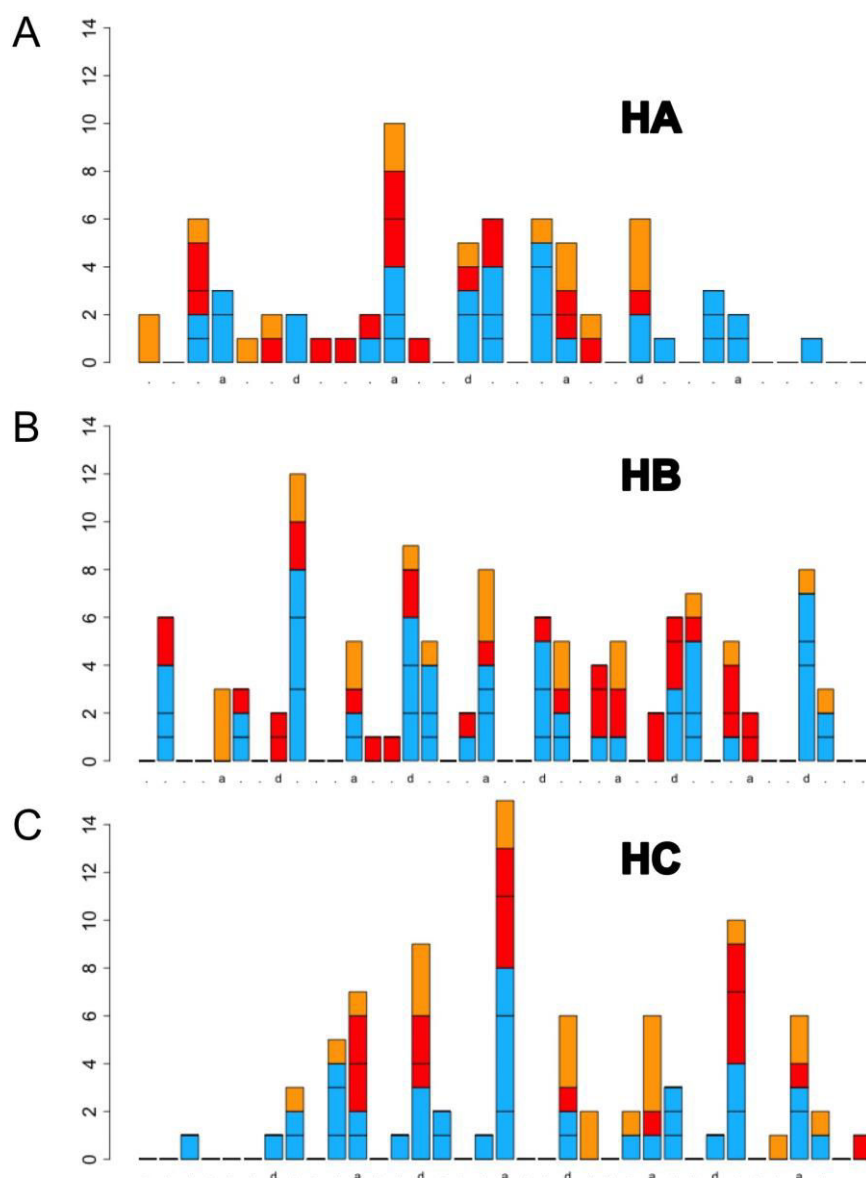


Fig. 4 : A, B and C. Cumulative count for amino acid involvement in an hydrophobic contact according to its position in the sequence of helices HA, HB and HC, respectively. Bar color indicate the dystrophin repeat (R1 in blue, R2 in red and R3 in orange) in which the inter-helix hydrophobic contacts were counted and their size is proportional to the number of atoms involved in the contact.

structural set, their N-terminal end showed a poor involvement in hydrophobic interaction network (Figure 4C). However, as observed for helix A, an essential *a* position emerged at the centre of the helix C (a highly conserved tryptophan residue), in the close vicinity of the bend of helix B.

## DISCUSSION

Since the dystrophin central domain has been initially proposed to be a rod-shaped structure<sup>36</sup>, this prediction had been mitigated by a finer analysis of its sequence, showing that the coiled-coil heptad pattern is less conserved in dystrophin than in spectrin<sup>11</sup>. Moreover, unlike many coiled-coil proteins forming multimeric structures among them spectrin<sup>6</sup>, dystrophin is known to be monomeric and this specificity appears to be partly related to this loosely conserved heptad pattern. Finally, while the first structure prediction computed for dystrophin and based on spectrin patterns revealed only a majority of straight linker regions<sup>19</sup>, it was later observed experimentally that the dystrophin central domain was presenting many kinks at linker regions<sup>27</sup>. Our study of hydrophobic contacts through the first coiled-coil repeats of dystrophin, confirmed its singularity towards both mathematical model of triple-helix coiled-coil<sup>7</sup> and spectrin regular pattern, confirming previous observations<sup>11,12</sup>. Indeed, the definition of the dystrophin coiled-coil repeat seems related to a dense, asymmetric and highly adaptable network of inter-helices hydrophobic contacts, a network also strengthened by local bending of helices B.

By firstly focusing on the hydrophobic interaction network achieved between the helices of repeat R1 in different conformational states, we highlighted the frequent use of hydrophobic residues at non-canonical *e* and *g* heptad positions. Moreover, an analysis of hydrophobic network achieved by any given helix across different dystrophin coiled-coil repeats revealed that these alternative residues are also often involved through a tandem usage with their

canonical *a* and *d* neighbours. We propose that the possible multiple arrangements of inter-helices contacts could be achievable upon longitudinal stretching of the dystrophin filament. This should confer an increased plasticity to the dystrophin three-helices coiled-coil repeat in comparison to the spectrin one, even if monomeric spectrin coiled-coil repeats were already described as more dynamic than previously thought<sup>15,25</sup>. From our analysis, AB and BC loops do not seem to play a crucial role neither in the coiled-coil stabilization or in the linker organisation in contradiction to what was observed in crystal structures of spectrin tandem repeats<sup>15</sup>. Linker region of dystrophin first repeats appeared to be mainly defined through the short two-helices coiled-coil made by helix B of a repeat and helix A of the following repeat.

In the set of structural data we analysed, it astonishingly appeared that the hydrophobic interaction network made by the helix B of repeat R1 does not seem impaired by the kink changes at the linker region between repeats R1 and R2. In addition, we were able to identify a distribution in the hydrophobic interaction network that can be mostly associated to a kinked (R1-R2) or a straight (R2-R3) linker region. At this stage, it is still unclear how we could predict the presence or the absence of a kink at the linker region between two dystrophin repeats from the simple primary sequence. However, we here propose a simple analysis of a unique tri-dimensional structure of a coiled-coil tandem repeat allowing the prediction of the flexibility of a linker region from the distribution and the density of its hydrophobic interactions network. The dystrophin elastic and flexible global properties would thus not only derive from the nature of a homogenous filament composed of nested coiled-coil repeats<sup>11</sup> but also mainly from the behaviour of independent sub-domains distributed all along a non-monotonous dystrophin central domain<sup>27</sup>, boundaries of these sub-domains being defined by the structural organisation of the linker regions.

From an exhaustive biochemical characterization of the dystrophin central domain, Mirza *et al.* did already show that the R2 and R3 repeats of dystrophin were defining an integrated site<sup>37</sup>.

This was structurally confirmed in our study of dystrophin fragments by SAXS, by the observation of various large kinks between R1 and R2 but not between R2 and R3<sup>27</sup>. Analysis of the hydrophobic networks involving the R1-R2 and R2-R3 linkers brings some clues about the involvement of specific residues in these differentiated organisations. Indeed, the interlacing of inter-repeat and intra-repeat hydrophobic networks, and the more regular use of hydrophobic residues at canonical heptad positions in repeats R2 and R3, could be linked to the better stability of R2-3 tandem repeat towards R1-2 tandem repeat<sup>37</sup>. All these observations indicate that the tandem repeat R2-R3 is relatively independent from repeat R1 structural and dynamic behaviour. We demonstrated earlier that the R1-3 dystrophin fragment was interacting *in vitro* with phospholipids<sup>38,39</sup> and this was in part confirmed *in vivo*, this fragment being associated with the membrane of muscle cells<sup>40</sup> and highlighting its crucial biological role. More recently, we showed from neutron scattering experiments that, in the R1-3 fragment, dystrophin repeat R1 unfolds its helices A and B to achieve interaction with bicelles (Dos Santos-Morais, submitted). The analysis of the hydrophobic network achieved by helix B of repeat R1 indicates the non-canonical involvement of a phenylalanine residue (Phe379) at a *e* position. This last interaction feature could be interpreted as a structural regulator of the accessibility of the putative lipid-packing sensor motif predicted in dystrophin<sup>41</sup>.

Clinical analysis indicates that some amino acids that play an essential role in the interaction network, as shown in the current study, are related to muscular dystrophies. For instance, a single mutation of amino acids involved in multiple hydrophobic contacts, as Leu427 or Leu575, have been observed in Becker patients<sup>42,43</sup>. Moreover, a double mutation (Met450Ile, Asp451Tyr) affecting the linker between repeats R1 and R2 has been associated to an Intermediate Muscular Dystrophy<sup>44</sup>. These mutations should increase the hydrophobic patch in this region and thus, rigidify the kink we characterized. Finally, the description of dystrophin first coiled-coil repeats we reported here, in their detailed structure and plasticity, could be

crucial for the understanding of the R1-3 importance in the design of functional mini-dystrophin used for gene therapy strategies<sup>28-30</sup>.

## MATERIAL AND METHODS

### *Coiled-coil repeats of dystrophin.*

The structure and SAXS-based models used for this study are listed in the Table S1. All fragments do contain the first coiled-coil repeat of dystrophin in different structural contexts.

*X-ray crystallographic structure of dystrophin repeat R1.* The first coiled-coil repeat R1 is the only experimentally defined structure of the dystrophin central domain (residues 338 to 456; PDB\_ID 3UUN<sup>18</sup>). The structure was defined for a forced homodimer through addition of a disulfide bridge between cysteine residues C433 located in the helix C of each R1 repeat. In order to standardize our data set and to lower the structural effects due to crystal packing and/or forced dimerization, we performed a geometry optimization of R1 monomer using a 10k steps steepest descent minimization using NAMD<sup>45</sup>, as it was done for SAXS-based models.

*Canonical model of dystrophin repeat R1.* To evaluate the specificity of a dystrophin coiled-coil repeat towards a theoretical three-helices coiled-coil, we generated a canonical model of dystrophin repeat R1 by submitting its primary sequence to the Coiled-Coil Builder (CCB) web-server<sup>46</sup> ([http://coiledcoils.chm.bris.ac.uk/app/cc\\_builder/](http://coiledcoils.chm.bris.ac.uk/app/cc_builder/)). The helices boundaries were defined following the repeat R1 structure: <sup>338</sup>EVN...QAQ<sup>365</sup> for helix A, <sup>372</sup>VEV...IGT<sup>409</sup> for helix B and <sup>416</sup>EDE...NLH<sup>446</sup> for helix C. In the final theoretical model provided by Coiled-Coil Builder (CCB model), helix structures were completed with loop AB and loop BC from the experimental structure before energy minimization (same as previously described).

*R1-2 and R1-3 SAXS-derived models of dystrophin fragments including repeat R1.* SAXS-derived models of the dystrophin fragments R1-2 and R1-3 were optimized following the previously described procedure based on interactive flexible fitting simulations driven by low-resolution molecular shapes<sup>27,47</sup>. The theoretical SAXS curves computed from these atomic molecular models were all fitting to the experimental SAXS data at a high level of confidence. Final SAXS-based models have been deposited in the SASBDB database (<http://www.sasbdb.org/>) after validation by structural biology standard (SAVES v5.0 <http://servicesn.mbi.ucla.edu/SAVES/>).

***Software for the mapping of structure and interaction features.***

*Py-interaction (Pynte) software.* Whereas H-bonds and salt-bridges are often identified on an atomic basis, hydrophobic interactions are more commonly detected assuming qualitative amino acid properties. We here developed an approach aiming to build atomic mapping of interactions via user's defined parameters (cut-off distances, angles) and to plot basic graphical representation of these interactions from a PDB-derived file generated by the PLATINUM program<sup>32,33</sup>. The structural information, like boundaries of helices of dystrophin were derived from DSSP program<sup>48,49</sup> measurements. Note that as we defined all helices and loops as independent subdomains, the contact measurements were simply achieved between user-defined subdomains. The Pynte approach could thus be extended for any quantitative measurement and graphical representation of the interactions measured between protein subdomains or between different biomolecules in an assembly. The Pynte software uses python scripts as basic input, R<sup>50</sup>, Circos configuration files<sup>51</sup> and VMD<sup>52</sup> for graphical representation output. The source code of Pynte is available on request from authors.

*Hydrophobic interactions.* In this study, atomic hydrophobic interactions have been defined between two atoms distant less than 4.2 Å and that present a positive Molecular Hydrophobic

Potential (MHP) computed by the PLATINUM program<sup>32,33</sup> using the parameters developed by Ghose et al.<sup>31</sup>. The cut-off distance could be chosen from a simple rule arguing that two hydrophobic atoms ( $r \sim 1.0 \text{ \AA}$ ) are close enough to interact if they are not separated by a water molecule ( $d_{\text{H}_2\text{O}} \sim 2.8 \text{ \AA}$ ). Nevertheless, we quantitatively checked that, in our structure set, the distribution of inter-atomic distances measured between two hydrophobic atoms present an effective minimum around  $4.2 \text{ \AA}$  (Figure SM1A), meaning that distances over this value could probably not be related to a direct atomic contact. As amino acids residues with only few atoms at the limit of the cut-off distance could lead to the detection of unrelevant interactions, we selected inter-residue contacts involving at least 6 inter-atomic interactions. In this study, a hydrophobic contact was considered strong in the case of 25 atomic cross-interactions or more measured between two different residues.

*H-bonds and salt bridges.* To complete our automated structure analysis, Hydrogen bonds have been defined between two heavy atoms including a hydrogen bond donor (D, bearing a hydrogen atom H) and an acceptor (A). Following standard parameters<sup>52</sup>, the H-bond is effective if the distance  $\|D-A\|$  is lower than  $4.2 \text{ \AA}$  and the angle  $\widehat{DHA}$  lower than  $40^\circ$ . In the complement, we defined salt bridges between basic and acidic residues forming H-bonds.



**SUPPLEMENTARY MATERIAL**

PDF file including supplemental Table ST 1 and three Figures SM 1 to SM 3.

**ACKNOWLEDGMENTS**

This research was supported by research grants from the AFM–Téléthon (Application #018211A) and the french GENCI computing program (DYSIM project). DML was recipient of a PhD fellowship from AFM-Téléthon (grant 19409). We also wanted to acknowledge Tim Pyrkov for providing a stand-alone version of Platinum. The authors are also grateful to Giulio Gambarota for his careful reading of the manuscript.

## REFERENCES

- (1) Barbara KE, Willis KA, Haley TM, Deminoff SJ, Santangelo GM (2007) Coiled coil structures and transcription: an analysis of the *S. cerevisiae* coilome. *Mol Genet Genomics* 278:135–147.
- (2) Lou X, Shin YK (2016) SNARE zippering. *Biosci Rep* 36:e00327.
- (3) Munro S (2011) The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harb Perspect Biol* 3:a005256.
- (4) Armony G, Jacob E, Moran T, Levin Y, Mehlman T, Levy Y, Fass D (2016) Cross-linking reveals laminin coiled-coil architecture. *Proc Natl Acad Sci U.S.A.* 113:13384–13389.
- (5) Severin K, Lee DH, Kennan AJ, Ghadiri MR (1997) A synthetic peptide ligase. *Nature* 389:706–709.
- (6) Lupas AN, Bassler J (2017) Coiled coils – a model system for the 21st Century. *Trends Biochem Sci* 42:130–140.
- (7) Crick FHC (1953) The Fourier transform of a coiled-coil. *Acta Crystallogr* 6:685–689.
- (8) Crick FHC (1953) The packing of  $\alpha$ -helices: simple coiled-coils. *Acta Crystallogr* 6:689–697.
- (9) Crick FHC (1952) Is  $\alpha$ -Keratin a Coiled Coil? *Nature* 170:882–883.
- (10) Parry DAD, Squire JM (2005) Fibrous proteins: new structural and functional aspects revealed. *Adv Protein Chem* 70:1–10.
- (11) Winder SJ, Gibson TJ, Kendrick-Jones J (1995) Dystrophin and utrophin: the missing links! *FEBS Lett* 369:27–33.
- (12) Nicolas A, Delalande O, Hubert JF, Le Rumeur E (2014) The spectrin family of proteins: A unique coiled-coil fold for various molecular surface properties. *J Struct Biol* 186:392–401.
- (13) Delalande O, Czogalla A, Hubert JF, Sikorski A, Le Rumeur E, Dystrophin and spectrin, two highly dissimilar sisters of the same family. In: Parry DAD & Squire JM, Ed. (2017) *Fibrous Proteins: Structures and Mechanisms, Sub-cellular Biochemistry*, Springer, pp 373–403.
- (14) Parry DA, Dixon TW, Cohen C (1992) Analysis of the three- $\alpha$ -helix motif in the spectrin superfamily of proteins. *Biophys J* 61:858–867.
- (15) Grum VL, Li D, MacDonald RI, Mondragón A (1999) Structures of two repeats of spectrin suggest models of flexibility. *Cell* 98:523–535.
- (16) Yan Y, Winograd E, Viel A, Cronin T, Harrison SC, Branton D (1993) Crystal structure of the repetitive segments of spectrin. *Science* 262:2027–2030.
- (17) Yläñne J, Scheffzek K, Young P, Saraste M (2001) Crystal structure of the  $\alpha$ -actinin rod reveals an extensive torsional twist. *Structure* 9:597–604.
- (18) Muthu M, Richardson KA, Sutherland-Smith AJ (2012) The crystal structures of dystrophin and utrophin spectrin repeats: implications for domain boundaries. *PLoS ONE* 7:e40066.
- (19) Legrand B, Giudice E, Nicolas A, Delalande O, Le Rumeur E (2011) Computational study of the human dystrophin repeats: interaction properties and molecular dynamics. *PLoS ONE* 6:e23819.
- (20) Bloch RJ, Pumplin DW (1992) A model of spectrin as a concertina in the erythrocyte membrane skeleton. *Trends Cell Biol* 2:186–189.
- (21) Speicher DW, Marchesi VT (1984) Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature* 311:177–180.
- (22) Lux SE, Palek J, Disorders of the red cell membrane, in Handin RI, Lux SE, Stossel TP, Ed. (1995) *Blood: Principles and Practice of Hematology*, JB Lippincott Co., Philadelphia, pp 1701–1818

- (23) Nicolas G, Pedroni S, Fournier C, Gautero H, Craescu C, Dhemy D, Lecomte MC (1998) Spectrin self-association site: characterization and study of beta-spectrin mutations associated with hereditary elliptocytosis. *Biochem J* 332:81–89.
- (24) Calvert R, Kahana E, Gratzner WB (1996) Stability of the dystrophin rod domain fold: evidence for nested repeating units. *Biophys. J.* 71:1605–1610.
- (25) Mirijanian DT, Voth GA (2008) Unique elastic properties of the spectrin tetramer as revealed by multiscale coarse-grained modeling. *Proc Natl Acad Sci U.S.A.* 105:1204–1208.
- (26) Mirijanian DT, Chu JW, Ayton GS, Voth GA (2007) Atomistic and coarse-grained analysis of double spectrin repeat units: the molecular origins of flexibility. *J Mol Biol* 365:523–534.
- (27) Delalande O, Molza AE, Dos Santos-Morais R, Chéron A, Pollet E, Raguènes-Nicol C, Tascon C, Giudice E, Guilbaud M, Nicolas A, Bondon A, Leturcq F, Férey N, Baaden M, Perez J, Roblin P, Piétri-Rouxel F, Hubert JF, Czjzek M, Le Rumeur E (2018) Dystrophin's central domain forms a complex filament that becomes disorganized by in-frame deletions. *J Biol Chem* 293:6637–6646
- (28) Harper SQ, Hauser MA, DelloRusso C, Duan D, Crawford RW, Phelps SF, Harper HA, Robinson AS, Engelhardt JF, Brooks SV, Chamberlain JS (2002) Modular flexibility of dystrophin: Implications for gene therapy of Duchenne muscular dystrophy. *Nat Med* 8:253–261.
- (29) Judge AD, Bola G, Lee ACH, MacLachlan I (2006) Design of non-inflammatory synthetic siRNA mediating potent gene silencing in vivo. *Mol Ther* 13:494–505.
- (30) Gregorevic P, Blankinship MJ, Allen JM, Chamberlain JS (2008) Systemic microdystrophin gene delivery improves skeletal muscle structure and function in old dystrophic mdx mice. *Mol Ther J Am Soc Gene Ther* 16:657–664.
- (31) Ghose AK, Viswanadhan VN, Wendoloski JJ (1998) Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: an analysis of ALOGP and CLOGP methods. *J Phys Chem A* 102:3762–3772.
- (32) Efremov R, Chugunov A, Pyrkov T, Priestle J, Arseniev A, Jacoby E (2007) Molecular lipophilicity in protein modeling and drug design. *Curr Med Chem* 14:393–415.
- (33) Pyrkov TV, Chugunov AO, Krylov NA, Nolde DE, Efremov RG (2009) PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organisation of biomolecular complexes. *Bioinformatics* 25:1201–1202.
- (34) Taylor WR (1986) The classification of amino acid conservation. *J Theor Biol* 119:205–218.
- (35) Tina KG, Bhadra R, Srinivasan N (2007) PIC: Protein Interactions Calculator. *Nucleic Acids Res* 35:W473–W476.
- (36) Koenig M, Monaco AP, Kunkel LM (1988) The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* 53:219–228.
- (37) Mirza A, Sagathevan M, Sahni N, Choi L, Menhart N (2010) A biophysical map of the dystrophin rod. *Biochim Biophys Acta BBA - Proteins Proteomics* 1804:1796–1809.
- (38) Legardinier S, Hubert JF, Le Bihan O, Tascon C, Rocher C, Raguènes-Nicol C, Bondon A, Hardy S, Le Rumeur E (2008) Sub-domains of the dystrophin rod domain display contrasting lipid-binding and stability properties. *Biochim Biophys Acta BBA - Proteins Proteomics* 1784:672–682.
- (39) Vié V, Legardinier S, Chieze L, Le Bihan O, Qin Y, Sarkis J, Hubert JF, Renault A, Desbat B, Le Rumeur E (2010) 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 BBA - Biomembr* 1798:1503–1511.
- (40) Zhao J, Kodippili K, Yue Y, Hakim CH, Wasala L, Pan X, Zhang K, Yang NN, Duan D, Lai Y (2016) Dystrophin contains multiple independent membrane-binding domains. *Hum Mol Genet* 25:3647–3653.

- (41) Drin G, Casella JF, Gautier R, Boehmer T, Schwartz TU, Antony B (2007) A general amphipathic  $\alpha$ -helical motif for sensing membrane curvature. *Nat Struct Mol Biol* 14:138–146.
- (42) Acsadi G, Moore SA, Chéron A, Delalande O, Bennett L, Kupsy W, El-Baba M, Le Rumeur E, Hubert JF (2012) Novel mutation in spectrin-like repeat 1 of dystrophin central domain causes protein misfolding and mild Becker Muscular Dystrophy. *J Biol Chem* 287:18153–18162.
- (43) Veerapandiyar A, Shashi V, Jiang YH, Gallentine WB, Schoch K, Smith EC (2010) Pseudometabolic presentation of dystrophinopathy due to a missense mutation. *Muscle Nerve* 42:975–979.
- (44) Juan-Mateu J, Gonzalez-Quereda L, Rodriguez MJ, Baena M, Verdura E, Nascimento A, Ortez C, Baiget M, Gallano P (2015) DMD Mutations in 576 dystrophinopathy families: a step forward in genotype-phenotype correlations. *PLoS ONE* 10:e0135189
- (45) Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.
- (46) Wood CW, Bruning M, Ibarra AA, Bartlett GJ, Thomson AR, Sessions RB, Brady RL, Woolfson DN (2014) CCBUILDER: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* 30:3029–3035.
- (47) Molza AE, Férey N, Czjzek M, Le Rumeur E, Hubert JF, Tek A, Laurent B, Baaden M, Delalande O (2014) Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly. *Faraday Discuss* 169:45–62.
- (48) Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.
- (49) Joosten RP, te Beek TAH, Krieger E, Hekkelman ML, Hooft RWW, Schneider R, Sander C, Vriend G (2011) A series of PDB related databases for everyday needs. *Nucleic Acids Res* 39:D411–D419.
- (50) R: A language and environment for statistical computing, ISBN 3-900051-07-0.
- (51) Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–45
- (52) Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38.

**SUPPLEMENTARY FIGURES AND TABLES**

Table S1: List of studied crystal structure and mathematical or SAXS-based models.

Fragment	Origin	# of repeats	# of residues	Position in dystrophin
R1 CCB	Coiled-coil builder software	1	109	338-446
3UUN	Crystallography	1+7 residues of R2	116	338-453
SAXS R1-2	SAXS	2	228	338-565
SAXS R1-3	SAXS	3	333	336-668

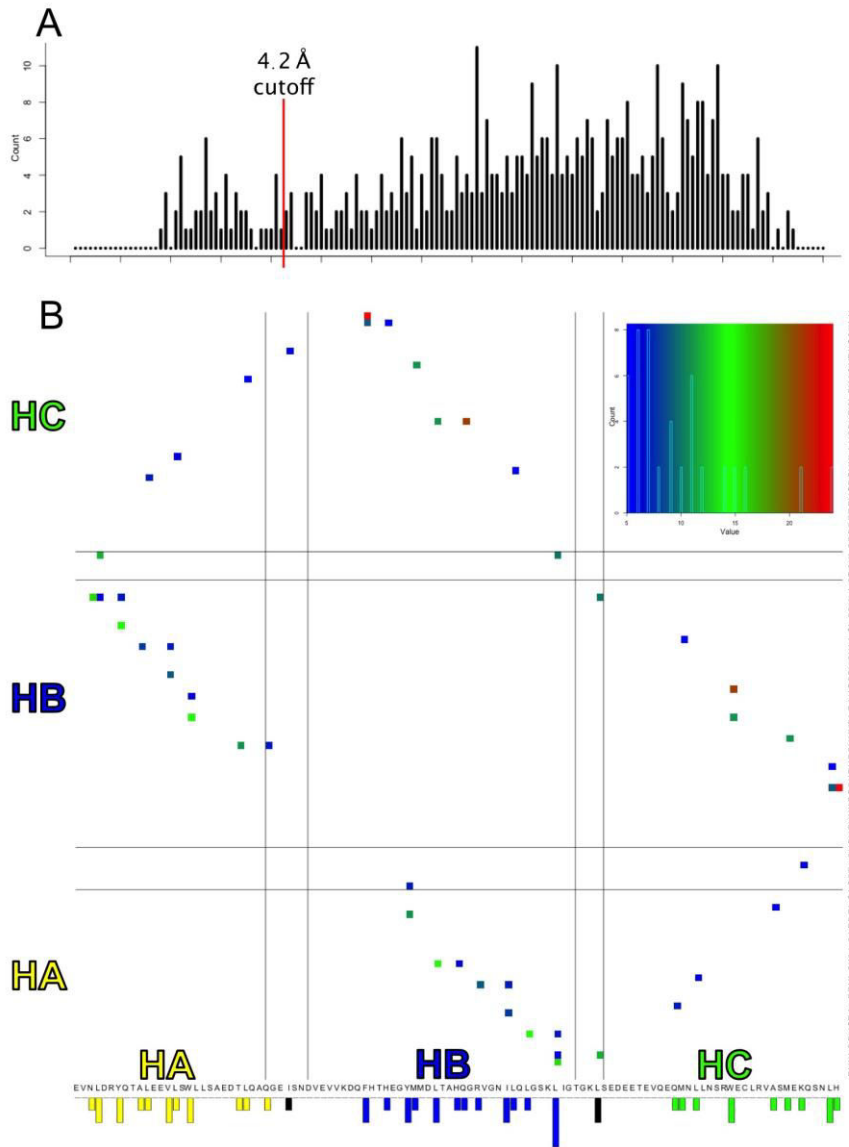


Figure SM1: (A) Distance distribution of inter-atomic distances in repeat R1 crystal structure (PDB ID: 3UUN) for the definition of hydrophobic interactions. (B) (top) Contact map of hydrophobic interactions measured in the Dystrophin repeat 1 structure (PDB ID: 3UUN), a point indicating an interaction between two residues (primary sequence on both axis) belonging to two different alpha-helices (HA, HB and HC). Color of the dot corresponds to the number of atoms involved in this interaction following top right scale. (bottom) Bar plot indicates the number of interactions involving each residue of the primary sequence (HA in yellow, HB in blue, HC in green and loops in black). It appears that all interactions are well distributed all along the sequence, but we can notice an asymmetry in the helices partnerships. At first, HA and HC are poorly interacting together as both mainly interact with HB. Then, HA interacts with the C-terminal end of HB through numerous interactions, while HC interacts with the N-terminal end of HB through fewer but stronger interactions (more atoms involved per residue). Finally, HA is achieving strong interactions with HB thanks to its N-terminal to middle part as HC is more strongly interacting with HB through its middle to C-terminal part (green to red dots).

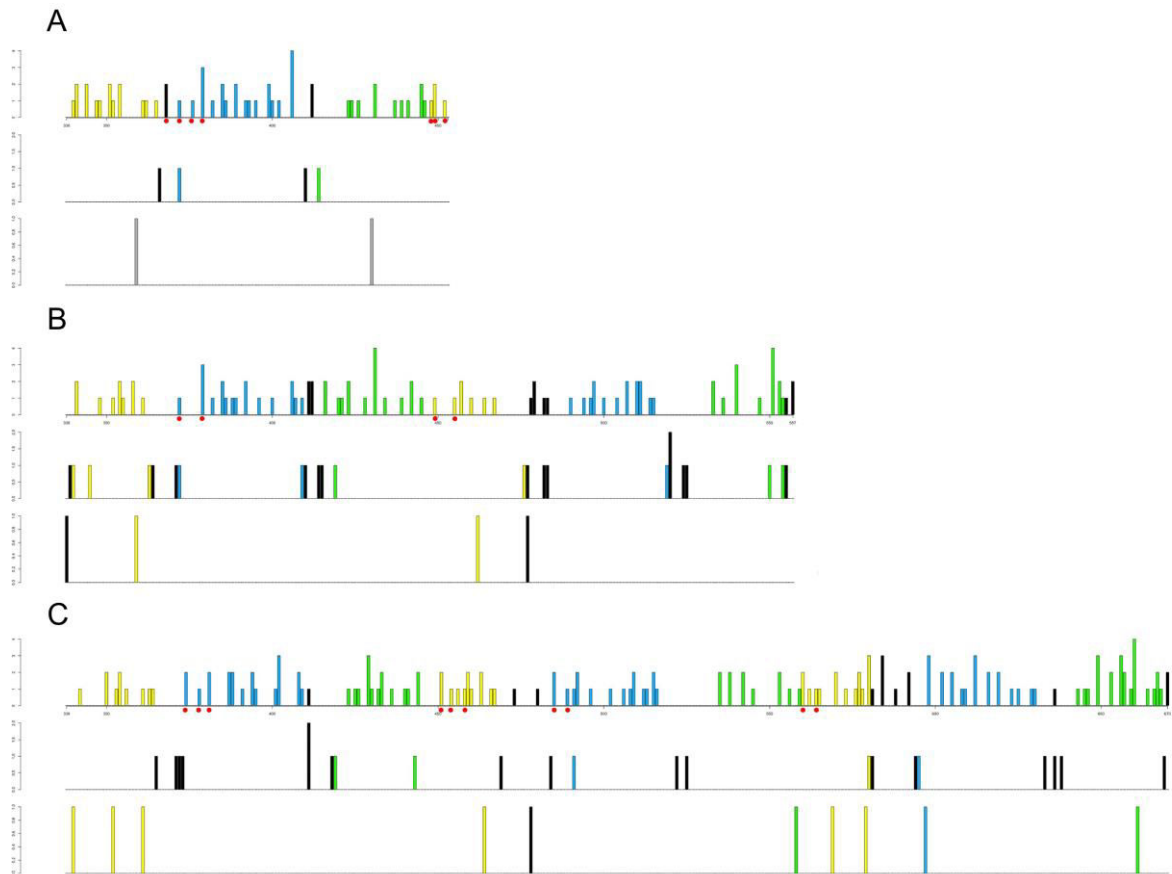


Figure SM 2: Barplot of hydrophobic interactions, hydrogen bond and salt bridge along sequence of structure or models (A: R1 crystallographic structure; B: R1-2 SAXS model; C: R1-3 SAXS model) and coloured following the secondary structure belonging: HA in yellow, HB in blue, HC in green and loops in black. For each fragment, the three plots are corresponding to hydrophobic interactions (top), hydrogen bond (middle) and salt bridge (bottom). Red dots indicate residues involved in inter-repeats interactions. No interactions other than hydrophobic are found between repeats.

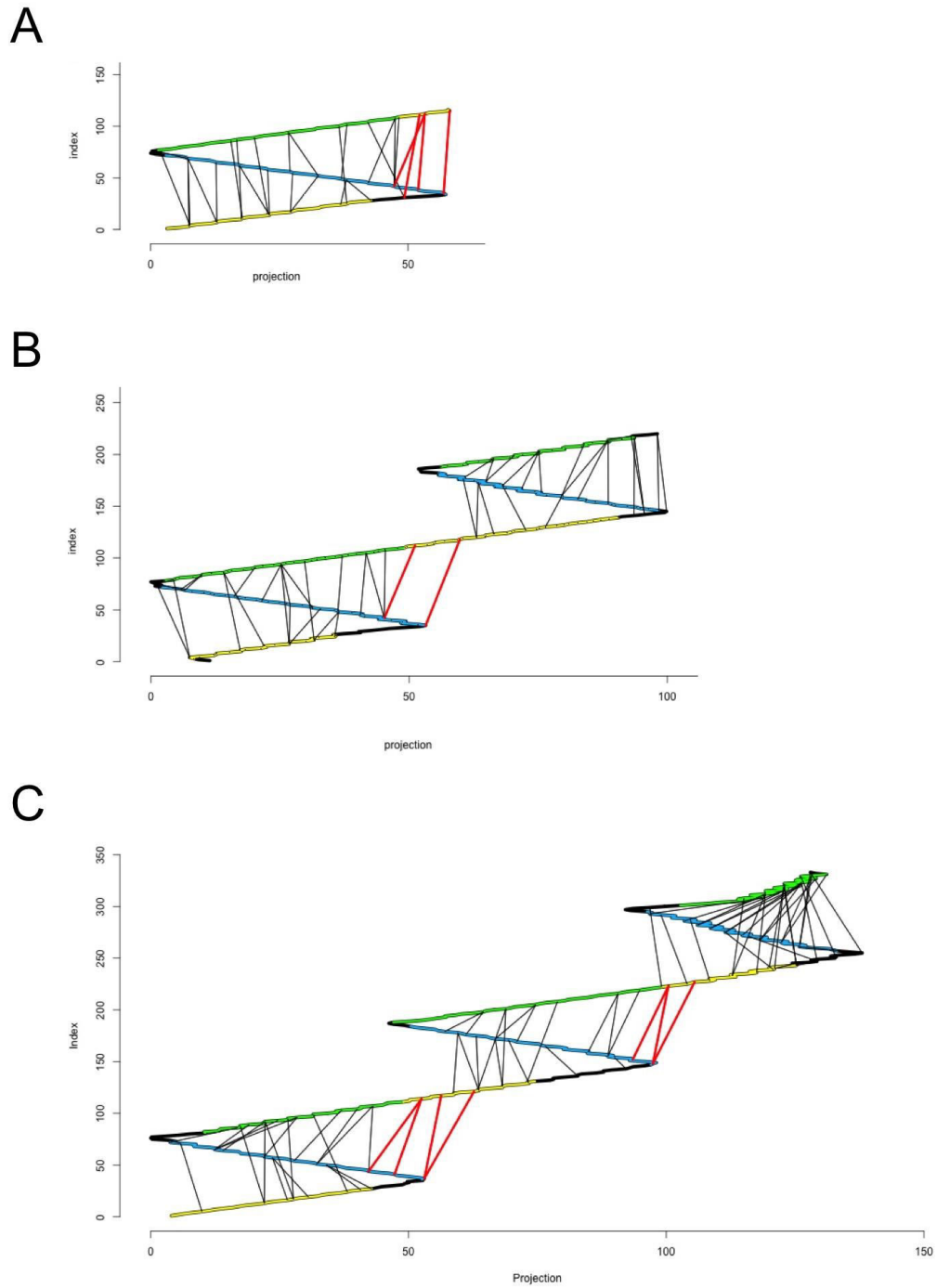
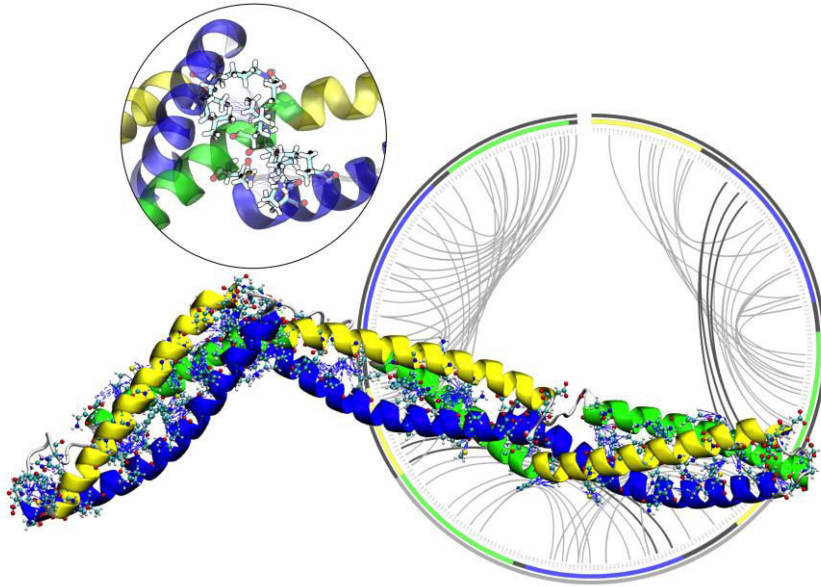


Figure SM 3: 2D projection of hydrophobic interactions along main axis and primary sequence of structure and models (A: R1 crystallographic structure; B: R1-2 SAXS model; C: R1-3 SAXS model). Red bold lines indicate interactions involving residues of different repeats, notably the one involved in the two-helices coiled-coil bridging successive coiled-coil repeats; yellow segments: HA; blue segments : HB; green segments : HC. It has to be noticed that the crystal structure of dystrophin repeat R1 contains the start of the following repeat R2, accounting for the presence of inter-repeat interactions in this structure.



## COVER ARTWORK



## C. Discussion

### 1. Compléments à l'analyse d'une répétition de dystrophine

Il était anticipé que la structure du domaine central de la dystrophine est principalement stabilisée par des interactions hydrophobes entre les hélices des répétitions repliées en faisceau d'hélices. Grâce à notre étude, nous montrons que dans la structure cristallographique de la répétition R1 (PDBID 3UUN) (Muthu et al., 2012) ou dans les répétitions R1 des modèles SAXS, ces interactions inter-hélices impliquent les résidus situés à des positions *e* ou *g*, à des fréquences parfois plus élevées que les positions canoniques *a* ou *d*. Nous proposons que ces variations observées dans la répétition R1 puissent être retrouvées dans d'autres répétitions du domaine central de la dystrophine. Ces particularités pourraient permettre d'adopter plusieurs arrangements d'interactions inter-hélices, d'alterner entre plusieurs réseaux d'interactions, conférant au domaine central de la dystrophine une certaine plasticité, et ainsi de résister aux contraintes mécaniques longitudinales.

D'autre part, des schémas d'interactions hydrophobes différents ont été observés au niveau des régions « *linker* » de R1-2 et R2-3, quand l'hélice B d'une répétition forme un court faisceau à deux hélices avec l'hélice A de la répétition suivante. Dans un premier cas, le coude observé entre les répétitions R1 et R2 est associé à un faible chevauchement des interactions entre le faisceau à deux hélices du *linker* et le faisceau à trois hélices de la répétition R2 (Figure 30A). Au contraire, dans le second cas, les résidus hydrophobes de l'extrémité N-terminale de l'hélice A de la répétition R3 alternent des interactions hydrophobes avec l'hélice B de la répétition R2 et l'hélice B de la répétition R3. Localement, il y a donc une unité entre le faisceau à deux hélices de la région *linker* R2-3 et le faisceau d'hélice à trois hélices de la répétition R3. On observe dans les différents modèles publiés contenant les répétitions R1 et R2 des variations de l'amplitude du coude du *linker* R1-2 avec des valeurs d'angle de 40° à 70° dans les modèles SAXS R1-2 ou R1-3, respectivement (Delalande et al., 2018). En parallèle, on note une continuité dans l'utilisation des résidus formant des interactions hydrophobes inter-hélices au niveau du *linker* R2-R3, contrairement au *linker* R1-2. L'absence de chevauchement des interactions hydrophobes entre R1 et R2, associée avec une plus grande disparité spatiale des résidus impliqués dans le réseau d'interaction pourrait être associée à une flexibilité accrue

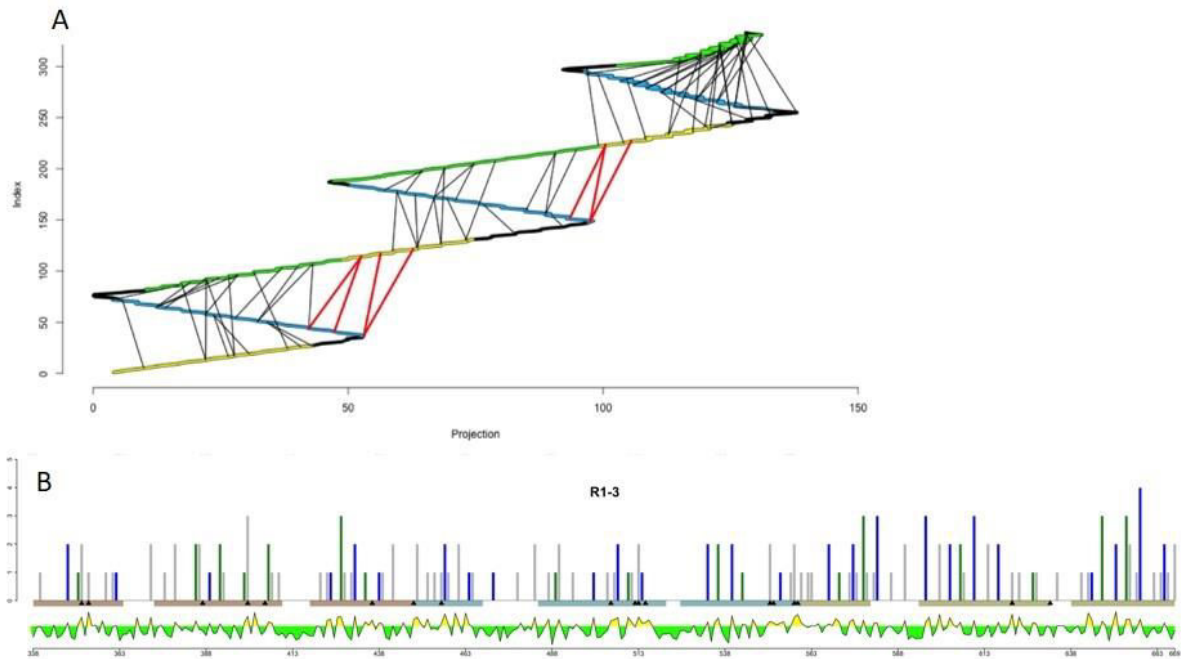


Figure 30 : Etude des interactions hydrophobes le long du fragment R1-3 ; A) Projection suivant x : l'axe principal du fragment R1-3 et y : la position des acides aminés dans la séquence; les traits épais représentent la séquence primaire de R1-3 : R1 (rouge), R2 (bleu) R3 (jaune) ; boucles (noires); Les traits fins représentent les interactions hydrophobes entre résidus (noir) d'une même répétition (rouge) de répétitions différentes. B) Haut : Graphique en barre des résidus impliqués dans des interactions hydrophobes inter-hélices ; la hauteur de la barre est proportionnelle au nombre de résidus en contact à une position donnée ; Les segments horizontaux marquent les hélices des répétition R1 (rouge), R2 (bleu) et R3 (jaune) ; Les résidus exposés au solvant et avec un MHP global positif sont marqués par une flèche noire ; Bas : Tracé du MHP des acides-aminés (somme des MHP atomiques) ; en vert sont marqués les résidus au MHP négatif (hydrophile), et en jaune, ceux au MHP positif (hydrophobe).

du linker R1-2, qui ne serait pas permise par le réseau d'interactions hydrophobes continu entre R2 et R3.

Parmi les autres disparités entre les répétitions du fragment R1-3, on remarque une accumulation d'interactions dans la répétition R3, en particulier entre les hélices B et C. Bien que R1 présente une densité d'interactions hydrophobes plus faible que R3, deux résidus de l'hélice B de R1 se distinguent comme étant fortement impliqués dans la stabilisation de l'interaction inter-hélices. Dans toutes les structures et modèles étudiés, une phénylalanine située sur une position *e* au début de l'hélice B de R1 (Phe379) est systématiquement retrouvée impliquée dans de nombreuses interactions hydrophobes avec l'extrémité C-terminale de l'hélice C de la répétition R1. L'autre résidu (Leu406), situé à l'extrémité C-terminale de l'hélice B de R1, a un comportement similaire avec l'extrémité N-terminale de l'hélice A de R1. Ces deux positions semblent former des verrous structuraux, maintenant le faisceau d'hélices dans un état stable, fermé. Un plein accès par des lipides au motif ALPS

prédit sur l'hélice B de la répétition R1 nécessiterait alors de rompre les interactions formées par ces résidus.

L'observation du potentiel hydrophobe moyen des résidus le long de la séquence du fragment R1-3 montre de nouvelles disparités entre les régions *linker* R1-2 et R2-3 (Figure 30B). Entre les répétitions R1 et R2, la région précédant la jonction est globalement hydrophile et la région suivant la jonction est globalement hydrophobe. Au contraire, la région *linker* R2-3 est bien plus homogène, en étant globalement hydrophobe.

Enfin, on observe que dans le fragment R1-3, les résidus globalement hydrophobes et exposés au solvant sont répartis tout le long du fragment, laissant apparaître des patches hydrophobes clairsemés sur la surface, ne formant pas de motifs topologiques réguliers ou identifiables sur la protéine. L'existence d'un sillon hydrophobe s'étendant sur toute la longueur du filament et propice à l'interaction avec des lipides n'est ainsi pas observable. De plus, les boucles sont globalement hydrophiles, limitant leur implication dans le maintien du faisceau d'hélice via des interactions hydrophobes.

## 2. Fragment R1-3 et interaction avec un modèle de bicelle

Pour étudier le comportement du fragment R1-3 en présence de lipides, et de manière à modéliser les données issues des expériences de SANS et de spectrométrie de masse, nous avons procédé à des simulations de dynamique moléculaire gros grains (voir chapitre // *Méthodes*) de modèles de fragments R1-3 avec des modèles de bicelles zwitterioniques ou anioniques (Figure 31A). Les simulations ont été effectuées en utilisant GROMACS 5 (Abraham et al., 2015) avec le champ de force MARTINI (Marrink et al., 2007). Nous allons tour à tour détailler ici le protocole de simulation qui n'est que succinctement décrit dans la publication scientifique associée (voir Annexe). Dans un second temps, nous apporterons une analyse complémentaire et non-publiée mettant en lien les deux articles présentés dans cette partie.

La cellule de simulation a été créée en deux étapes car les outils permettant de créer une cellule de simulation contenant des lipides ordonnés se destinent généralement à la construction de bicouches (Figure 31). Ainsi, la première étape consiste à modéliser une bicouche lipidique, avec des dimensions permettant d'insérer le nombre souhaité de lipides. Le ratio entre les différents lipides a été choisi pour être le plus proche possible de celui présent dans les bicelles obtenues *in vitro*. (Figure 31A). Cette bicouche a dans un second

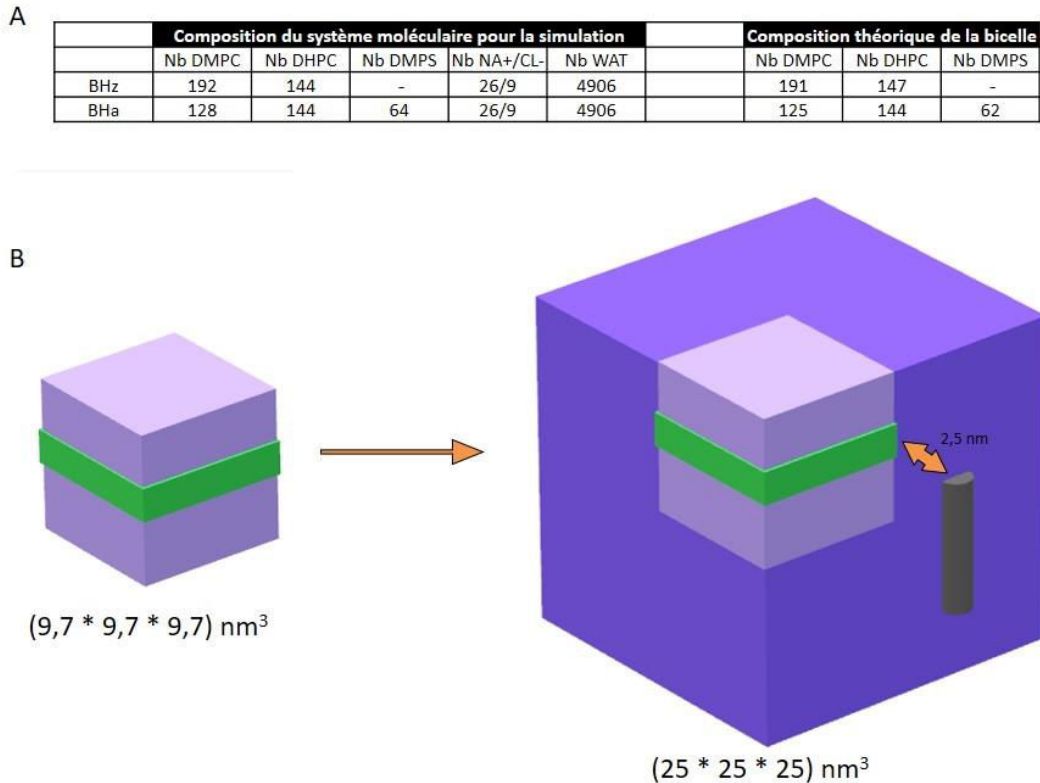


Figure 31 : Composition des systèmes de simulation de dynamique moléculaire; A) Gauche : Composition des différentes cellules de simulation en lipides, ions et eau ; Droite : Comparaison de la composition aux modèles théoriques de bicelles ; BHz : bicelle zwitterionique ; BHa : bicelle anionique; B) Principe de formation du système de simulation des bicelles ; Gauche : une cellule de simulation de bicouche lipidique est créée et complétée avec de l'eau ; les dimensions sont choisies de telle manière à ce que le nombre de lipides présents corresponde à celui attendu d'après les modèles théoriques (composition théorique de la bicelle en A). Droite : la cellule de simulation est agrandie, le fragment de dystrophine est ajouté à 2,5nm de la bicouche ; La cellule est ensuite remplie d'eau. Violet et bleu : eau ; Vert : lipides ; Gris : fragment R1-3 de dystrophine.

temps été incorporée à une cellule de simulation plus grande, contenant le fragment de dystrophine, dans laquelle il a été finalement ajouté de l'eau et des ions (Figure 31B). Les deux systèmes ont été simulés durant 20  $\mu$ s à 303 K.

Au début de la simulation, la bicelle est un pavé droit dans lequel les queues hydrophobes des lipides situés sur son bord sont exposées au solvant (Figure 32A et B). Toutefois, dans les premières 200 ns (sur les 20  $\mu$ s simulées au total), les lipides à chaînes courtes (DHPC) ont migré à la périphérie de la bicelle (ou tore) pour couvrir les chaînes hydrophobes initialement situées sur les côtés des bicelles (Figure 32B). En mesurant la fonction de distribution radiale (*rdf*, *radial distribution frequency*) des lipides (Figure 32A), on observe que cet état de distribution des lipides est par la suite stable sur le reste du temps de simulation. Les fragments protéiques sont recrutés après 320 ns par la bicelle anionique et après 400 ns par la bicelle zwitterionique (Figure 32C).

Dans les deux simulations, la protéine est spontanément recrutée sur la partie toroïdale des bicelles, là où des défauts de compacité peuvent être observés (Figure 32D). Cette prédiction de localisation s'est révélée être en parfait accord avec les résultats issus d'acquisitions de microscopie électronique d'échantillons contenant des bicelles et le fragment R1-3 de la dystrophine marqué par des nanoparticules d'or (à son extrémité N-terminal) (Figure 32D et E). D'autre part, la taille des bicelles obtenues en simulation (environ 12 nm de diamètre) est très proche de celle observée dans ces mêmes expériences de microscopie électronique, validant *a posteriori* nos calculs de compositions bicellaires basés sur des équations empiriques détaillées dans l'article de ce chapitre (Figure 32D et E).

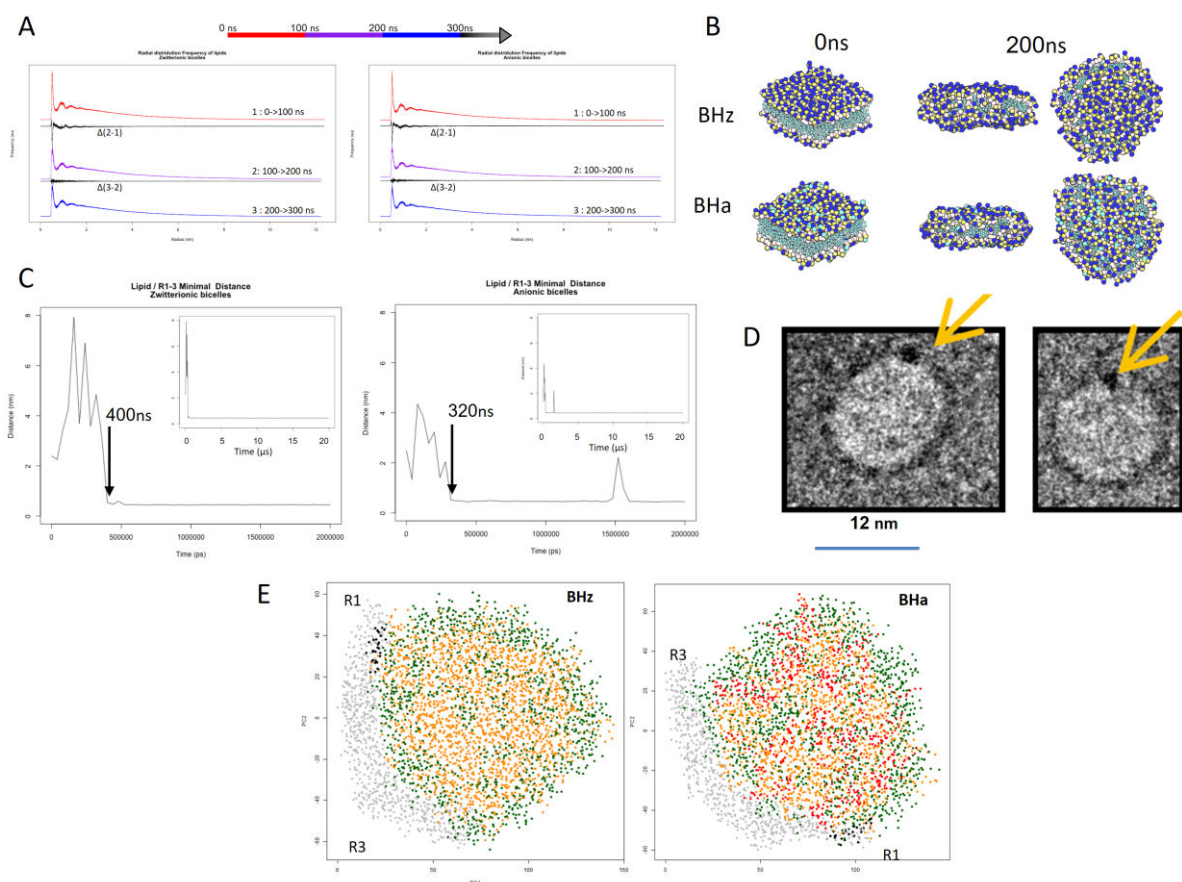


Figure 32 : Simulation de dynamique moléculaire gros-grains R1-3 / Bicelles zwitterioniques (DMPC / DHPC) et R1-3 / Bicelles anioniques (DMPC / DHPC / DMPS). **A**) Fonction de distribution radiale inter-grains durant les 300 premières nano-secondes; rouge : RDF de 0 ns à 100 ns ; violet : RDF de 100 ns à 200ns ; bleu : RDF de 200 ns à 300 ns ; les tracés en noir correspond à la différence des RDF les entourant ; gauche : simulation r1-3 / bicelles zwitterioniques, droite : simulation r1-3 / bicelles anioniques. **B**) Etat des bicelles à (gauche) 0 ns et à (centre et droite) 200 ns; haut : bicelles zwitterioniques, bas : bicelles anioniques. **C**) Distance minimale entre le fragment r1-3 et les lipides entre 0 $\mu$ s et 2 $\mu$ s ; insert : entre 0 $\mu$ s et 20 $\mu$ s ; la flèche noire indique le moment du rapport du premier contact entre les deux partenaires ; gauche : simulation r1-3 / bicelles zwitterioniques, droite : simulation r1-3 / bicelles anioniques. **D**) micrographies de cryo-microscopie électronique représentatif des complexes r1-3/bicelles anioniques. Les lysines de la protéine ont été marquées avec une nano-particule d'or (flèche orange) (1.4 nm mono-sulfo-nhs-nanogold<sup>®</sup>, nanoprobes). Extrait de la thèse de Raphael Dos Santos Morais. **E**) Projection selon les deux axes principaux définis par **acp** (analyse en composante principale) de l'état final (20  $\mu$ s) des simulations ; haut : simulation R1-3 / bicelles zwitterioniques, bas : simulation R1-3 / bicelles anioniques. Gris : projections des grains appartenant à la protéine ; noire : alps ; vert : DHPC (chaîne courte) ; orange : DMPS (chaîne longue) ; rouge : DMPS (chaîne longue).

Dans les expériences de SANS, on peut mesurer une augmentation nette (40%) de la longueur maximale ( $D_{\max}$ ) du fragment protéique R1-3 en présence d'une bicelle anionique (Figure 33A et B). Dans la simulation de dynamique moléculaire correspondant à ce système, nous avons observé des mouvements de grande amplitude (supérieurs à  $50^\circ$ ) de l'hélice A de la répétition R1 (Figure 33C), qui ne permettent cependant pas de proposer une explication quant au large changement de conformation identifié expérimentalement par SANS. Après un regroupement des conformations proches suivant la méthode de Jarvis, nous avons isolé deux modèles représentatifs de la trajectoire de dynamique moléculaire. Dans la première moitié de la simulation, l'hélice A de la répétition R1 s'écarte des autres hélices du faisceau (Figure 33C et D). Dans un second temps, cette hélice se rapproche de nouveau du faisceau initial, mais en glissant légèrement vers la répétition R2. Dans la publication présentée dans ce chapitre, nous avons décrit la possibilité d'un tel mouvement par un décalage d'une heptade permettant de conserver les interactions entre résidus situés en positions *a* et *d*. Toutefois, dans les modèles obtenus par simulations de dynamique moléculaire, le mouvement est deux fois plus important, créant un décalage final de deux heptades complètes (Figure 33C).

Les simulations de dynamique moléculaire ne durent que  $20 \mu\text{s}$ . Il est possible que celles-ci ne permettent pas d'observer des événements de plus grande amplitude dans un temps de calcul raisonnable en termes de délai et de coût. De plus, il semble également possible que ces deux états que nous avons isolés ne soient que des artefacts de simulation. Toutefois, pour produire cette simulation basse-résolution, nous avons utilisé le champ de force MARTINI, initialement développé pour simuler les lipides (Marrink et al., 2007) puis étendu à la simulation des systèmes protéines-lipides (Monticelli et al., 2008). Les interactions hydrophobes y sont donc très bien prises en compte, ce qui doit permettre une prédiction efficace quant à l'agencement des heptades d'un *coiled-coil* et proposer une description fiable du mouvement putatif de l'hélice A au sein du faisceau d'hélices. L'écartement de l'hélice A pourrait donc être un micro-événement pertinent de la simulation, mais qui n'explique pas les observations de SANS, tout ceci étant probablement dû à un décalage dans les échelles de temps de l'une et l'autre méthode pour la description d'un même système moléculaire.



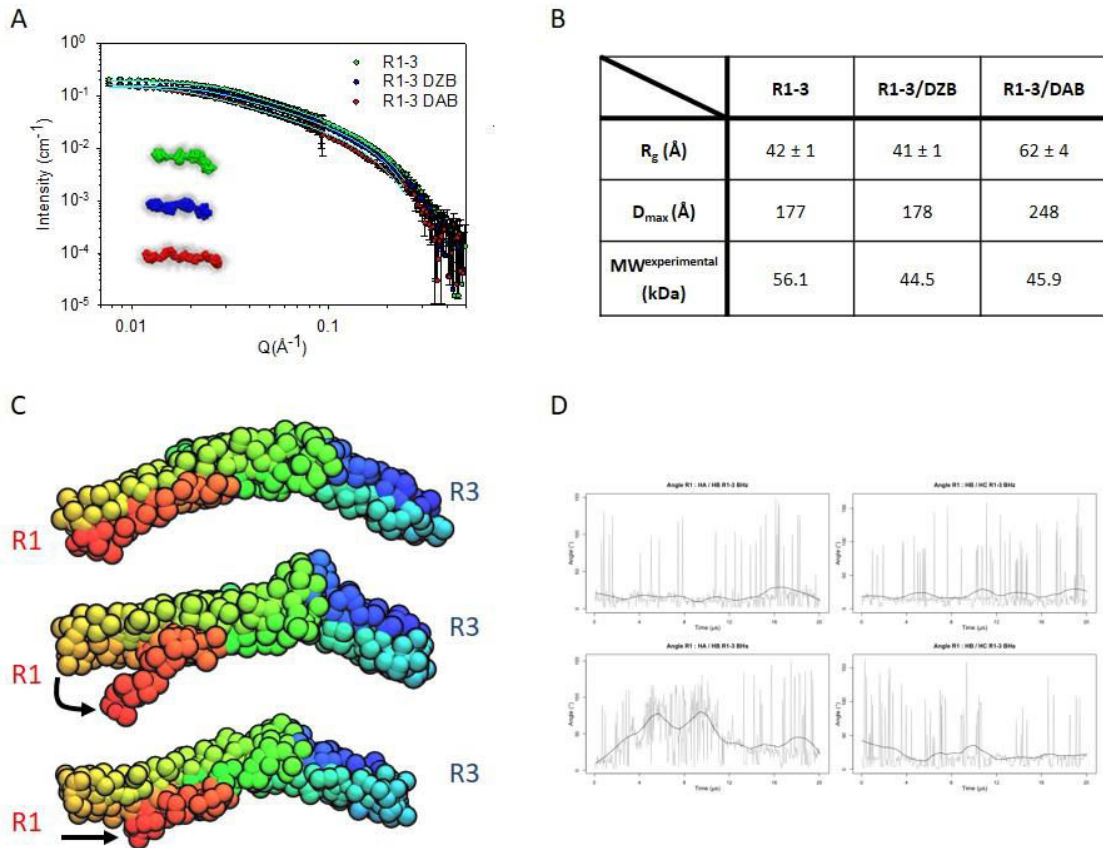


Figure 33 : Résultats des analyses SANS et confrontation aux modèles de simulations gros-grain. A) Courbes de diffusion SANS ; vert : R1-3 en solution ; bleu : R1-3 + bicelles zwitterioniques deutérés ; rouge : R1-3 + bicelles anioniques deutérés ; Insert : Modèles GSBOR correspondants. B) Paramètres des modèles déduit de l'analyse des données SANS ;  $R_g$  ;  $D_{max}$  ;  $MW$  : Poids moléculaire ; La diminution du poids moléculaire est un artefact causé par la présence de deutérium dans l'échantillon ; C) Modèles de R1-3 extrait de la simulation gros-grain en présence d'une bicelle anionique ; Haut : état initial ; Milieu : état à  $5 \mu s$  ; Bas : état à  $20 \mu s$ . D) Mesure de l'angle entre (gauche) l'hélice A et B de R1 (angle entre les carbones  $C\alpha$  du premier résidu N-terminal de l'hélice A, du dernier résidu C-terminal de l'hélice A et le dernier résidu C-terminal de l'hélice B) (droite) l'hélice B et C de R1 (angle entre les carbones alpha du premier résidu N-terminal de l'hélice B, du dernier résidu C-terminal de l'hélice B et le dernier résidu C-terminal de l'hélice C) ; Haut : Simulation R1-3 / Bicelles zwitterioniques, Bas : Simulation R1-3 / Bicelles anioniques ; Gris : valeur de l'angle toutes les 40ns ; Noir : valeur de l'angle lissée au cours du temps.

Au contraire de ce qui est observé dans le système moléculaire décrit précédemment, la longueur maximale mesurée expérimentalement par SANS pour le fragment R1-3 de dystrophine en présence de bicelles zwitterioniques est identique à celle du même fragment seul en solution (Figure 33A et B). Dans nos simulations basse-résolution du même système R1-3/bicelles zwitterioniques, aucun changement conformationnel majeur n'a été observé au niveau de la protéine, notamment au niveau de l'hélice A de la répétition R1 (Figure 33D). Nous nous sommes donc par la suite concentrés sur l'optimisation du modèle R1-3/bicelle anionique de manière à expliquer les changements de conformations, mis en évidence par les expériences de SANS.



Pour tenter de modéliser les paramètres géométriques mesurés par SANS, nous avons étendu la simulation basse-résolution R1-3/bicelle anionique par une approche de dynamique moléculaire interactive. Les techniques de simulations interactives (IMD) (décrites dans le chapitre ) (Rapaport, 1997) permettent à un utilisateur d'interférer avec une simulation de dynamique moléculaire classique afin d'en altérer le cours. Dans notre cas, nous avons utilisé l'IMD pour forcer l'ouverture complète des hélices des faisceaux formant les répétitions R1 ou R3, et ainsi augmenter la valeur de  $D_{max}$ , conformément à ce qui a été observé par SANS. Quatre scénarios d'ouverture ont été envisagés (Figure 34A) : l'ouverture **(a)** de l'hélice A de la répétition R1, **(b)** des hélices A et B de la répétition R1, **(c)** des hélices B et C de la répétition R3 et **(d)** de l'hélice C de la répétition R3. Les scénarios **(a)** et **(d)** ont rapidement été écartés car ils correspondent à l'écartement d'une hélice de son faisceau d'origine et à son rapprochement de R2. Or, un tel remaniement ne permet pas d'envisager une augmentation de la valeur de  $D_{max}$  de l'objet. Cette hypothèse a été validée par des simulations d'IMD

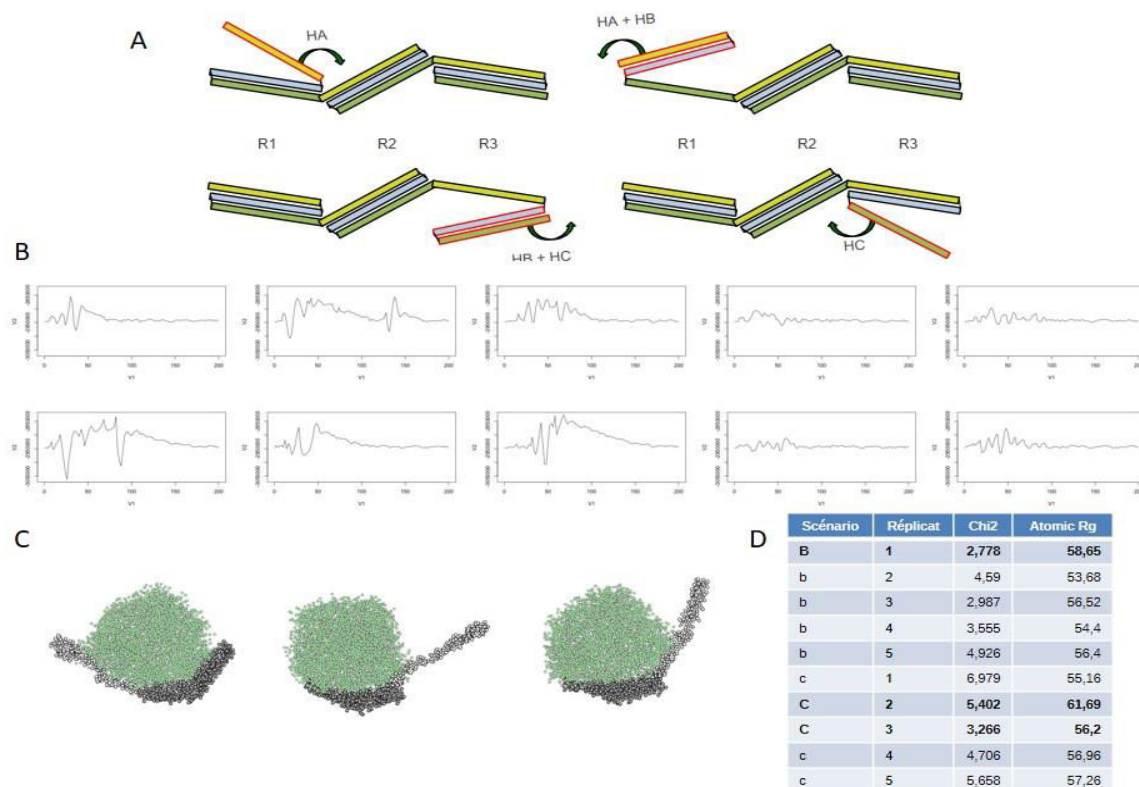


Figure 34 : IMD du fragment R1-3 avec une bicelle anionique (DMPC/DHPC/DMPS) ; A) Scénarios d'ouverture des faisceaux d'hélices ; jaune : hélices A, bleu : hélices B, verts : hélices C ; Contour rouge : Hélices considérées lors du scénario ; flèche verte : ouverture effectuée au cours du scénario. B) Evolution de l'énergie du système au cours de l'IMD ; Haut : Scénario b ; Bas : Scénario c. C) Meilleurs modèles R1-3/Bicelles anioniques issu de l'IMD après 2ns de relaxation ; Gauche : réplicat 1 du scénario b ; Milieu : réplicat 2 du scénario c ; Droite : réplicat 3 du scénario c ; Vert : bicelle ; Gris : R1-3 ; gris clair : Région perturbée lors de l'IMD. D) Paramètres calculés à partir du modèle relaxé après IMD ;  $\chi^2$  : corrélation courbe SANS expérimentale/courbe SANS théorique ; Rg : indicateur de la taille globale de l'objet moléculaire.

préliminaires et qui ne seront pas détaillées dans cette thèse. À l'inverse, les scénarios **b** et **c**, par l'écartement de deux hélices vers l'extérieur du fragment, permettent de rendre compte de l'élongation du fragment R1-3 de la dystrophine. De plus, ces deux scénarios ont l'avantage de limiter le mouvement des extrémités du fragment, et donc de la portion de protéine située avant la répétition R1 (*hinge* H1) ou après la répétition R3 (*hinge* H2). Les mouvements de l'hélice A de la répétition R1 ou de l'hélice C de la répétition R3, s'ils devaient se réaliser suivant les scénarios retenus, seraient associés à de difficiles déplacements de grande amplitude du reste de la protéine.

Les simulations d'IMD ont été effectuées sur une période de 200 ps, l'utilisateur intervient durant cette période (correspondant à environ 10 minutes réelles) jusqu'à obtenir une élongation du filament complète suivant le scénario suivi, puis en laissant le système se stabiliser jusqu'à la fin de la production (MD classique). Pour chaque scénario testé, cinq réplicas ont été effectués pour diversifier les solutions obtenues et le parcours emprunté pour y parvenir, tels que peuvent le montrer les tracés de l'énergie des systèmes au cours de la procédure (Figure 34B). L'énergie finale obtenue pour le système après dépliement du fragment protéique au contact de la bicelle est relativement proche de l'énergie initiale de la simulation basse-résolution, avant intervention de l'utilisateur et pour une protéine repliée adsorbée sur le tore de la bicelle (variation de moins de 1%). Ceci nous permet de suggérer que les simulations d'IMD ont permis au système de franchir une barrière énergétique pour atteindre une géométrie alternative et stable. Dans ce nouvel état du système, la protéine, en partie dépliée sur une répétition, compense cette ouverture par une interaction forte avec la bicelle lipidique. Notons que bien que les énergies déployées en cours d'IMD par l'utilisateur sont très importantes en valeur absolue, elles ne correspondent en aucun cas à une réalité physique et ne peuvent donc pas être comparées à des échelles de forces produites ou mesurées lors d'expériences de dépliement protéique réalisées sur des filaments de spectrine par Microscopie de Force Atomique (AFM, *Atomic Force Microscopy*) (Altmann et al., 2002; Paci and Karplus, 2000; Zhu and Asaro, 2008). La gamme des forces induites par l'utilisateur en cours d'IMD est à relier au temps de simulation sur lequel doit se dérouler ce type de simulations: en effet le changement géométrique imposé au système moléculaire est très grand sur un très petit laps de temps de simulation. Une première analyse qualitative repose

sur l'appréciation de l'expérimentateur, pour qui il semble qu'il est plus facile d'ouvrir R1 que R3.

Chaque simulation d'IMD a été suivie d'une simulation de dynamique moléculaire classique, toujours en utilisant le champ de force MARTINI (Marrink et al., 2007) et pour une trajectoire de 2 ns, ceci afin de laisser le système atteindre un état pleinement relaxé. À l'issue de cette courte MD, le modèle de protéine a été isolé et converti en modèle à l'échelle atomique en suivant le protocole standard proposé par le programme Backward (Wassenaar et al., 2014). À partir de ce nouveau modèle haute-résolution, une courbe SANS théorique a été simulée (programme Cryson de la suite ATSAS) et confronté à la courbe SANS expérimentale (Franke et al., 2017). De cette comparaison, une valeur de convergence des deux courbes ( $\chi^2$ ) a été calculée et celle-ci est d'autant plus petite que les deux courbes sont en accord. Nous avons également comparé le rayon de gyration expérimental ( $R_g$ , indication de la taille globale de l'objet moléculaire) aux valeurs calculées pour les différents modèles du fragment R1-3 de la dystrophine (Figure 34D). Le modèle le plus en accord avec les données expérimentales correspond à un réplica du scénario **b** (ouverture des hélices A et B de la répétition R1 vers l'extrémité N-terminale du fragment) (Figure 34C et D).

Pour la suite de ce chapitre, et en accord avec l'article portant sur l'étude couplée SANS/dynamique moléculaire interactive gros-grain (Dos-Santos Morais et al., 2018 ; Article en révision) (Annexe : Dos Santos Morais et al, 2018), nous ne conserverons que le modèle rendant le mieux compte de l'interaction du fragment R1-3 avec une bicelle anionique après l'ouverture du faisceau par IMD (scénario de l'ouverture des hélices A et B de la répétition R1), en accord avec les données SANS. Nous confronterons également la simulation de dynamique moléculaire ayant permis la relaxation de ce modèle à celle du système fragment R1-3/bicelle zwitterionique.

De manière remarquable, l'ouverture du faisceau ne modifie pas significativement l'interaction entre les lipides de la bicelle anionique et la dystrophine au niveau de sa région ALPS prédite (voir Introduction de ce chapitre) (Figure 35). Ainsi, contrairement à ce que nous supposions, l'interaction du motif ALPS prédit avec des membranes présentant des défauts de compacité lipidique pourrait précéder l'ouverture du faisceau d'hélice.

Nous proposons ainsi que l'interaction de l'ALPS avec les lipides permettrait de créer un intermédiaire déstabilisant les interactions hydrophobes entre les hélices B et C de R1, facilitant l'ouverture du faisceau.

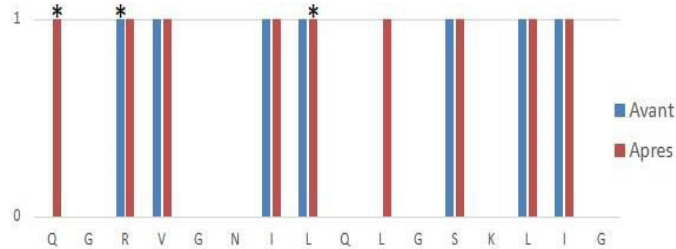


Figure 35 : Résidus du motif ALPS prédit retrouvés au contact des queues des lipides de la bicelle anionique ; Bleu : Avant IMD ; Rouge : Après IMD. Les barres marquées par une étoile indiquent les positions impliquées également avec la tête des lipides.

On observe cependant des différences notables entre les deux systèmes bicelle anionique et zwitterionique quant à ces mêmes résidus situés à proximité des lipides (Figure 36). Ainsi, en présence d'une bicelle anionique, on peut identifier quatre zones majeures d'interaction: la première située entre le milieu de l'hélice B de la répétition R1 jusqu'à la fin de l'hélice A de la répétition R2 ; la deuxième située sur l'hélice B de la répétition R2 ; la troisième correspondant à l'hélice A de la répétition R3 et la dernière étant située sur la première moitié de l'hélice C de la répétition R3. Ces régions de contacts sont retrouvées tant à proximité des têtes (effet des charges) que des queues (effet hydrophobe) des lipides. On peut observer des différences en fonction du type de lipides. Ainsi, la protéine, localisée au niveau du tore, est principalement à proximité des lipides à chaînes courtes (DHPC), eux-mêmes situés principalement sur le pourtour de la bicelle. Pour les lipides à chaînes longues, la fréquence de contact plus faible avec le DMPS qu'avec le DMPC reflète leur différence de proportion dans la composition en lipides de la bicelle (Figure 31).

Dans le cas d'une interaction avec une bicelle zwitterionique, la dystrophine semble en contact régulier avec l'objet sur toute sa longueur, à l'exception notable d'une large région comprise entre l'extrémité N-terminale de l'hélice C de la répétition R1 et la fin de l'hélice A de la répétition R2. Si les interactions avec les têtes des lipides suivent un motif similaire entre R1 et R2, on remarque cependant que les interactions avec les queues lipidiques sont limitées à certaines régions de la protéine: les hélices B des répétitions R1, R2 et R3, et le *linker* R2-R3. Ce dernier a été caractérisé comme rectiligne dans les modèles publiés (Delalande et al., 2018), et est lié au réseau dense d'interactions hydrophobes formant localement un faisceau à 2 hélices (voir article de ce chapitre).

L'interaction avec les queues hydrophobes des lipides à chaînes longues, situés plus au centre de la bicelle, est favorisée par l'espacement des lipides à chaînes courtes, plus présents sur le tore. En complément des expériences de SANS et des simulations de MD, nous avons effectué des mesures de spectrométrie de masse sur des complexes réticulés de dystrophine et de bicelles. Lors de la préparation des échantillons qui devaient être par la suite analysés par MS, nous avons ajouté aux bicelles zwitterioniques et anioniques, un lipide (le 16 :0 PacFA) dont un groupement diazirine situé à l'extrémité d'une de ses queues hydrophobes est capable après activation par des ultra-violets de créer des liaisons covalentes avec les résidus protéiques situés à proximité (Dos-Santos Morais et al., 2018 ; Article en révision) (Annexe : Dos Santos Morais et al, 2018). La protéine est ensuite fragmentée par un traitement à la trypsine et l'analyse MS permet d'identifier les résidus protéiques associés au lipide PacFA.

Trois fragments communs aux deux systèmes ont été identifiés (Dos-Santos Morais et al., 2018 ; Article en révision): le premier correspond à la moitié N-terminale de l'hélice C de la répétition R1; le deuxième et le troisième aux hélices A et C de la répétition R3. Dans les modèles résultant de la simulation du système R1-3/bicelle anionique, ces trois régions sont toutes retrouvées au contact des queues des lipides (et les têtes dans une moindre mesure), et cela tout au long de la simulation. Elles sont toutes retrouvées au contact des DHPC, mais seules les deux premières régions semblent interagir avec les DLPC et les DMPS. Dans nos simulations R1-3/bicelle zwitterionique, les deux premières régions sont faiblement retrouvées au contact des lipides, mais pas la troisième (hélice C de la répétition R3). Il est à noter que les zones de contacts identifiées par spectrométrie de masse sont localisées dans des régions globalement polaires (Figure 36), mettant en exergue le besoin d'étude des interactions hydrophobes à l'échelle atomique tel que présenté dans ce chapitre. D'autre part, alors que l'ajout de lipides anioniques (DMPS) modifie significativement la configuration de l'interaction de la protéine, on ne remarque pas de lien entre les régions au contact des lipides et les propriétés électrostatiques du fragment, en particulier avec les zones chargées positivement (Figure 36).

La présence de DMPS change le comportement de la protéine vis-à-vis de son interaction avec les lipides, en particulier au niveau des zones identifiées par spectrométrie de masse. D'une part, les contacts de la région R1 avec le DMPS entraînent l'ouverture du faisceau d'hélice R1. D'autre part, cette interaction privilégiée avec l'extrémité N-terminale

influe sur la capacité à lier les lipides de régions C-terminales, comme la troisième région identifiée par spectrométrie de masse (hélice C de R3). Cette région, très peu retrouvée au contact des DMPS, interagit de manière accrue avec les DHPC dans les bicelles contenant du DMPS en comparaison aux bicelles qui en sont dépourvues.

La région ALPS prédite sur la dystrophine est retrouvée au contact des lipides dans les simulations des deux systèmes, et semble notamment très impliquée dans l'interaction de la protéine avec les queues hydrophobes des lipides au niveau du tore de la bicelle.

Dans ce chapitre, nous avons présenté les résultats de notre exploration des propriétés physico-chimiques du fragment R1-3 en relation avec sa capacité à lier les membranes lipidiques. L'objectif était dans un premier temps de comprendre comment les irrégularités dans la répartition des interactions hydrophobes entre les hélices des faisceaux d'hélices étaient liées aux différences de conformations des régions *linkers* entre R1-R2 et R2-R3. Notre approche de cartographie atomique des interactions hydrophobes a également mis en évidence un réseau d'interactions entre les hélices de R1 différent de celui de R2 ou R3. En particulier, dans l'hélice B de R1 deux résidus forment un verrou structural. Ceux-ci auraient pour fonction de maintenir le faisceau d'hélices dans un état fermé en absence de lipides.

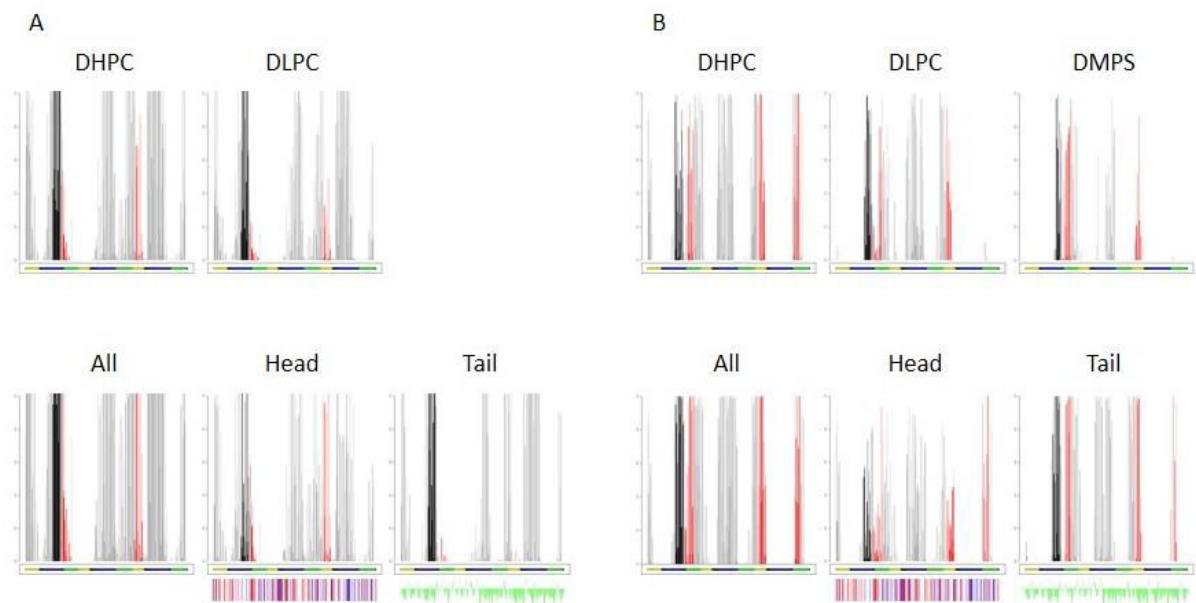


Figure 36 : Résidus en interactions avec les lipides dans les simulations de dynamique moléculaire. A) Bicelle zwitterionique après 20 $\mu$ s de dynamique moléculaire; B) Bicelle anionique après IMD et minimisation ; Graphique en barre : gauche) Charge électrostatique du fragment. (Bleu) Charge positive ; (Rouge) Charge négative ; droite) Potentiel hydrophobe. (Vert) Hydrophile ; (Jaune) Hydrophobe.

Le second objectif de ce chapitre était d'étudier le comportement du fragment R1-3 en présence de lipides. Ainsi, la protéine se localise au niveau du pourtour de la bicelle, ce qui indique une sensibilité au défaut de compaction des lipides dans cette région. De plus, il a été exploré *in silico* l'élongation du fragment R1-3 qui avait été mesurée expérimentalement en présence de bicelles anioniques. Les approches de dynamique moléculaire gros grains classiques ayant échoué à apporter une solution à cette observation, nous avons mis en place un protocole de simulation interactif afin de tester des hypothèses quant à son origine. La solution identifiée comme étant la plus en accord avec les données expérimentales correspond à l'ouverture des hélices A et B de la répétition R1 vers l'extrémité N-terminale du fragment. Ce mouvement nécessiterait la rupture des deux verrous structuraux identifiés sur l'hélice B de R1.

Enfin, nous avons mis en évidence que l'ajout du lipide anionique DMPS, en plus d'entraîner l'ouverture de R1, modifie le comportement d'interaction avec les autres lipides (DHPC, DMPC) à l'autre extrémité de la protéine (hélice C de R3) dans une région interagissant très peu avec les molécules de DMPS.

La connaissance fine de la structure et du comportement de ce fragment est un élément clé de la mise en place de nouvelles thérapies basées sur les mini- ou micro-dystrophines. L'interaction avec les lipides du sarcolemme étant un paramètre essentiel de la fonction de la dystrophine, la conception de protéines raccourcies mais fonctionnelles nécessite la conservation des propriétés de liaison aux lipides. De plus, l'ouverture du faisceau ne semblant pas essentielle à la liaison aux lipides, des études ultérieures sont nécessaires afin de déterminer le rôle de ce changement de conformation dans la fonction de la protéine. Une hypothèse que nous avançons est que cette ouverture qui expose les résidus hydrophobes normalement dévolus au maintien du faisceau d'hélice aux queues hydrophobes des lipides, permet d'augmenter la force de l'interaction entre la protéine et les lipides. Dans ce cas de figure, la conception de mini-dystrophines devra prendre en compte la présence du verrou maintenant le faisceau d'hélice fermé en absence de lipides.

#### D. Perspectives

Dans cette partie, nous discuterons des limites de notre présente étude et nous proposerons des éléments complémentaires pour la poursuite de ces travaux. Tout d'abord,

concernant les différences de comportement structural du fragment R1-3 de la dystrophine avec des bicelles anioniques ou zwitterioniques: il apparaît que la différence majeure entre les bicelles zwitterioniques et anioniques tient à la présence d'un lipide (DMPS) chargé négativement et qui est capable de créer des liaisons hydrogène. L'élongation de R1-3 observée par SANS, et que nous avons prédite comme étant causée par l'écartement des hélices A et B de la répétition R1, pourrait être initiée puis stabilisée par la formation de liaisons hydrogène entre la protéine et des DMPS. Cependant le champ de force utilisé pour nos simulations de dynamique moléculaire (gros-grain MARTINI), ne prend pas en compte ce type d'interaction et il faudrait donc simuler nos systèmes à haute-résolution pour avoir une meilleure description du mécanisme conduisant à l'ouverture du faisceau d'hélices de la répétition R1 de la dystrophine.

En parallèle, des approches FRET (*Fluorescence Resonance Energy Transfer*) sont en cours de mise en place pour étudier *in vitro* l'élongation du fragment R1-3 mesurée par SANS. Dans ces approches, deux protéines fluorescentes sont utilisées, choisies de telle manière à ce que la longueur d'onde d'émission de la première (le donneur) corresponde à la longueur d'onde d'excitation de la seconde (l'accepteur) (Figure 37A). À chaque étape d'absorption / émission, la longueur d'onde de la lumière est décalée vers le rouge. Si les deux marqueurs sont distants de moins de 10 nm, l'excitation lumineuse du donneur produit des photons qui sont absorbés par l'accepteur et qui émet à son tour de la lumière. Ce double cycle d'excitation/émission entraîne donc l'émission de photons à une longueur d'onde plus grande que celle attendue par la simple excitation du donneur. Par cette méthode, on obtient des informations sur la proximité spatiale des deux fluorophores.

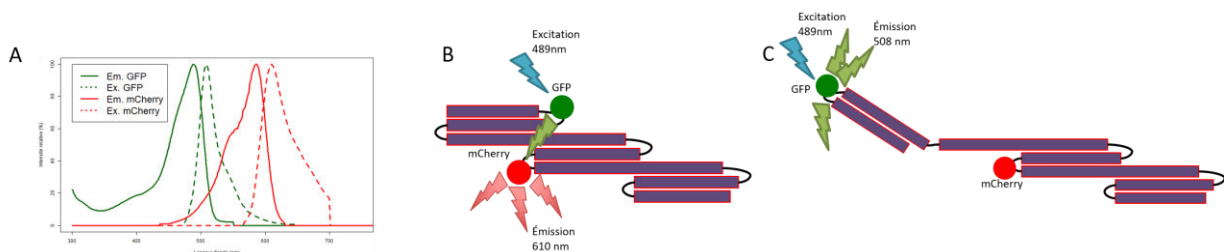


Figure 37 : Principe du FRET. A) Spectres (traits pleins) d'émission et (traits pointillé) d'excitation de la (vert) GFP et (rouge) mCherry. B) Principe du transfert de photon lorsque les deux fluorophores sont proches. C) Absence de transfert de fluorescence lorsque les fluorophores sont distants. Violet : Fragment de dystrophine R1-3 ; Vert : Protéine GFP ; Rouge : protéine mCherry.



Pour adapter cette approche à notre thématique, notre équipe travaille actuellement à la construction un fragment R1-3 dont deux boucles contiendraient deux protéines fluorescentes (cpGFP et mCherry) (Figure 37B). Par l'approche FRET, nous espérons pouvoir obtenir *in vitro* et *in cellulo* de nouvelles informations sur l'élongation du filament de dystrophine dans différentes conditions de courbure membranaire pour valider nos modèles issus de l'approche couplée SANS/dynamiques moléculaires gros-grain interactive. De plus, ces constructions fourniront un nouvel outil très puissant pour étudier la physico-chimie de R1-3 dans différents environnements lipidiques et dans le contexte de mutations affectant ce fragment.





#### IV. Modélisation de complexes de la dystrophine (actine, microtubules, lipides)

##### A. Introduction

Le réseau d'interaction entre protéines et lipides bâti autour de la dystrophine est essentiel à la stabilisation du sarcolemme lors des contractions musculaires répétées. Ainsi, des expériences menées dans notre équipe suggèrent que la région R11-15 du domaine central de la dystrophine augmenterait la résistance du sarcolemme aux forces de cisaillement via des interactions simultanées avec l'actine cytosquelettique et les lipides membranaires (Sarkis et al., 2013).

La détermination des interactions entre la dystrophine et ses partenaires est un élément essentiel à la compréhension de son fonctionnement chez les individus sains, et à l'origine de la pathologie chez les personnes atteintes de DMD et BMD. Des modèles d'interactions entre des fragments de la dystrophine et certains de ses partenaires, basés sur des données expérimentales, sont déjà disponibles dans la littérature scientifique, entre le domaine WW de la dystrophine et le  $\beta$ -dystroglycane (Huang et al., 2000), ou entre le domaine R16-17 et le domaine PDZ de la nNOS (Molza et al., 2015).

Dans ce chapitre, nous présenterons les différentes approches mise en place pour modéliser les complexes que forme le domaine central de la dystrophine avec les microtubules, avec l'actine-F, et avec les lipides membranaires.

## B. Microtubules

### 1. Introduction

Comme présenté dans l'introduction (Chapitre I), la dystrophine interagit avec différents éléments du cytosquelette via des domaines partagés avec d'autres protéines:

- l'actine, qui entre en contact avec le domaine ABD1 (composé de deux domaines CH (Calponin Homology)) à l'extrémité N-terminale de la dystrophine. Ces domaines sont retrouvés par exemple dans la calponine (Hemmings et al., 1992) ou la filamine (Duff et al., 2011).
- des protéines de filaments intermédiaires, dont les synamines (Bhosle et al., 2006) dans une région située entre la répétition R11 et l'extrémité C-terminale, et la cytokératine 19 (Stone et al., 2005) via le domaine ABD 1. Peu de travaux ont été publiés sur les liens entre dystrophine, DMD et filaments intermédiaires (36 réponses à la requête PubMed « (dystrophin[title/abstract] or DMD[title/abstract]) and "intermediate filaments" [title/abstract] »).

La dystrophine présente des similitudes avec d'autres *cytolinkers*, tels que les plakines (Fuchs and Karakesisoglou, 2001), qui ont pour fonction d'organiser spatialement et temporellement les différents éléments du cytosquelette. Cependant, contrairement à ces *cytolinkers*, la dystrophine ne présente pas de domaines connus de liaison aux microtubules (ex. : Gas-2 ou glycine-sérine-arginine (GSR) (Leung et al., 2002)).

Dans leur étude de 2007, Percival et al. (Percival et al., 2007) ont étudié l'effet de la déficience en dystrophine des souris *mdx* (Bulfield et al., 1984) sur l'organisation de l'appareil de Golgi. En parallèle d'anomalies d'organisation de l'appareil de Golgi dans les fibres musculaires (perte de l'uniformité de la distribution et de l'orientation, localisation, taille, ...), il est aussi observé une restructuration du réseau de microtubules. Celui-ci passe d'un maillage orthogonal et régulier à un réseau désorganisé (Figure 38). Ce travail fut le premier à montrer un lien entre dystrophine et microtubules, sans éclairer les bases moléculaires sous-jacentes.

### 2. Analyses des travaux de l'équipe du Pr.Ervasti

Depuis la publication du travail de Percival *et al.*, trois publications ont présenté des évidences biochimiques d'interactions directes entre la dystrophine et les microtubules.

Dans le premier article datant de 2009, Prins *et al.* ont mené des analyses d'immunofluorescence dans des fibres musculaires isolées de modèles murins *mdx* (Prins *et al.*, 2009). Il y est mis en évidence la co-localisation partielle des microtubules avec la dystrophine, au niveau des bandes I et des lignes M, les microtubules formant en plus un lien orthogonal entre les bandes I et lignes M. Il y est aussi observé la désorganisation du réseau de microtubules précédemment décrite par Percival *et al.* (Percival *et al.*, 2007). Cette perte d'organisation n'est cependant pas retrouvée en absence d'utrophine, mettant en évidence que l'organisation des microtubules est une fonction spécifique de la dystrophine. L'utrophine est une protéine exprimée à haut niveau durant le développement du muscle. C'est une protéine homologue à la dystrophine, avec une grande conservation de la séquence en acides aminés (Tinsley *et al.*, 1992), et avec une redondance dans les partenaires, comme notamment l'actine-F (Galkin *et al.*, 2002). Certains auteurs ont proposé que l'utrophine pourrait être surexprimé chez des patients déficients en dystrophine pour inverser les effets de l'absence de cette dernière (Tinsley and Davies, 1993). C'est pour cette raison que les propriétés de la dystrophine sont régulièrement comparées à celles de l'utrophine.

Des expériences de co-sédimentation ont montré une interaction directe entre la dystrophine et les microtubules. La région d'interaction dystrophine / microtubules a pu être localisée sur la répétition R24 grâce à l'utilisation de formes raccourcies de la dystrophine, qui co-sédimentent (de R10 à l'extrémité C-terminale, ou Dp260) ou non (De la fin de R24 à l'extrémité C-terminale) avec les microtubules. De plus, pour le fragment Dp260, une constante de dissociation d'environ 0,66  $\mu$ M a pu être déterminée.

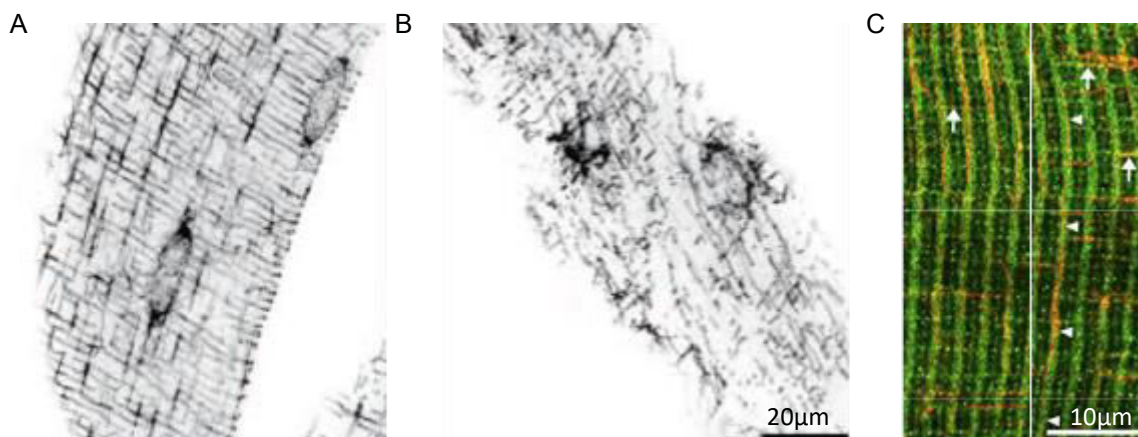


Figure 38 : Réseau de microtubules dans des fibres musculaires. A et B) Immuno-marquage du squelette de microtubule de fibres musculaires gastrocnémiennes de souris (A) sauvage ; (B) *mdx*. ; D'après Percival *et al.*, 2007. C) Immuno-marquage de la dystrophine (vert) et des microtubules (rouge) dans des fibres musculaires de souris ; Flèches : Microtubules longitudinaux ; Têtes de flèches : Microtubules transversaux ; D'après Prins *et al.*, 2009.

Dans une seconde publication de 2014 (Belanto et al., 2014), la zone d'interaction entre la dystrophine et les microtubules a été localisée entre les répétitions 20 à 23 du domaine central. Les constantes de dissociation ont pu également être déterminées entre les partenaires, et varient de 0,3  $\mu\text{M}$  à 0,66  $\mu\text{M}$ . Des kymographies ont été effectuées *in vitro* entre des microtubules dynamiques et des micro-particules recouvertes de dystrophine. Lors de ces expériences, il a été observé que la cinétique de polymérisation des microtubules était momentanément arrêtée lorsqu'ils rencontraient une micro-particule recouverte de dystrophine. Cette pause n'a pas été observée avec des micro-particules recouvertes d'utrophine. Avec la désorganisation du réseau de microtubule observée précédemment, les auteurs suggèrent que la dystrophine est un guide ou un aiguillage permettant l'organisation des microtubules au niveau du sarcolemme.

Dans la dernière étude, publiée par Belanto et al. (Belanto et al., 2016), l'objectif était de définir le domaine minimal permettant l'interaction directe entre la dystrophine et les microtubules. Des constructions de dystrophines, dont les domaines R20-24 ont été échangés avec leurs homologues R18-22 de l'utrophine montrent l'aspect nécessaire mais non-suffisant de R20-24 pour la liaison aux microtubules. Pour affiner l'importance individuelle des répétitions 20 à 24 de la dystrophine, des protéines de Dp260 avec le remplacement de répétitions de la région R20 à R24 par les équivalents R18 à R22 de l'utrophine ont été construites. Il a ainsi été montré qu'aucune répétition n'était individuellement responsable de l'interaction avec les microtubules.

Pour déterminer l'importance fonctionnelle de l'interaction dystrophine / microtubules, des souris *mdx* ont reçu un gène codant une dystrophine hybride dont les répétitions 20 à 24 ont été remplacées par leur homologues R18-22 de l'utrophine. Ce transgène est sous contrôle du promoteur de l'actine  $\alpha$ , exprimé seulement dans les muscles squelettiques. Chez ces souris *mdx-Dys<sup>ΔMTB</sup>*, la dystrophine et ces complexes associés (dystroglycane, sarcoglycane, etc) sont correctement localisés mais en plus faible quantité que chez des souris *wt*, mais absents chez les souris *mdx*. De plus, le maillage orthogonal de microtubules est restauré dans des paramètres de densité et de direction similaires à ceux de muscles de souris au phénotype sauvage.

Les trois publications présentées ici montrent qu'il existe une interaction directe entre la dystrophine et les microtubules, via une région comprenant les répétitions 20 à 24 du

domaine central, et s'étendant jusqu'au domaine C-terminal. Aucune partie de cette région n'est suffisante individuellement pour assurer la liaison aux microtubules. L'interaction de la dystrophine avec les microtubules serait plutôt assurée par un effet coopératif entre les différentes parties de ce fragment. Cette région a aussi été montrée comme essentielle au maintien d'un maillage de microtubules au sarcolemme.

Différents points n'ont pas été abordés dans ces études, comme l'effet de la désorganisation du réseau de microtubules sur la fonctionnalité du muscle et la capacité du sarcolemme à résister aux blessures induites par les contractions. En particulier, dans la publication la plus récente (Belanto et al., 2016), on peut noter l'absence de mesures comparatives de la résistance du sarcolemme entre muscle normaux et muscles de souris *mdx-Dys<sup>ΔMTB</sup>*. De plus, l'interaction directe entre les microtubules et la dystrophine (de R20 jusqu'à l'extrémité C-terminale) a été observée *in vitro*. Cependant, cette région de la dystrophine a aussi été identifiée comme interagissant avec d'autres partenaires liant eux aussi les microtubules ( $\beta$ -dystroglycane, plectine, synemine, ankyrine). La désorganisation du réseau de microtubules observée *in vivo* associée à la perte de la dystrophine pourrait ainsi être médiée par la perte d'un de ces intermédiaires potentiels. D'autres sources ont rapporté un tel lien indirect entre la dystrophine et les microtubules lors d'analyses protéomiques du réseau d'interaction de la dystrophine (Murphy and Ohlendieck, 2016; Murphy et al., 2018). Ainsi, alors que l'interaction entre dystrophine et microtubules semble admise, il existe toujours des incertitudes sur le lien, direct ou indirect, entre ces protéines.

### 3. Reproduction des résultats

Les travaux montrant l'interaction directe entre la dystrophine et les microtubules mentionnés précédemment ont la particularité de tous être issus du laboratoire du Dr James Ervasti, de l'Université du Minnesota. De manière générale, peu d'équipes dans le monde étudient la dystrophine via des approches biochimiques. De ce fait, il n'existe pas à notre connaissance dans la littérature d'autres sources que celles provenant du laboratoire du Dr. Ervasti rapportant des preuves d'interactions directes entre la dystrophine et les microtubules.

L'un des objectifs majeurs de cette thèse consistait à étudier les bases physico-chimiques des interactions entre la dystrophine et les microtubules par la combinaison



d'approches de simulations numériques et de données structurales obtenues par microscopie électronique.

La première étape de ce travail consistait donc à reproduire les résultats du Dr Ervasti pour déterminer les conditions optimales d'association entre les partenaires avant l'observation du complexe par microscopie électronique. Nous avons entrepris des expériences de co-sédimentation avec les fragments R16-24 et R20-24.

Toutes nos expériences de co-sédimentations avec les microtubules ont été effectuées à des concentrations variant de 10  $\mu\text{M}$  à 20  $\mu\text{M}$  de fragments de dystrophine. Les microtubules ont été préparés à la concentration standard de 10  $\mu\text{M}$  de dimère  $\alpha/\beta$ -tubuline dans le tampon Brb80 (80 mM PIPES pH 6.8, 1 mM  $\text{MgCl}_2$ , and 1 mM EGTA) et stabilisés par 10  $\mu\text{M}$  de taxol (Figure 39) (Guesdon et al., 2016). Nos premières expériences de co-sédimentations entre la dystrophine et les microtubules ont été effectuées avec le fragment R20-24 de la dystrophine.

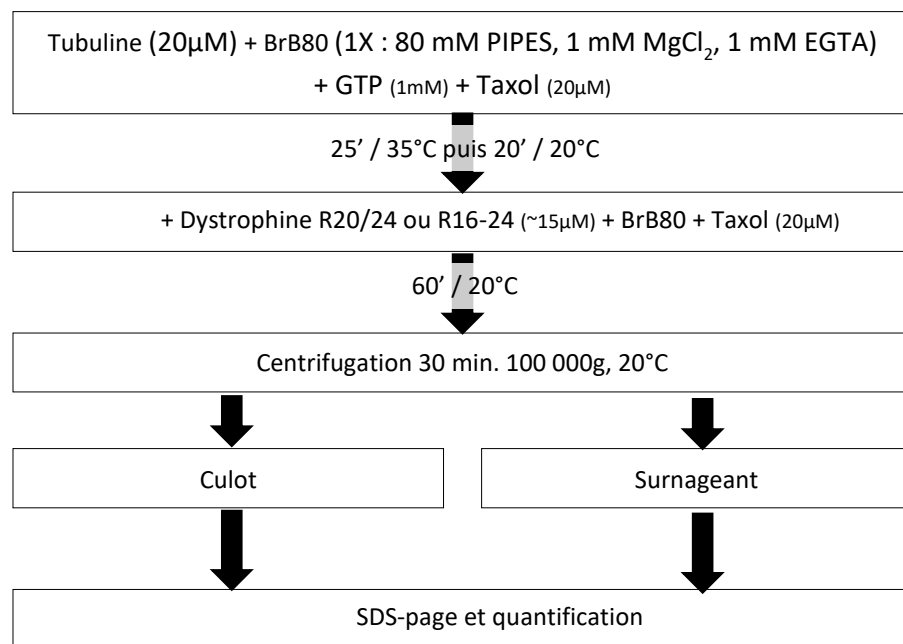


Figure 39 : Protocole de co-sédimentation microtubules / dystrophine.

La dystrophine a été incubée une heure avec des microtubules préformés et stabilisés par du taxol. L'échantillon a ensuite été centrifugé à 100 000g 30 minutes à 20°C. Les résultats, présentés Figure 40, montrent que le fragment R20-24 ne co-sédimente pas avec les microtubules (Figure 40B et C). Ces expériences ont été répétées à deux reprises, aboutissant aux mêmes conclusions.

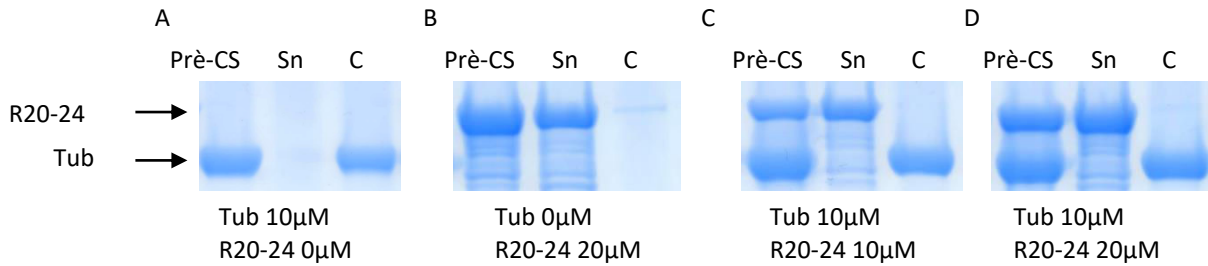


Figure 40 : Analyse SDS-Page des échantillons après co-sédimentation entre le fragment de dystrophine R20-24 et les microtubules. Prè-CS : Avant centrifugation; SN : Surnageant après centrifugation; C : Culot après centrifugation; Tub : concentration en dimère Alpha/beta de tubuline ; R20-24 : Concentration en fragment R20-24 de dystrophine. A) Contrôle Microtubules ; B) Contrôle Dystrophine R20-24 ; C) Co-sédimentation Microtubules / R20-24 à 10µM de dystrophine ; D) Co-sédimentation Microtubules / R20-24 à 10µM de dystrophine.

Au regard des résultats obtenus avec R20-24, nous avons choisi de répéter cette expérience en utilisant à la place le fragment R16-24. Pour écarter un éventuel effet du tampon Brb80, nous l'avons remplacé par de l'HEPES dans une partie de la manipulation (Figure 41). Quelque soit nos conditions expérimentales, nous n'avons observé aucune différence de sédimentation du fragment R16-24 en présence ou en absence de microtubules.

Dans ce chapitre, nous nous sommes basés sur les travaux de Belanto et Prins (Belanto et al., 2014, 2016; Prins et al., 2009) qui montrent, par co-sédimentation, une interaction directe entre la dystrophine et les microtubules. Dans ces travaux, la zone d'interaction de la dystrophine y est tour à tour définie sur R24, puis entre R20 et R23. L'utilisation des fragments R20-24 et R16-24 était donc en accord avec ces résultats, et a fait l'objet de co-sédimentations avec des microtubules dans quatre expériences indépendantes. Alors que le lien entre la

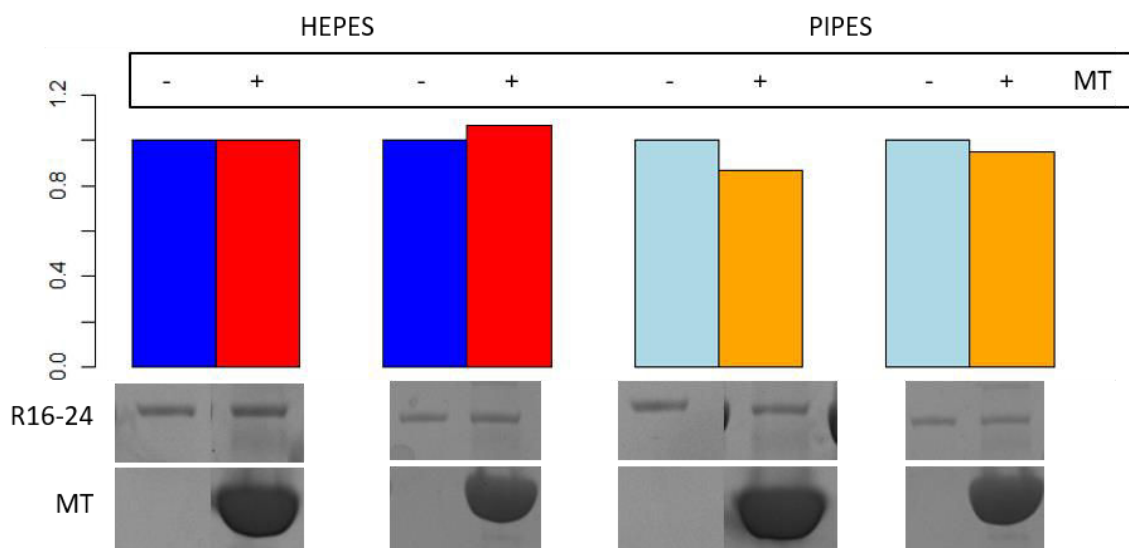


Figure 41 : Analyse et quantification SDS-Page de la fraction sédimentée des échantillons après co-sédimentation entre le fragment de dystrophine R16-24 et les microtubules. (bleu/rouge) Tampon HEPES (cyan/orange) Tampon PIPES. La quantité de dystrophine récupérée dans les culôts ne varie pas en fonction de la présence ou non de microtubules.

dystrophine et les microtubules a été rapporté par plusieurs sources, ces expériences ne permettent pas d'apporter d'éléments quant à la nature de l'interaction, directe (Belanto et al., 2016) ou indirecte (Murphy and Ohlendieck, 2016; Murphy et al., 2018), entre elle.

Dans ce cadre, et en dépit de ce qui était prévu initialement, nous n'avons pas pu acquérir des images par microscopie électronique de complexes dystrophine/microtubules avec les fragments de dystrophine à notre disposition. Nous avons donc décidé de ne pas continuer cette partie de mon projet de thèse.

## C. Actine

### 1. Introduction

Suite aux premiers séquençages du gène *DMD* (Hoffman et al., 1987b; Koenig et al., 1987), il a été observé de grandes similitudes entre l'extrémité N-terminale de la dystrophine et le domaine N-terminal de l' $\alpha$ -actinine liant l'actine (Hammonds, 1987). Par la suite, la capacité de la dystrophine à lier directement à l'actine-F via son domaine N-terminal ABD1 a été validée par des expériences de co-sédimentation (Hemmings et al., 1992).

Un second domaine de liaison à l'actine (ABD2) a ensuite été identifié dans le domaine central de la dystrophine, entre les répétitions R11 et R17 (Amann et al., 1998; Rybakova et al., 1996). Cette région a été par la suite réduite au fragment R11-15, la région R16-17 ayant été identifiée comme le site de liaison de la nNOS (Lai et al., 2009). De plus, il a été montré dans notre équipe que le fragment R11-15 co-sédimentait avec l'actine-F (Sarkis et al., 2013).

La constante d'affinité entre l'actine et la dystrophine entière est inférieure au micromolaire ( $0,32 \pm 0,04 \mu\text{M}$  (Belanto et al., 2014),  $0,5 \mu\text{M}$  (Rybakova et al., 1996)), alors qu'elle est mesurée à  $14,2 \pm 5,2 \mu\text{M}$  pour R14-16 (Amann et al., 1998), et  $3,92 \pm 2,18 \mu\text{M}$  pour ABD1 (Henderson et al., 2012). Dans cette même étude d'Henderson et al, il est montré que l'extrémité C-terminale de la dystrophine (à partir de la répétition R17), accroît l'affinité du duo ABD1/2 pour l'actine, bien que ne se liant pas elle-même directement à l'actine. L'affinité passe de  $0,60 \mu\text{M}$  pour le fragment s'étendant de ABD1 jusqu'à ABD2 à  $0,11 \mu\text{M}$  pour la dystrophine complète.

En plus des preuves obtenues par ces expériences de co-sédimentation, il existe aussi dans la littérature d'autres approches ayant mis en évidence l'interaction entre ces deux

protéines. Ainsi, dans une étude de notre équipe (Sarkis et al., 2013), des mesures *in vitro* de la résistance aux forces de cisaillement de modèles de membranes ont été effectuées en présence de dystrophine et/ou d'actine-F. Il a été mis en évidence que l'association actine-F/Dys-ABD2 permet d'accroître d'un facteur 6 à 10 la résistance de la membrane aux forces de cisaillements.

Bien que cette observation soit relativement peu documentée, les preuves expérimentales de l'interaction directe entre l'actine-F et le domaine ABD2 proviennent de plusieurs laboratoires, contrairement aux données concernant le complexe dystrophine / microtubules (voir C-Actine). L'un des objectifs de ma thèse était d'établir les bases physico-chimiques de l'interaction dystrophine-ABD2/Actine-F par la construction, puis l'analyse de modèles de complexes issus de l'exploitation d'images de cryo-microscopie électronique. Nous avons choisi d'utiliser le fragment R11-15 de la dystrophine.

## 2. Co-sédimentation

La première étape de ce travail d'observation a consisté à déterminer par co-sédimentation les paramètres de formation du complexe ABD2/actine (Figure 42). Dans cette technique, le fragment de dystrophine est incubé avec de l'actine-F. Le mélange est ensuite centrifugé à haute vitesse (100 000g) : l'actine-F sédimente, emportant avec elle les protéines qui lui sont liées. Au contraire, les protéines non liées à l'actine-F restent en solution. Le contenu des différents compartiments (culot et surnageant) est ensuite analysé par SDS-PAGE. La quantification de la coloration de chaque bande sur le gel permet de mesurer la quantité de dystrophine liée à l'actine-F. La constante d'affinité entre les deux partenaires peut être déterminée en répétant la manipulation avec une gamme de concentrations du fragment R11-15, à concentration constante en actine. La méthode consiste dans un premier temps à se placer à une concentration en R11-15 initiale ( $C_{ini}$ ) telle que les sites de liaisons de l'actine soient saturés et que la quantité de dystrophine sédimentée ( $C_{sed}$ ), soit constante ( $C_{plateau}$ ), formant un plateau sur la courbe d'association (Figure 43). La constante d'association (Kd) du fragment R11-15 avec l'actine est alors définie par la concentration initiale en dystrophine  $C_{ini} = Kd$  tel que  $C_{sed} = C_{plateau}/2$ .

La concentration d'actine-G permettant d'obtenir un équilibre entre la polymérisation et la dépolymérisation des filaments d'actine (ou concentration critique) est de l'ordre de 0,1  $\mu\text{M}$  (Heier et al., 2017; Lodish et al., 2000). Pour s'assurer une polymérisation correcte de l'actine-G en filaments, nous avons choisi de travailler systématiquement à une concentration finale d'actine-G de 10 $\mu\text{M}$  (~100x la concentration critique). L'association dystrophine/actine a été testée dans une gamme de concentration en R11-15 variant de 1 $\mu\text{M}$  à plus de 100 $\mu\text{M}$ . Cette seconde limite a été définie afin d'être environ d'un ordre de grandeur supérieure aux valeurs de  $K_d$  de ce fragment qui ont été précédemment rapportées dans la littérature (voir *Introduction*).

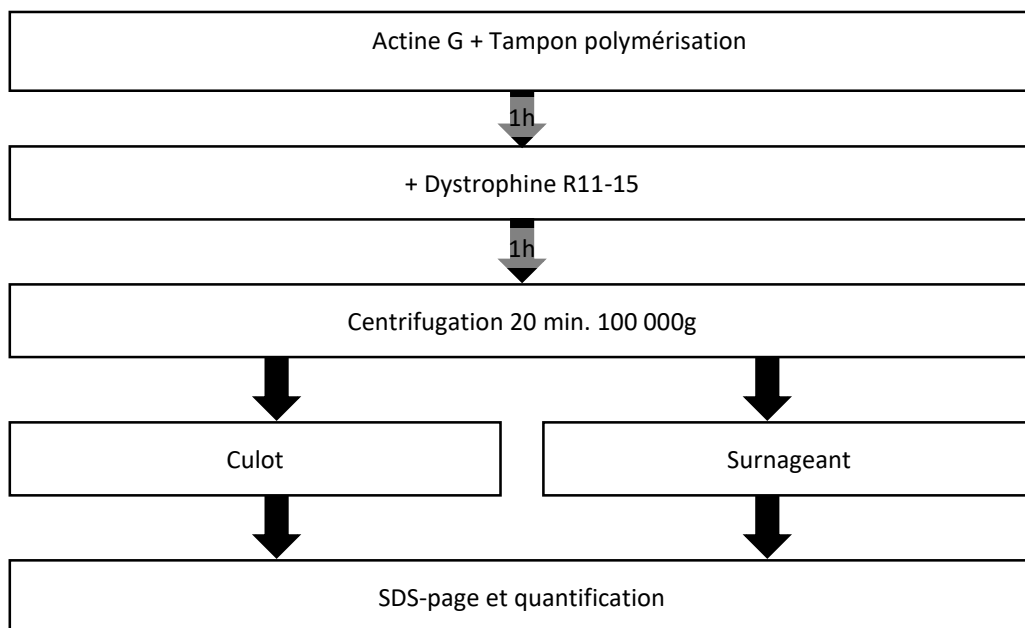


Figure 42 : Protocole de préparation et de co-sédimentation du complexe actine-F/ABD2.

Dans une première expérience de co-sédimentation (Figure 43), la gamme de concentration en R11-15 s'étendait de 3  $\mu\text{M}$  à 110  $\mu\text{M}$ . Il a été mesuré un  $K_d$  de 10  $\mu\text{M}$ , en accord avec les valeurs de la littérature.

De plus, il a été calculé un ratio dystrophine:monomère dans l'actine F de 1:3, c'est-à-dire qu'un fragment R11-15 de dystrophine est retrouvé tous les 3 monomères d'actine-G.

L'expérience a été répétée dans une seconde série de co-sédimentations, avec une gamme de dystrophine R11-15 variant de 2  $\mu\text{M}$  à 120  $\mu\text{M}$  (Figure 43). Le  $K_d$  a été mesuré à 20  $\mu\text{M}$ , mais avec un ratio dystrophine/actine-F de 1:5. Cette valeur est du même ordre de

grandeur que celle obtenue précédemment. Notre système expérimental permet donc d'obtenir des complexes R11-15-actine-F dans des conditions d'associations attendues.

Afin de s'assurer de l'association maximale entre l'actine et le fragment R11-15 lors des acquisitions expérimentales par CET, les grilles ont été préparées avec une concentration finale en dystrophine 5 à 10 fois supérieure au Kd, c'est-à-dire entre 50  $\mu\text{M}$  et 100  $\mu\text{M}$ . En se plaçant à ces valeurs, nous nous approchons d'une saturation de l'actine en dystrophine, l'objectif étant de maximiser le nombre de complexes observables lors des acquisitions.

### 3. Acquisitions CET

Les échantillons contenant les complexes moléculaires ont été préparés suivant un protocole similaire à celui utilisé lors de la préparation des échantillons des expériences de co-sédimentations. Juste avant le dépôt sur les grilles, des nano-particules d'or de 10 nm ont été ajoutées à l'échantillon. Ces billes d'or sont opaques aux électrons et servent de points de

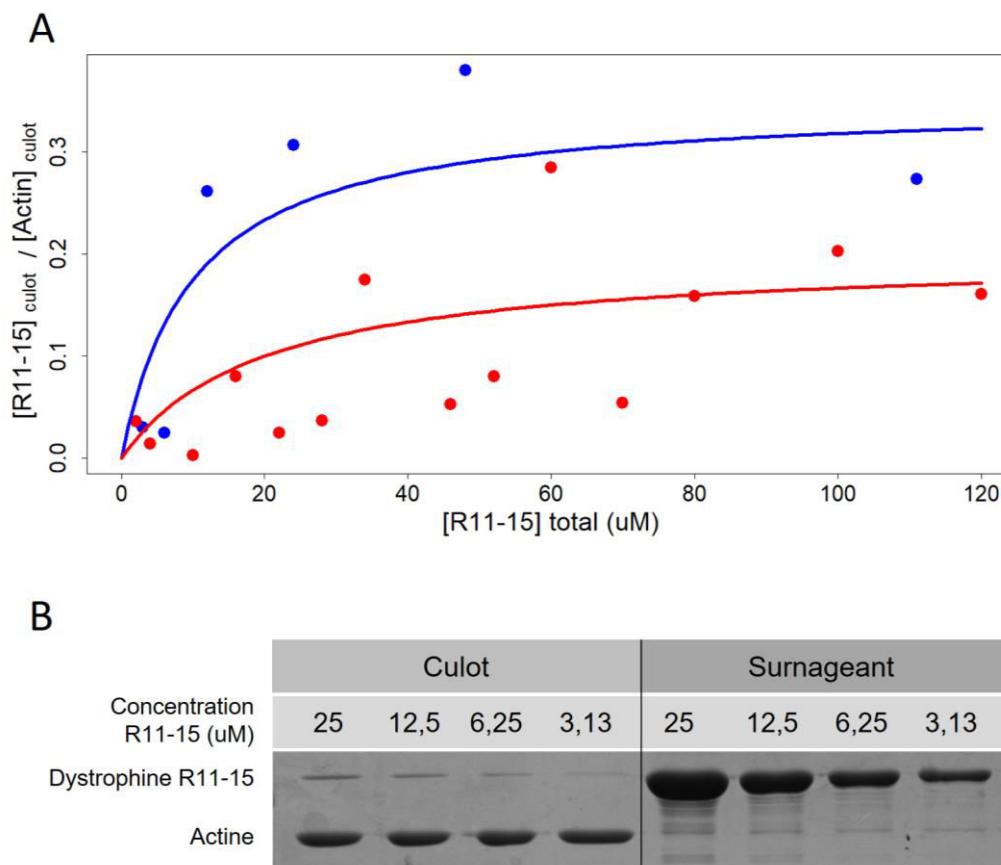


Figure 43 : A) Courbes de co-sédimentations; En bleu : premières mesures; en rouge : secondes mesures. B) Exemple d'analyse SDS-Page d'expériences de co-sédimentations R11-15 / Actine-F;

repère fixes (ou marqueurs fiduciaires) lors de la reconstruction des tomogrammes en 3D à partir des acquisitions en 2D.

Lors de l'analyse des acquisitions par CET, il a été régulièrement observé une zone de densité à proximité de l'actine faisant penser aux fragments de dystrophines (Figure 44). Cependant, ces événements restent rares, et difficilement exploitables à cause de la qualité limitée des images. Nous attribuons ce problème à la concentration élevée de dystrophine R11-15 non-liée. En effet, à 100 $\mu$ M de dystrophine, et à 10 $\mu$ M d'actine-G que l'on peut considérer comme entièrement impliquée dans des filaments d'actine, seulement 1  $\mu$ M à 2  $\mu$ M de R11-15 est impliqué dans un complexe avec l'actine, il en reste 98 $\mu$ M libre. Comme il est attendu que la protéine libre se localise aux interfaces air/eau, là où se retrouvent les marqueurs fiduciaires, la concentration élevée en fragment de dystrophine non-lié à l'actine saturerait l'interface. L'excédent de protéines est alors retrouvé dans la glace et en dégrade la qualité.

Dans l'objectif de pouvoir identifier des molécules uniques de R11-15 dans des complexes actine-dystrophine préparés à partir de concentrations faibles de R11-15, nous

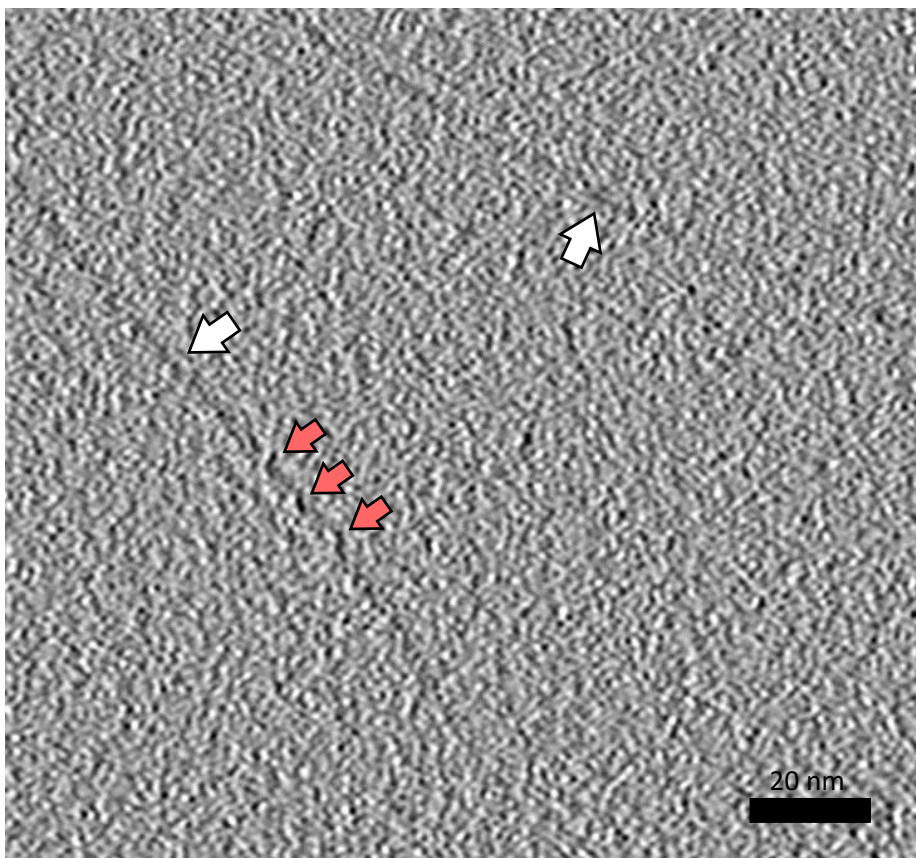


Figure 44 : Acquisition CET. (Flèches blanches) Filaments d'actine; (Flèches rouges) Densités électroniques atypiques le long du filament d'actine.

avons entrepris de travailler avec des molécules de R11-15 couplées à une ou plusieurs nanoparticules d'or.

Le fragment de dystrophine R11-15 contient 7 histidines à son extrémité N-terminale. Celles-ci forment une étiquette histidine (*His-Tag*) utile pour sa purification. Le marquage de la protéine consiste à utiliser des nano-particules d'or fonctionnalisées par un unique groupement Ni-Tri-NTA pour lier le His-Tag de la protéine. Brièvement, les particules d'or sont mises en présence de HS-C<sub>16</sub>-EG<sub>3</sub>-Tris-NTA (Tri-NTA) selon un ratio permettant qu'il ne soit greffé qu'un unique Tris-NTA au maximum par particule d'or (Duchesne et al., 2008). Les billes fonctionnalisées, représentant environ 10% de la quantité initiale de particules, sont isolées par chromatographie d'affinité sur une colonne histidine, et finalement concentrées à environ 130 nM dans un volume d'échantillon de 500  $\mu$ L.

Le marquage de la dystrophine R11-15 consiste à mettre en présence les nano-particules mono-fonctionnalisées et la protéine selon un ratio de 1:100 (mol/mol). Ce rapport des concentrations a été préalablement déterminé dans l'objectif de ne pas saturer le champ de prise de vue en nano-particules d'or. L'excédent de protéines non marquées est éliminé par centrifugation (17 000 g, 10 minutes) : les protéines marquées sédimentent avec les nano-particules d'or tandis que les protéines non-marquées restent dans le surnageant. Après incubation avec de l'actine-F et dépôt sur grilles de CET, la dystrophine marquée est à une concentration finale de 100nM.

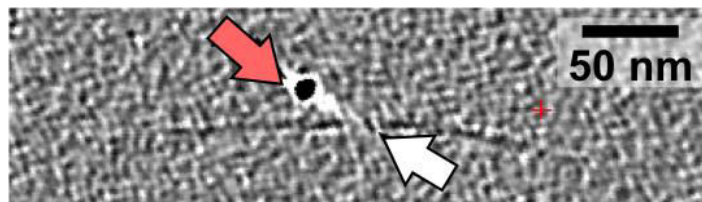


Figure 45 : Image issue de la reconstruction d'acquisitions CET de fragment R11-15 marqué par des nano-particules d'or de 10 nm de diamètre ; (Flèche blanche) Filament d'actine; (Flèche rouge) Particule d'or.

Malgré la faible concentration en dystrophine marquée, inférieure à la valeur du Kd, les premières acquisitions CET permettent d'observer à plusieurs reprises des nano-particules à proximité des filaments d'actine (Figure 45). Ces résultats montrent la pertinence des approches de marquages de protéines par des nano-particules d'or. Cependant, les particules de 10nm ont un contraste très fort comparé à celui des objets biologiques étudiés. La procédure d'acquisition crée ainsi une zone d'ombre autour des particules. Dans cette région,



le signal est dégradé et empêche d'étudier le détail de la proximité entre la bille d'or liée à la dystrophine et l'actine (Figure 46).

Pour limiter l'effet de la taille des particules sur la qualité des acquisitions dans leur entourage, une seconde approche de marquage a été mise en place avec des nano-particules d'or de 1,4 nm de diamètre (Nanoprobes 1,4 nm Nanogold®) (Figure 46). Ces particules sont mono-fonctionnalisées par des groupements chimiques permettant de les associer de manière covalente aux protéines. Selon la fonction chimique greffée sur la particule d'or, différents sites de la protéine peuvent être ciblés.

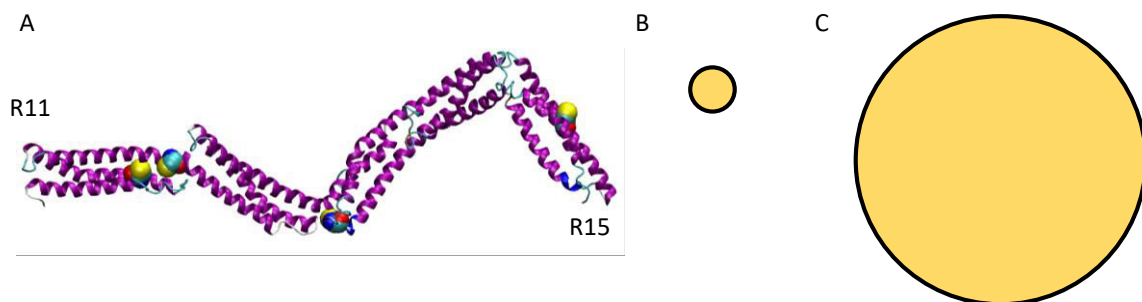


Figure 46 : Représentation à l'échelle de la dystrophine et des nano-particules d'or ; A) Modèle sous contraintes SAXS du fragment R11-15 de la dystrophine (d'après Delalande et al., 2018). Les sphères représentent les cystéines présentes dans ce fragment. Dans ce modèle, l'atome de soufre (jaune) du résidu cystéine situé sur la répétition R15 est le seul à être accessible au solvant; B) Représentation schématique à l'échelle d'une nano-sonde d'or de 1,4 nm de diamètre; C) Représentation schématique à l'échelle d'une nano-particule d'or de 10 nm de diamètre.

Deux stratégies de marquages ont été utilisées, ciblant deux régions différentes. La première utilise des nano-particules de 1,4 nm équipées d'un groupement chimique sulfo-N-hydroxysuccinimide ester (Sulfo-NHS). Celui-ci est réactif aux amines primaires, tel que celles situées à l'extrémité N-terminale des protéines. La seconde stratégie consiste à marquer les cystéines via des nano-particules d'or portant chacune un unique groupe maléimide. Le fragment R11-15 contient 4 résidus cystéines (Figure 46), et les modèles actuellement disponibles (Delalande et al., 2018) semblent montrer qu'au moins une cystéine est accessible et peut donc potentiellement être marquée par ces nano-probes.

Les protocoles de marquages sont sensiblement similaires dans les deux stratégies (Hainfeld and Furuya, 1992; Hainfeld and Powell, 1999). Brièvement, les nanoparticules sont mises en excès en présence de la protéine à marquer (ratio 4 :1). Après un temps de réaction variant de 1h (à température ambiante) à 12h (à 4°C), les nanoparticules non-liées à des protéines sont séparées par chromatographie d'exclusion de taille. Il est à noter que cette étape ne permet pas de séparer les protéines marquées des protéines non-marquées.

Cependant, un dosage par spectrophotométrie permet de conclure que l'ensemble de la dystrophine récupérée est marquée par des nano-probes.

Après marquage, les échantillons marqués (à l'extrémité N-terminale ou sur les cystéines) sont à des concentrations d'environ 10 $\mu$ M, du même ordre de grandeur que le Kd de l'interaction actine/R11-15 rapportés dans la littérature (Amann et al., 1998) ou déterminé dans nos propres expériences. La préparation des grilles de microscopie électronique consistait en l'incubation du fragment de dystrophine marqué avec l'actine-F, le tampon de polymérisation de l'actine et les marqueurs fiduciaux (nano-particules d'or de 10 nm de diamètre permettant à la construction des tomographies). Durant cette étape, la dystrophine subissait une dilution d'un facteur 2, pour atteindre une concentration finale de 5  $\mu$ M.

Après acquisition des micrographies et reconstruction des volumes par tomographie de complexes avec le fragment de dystrophine aux cystéines marquées, on peut observer de manière sporadique des nano-particules à proximité des filaments d'actine (Figure 47). Cependant, ces événements sont rares, difficilement discernables du bruit de fond, et ne permettent pas de rassembler suffisamment d'information pour espérer pouvoir créer un modèle tridimensionnel du complexe. Aucune sonde à proximité de l'actine n'a pu être observée avec la dystrophine marquée à son extrémité N-terminale.

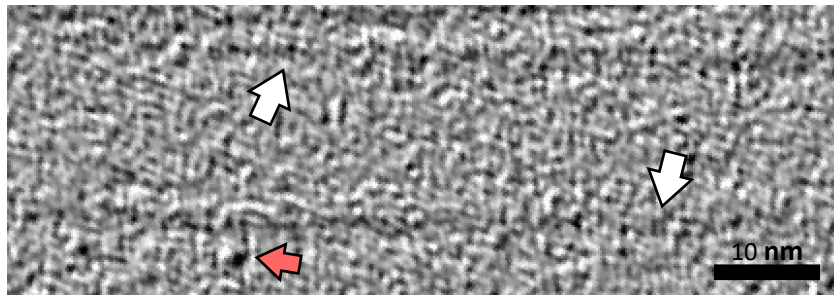


Figure 47 : Acquisitions par CET de complexes Actine-F / ABD2 aux cystéines marquées par des nano-sondes d'or de 1,4nm de diamètre. (Flèches blanches) Filaments d'actine ; (Flèche rouge) Nano-sonde d'or de 1,4 nm de diamètre.

Dans une dernière tentative pour pouvoir observer le complexe R11-15/Actine-F, nous avons pris le parti de « forcer » l'association entre l'actine et la dystrophine par l'utilisation d'un agent chimique réticulant (*cross-linker*), l'EDC. Des tests préliminaires par coloration négative d'échantillons de complexes réticulés observés par microscopie électronique à transmission ont cependant montré que l'actine formait des faisceaux très larges, empêchant toute acquisition de fibres isolées permettant d'observer potentiellement les complexes dystrophine/actine.

Toutes les approches présentées dans ce chapitre se sont avérées infructueuses. La principale difficulté provient de la valeur de la constante d'affinité entre le fragment R11-15 de la dystrophine pour l'actine F. Ainsi, les concentrations d'ABD2 permettant une décoration régulière du filament d'actine ne sont pas compatibles avec l'acquisition de micrographies de qualité, permettant d'y isoler les complexes avant reconstruction 3D. Des approches de marquages par des particules d'or n'ont pas permis de surmonter cette contrainte. Une alternative à ces méthodes serait d'effectuer les acquisitions sur des échantillons moins concentrés en dystrophine, de l'ordre de  $10\mu\text{M}$ , grâce à des microscopes électroniques de nouvelle génération (par exemple FEI Titan), associé à des caméras plus performantes (améliorant les résolutions spatiales et temporelles). Il serait ainsi possible d'identifier des complexes plus épars le long des filaments. Cependant, ces dispositifs restent rares et peu disponibles. Une autre stratégie pourrait consister à l'utilisation d'un fragment de dystrophine s'étendant de l'extrémité N-terminale (incluant le domaine ABD1) jusqu'au domaine ABD2. Comme rapporté dans l'introduction, la constante d'affinité de ce fragment avec l'actine-F a été déterminée inférieure au micro-molaire (Henderson et al., 2012). Il serait ainsi possible de préparer des échantillons à une concentration de dystrophine de 5 à  $10\mu\text{M}$ , en faisant l'hypothèse qu'elle ne gêne pas la qualité des acquisitions. D'autre part, la taille élevée du fragment (1973 acides aminés), comparée à la taille de R11-15 (515 acides aminés), facilitera son identification le long de l'actine. Cependant, un fragment aussi long doit être produit dans des systèmes eucaryotes (cellules d'insectes SF6), bien moins performants en termes de quantité produite, comparés aux systèmes bactériens utilisés pour produire R11-15. Il n'est ainsi pas certain que nous aurions pu produire assez de protéines pour permettre des conditions d'acquisition optimales.

Du fait des difficultés rencontrées dans ce volet du projet, et du temps que nécessiterait la mise en place de nouvelles stratégies, nous avons fait le choix de ne pas aller plus loin dans ce projet le cadre de cette thèse.

D. Article 2 : How the central domain of dystrophin bridges F-actin to sarcolemma lipids

Des travaux effectués dans notre équipe (Molza et al., 2014) ont permis de proposer les premiers modèles du complexe ternaire nNOS / dystrophine / actine-F (Figure 48). Après une première étape d'ajustement flexible pour la détermination de la conformation d'un fragment R11-15 de dystrophine (voir *II-Méthode*), la modélisation du complexe R11-15 / actine-F a été effectuée par la combinaison d'approches rigides quasi-exhaustives et d'approches flexibles et interactives, et fait l'objet d'une publication en cours de soumission.

Dans un premier temps, les modes d'associations entre les protéines ont été explorés sans *a priori* par une méthode de *docking* rigide basse-résolution (voir *II-Méthode*). Les 40506 résultats d'amarrage calculés par le programme Attract (Zacharias, 2005) ont été groupés selon leur proximité (RMSD) et les meilleures solutions ont été isolées selon l'énergie du complexe calculée par le logiciel et le nombre de solutions regroupées ensemble (plus un *cluster* est grand, plus la solution qu'il représente est favorable). Ces meilleurs modèles ont ensuite servi de base pour déterminer les zones de contacts entre les partenaires.

Les complexes ont ensuite été reconstruits en utilisant une approche d'amarrage flexible et interactive. Ces méthodes permettent de combiner plusieurs sources d'informations pour guider et contraindre l'amarrage (potentiel électrostatique, carte de contact, ...) tout en profitant de l'expertise de l'utilisateur, en particulier pour obtenir des modèles de complexes n'interdisant pas le prolongement du filament de dystrophine au-delà des extrémités du fragment R11-15. Pour des raisons de performances de calculs, la simulation a reposée sur la réduction préalable du modèle de dystrophine en gros-grain puis à l'élaboration d'un réseau de ressorts similaire à celui utilisé lors de l'ajustement flexible. Le modèle obtenu (Figure 48A) montre une dystrophine qui « s'allonge » sur l'actine, tout en laissant les extrémités de R11 et de R15 libres vers des positions accessibles aux répétitions R10 et R16, respectivement.

En suivant la même démarche, nous avons étudié le complexe tripartite lipide / R11-15 / actine-F, tout en guidant au maximum les étapes de modélisation par des éléments structuraux, qualitatifs ou quantitatifs, issus d'approches expérimentales. La démarche mise en place visait ainsi à mieux comprendre comment le domaine ABD2 de la dystrophine pouvait



# How the central domain of dystrophin bridges F-actin to sarcolemma lipids

*Dominique Mias-Lucquin,<sup>†</sup> Raphaël Dos Santos Morais,<sup>†#</sup> Angélique Chéron,<sup>†</sup> Thomas Chenuel,<sup>†</sup> Javier Pérez, Mélanie Lagarrigue, Steve Winder, Guillaume Alviset,<sup>†</sup> Elisabeth Le Rumeur,<sup>†</sup> Mirjam Czjzek, Sophie Combet,<sup>#</sup> Jean-François Hubert,<sup>†</sup> and Olivier Delalande<sup>\*,†</sup>*

<sup>†</sup> Univ Rennes, CNRS, IGDR - UMR6290, F-35000 Rennes, France

<sup>#</sup> Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, F-91191 Gif-sur-Yvette CEDEX, France

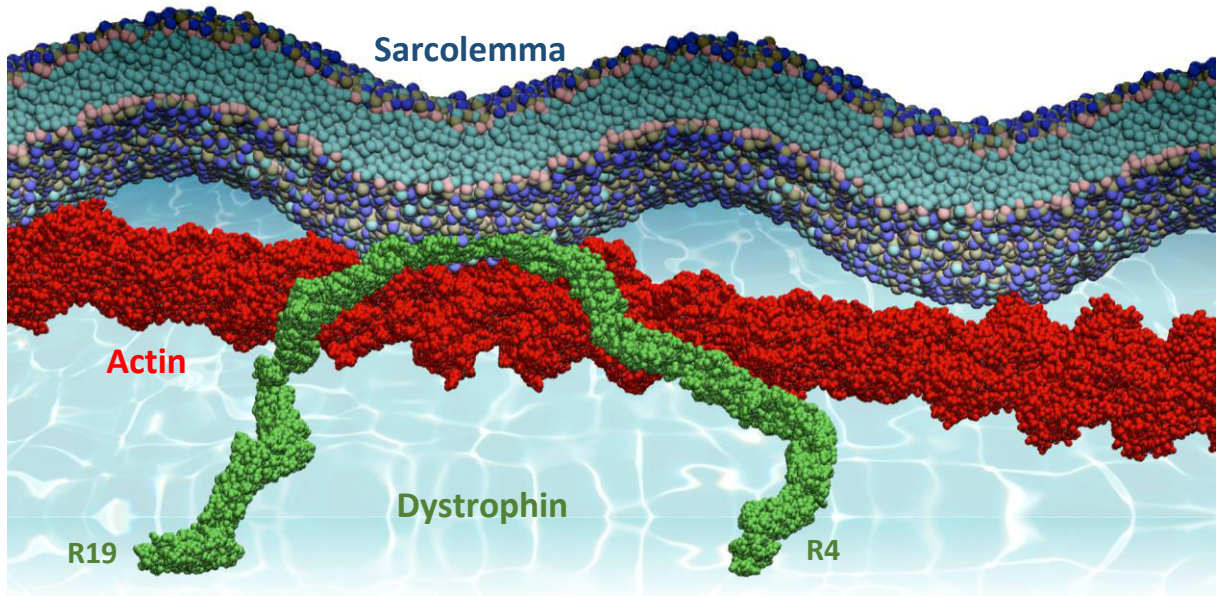
SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP 48, Saint-Aubin, F-91192 Gif-sur-Yvette, France

Inserm U1085, Protim-Plate-forme Protéomique

**KEYWORDS:** dystrophin; F-actin; bicellar membrane models; SAXS; SANS; click-chemistry

**ABSTRACT**

## GRAPHICAL ABSTRACT



## ABSTRACT

Dystrophin is a large intracellular protein providing a mechanical link between the intracellular actin cytoskeleton and the transmembrane dystroglycan complex and thus prevents sarcolemma ruptures. Previous works showed that concomitant interaction of the R11-15 fragment of the central domain of the dystrophin with actin and membrane lipid can greatly increase membrane stiffness. Based on a combination of SAXS and SANS measurements, mass spectrometry of cross-linked complexes, interactive modelling and low-resolution molecular dynamic simulations, we explored the physical properties that underlies the formation of ABD2-actin and ABD2-membrane complexes. We thus identified two sub-domains of interaction with actin, located on R11 and R15, surrounding the lipid binding domain on R12. Moreover, the relative orientation of the bicelle and actin at the contact of dystrophin, obtained independently, is compatible with a prospective tripartite actin / dystrophin / membrane lipids complex. The ABD2-actin complex was extended toward a R4-19-actin complex, opening the way to the modelisation of even bigger macromolecular complexes.

## INTRODUCTION

Dystrophin is a large intracellular protein that contributes to support the stretching or bending stresses developed within the membrane during muscle contraction and elongation. It provides a mechanical link between the intracellular actin cytoskeleton and the transmembrane dystroglycan complex and thus prevents sarcolemma ruptures (Campbell and Kahl, 1989). It is a long filamentous 427kDa protein made of four main domains separated by hinges. In muscle cells, the continuity between the internal filamentous actin cytoskeleton and the extracellular matrix is ensured by binding of dystrophin at its N-terminus to actin filaments through a tandem calponin homology domain constituting an actin binding domain (ABD1) and at its C-terminus, to  $\beta$ -dystroglycan. The latter protein spans the membrane and anchors the DGC with proteins from the extracellular matrix (Henry and Campbell, 1999). While essential regions of the dystrophin protein are located at both ends of the protein, the central domain is also thought to play a key role in regulating the membrane cytoskeleton and the extracellular matrix scaffold. This central domain of dystrophin comprises 24 spectrin-like repeats of approximately 100 residues each in which some specialised regions have been shown to interact with numerous



partners such as n-Nitric oxide synthase (Lai et al., 2009), F-actin (Amann et al., 1998; Way et al., 1992) and lipids (DeWolf et al., 1997). Previously qualified as a “rod domain”, *in silico* (Legrand et al., 2011), *in vitro* analyses (Delalande et al., 2018) and *in vivo* (Zhao et al., 2016) analyses showed that the central domain cannot be longer considered to be a passive linker between ABD1 and the C-terminus; rather, it is believed to have important function in scaffolding muscle cells.

Mutations in dystrophin are the cause of invalidating muscular dystrophies. Duchenne muscular dystrophy (DMD), a life-limiting, rapidly progressive form of muscular dystrophy is an X-linked recessive disorder affecting 1:3,500 boys in which loss of dystrophin destabilizes muscle membranes, Internal deletion mutations in the dystrophin gene (*DMD*) that preserve the amino- and carboxyl-termini of the protein but eliminate various parts of the central domain cause Becker muscular dystrophy (BMD), a relatively mild muscle disease (ref). AS seen in BMD patients, the dystrophin protein therefore can tolerate internal deletions that maintain a shortened central domain with intact amino- and carboxyl-termini regions, resulting in a more or less severe loss of muscle function. The severity of the disease is not linked to the length of the protein but to much more sophisticated and yet largely unknown functional patterns of the shortened dystrophin (Nicolas et al., 2014; Wein et al., 2015). Therapeutic strategies aim to convert DMD to the mildest BMD by skipping or deleting out-of-frame exons and restoring the expression of truncated forms of dystrophin protein (Yokota et al., 2012). In parallel, shortened forms of dystrophin, the so called mini- or microdystrophins, are developed for gene therapy (Belanto et al., 2014). However, the precise consequences of in-frame deletions on the stability and function of dystrophin are not to date predictable *a priori*, this largely due to the lack of knowledge on the molecular bases of dystrophin structure and interactions. Thus, it appears crucial to analyze first the structure to function relationship of native full length dystrophin and secondly the dystrophin protein products generated from in-frame deletions before reaching conclusions regarding their potential therapeutic effects.

Binding of dystrophin to F-actin is ensured through two domains; the ABD1 at the N terminus is well described with the calponin homology domains that can bind the f-actin on its surface (Singh et al., 2017). This is a crucial event since mutations of dystrophin that disrupt it represent the second-most common cause of DMD (Henderson et al., 2010).

On the other hand much less is known about the second actin binding domain (ABD2). Early work from the Ervasti group demonstrated that in the *mdx* mouse, a rodent model of muscular dystrophy that lacks dystrophin, the entire  $\gamma$ -actin network is poorly connected to the sarcolemma (Sonnemann et al., 2006). This suggested a role for dystrophin in maintaining the sub-sarcolemmal F-actin network in close register with the membrane thus providing mechanical stability. Later experiments demonstrated that a proteolytic product of dystrophin comprising amino acids 1416-1880, from repeats 11 to 15 (DYS R11-15) (Lai et al., 2009; Sarkis et al., 2013) or slightly enlarged to repeat 17 (Amann et al., 1998; Rybakova et al., 1996) exhibited actin binding properties constituting a second actin-binding domain (ABD2), which could serve to reinforce the association of actin with membrane. These studies established that dystrophin binds laterally along actin filaments through the synergistic effect of two distinct low affinity binding sites which may hold important clues to its mechanical function *in vivo* (Henderson et al., 2012). It was shown in addition that the dystrophin monomer is able to bind along the length of one single actin filament (Kumar et al., 1997). Therefore the presence of two low-affinity actin-binding sites in dystrophin may allow for a more elastic response of the actin–dystrophin–sarcolemma linkage to muscle mechanical stress.

In other studies, we investigated the lipid binding properties of the spectrin-like repeats of the central domain and defined that repeats 1 to 19 constitute are able to bind membrane models *in vitro* (Vié et al., 2010). We moreover showed that the sub domain comprising repeats 11 to 15 (DYS R11-15) exhibits unique lipid-binding properties, namely the ability to interact with anionic as well as zwitterionic lipids, and to adapt its interaction and organization depending on lipid packing and lipid type (Sarkis et al., 2011). From these studies, we suggested that the physiological role of the central domain of dystrophin on sarcolemma scaffolding could partly occur through modulation of lipid-protein interactions. Recent *in vivo* data are in line with our *in vitro* findings and demonstrated that at least the R1-3 and R 10-12 fragments are targeted to the sarcolemma (Zhao et al., 2016). It was in addition concluded from computing studies (Legrand et al., 2011) that DYS R11-15 is delimited by identified flexible junctions and therefore can be considered by itself as a subdomain of dystrophin.

Considering the above findings, we hypothesised that both sarcolemma and actin filaments could bind simultaneously to dystrophin through DYS R11-15, which may explain the

biological significance of this region *in vivo*. By surface shear rheology we showed that the central domain of dystrophin is able to create a bridge between F-actin and membrane lipids, dramatically modifying the membrane resistance to shear stress (Sarkis et al., 2013). DYS R11-15 thus creates a mechanical link which likely contributes *in vivo* to the protection of the sarcolemma from ruptures and to the shock absorber function of dystrophin in muscle cells.

Given that gene therapies for DMD aim to induce the production of shortened dystrophins, it appears mandatory to be able to predict their properties compared with these from the native full length molecule. This however remains dependent upon exhaustive knowledge of how wild type dystrophin interacts in muscle cells. Particularly, a full understanding at the molecular level of how the interaction of ABD2 with F-actin and membrane lipids could be ensured is required. In recent studies we obtained for the first time structural models for ABD2 following SAXS measurements and molecular modelling (Delalande et al., 2018; Molza et al., 2014). Moreover, we used these tools to build from these experimentally based models a first structural approach of complexes made of repeats R11-15 and actin or also the repeats R16-17 and nNOS (Molza et al., 2014, 2015).

We recently successfully developed the combination of SAXS, SANS contrast matching (Dos Santos Morais et al., 2017), click chemistry, MS and simulation methods to characterise for the first time lipid protein interactions at the molecular level of repeats 1-3 of dystrophin with bicelles. We show the opening of repeat 1 upon binding to anionic model membranes, that may explain the behaviour of dystrophin central domain in cells (Dos Santos *in press*).

In the present study, we used our experimentally driven molecular modelling tools to characterise the interactions of dystrophin ABD2 with F-actin and with membrane lipids. We conclude that both partners can together contribute to bridging cytoskeleton to sarcolemma.

## MATERIAL AND METHODS

**R11-15, R8-15, R11-19, production.** All multirepeat fragments were designed according to the alignment of Winder (Winder et al., 1995a).

Name	Recombinant protein sequence	MW(Da)	Number of residues	$\epsilon$ ( $M^{-1}.cm^{-1}$ )
R11-15	<i>GS FQKPA...LNFAQ</i>	60	515	45,950
R8-15	(HIS <sub>6</sub> tev) RKEAL...LNFAQ	100.2	861	92,360
R11-19	(HIS <sub>6</sub> tevthb) FQKPA...LLQEL	116.8	997	113,245

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

### Cloning and expression in *E. coli* for R11-15

Briefly, the R11-15 fragment was PCR amplified including a 6xHis tag sequence in N-terminal region and a thrombin recognition site and cloned into pET21 vector at NdeI/XhoI restriction sites. The His-tagged recombinant protein was produced in *E. coli* BL21(DE3) strain after 1mM IPTG induction for 4 hours at 37°C and purified by immobilized metal affinity chromatography (IMAC) on Ni-Sepharose column (HisTrap, GE Healthcare) according to the manufacturer 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. The concentration was determined spectrophotometrically using molar extinction coefficients at 280 nm of 45,950  $M^{-1}.cm^{-1}$ . Buffer exchange for SANS experiments was performed by three successive diafiltrations with Amicon Ultra-15 (Millipore, MWCO 10 kDa), followed by a step with a desalting column (NAP5, GE Healthcare).

### Cloning and expression in insect cells:

Extended (8 extra aa) R8-15 (R1148\_Q1973) and R11-19 (F1461-L2420) fragments were amplified by PCR from plasmid DNA pTG11025 Dp427m human ORF (a courtesy gift by Transgene SA, Strasbourg France) with designed primers including a BamHI or XhoI restriction cassette (Integrated DNA Technologies). PCR products were amplified with Q5 HiFi DNA polymerase (NEB), digested, purified (Macherey-Nagel NucleoSpin kit). The R8-15 fragment was inserted at the NcoI and XhoI sites with T4 DNA ligase (NEB) into pFastBacHT vector (Invitrogen<sup>TM</sup>, Life technologies). The R11-19 fragment was inserted at the BamHI and XhoI sites of a previously modified pFastBacHT including a thrombin recognition site in NcoI/BamHI. The constructs were verified by sequencing (Big Dye Terminator V3, Applera) and transferred into DH10BAC cells (Invitrogen<sup>TM</sup>, Life technologies) for recombination and bacmid DNA production. Sf9 insect cells were transfected with 1  $\mu$ g bacmid DNA thanks to

Cellfectin-II reagent according to BAC-to-BAC instruction manual (Invitrogen™, Life technologies). Resulting baculovirus were amplified up to 3 passages.

Recombinant proteins were produced in Sf9 cells infected with baculovirus for 3 days. Cells were lysed for one hour at room temperature in phosphate buffer saline (sodium phosphate 25mM pH8, NaCl 300 mM) containing protease and phosphatase inhibitors (Pierce), 50 µL/mL insect PopCulture® reagent (Novagen), 10 U/mL benzonase (Novagen) and 5 mM imidazole. Lysates were clarified by centrifugation at 15000g, 30 min, 4°C; loaded onto a HisTrap excel™ Ni-NTA column (GE Healthcare, Äkta start instrument), washed with PBS pH8 + 5mM imidazole. Purified proteins were eluted in phosphate buffer saline pH8 containing 50mM imidazole then subjected to gel filtration onto a HiPrep 26/60 Sephacryl S-200HR column (GE Healthcare, Äkta FPLC instrument) and stored in Tris-HCl 20mM, NaCl 150mM, EDTA 0.1mM pH7.5 (TNE buffer) as described above. Quality control of all the purified proteins was assessed by circular dichroism with a JASCO J-810 (Nantes, France) spectropolarimeter with protein concentration of about 1 µM against the TNE buffer. Spectra were acquired in the range 200 – 250 nm at 20°C with a path length of 0.2 cm. The percentage of  $\alpha$ -helix was obtained using a 100%  $\alpha$ -helix value of -36000 deg .cm<sup>2</sup>.dmol<sup>-1</sup> at 222 nm as previously described (Legardinier JMB 2009).

**Bicelle preparation.** Bicelles were prepared as described in a previous work (Dos Santos-Morais 2017). The 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 all purchased from Avanti Polar Lipids and used without any further purification. D<sub>2</sub>O, 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<sub>2</sub>, thawed 10 min at 40°C, vigorously shaken for 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.

**Intrinsic tryptophan (Trp) fluorescence.** Fluorescence measurements were carried out on a Fluorolog spectrofluorometer (Jobin-Yvon). Tryptophan fluorescence emission spectra were recorded in low volume quartz Hellma cells (120  $\mu$ L) 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  $\mu$ M and the total lipid concentration was 50 mM (~150  $\mu$ M of bicellar objects). The measurement temperature was 26°C.

**Microscale thermophoresis (MST).** Bicelles were labeled with 1  $\mu$ M 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  $\mu$ M of bicellar objects) at  $q_{\text{eff}} = 1.3$  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 R11-15 from 338 to 4.5  $\mu$ M. The 16 solutions were then loaded into Monolith NT premium capillaries (NanoTemper) and thermophoresis was measured with a Monolith NT.115 instrument (NanoTemper). 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 (NanoTemper). From the fit of the MST data, the concentration of the binding site ranges from 50 to 100  $\mu$ M. 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  $q_{\text{eff}}$ . Indeed, working at 15 mM of lipids tends to be close to the  $[\text{DHPC}]_{\text{free}}$  that could drastically affect  $q_{\text{eff}}$  even for a small misestimate of the total lipid concentration. The measurement temperature was 26°C.

### **Actin polymerization and sedimentation assays**

Globular actin (G-actin) was prepared from rabbit muscle, as described previously (Winder et al., 1995b). Cosedimentation assays were done as described in (sarkis et al faseb). Briefly, monomeric G-actin stored in G buffer (2 mM Tris, pH 8.0; 0.2 mM ATP; 0.5 mM dithiothreitol;

0.2 mM CaCl<sub>2</sub>; and 1 mM NaN<sub>3</sub>) was diluted in polymerization buffer (F buffer; final composition: 20 mM HEPES, pH 7.5; 150 mM NaCl; 2 mM ATP; 50 mM KCl; 2 mM MgCl<sub>2</sub>; and 1 mM dithiothreitol) and allowed to polymerize at room temperature for 1 h. Actin polymerization was assessed by light scattering at 500 nm. For sedimentation assays, 50-100- $\mu$ l samples were prepared in F buffer in ultracentrifuge tubes. The final concentrations of actin were 5  $\mu$ M and 0 to 150  $\mu$ M for actin and for DYS R11–15 respectively. Following 30 min of incubation, assays were centrifuged for 30 min at 100,000 g at 20°C. Supernatants and pellets were analyzed by SDS-PAGE (Figure SM2). Gels were stained with Coomassie blue and the amounts of proteins quantified by imageJ software.

**Click-chemistry, chemical cross-links and liquid chromatography coupled with mass spectrometry (LC-MS/MS). R11-15 - bicelles :** 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 (Haberkant and Holthuis, 2014). The modified bicelles were mixed with R11-15 to reach a final concentration of 50 mM of lipids and 15  $\mu$ M 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 SM2A), just above the band corresponding to R11-15, were cut into small pieces. R-11-15 – F actin: F-actin and R1-15 at 5  $\mu$ M concentrations were incubated in polymerisation buffer for 30 min at 20°C and then EDC ((1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride-thermofisher scientific) was added at a concentration of x mM for x min at 20°C. Crosslinked proteins were analysed by SDS PAGE stained by Coomassie blue and the bands corresponding to one actin monomer-one R11-15 monomer complexes were excised and cut into small pieces (gel fig S X+1). For both lipid-R11-15 and actin R11-15 complexes, proteins contained in the gel pieces were reduced, alkylated, and then digested with trypsin (Promega), and the resulting peptides were extracted as previously described (Lavigne et al., 2012). The peptide extract was loaded (10  $\mu$ L) on a nano-HPLC system (LC Packings Ultimate 3000, Dionex) equipped with a trapping precolumn (5 mm  $\times$  300  $\mu$ m id, 300 Å pore size, Pepmap C18, 5  $\mu$ m, ThermoScientific) and an analytical column (15 cm  $\times$  75  $\mu$ m id, 300 Å pore size, Pepmap C18, 5  $\mu$ m, ThermoScientific). Reversed-phase separation was performed with the same gradient as described previously (Lavigne et al., 2012). 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. MS data were saved in the RAW file format with Xcalibur 2.0.7. For the lipid-R11-15 complex, the RAW file was 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 R11-15 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<sub>39</sub>H<sub>72</sub>NO<sub>8</sub>P, 713.4996 Da) was manually defined in the Mascot configuration editor with a neutral loss of phosphocholine (C<sub>5</sub>H<sub>14</sub>NO<sub>4</sub>P, 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) carapito et al. For the actin-R11-15 complex, the RAW file was converted in the Mascot generic file format using Mascot Distiller and imported in the StavroX 3.6.6.5 (PMID: 22038510) software used to identify EDC cross-linked peptides. Cystein carbamidomethylation was selected as static modification whereas methionine oxidation was selected as variable modification. Precursor and fragment tolerances were respectively set to 5 ppm and 0.5 Da. K and R (both blocked by P) were defined as protease sites. Reported results corresponded to a false discovery rate lower than 0.05 (Figure SM2B).

### **Small-angle X-ray scattering (SAXS) experiments.**

SAXS experiments were performed as described previously (Delalande 2018). Measurements were conducted at the French synchrotron SOLEIL (St. Aubin, France) on the SWING beamline. All experiments were performed at 15°C. The scattering vector is defined as  $q = 4\pi/\lambda \sin\theta$ , where  $2\theta$  is the scattering angle. Data were collected in a q-range of 0.005–0.5 Å<sup>-1</sup>. Data were recorded using an AVIEX170170 CCD detector at the distance of 1.807 to 1.845 m ( $\lambda = 1.033$  Å). For the constructs R8-15 and R11-19, a stock solution of each fragment was prepared at a final concentration between 5 and 8 mg/ml. A volume between 50 to 70 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 elution buffer consisted of TNE buffer adjusted to 300mM NaCl, supplemented with 2% glycerol. 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 and were used for further characterization. Data reduction to absolute units and frame averaging and subtraction were performed in FOXTROT.

**SAXS data analysis.** As described in ref Delalande *et al.* (2018) briefly, all subsequent data processing and analysis were conducted with the Scatter (<http://www.bioisis.net/>), SASFit (Breßler *et al.*, 2015) or PRIMUS software (Konarev *et al.*, 2003) and other programs of the ATSAS suite (Franke *et al.*, 2017). Guinier plots assess that all samples were not aggregated. The distance distribution function  $P(r)$  and the maximum particle diameter ( $D_{MAX}$ ) were calculated using the GNOM program. The overall *ab initio* models of the protein fragment R11-15 were obtained from the SAXS experimental data presented in Delalande *et al.*, 2018 (Delalande *et al.*, 2018). *Ab initio* models for others fragments (R8-15 and R11-19) were obtained using the GASBOR program, using 50 harmonics, the scattering profiles were fitted to a  $q_{max} = 0.45 \text{ \AA}^{-1}$  for all samples. Following a similar protocol to the one described in Delalande *et al.*, 2018. One hundred independent GASBOR computations were performed for each scattering profile, to factor the decrease in envelope convergence associated with the increased fragment size. For each fragment, the three best GASBOR computations, with the smallest  $\chi^2$ , were conserved and converted to volume grid constraints (molecular shape) to guide the interactive flexible fitting simulations, as described in previous work (Molza *et al.*, 2014). Beforehand simulation, the homologies model and the molecular shape are aligned along their longest axis. The model is then rotated 10 times around the long axis by  $36^\circ$  steps, followed by a  $180^\circ$  head-to-tail rotation to exchange N- and C-terminal end, before performing again 10 rotations. The 60 models (3 envelopes, 10 rotations, 2 directions) were afterward adjusted. All final models were refined by standard energy minimization (Molza *et al.*, 2014). The best models were selected upon evaluation of the final structural models was performed using PROCHECK (Laskowski *et al.*, 1993) and DSSP (Kabsch and Sander, 1983; Touw *et al.*, 2015) and by validation after calculating their theoretical SAXS curves with the CRY SOL program (Franke *et al.*, 2017).

**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<sup>-1</sup> for the largest one, where  $Q = \frac{4\pi \sin \theta}{\lambda}$  is the momentum transfer,  $\lambda$  is the wavelength, and  $2\theta$  is the scattering angle. All measurements were performed in 1 mm thickness Hellma  $QS$  cells, and the intensities obtained are in absolute units (cm<sup>-1</sup>). 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. The protein concentration was 103 μM (4.2 g/L).

**SANS data analysis.** SANS data were analyzed like SAXS data with the ATSAS suite and the Scatter software, following guidelines unless otherwise indicated. Bicelle data were fitted with the “core-shell cylinder” form factor models. Twenty *ab initio* models were generated with the GASBOR software with the pair-distribution function  $P(r)$  obtained from GNOM analysis and by considering the data up to  $Q = 0.25$  Å<sup>-1</sup>. 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.

**Low-resolution docking.** Rigid docking was performed using ATTRACT software (Schneider, 2012) with the previously selected 13 models of R11-15 from SAXS-driven flexible fitting. Given the large size of the 12-mer actin model (3G37, (Murakami et al., 2010)), it was considered as the fixed receptor, and R11-15 the ligand. The use of 13 models allows to simulate the ligand modulation that normally occur during docking. The models were reduced according to the Zacharias force field, with one amino acid being described by one to three coarse grains (Schneider and Zacharias, 2012). The R11-15 starting position was defined using a density grid of 10 Å around the two central F-actin monomers, corresponding to 80 starting positions. For each model, 20 640 theoretical complexes were calculated, corresponding to 80 translations time 258 rotations to modulate the relative orientation of ligand protein (dystrophin) towards

the receptor (F-actin). The rigid docking results were analyzed through an energy evaluation of the theoretical complexes and by clustering procedure using a root mean square deviation of 10 Å according to the ligand final position. The best theoretical complexes *are defined as the* most populated clusters with the lowest ATTRACT energy for the complex. The contact frequencies between the two proteins were calculated with a cut-off distance of 5 Å. Peptides identified by MS analysis of the cross-linked complexes were used to filter theoretical complexes with a cut-off distance between cross-linked fragment lower than 5 Å. Interactive semi-flexible docking were performed thanks to BioSpring program and following the protocol described previously (Molza et al., 2014).

**Coarse-grained molecular dynamics (CG-MD) simulations.** CG-MD simulations were run using the GROMACS 5.0 program with the Martini coarse-grained force field. Two protein/lipid systems were simulated in a similar way to what was done previously for the R1-3 dystrophin fragment with either zwitterionic bicelle or anionic bicelle. The R11-15 model is the one validated from our previous study and accessible through the online database SASBDB at <http://www.sasbdb.org/data/SASDB63/>. All CG-MDs were performed at a constant temperature of 303 K and under NPT conditions for the production of final CG-MD trajectories of 20  $\mu$ s. Analysis was achieved with GROMACS 5.0 tools.

## RESULTS

***Dystrophin actin-binding domain 2 in solution is a topologically well-structured and organized region.*** In the aim to explore the association modes of the dystrophin actin binding domain 2 with F-actin in a more detailed approach than the one we previously deployed [Molza 2014], we decided to improve the analysis of our SAXS data acquired for the R11-15 dystrophin fragment in solution (Figure 1). *De facto*, the use of multiple *ab initio* models (Figure 1A) to enlarge the conformational space exploration of the R11-15 fragment in addition to a deeper and systematic protocol for flexible fitting simulations (Figure 1B) allowed us to produce new high quality models of dystrophin ABD2. The production of multiple structural models has been thought to propose diversity in the ligand geometry throughout rigid docking simulations, a point that has shown to be a limit in our previous study (Molza et al., 2014). Among the best models we obtained over 60 simulated structures, we selected the 13 firstly ranked R11-15 models (Figure 1C) verifying the four quality criteria presented in the Material and Methods section. After clustering, the four best models are defining two groups that are showing a

remarkable convergence in terms of global orientation of the fragment towards topological position of kinks in the filament. This is probably partly due to the fact that these models derived from an identical orientation in a unique molecular envelope (top in Figure 1A); however, this was not so predictable from raw data analysis as over the selected models (total of 13), but to the exception of two additional similar models, all others are corresponding to flexible fitting results either coming from the other orientation in the same envelope or to other envelopes. The position we already identified for the major kinks present in the ABD2 filament are confirmed, notably between R12 and R13 and between R14 and R15 repeats. However, the relative position of the R11-12 tandem towards the rest of the fragment, and thus, the kink defining the R12-13 linker organization, seems highly modular.

***Actin-binding domain 2 is contacting F-actin with its both extremities.*** Docking simulations were achieved by using all (13) selected ABD2 SAXS-based models as the ligand and with a F-actin crystal structure as the receptor (Figure SM1A). A global analysis of the contact frequencies of dystrophin on the F-actin indicates that the ABD2 fragment could interact with a single F-actin face (Figure 2A) and in a preferential manner with subdomains 1 and 3 of a G-actin unit (projection of all contacts on a single actin monomer), even if subdomains 2 and 4 are also participating to the association of both filaments to a lesser degree (Figure SM1B). The highest contact frequencies that were observed for the dystrophin ABD2 interacting with F-actin essentially concern the repeats R11 and R15 (Figure SM1C). An experimental approach was designed to support theoretical mapping. The R11-15 protein was incubated with F actin, cross-linked by EDC and the peptides in contact were identified by mass spectrometry. From these data, it clearly appears that region of actin in close contact with dystrophin ABD2 is principally belonging to the subdomain 1 as the two peptides isolated by mass spectroscopy with a high confidence are delimited by Ala21 and Arg30 or Val98 and Lys115, respectively (Figure SM2B). Corresponding results defining the contact mapping on the dystrophin ABD2 fragment characterize multiple regions to be involved in the interaction with F-actin but interestingly all are included in repeats R11 and R15. These fragments are delimited from Phe1461 to Arg1470, from Ser1517 to Arg1527, from Lys1533 to Asn1537 and from Leu1550 to Lys1568 in repeat R11 and from Lys1922 to Arg1930 in repeat R15.

In a second analysis step of the simulations, docking results were filtered and only the ABD2–F-actin theoretical complexes that were in accordance with the MS results were further analysed (Figure 2A, Figure SM1D and E). This allowed us to evaluate the great prediction power of the non-filtered docking results and to identify false-positive regions firstly characterized as interacting spots, notably on actin. Finally, it allowed confirming that dystrophin should contact F-actin at a site far from the interface defined by the contact region of subdomains 3 and 4 of different actin subunits. After an energy-based ranking and geometrical clustering of R11-15–F-actin theoretical complexes, the best structural models were isolated (Figure 2B) and it appeared that no model could verify both MS contact measured for both repeats R11 and R15. As both dystrophin and actin filaments relative orientation seemed unsuitable for double contact verification, the approach consisting in the multiplication of dystrophin (ligand) conformers in the rigid docking simulation revealed to be not sufficient. Nevertheless, rigid docking simulations have provided high quality models for the positioning of each R11 and R15 repeats relative to the F-actin structure, geometries that will serve as starting structures for the models improvement.

***Actin-binding domain 2 is contacting F-actin on two different faces of the microfilament.***

Interactive flexible docking is a fast and efficient technique to improve docking models as we already discussed earlier (Molza et al., 2014). Through interactive simulations, we now explored the multiple possibilities in the positioning of the repeats R11 and R15 relatively to each other, both in terms of distance (number of actin units separating each contacting monomer) and topology (same face of the microfilament or not), by fixing either the position of the repeat R11 or of the repeat R15 according to the best models from rigid docking (principle is shown in Figure 3A). From the refined models that were obtained by this approach, we selected ABD2–F-actin models on their internal and binding energies or on the secondary structure conservation of dystrophin. Among the best models, all derived from the simulations in which repeat R11 was restrained and the rest of the dystrophin ABD2 fragment was driven to verify the contacts experimentally mapped on the repeat R15 (Figure 3A and B). By closely characterizing the main properties of the molecular surfaces involved in the dystrophin ABD2 – F-actin interaction, it appears that electrostatics should play a key role in the association (Figure 3C). However, some hydrophobic patches could also be crucial for the later stabilization of the complex, and more specifically a topological region of F-actin defining the side of a groove that accommodate the repeat R15 of dystrophin (Figure 3D). The correlation with

experimental mapping results and the contact analysis of the models produced from the six best interactive simulations revealed that repeat R11 and R15 are ideally located on opposite faces of F-actin (Figure 3E). Indeed, simulations 1 and 2 shown on the contact analysis and bearing the best evaluation criteria (see Material and Methods) are corresponding to models with repeats R11 and R15 distant from 5 actin subunits. Other models of reliable quality correspond to similar simulations with the two dystrophin ABD2 repeats distant from 3 actin subunits (experiments 3 to 6, depending on the level of wrapping of the dystrophin around F-actin filament). Consequently, these results support that dystrophin filament should not follow the F-actin helicity through its double contact with opposite faces of the microfilament.

***Lipids interaction with dystrophin slightly alters its structural organization but does not overlap with F-actin contact regions.*** In a similar approach to what was achieved previously for the identification of putative conformation changes of the protein upon its binding to the lipids (Dos Santos-Morais 2018), small-angle neutron scattering experiments were performed on dystrophin ABD2 fragment in the presence of zwitterionic or anionic bicelles (Figure 4A). Thanks to the contrast matching approach (Dos Santos Morais et al., 2017), SANS measurements show that the interaction of dystrophin ABD2 does alter the fragment ternary structure upon binding to anionic bicelles (increase of 8 to 18% of the fragment length towards free fragment) whereas no change is observed in the case of dystrophin ABD2 interaction with zwitterionic bicelles. This observation is of a high relevance given that the inner leaflet of the sarcolemma is a layer rich in anionic phospholipids. Moreover, the best *ab initio* models computed from the SANS experiments for the dystrophin ABD2 fragment are showing a very high degree of similarity by comparison to the models obtained from the SAXS experiments (Figure 1A and Figure 4A). Finally, several experiments were also achieved to better describe this interaction of dystrophin ABD2 with bicelles as was done in our previous works for the R1-3 subdomain of dystrophin (Dos Santos-Morais 2018). First, the protein/bicelle interactions were highlighted by Trp intrinsic fluorescence measurements in the presence of zwitterionic or anionic bicelles that increases slightly but significantly (Figures SM3A), confirming the interaction but without an important change in the Trp environment. Then, we used circular dichroism (CD) to probe potential secondary structure modifications of the dystrophin ABD2 during its interaction with both types of bicelles (Figures SM3B). Our first observation is that the CD spectra of the ABD2 fragment alone is typical of an overall  $\alpha$ -helical folding with the presence of two minima at 222 and 208 nm, supporting that the recombinant protein is properly

folded. Moreover, the  $\theta_{222}/\theta_{208}$  ratio is close to or above the unity, which is an evidence of a coiled-coil ternary structure. The CD spectra of the ABD2 alone or in the presence of both types of bicelles are very similar suggesting that the secondary structure of R11-15 is maintained during binding to bicelles. Finally, we used microscale thermophoresis (MST) to determine the dissociation constant ( $K_d$ ) of the complexes. No  $K_d$  could be determined in the protein concentration range that we used for the system formed by dystrophin ABD2 with zwitterionic bicelle as the same parameter was estimated to be approximately  $K_d \sim 10 \mu\text{M}$  (Figure SM3C) for the system formed by dystrophin ABD2 with anionic bicelle.

To explore molecular mechanism that could lead to such structural features, low-resolution simulation have been used to model both bicellar systems with molecular dynamics techniques. However, even if MD simulations do not bring clear evidences about the mechanism leading to an elongation of dystrophin ABD2 when bound to an anionic bicelle and not to a zwitterionic bicelle, some interesting points have to be emphasized. First, during MD simulations, the dystrophin ABD2 is lately recruited by zwitterionic bicelle whereas it is quickly and spontaneously bound to anionic bicelle (Figure SM4A and B). This contrasted behaviour could be in a way at the origin of the differences observed experimentally by MST. More interestingly, we observe a very different relative position of the dystrophin ABD2 fragment towards each bicelle type (Figure SM4C and D). As both interact with the core part of the disk, a region with an increase of packing density – and thus more hydrophobic, ABD2 is lying R11 to R13 repeats all along the anionic bicelle side but is only contacting zwitterionic bicelle through repeat R15. This last feature observed in our MD simulations could also be related to the conformational change observed by SANS, as the more expanded contact of ABD2 with anionic bicelle could diminish the kinks at linker region and thus result in a more elongated structure.

In addition to all these approaches, and for a more accurate mapping of the interaction between a dystrophin ABD2 fragment and an anionic bicelle, we reproduced the approach done previously for another central domain fragment, and combined click-chemistry techniques and mass spectroscopy (Dos Santos-Morais 2018). The identified peptide in contact with the bicelle *in vitro* belongs to the C-terminal end of the repeat R12 of the dystrophin (Figure 4B). This result is well in line with our previous work based on the sole trypsin digestion of ABD2 bound

to lipid vesicles (Sarkis et al., 2011). We should note that this R12 region is deeply contacting the membrane model in our MD simulations of the dystrophin ABD2 in the presence of an anionic bicelle (Figure SM4D).

***Central domain structure of the dystrophin can accommodate interaction with both F-actin and sarcolemma lipids.*** On the contrary of dystrophin ABD1 that is clearly an isolated subdomain, the dystrophin ABD2 is topologically dependant of a large part of the central domain. Our previous structural study of the central domain allowed us to describe a tortuous filament presenting kinks at specific linkers, the region bridging two dystrophin repeats. Some limits of this work were concerning the non-exhaustive characterization with the missing description of the R10 repeat – and consequently of R9-10 and R10-R11 linkers – and of the R15-16 linker. The present study that focuses on the dystrophin ABD2 interactions with partners could be suffering from these missing structural data at the R11-15 subdomain boundaries. In order to supplement our structural study of the central domain and to prevent artefacts that could invalidate the results we present here, we studied R8-15 and R11-19 dystrophin fragments by using small-angle X-ray scattering approach combined to interactive flexible fitting technique (Delalande et al., 2018; Molza et al., 2014). Global parameters measured from SAXS experiments indicate that the dystrophin R8-15 fragment is more elongated in solution than the R11-19 fragment, despite the supplemental repeat of this last region (radius of gyration of 100Å versus 91Å, respectively). This feature is well illustrated by the best *ab initio* models generated for each of the fragments (Figure 4A and B). High quality  $\chi^2$  values of their related theoretical SAXS curves (2.5 to 2.6 for R8-15 and 0.8 to 0.9 for R11-19) and the great convergence of their shapes (NSD<3.4Å for R8-15 and NSD<3.8Å for R11-19) support the relevance of using these *ab initio* models as a volume guide for flexible fitting experiments (Figure SM6A and B). The SAXS-based atomic models obtained for these two new central domain fragments were finally used to extend the R11-15 fragment in complex with F-actin to a whole R4-19 region. By maximizing the overlapping regions of the newly characterized fragments (R8-15 and R11-19) and the fragments previously deposited in the SASBDB (R4-9, R11-15 and R16-19, Figure 5A), we succeeded to propose a full domain of a R4-19 dystrophin fragment in great accordance in terms of distribution of the kinks located at the linker regions (Figure 5B and C). For instance, R8-9 and R11-12 linkers seem to bear only a moderate kink of their helical structure whereas R14-15 and R17-18 linkers should present



higher kinks. Nevertheless, we should nuance the kink presented between R15 and R16 repeats that is mainly imposed by the SAXS-derived model of R11-19 fragment, and whereas the structural behaviour of the mini-Hinge of 21 amino acids taking place of R15-16 linker could be revealed at least as highly flexible. Positioning of the complete R4-19 structural model according to the theoretical model of ABD2–F-actin has been achieved without steric clash and the position of R4-10 and R16-19 fragments. This integrative structural model revealed to be perfectly suitable with the experimental restraints related to the dystrophin structure and the ABD2 interaction patterns with F-actin and lipids (Figure 5D).

## DISCUSSION

***Dystrophin actin-binding domain 2 is defined by two interaction subdomains localized in repeat R11 and repeat R15.*** Until now, the boundaries of the dystrophin second actin binding domain were still under debate as authors firstly proposed an interacting region ranging from repeat R11 to repeat R17 (Amann et al., 1998; Rybakova et al., 1996), whereas later studies (Sarkis et al., 2013) shown that R11-15 was sufficient for the binding of dystrophin central domain to F-actin. We here precisely characterized the F-actin interacting regions and confirmed their localization to be confined to a large part of the repeat R11 and to a short region within the repeat R15. On the actin side, we described for the first time the subdomains 1 and 3 to be the main regions interacting with the dystrophin central domain. These results seem to be pretty coherent with the molecular surface potentials computed from structure and models of both protein partners. Indeed, F-actin is electrostatically highly negative and, even less so, the dystrophin repeats either as predicted from our systematic potential analysis proposed earlier (Legrand et al., 2011). But it is outstanding that R11 and R15 are two of the most zwitterionic repeats of the R11-15 dystrophin fragment, and this could make easier the association of both filaments. In addition, experimental mapping procedure, by nature, will provide peptides of both partners linked in most of the cases by salt-bridge interactions, as the chemical cross-linker is bridging acidic and basic residues. This probably emphasize artificially the importance of electrostatics in our case, therefore hydrophobic regions seem also play an essential role in the association of the dystrophin filament with the F-actin. Finally, we propose that dystrophin would use the two repeats R11 and R15 to contact two different faces of the actin microfilament with a distance of +5 actin subunits between them. Measurement of average distances in the models between R11 and R15 repeats and between two actin subunits distant

from +5 ( $\sim 170\text{\AA}$  and  $\sim 140\text{\AA}$ , respectively) seem to strengthen this hypothesis as other positions computed for dystrophin towards F-actin have shown to be irrelevant in terms of topological organization. It is also noticeable that stoichiometry assays we performed (1:3 for ABD2:G-actin) strongly support these results if we consider that dystrophin can fully decorate the F-actin on its two faces.

***Dystrophin actin-binding domain 2 is structurally allowing a non-overlapping interaction with F-actin and phospholipids.*** The dystrophin ABD2 subdomain can thus be surely defined as ranging from R11 to end to R15, but without any certitude concerning a putative extended and continuous contact of dystrophin central domain all along F-actin. Effectively, we characterized accurately the dystrophin repeat R12 as the major binding region with anionic bicellar systems, a result highly consistant with those results obtained recently by Zhao et al. that showed the R10-12 dystrophin fragment to be a lipid-binding domain *in vivo* (Zhao et al., 2016). It is noticeable that from our MD simulations, it appears that dystrophin preferably binds to the core part of the bicelle, a part rich in packing defects. This observation highlights that dystrophin interaction with phospholipids should be driven by hydrophobic interactions. This hypothesis is also supported by experimental mapping results as the position of the clickable chemical group on the UV-photoactivable lipids – ie at the very end of their aliphatic chain – is limiting dystrophin-lipid crosslink to the sole case of a close proximity between protein and packing defects. Indeed, the pacFa is a long-chain phospholipid that should be localized in the bilayer part of the bicelle and the end of its aliphatic chain should only be accessible to a protein on the core part of the bicelle. However, the specific behaviour of dystrophin towards anionic bicelles but not towards zwitterionic bicelles indicate that electrostatics or Hydrogen bonding could also play a role, even minor, in the recruitment or in the stabilization of dystrophin-sarcolemma interaction. The structural model we propose for the association of dystrophin ABD2 to F-actin remains accessible to the identified lipid-binding subdomain that we identified at the very end of repeat R12. It is even more remarkable that, in this study, we used a dystrophin ABD2 limited to the R11-15 fragment for the mapping of the interaction with F-actin. As a matter of fact, the present system does not take in account that the actin binding activity has been shown to be reduced when the C-terminal third of dystrophin is deleted. Indeed , this last

part of the dystrophin would affect allosterically the actin binding (Henderson et al., 2012). Nevertheless, the binding of the R11-15 dystrophin fragment to F-actin is with a  $K_d$  of  $10\mu\text{M}$  as we measured it by co-sedimentation assays. This value was also reported and as it was reported previously (Amann et al., 1998). Interestingly, we measured by microscale thermophoresis a similar  $K_d$  of  $10\mu\text{M}$  for the interaction of the same R11-15 fragment with anionic bicelles. We could hypothesize that a dynamical competition in the binding of one partner or the other by the dystrophin ABD2 could result from the similarity of both  $K_d$  values. The dystrophin ABD2 would show therefore a comparable affinity for two molecular representatives of its main cellular partners, namely the sarcolemma and cytoskeleton. This observation could be of a great importance if verified in a cellular context.

## **REFERENCES**

- Amann, K.J., Renley, B.A., and Ervasti, J.M. (1998). A Cluster of Basic Repeats in the Dystrophin Rod Domain Binds F-actin through an Electrostatic Interaction. *J. Biol. Chem.* *273*, 28419–28423.
- Belanto, J.J., Mader, T.L., Eckhoff, M.D., Strandjord, D.M., Banks, G.B., Gardner, M.K., Lowe, D.A., and Ervasti, J.M. (2014). Microtubule binding distinguishes dystrophin from utrophin. *Proc Natl Acad Sci U S A* *111*, 5723–5728.
- Breßler, I., Kohlbrecher, J., and Thünemann, A.F. (2015). SASfit: a tool for small-angle scattering data analysis using a library of analytical expressions. *J Appl Cryst, J Appl Crystallogr* *48*, 1587–1598.
- Campbell, K.P., and Kahl, S.D. (1989). Association of dystrophin and an integral membrane glycoprotein. *Nature* *338*, 259–262.
- Delalande, O., Molza, A.-E., Dos Santos-Morais, R., Chéron, A., Pollet, É., Raguene-Nicol, C., Tascon, C., Giudice, E., Guilbaud, M., Nicolas, A., et al. (2018). Dystrophin's central domain forms a complex filament that becomes disorganized by in-frame deletions. *J. Biol. Chem.*
- DeWolf, C., McCauley, P., Sikorski, A.F., Winlove, C.P., Bailey, A.I., Kahana, E., Pinder, J.C., and Gratzer, W.B. (1997). Interaction of dystrophin fragments with model membranes. *Biophysical Journal* *72*, 2599–2604.
- Dos Santos Morais, R., Delalande, O., Pérez, J., Mouret, L., Bondon, A., Martel, A., Appavou, M.-S., Le Rumeur, E., Hubert, J.-F., and Combet, S. (2017). Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS. *Langmuir* *33*, 6572–6580.
- 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. (2017). ATSAS 2.8: a

- comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Cryst, J Appl Crystallogr* 50, 1212–1225.
- Haberkant, P., and Holthuis, J.C.M. (2014). Fat & fabulous: bifunctional lipids in the spotlight. *Biochim. Biophys. Acta* 1841, 1022–1030.
- Henderson, D.M., Lee, A., and Ervasti, J.M. (2010). Disease-causing missense mutations in actin binding domain 1 of dystrophin induce thermodynamic instability and protein aggregation. *PNAS* 107, 9632–9637.
- Henderson, D.M., Lin, A.Y., Thomas, D.D., and Ervasti, J.M. (2012). The Carboxy-Terminal Third Of Dystrophin Enhances Actin Binding Activity. *J Mol Biol* 416, 414–424.
- Henry, M.D., and Campbell, K.P. (1999). Dystroglycan inside and out. *Current Opinion in Cell Biology* 11, 602–607.
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst, J Appl Crystallogr* 36, 1277–1282.
- Kumar, A., Crawford, K., Close, L., Madison, M., Lorenz, J., Doetschman, T., Pawlowski, S., Duffy, J., Neumann, J., Robbins, J., et al. (1997). Rescue of cardiac alpha-actin-deficient mice by enteric smooth muscle gamma-actin. *Proc. Natl. Acad. Sci. U.S.A.* 94, 4406–4411.
- 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. (2009). 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* 119, 624–635.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst, J Appl Crystallogr* 26, 283–291.
- Lavigne, R., Becker, E., Liu, Y., Evrard, B., Lardenois, A., Primig, M., and Pineau, C. (2012). Direct iterative protein profiling (DIPP) - an innovative method for large-scale protein detection applied to budding yeast mitosis. *Mol. Cell Proteomics* 11, M111.012682.
- Legrand, B., Giudice, E., Nicolas, A., Delalande, O., and Le Rumeur, E. (2011). Computational Study of the Human Dystrophin Repeats: Interaction Properties and Molecular Dynamics. *PLoS ONE* 6, e23819.
- Molza, A.-E., Férey, N., Czjzek, M., Le Rumeur, E., Hubert, J.-F., Tek, A., Laurent, B., Baaden, M., and Delalande, O. (2014). Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly. *Faraday Discuss.* 169, 45–62.
- Molza, A.-E., Mangat, K., Le Rumeur, E., Hubert, J.-F., Menhart, N., and Delalande, O. (2015). Structural Basis of Neuronal Nitric-oxide Synthase Interaction with Dystrophin Repeats 16 and 17. *J Biol Chem* 290, 29531–29541.

- Murakami, K., Yasunaga, T., Noguchi, T.Q.P., Gomibuchi, Y., Ngo, K.X., Uyeda, T.Q.P., and Wakabayashi, T. (2010). Structural Basis for Actin Assembly, Activation of ATP Hydrolysis, and Delayed Phosphate Release. *Cell* *143*, 275–287.
- 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. (2014). Becker muscular dystrophy severity is linked to the structure of dystrophin. *Hum. Mol. Genet.*
- Rybakova, I.N., Amann, K.J., and Ervasti, J.M. (1996). A new model for the interaction of dystrophin with F-actin. *J. Cell Biol.* *135*, 661–672.
- Sarkis, J., Hubert, J.-F., Legrand, B., Robert, E., Chéron, A., Jardin, J., Hitti, E., Le Rumeur, E., and Vié, V. (2011). Spectrin-like repeats 11-15 of human dystrophin show adaptations to a lipidic environment. *J. Biol. Chem.* *286*, 30481–30491.
- Sarkis, J., Vié, V., Winder, S.J., Renault, A., Rumeur, E.L., and Hubert, J.-F. (2013). Resisting sarcolemmal rupture: dystrophin repeats increase membrane-actin stiffness. *FASEB J* *27*, 359–367.
- Schneider, S., and Zacharias, M. (2012). Combining geometric pocket detection and desolvation properties to detect putative ligand binding sites on proteins. *J. Struct. Biol.* *180*, 546–550.
- Singh, S.M., Bandi, S., and Mallela, K.M.G. (2017). The N-Terminal Flanking Region Modulates the Actin Binding Affinity of the Utrophin Tandem Calponin-Homology Domain. *Biochemistry* *56*, 2627–2636.
- Sonnemann, K.J., Fitzsimons, D.P., Patel, J.R., Liu, Y., Schneider, M.F., Moss, R.L., and Ervasti, J.M. (2006). Cytoplasmic gamma-actin is not required for skeletal muscle development but its absence leads to a progressive myopathy. *Dev. Cell* *11*, 387–397.
- Touw, W.G., Baakman, C., Black, J., te Beek, T.A.H., Krieger, E., Joosten, R.P., and Vriend, G. (2015). A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* *43*, D364–368.
- Vié, V., Legardinier, S., Chieze, L., Le Bihan, O., Qin, Y., Sarkis, J., Hubert, J.-F., Renault, A., Desbat, B., and Le Rumeur, E. (2010). Specific anchoring modes of two distinct dystrophin rod sub-domains interacting in phospholipid Langmuir films studied by atomic force microscopy and PM-IRRAS. *Biochimica et Biophysica Acta (BBA) - Biomembranes* *1798*, 1503–1511.
- Way, M., Pope, B., Cross, R.A., Kendrick-Jones, J., and Weeds, A.G. (1992). Expression of the N-terminal domain of dystrophin in *E. coli* and demonstration of binding to F-actin. *FEBS Lett.* *301*, 243–245.
- Wein, N., Alfano, L., and Flanigan, K.M. (2015). Genetics and Emerging Treatments for Duchenne and Becker Muscular Dystrophy. *Pediatric Clinics of North America*.
- Winder, S.J., Gibson, T.J., and Kendrick-Jones, J. (1995a). Dystrophin and utrophin: the missing links! *FEBS Letters* *369*, 27–33.

Winder, S.J., Hemmings, L., Maciver, S.K., Bolton, S.J., Tinsley, J.M., Davies, K.E., Critchley, D.R., and Kendrick-Jones, J. (1995b). Utrophin actin binding domain: analysis of actin binding and cellular targeting. *J. Cell. Sci.* *108 ( Pt 1)*, 63–71.

Yokota, T., Duddy, W., Echigoya, Y., and Kolski, H. (2012). Exon skipping for nonsense mutations in Duchenne muscular dystrophy: too many mutations, too few patients? *Expert Opin Biol Ther* *12*, 1141–1152.

Zhao, J., Kodippili, K., Yue, Y., Hakim, C.H., Wasala, L., Pan, X., Zhang, K., Yang, N.N., Duan, D., and Lai, Y. (2016). Dystrophin contains multiple independent membrane-binding domains. *Hum Mol Genet* *25*, 3647–3653.

## FIGURES LEGENDS

Figure 1: Reconstruction protocol of high-resolution models in accordance to small angle X-ray scattering data. (A) Three best  $\chi^2$  *ab initio* models computed by the GASBOR method (Franke et al., 2017). (B) The generation of 20 interactive-fitting starting positions of the ABD2 homology model (in grey cartoon) into the SAXS molecular envelopes has been automated by the alignment following main axis in two opposite orientation and by  $36^\circ$  rotations of the model along this same axis. (C) Superposition of the four best high-resolution SAXS-based models for dystrophin ABD2 fragment. Colors correspond to the two groups of similar models obtained after clustering (RMSD of at least  $6.4\text{\AA}$  have been measured between carbon atom C $\alpha$  blue and green structures).

Figure 2: (A) Contact frequencies of the dystrophin ABD2 with the F-actin structure all over the docking simulations: raw data obtained by rigid low-resolution docking (left) and filtered data obtained after selection of the sole solutions in accordance to experimental mapping (right). (B) Best theoretical complexes computed for the association of dystrophin (purple and blue) ABD2 to the F-actin (yellow and orange). Two independent groups of solutions were verifying experimental contact mapping: one shows great accordance concerning interaction between repeat R11 and F-actin (left), the other one shows great accordance concerning interaction between repeat R15 and F-actin (right)

Figure 3: (A) Principle of the flexible docking. In this example, the dystrophin repeat R11 is fixed on the actin filament and the rest of the ABD2 fragment is driven thanks to an interactive simulation to fit the experimental mapping obtained for the dystrophin repeat R15. (B) Structure of the best model for the complex formed by the dystrophin ABD2 (yellow cartoon) and F-actin (grey volume) and verifying both contacts of the dystrophin repeats R11 and R15 with the actin filament. Peptides identified by experimental mapping are shown as a blue volume for F-actin and as an orange volume for dystrophin. (C) and (D) electrostatic and hydrophobic molecular potentials for the characterization of the interacting surfaces between both proteins; (C) Electrostatic potential is set for isovalues of  $\pm 50kTe$  (blue for positive and red for negative); (D) Molecular Hydrophobic Potential (MHP) with normalized values ranging from -1 (green, hydrophilic) to +1 (yellow, hydrophobic). (E) Projection of the intermolecular contacts in the

best dystrophin ABD2–F-actin models along the sequence of the dystrophin fragment. Orange and green dots (ref horizontal line) are delimiting the interaction regions experimentally identified with F-actin and lipids, respectively. Contacts are shown along the y-axis for the six best models obtained from interactive simulations, depending on the change in the relative position of dystrophin repeat R15 towards the fixed repeat R11 (number of actin units as a separation between both).

Figure 4: (A) SANS scattering curve of dystrophin R11-15 fragment in interaction with zwitterionic (DZB) and anionic (DAB) bicelles. SANS curve for free dystrophin ABD2 is colored green, but blue and red if in complex with zwitterionic or anionic bicelles, respectively. Inset: corresponding ab initio models corresponding to the GASBOR model with the smallest normalized spatial discrepancy (NSD) surrounded by the DAMAVER model. (B) Peptide experimentally identified to contact an anionic bicelle is colored green on the grey tube representing the backbone of the dystrophin ABD2 fragment.

Figure 5: Extension of the structural model of the dystrophin central domain in interaction with F-actin. (A) Schematic representation of the R4-19 fragment precisizing the nature of the different SAXS-based models used for its construction. Among them, R4-9, R11-15 and R16-19 structural models were detailed in a previous work (Delalande 2018). (B) Dystrophin R4-19 structural model after optimization, the colors are the same as used in A and highlight. Interactions domains are identified by characters (#) for F-actin binding domain, (\*) for lipids binding domain and (§) nNOS interaction domain. (C) Bar plot of the angles measured between the main axis of two consecutive repeats in the R4-19 fragment; colors correspond to the ones in (A). The asterisks highlight the values measured in the dystrophin R4-19 model. (D) R4-19/F-Actin model based on the merge of the R11-15/F-Actin theoretical complex and of the R4-19 fragment SAXS-based model. F-actin is shown as a grey volume and dystrophin as a purple cartoon. The lipid binding domain on R12 (orange spheres) is not involved in actin contact, and remain accessible for lipid interactions



## ASSOCIATED CONTENT

**Supporting Information.** PDF file including three supplemental Table S1 and Figures S1 to S3.

## AUTHOR INFORMATION

### Corresponding Author

\* To whom correspondance should be addressed: Dr Olivier Delalande, Institut de Génétique et Développement de Rennes, Faculté de Pharmacie, 2 avenue du Pr Léon Bernard 35043 Rennes Cedex, France; [olivier.delalande@univ-rennes1.fr](mailto:olivier.delalande@univ-rennes1.fr); Tel (+33)(0)2 23 23 30 07

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Funding Sources

This research was supported by research grants from the AFM–Téléthon (Application #018211A) and the french GENCI computing program (DYSIM project). DML was recipient of a PhD fellowship from AFM-Téléthon (grant 19409).

## ACKNOWLEDGMENT

We also wanted to acknowledge Tim Pyrkov for providing stand-alone version of Platinum. The authors are also grateful to Giulio Gambarota for his careful reading of the manuscript.

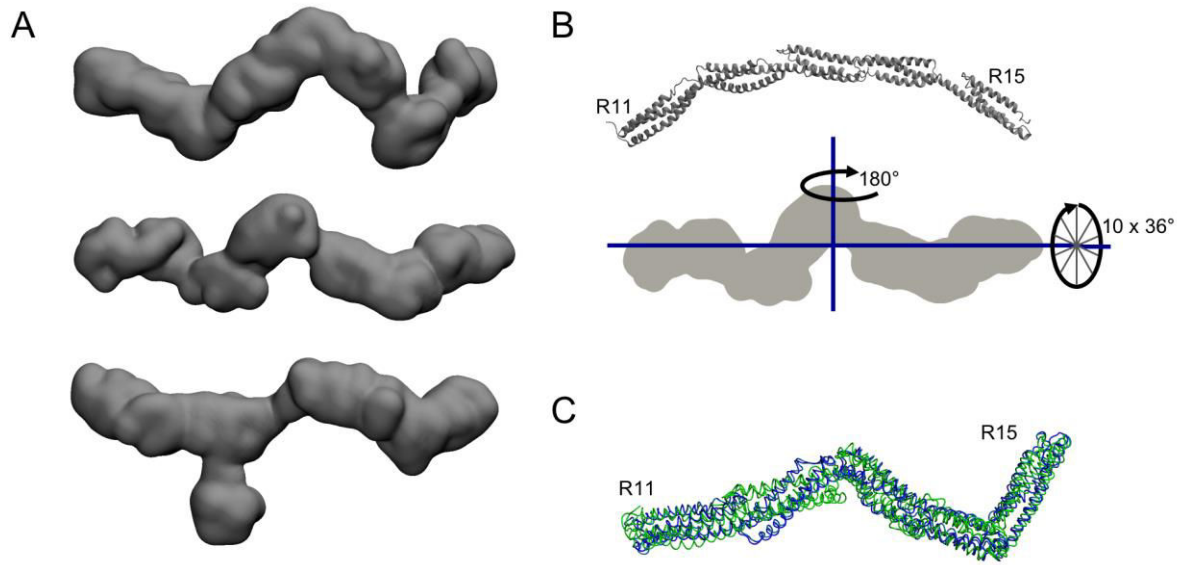
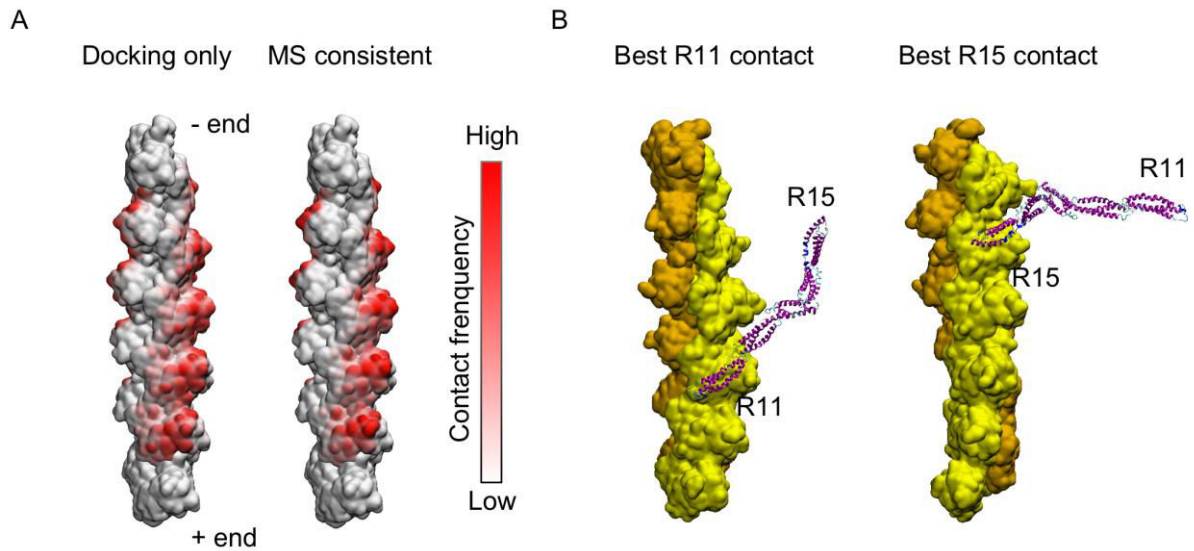


Figure 1

Figure 2



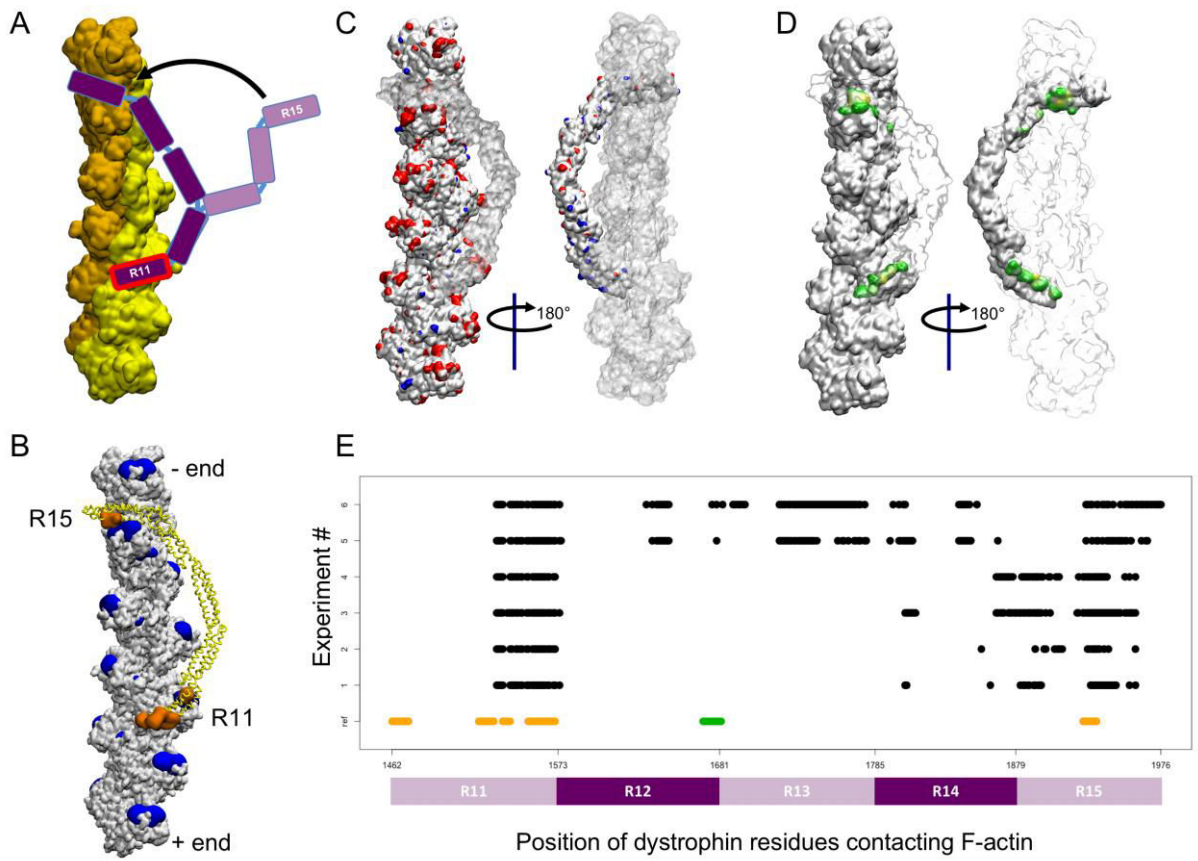


Figure 3

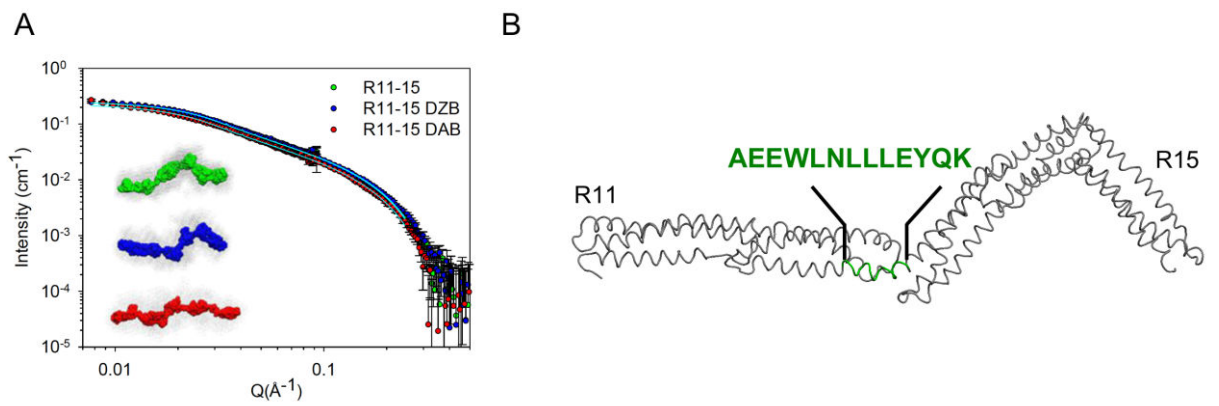
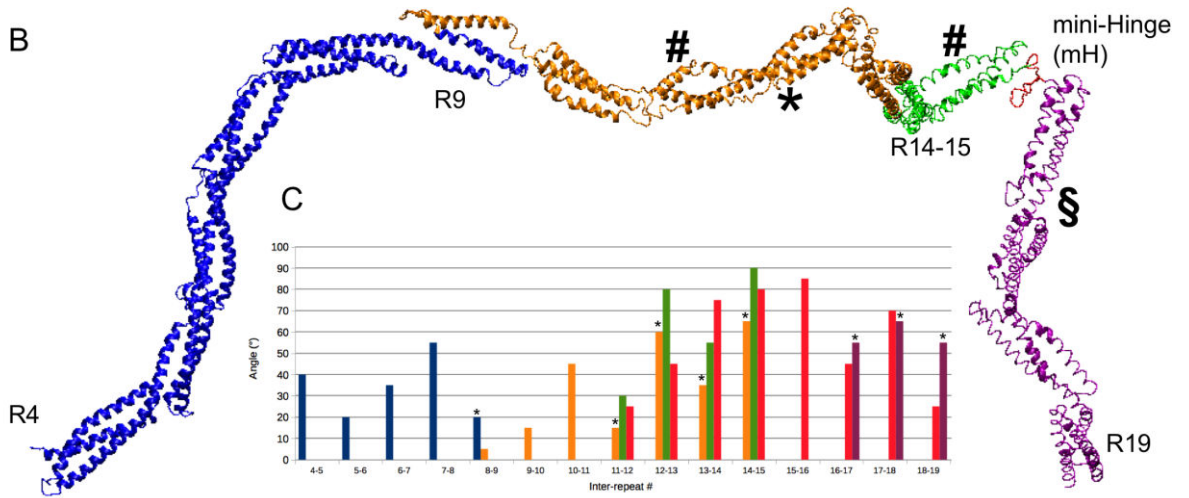


Figure 4

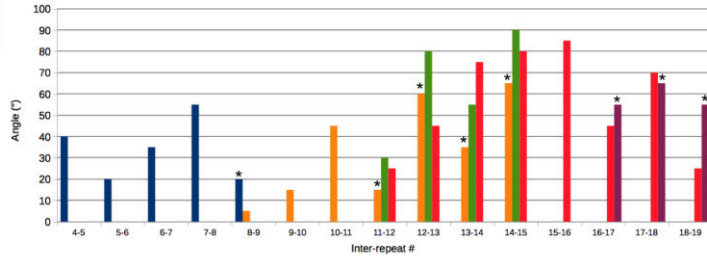
A

Repeat	4	5	6	7	8	9	10	11	12	13	14	15	mH	16	17	18	19
SAXS Model	R4-9					R8-15					R11-15			R11-19	R16-19		

B



C



D

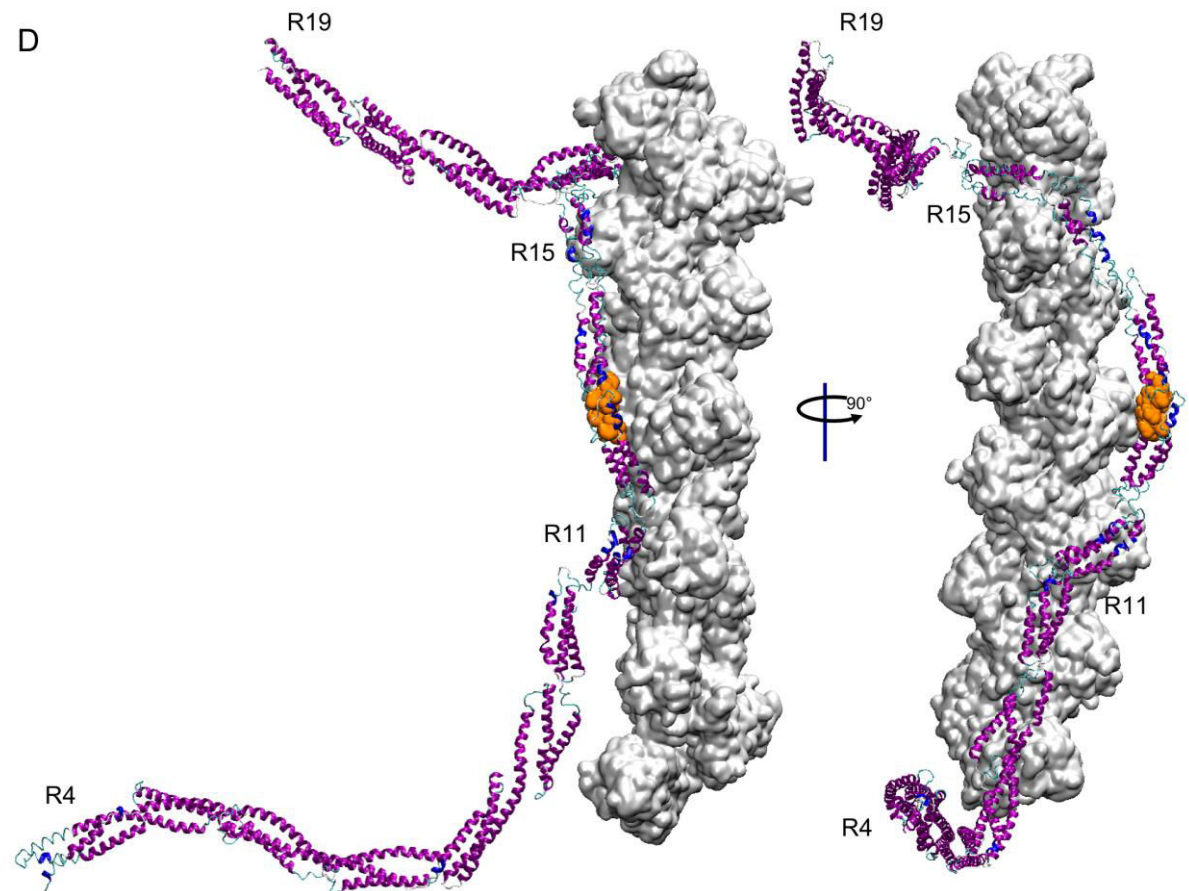


Figure 5

## Supplementary Figures

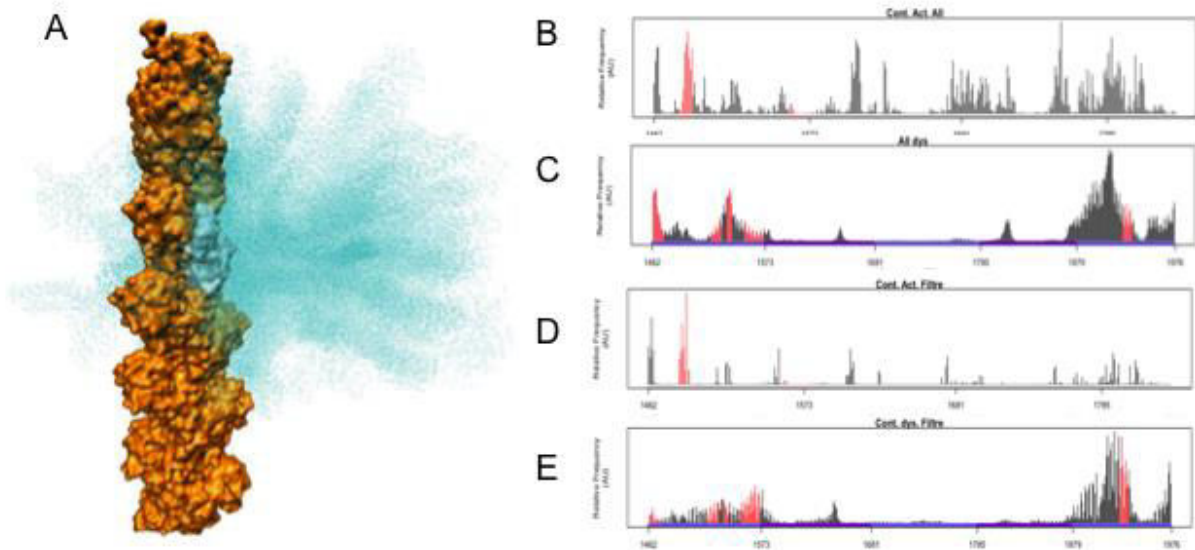
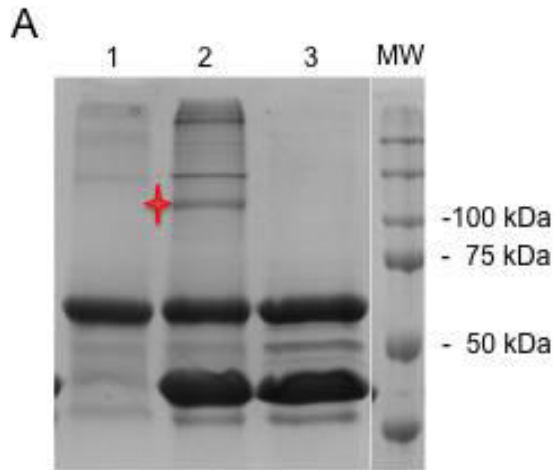


Figure S1: (A) All rigid docking results in accordance to experimental mapping of the contacts between dystrophin ABD2 (blue dots) and F-actin (orange volume). (B and C) Raw contact frequencies computed through low-resolution rigid docking methods; (B) contacts of dystrophin ABD2 all along the primary sequence of a unique actin unit; (C) contacts of F-actin all along the primary sequence of dystrophin ABD2. (D and E) Filtered contact frequencies from rigid docking solutions in accordance to experimental mapping; (D) contacts of dystrophin ABD2 all along the primary sequence of a unique actin unit; (E) contacts of F-actin all along the primary sequence of dystrophin ABD2.



**B**

Score	m/z	Peptide R11-15	From	To	Peptide Actin	From	To
176	789,38	[GSFQKPANFEQR]	0	12	[AGFAGDDAPR]	21	30
155	736,04	[SEVEMVIKTGR]	59	69	[AGFAGDDAPR]	21	30
118	837,19	[GSFQKPANFEQR]	0	12	[VAPEEHPTLLTEAPLNPK]	98	115
92	601,30	[KQTENPK]	75	81	[AGFAGDDAPR]	21	30
87	622,57	[LHYNELGAKVTER]	92	104	[AGFAGDDAPR]	21	30

Figure S2: Experimental mapping of the contact of the dystrophin ABD2 subdomain filamentous actin. (A) SDS-page for cross-linked F-actin/R11-15 complex by EDC. (Lane 1) R11-15+EDC. (Lane 2) R11-15 + F-actin +EDC. (Lane 3) R11-15 + F-actin controls. Red star : R11-15/F-actin complex. (B) Peptides identified by MS analysis after digestion of cross-linked complex formed by ABD2-actin-F



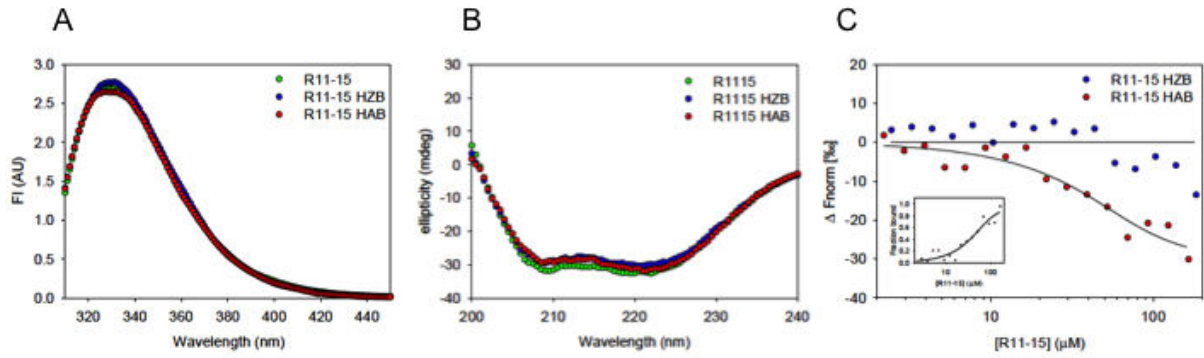


Figure S3: (A) Tryptophan fluorescence intensity (FI) of each dystrophin ABD2 in the absence (green) or in the presence of zwitterionic (HZB, blue) or anionic (HAB, red) bicelles. FI variation is due to a modification in the Trp environment attributed to lipid binding. (B) CD spectra highlight that the secondary and tertiary structure of dystrophin ABD2 fragment is maintained during lipid binding. (C) Microscale thermophoresis (MST) data: the dissociation constant ( $K_d$ ) was determined to be approximately 10  $\mu\text{M}$  for the ABD2/ anionic bicelle complex.

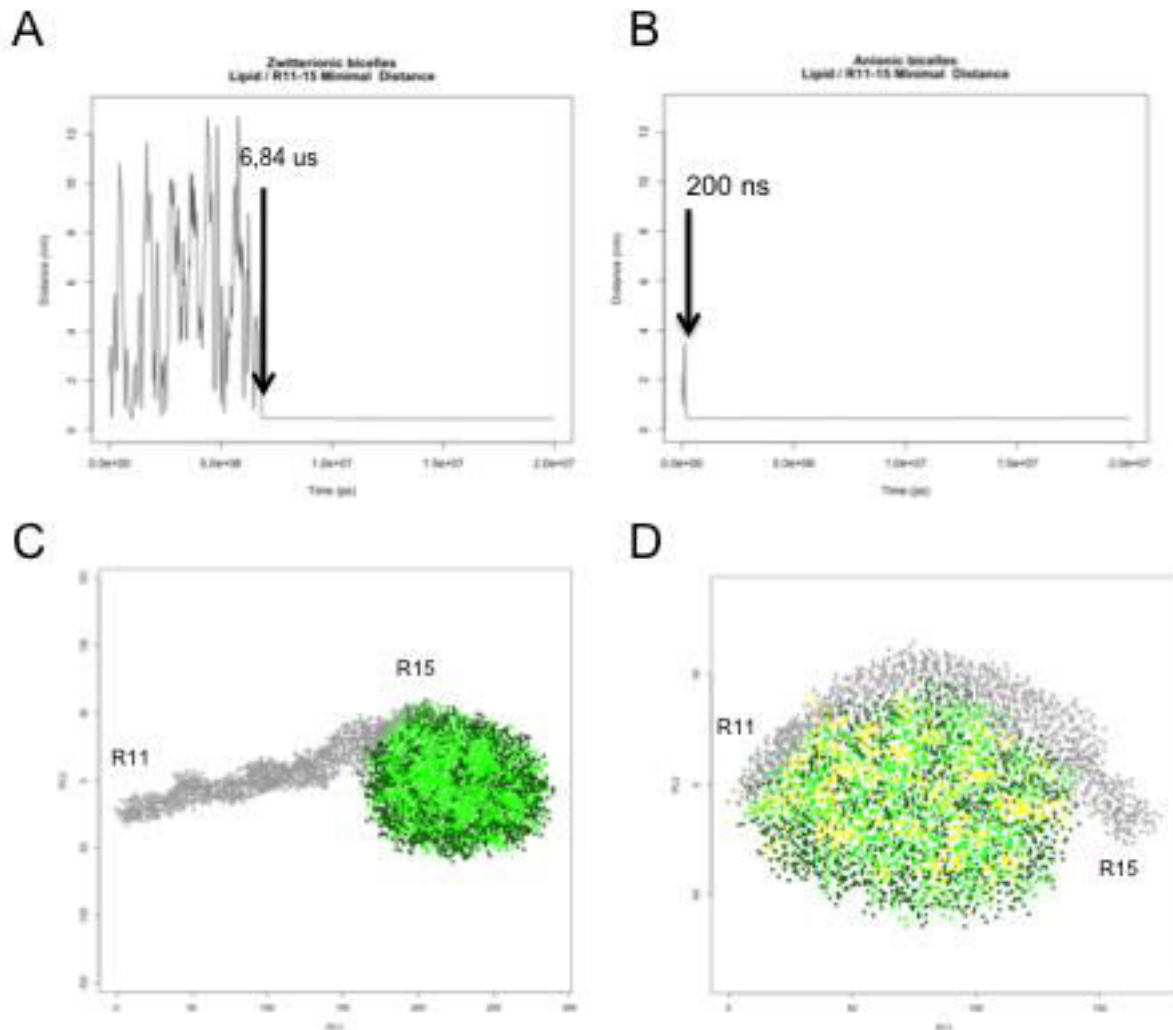


Figure S4: (A and B) Minimal distance between dystrophin R11-15 fragment and zwitterionic bicelle (A) and anionic bicelle (B) between 0 $\mu$ s et 20 $\mu$ s of simulation, black arrow indicates the time for a contact between protein and lipids. (C and D) Projection along the two principal axis defined by PCA (Principal Components Analysis) of the simulation final state for protein-zwitterionic bicelle (C) or protein-anionic bicelle complexes. Protein grains are shown in grey and lipids are coloured in light green for DHPC, dark green for DMPC and yellow for DMPS (for anionic bicelle only, D).



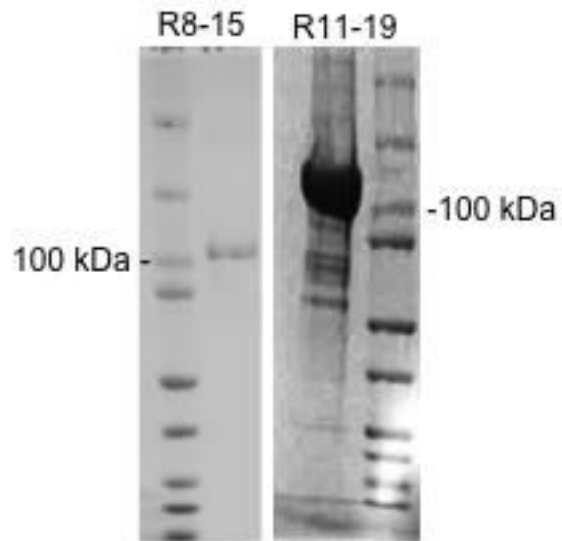


Figure S5: Purification of dystrophin fragments R8-15 and R11-19. Expression of the fragments of human dystrophin R8-15 and R11-19 were induced in insect cells. Purifications were performed by affinity chromatography followed by gel filtration. SDS-PAGE, using Coomassie blue staining.

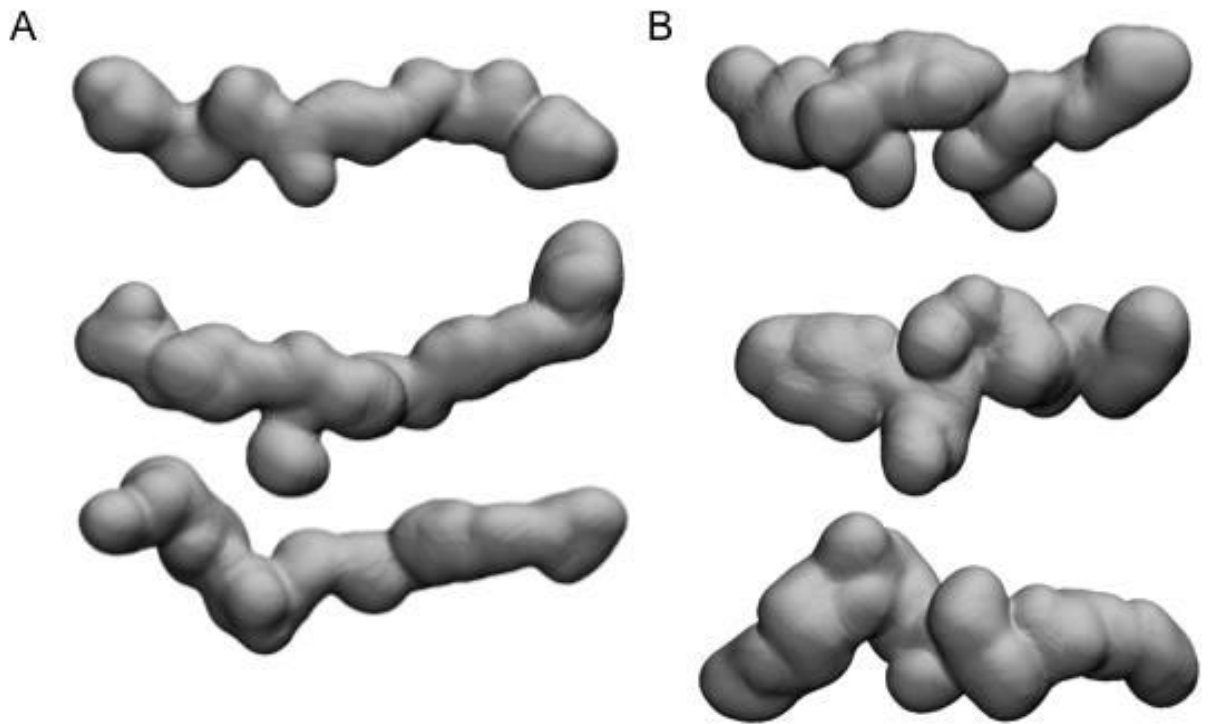


Figure S6: (A and B) Three best  $c^2$  *ab initio* models computed by the GASBOR method for the dystrophin R8-15 fragment (A) and R11-15 fragment (B).

## E. Discussion

Avant les travaux présentés dans ce chapitre, les régions d'interaction du fragment R11-15 avec l'actine n'étaient pas très précisément connues, en dehors d'expérience de digestion trypsique menée par notre équipe en 2011 (Sarkis et al., 2011). Grâce aux analyses de spectrométrie de masse de complexes réticulés, les zones précises d'interactions entre R11-15 et l'actine-F ont été identifiées au niveau des répétitions R11 et R15 de la dystrophine, alors que la région d'interaction aux lipides a été limitée à la fin de la répétition R12 incluse dans le sous-domaine ABD2.

Dans les modèles de Molza et al. (Molza et al., 2014), la dystrophine est retrouvée enroulée autour de l'actine-F, avec une zone de contact étendue sur tout le long du fragment, et d'une manière très similaire à certains des modèles présentés dans notre étude. Une telle conformation permettrait difficilement la liaison simultanée avec les lipides au niveau de la région identifiée par spectrométrie de masse. Dans le modèle que nous obtenons, la répétition R12 est écartée du faisceau, libre d'interagir avec les membranes.

Cependant, bien que le mode d'association entre les partenaires ne soit pas correctement prédit par notre étude théorique et préliminaire de l'association ABD2-actine-F (Molza et al., 2014), les orientations globales de la dystrophine et de ses répétitions R11 et R15 sur le filament sont similaires à celles que nous rapportons dans l'article présenté plus tôt. De même, on retrouve une grande similarité entre les zones d'interactions identifiées avec et sans filtrage par les données de spectrométrie de masse dans les modèles préliminaires qui avaient été obtenus précédemment. Ces résultats valident, selon nous, la grande fiabilité d'une approche combinant les méthodes de docking rigide quasi-exhaustif et celles de docking semi-flexible et interactif à basse résolution.

De manière remarquable, on constate que l'écart de 5 unités d'actine entre les deux répétitions liées à l'actine (R11 et R15) est compatible avec une décoration régulière du filament d'actine par la dystrophine, avec un rapport stœchiométrique d'environ un fragment de dystrophine pour cinq monomères d'actine, tel que rapporté dans les expériences de co-sédimentations (voir *C-Actine*).

Les approches d'extinction sélective du signal SANS développées durant la thèse de Raphaël Dos Santos-Morais ont été utilisées dans ce travail afin de déterminer la conformation du fragment R11-15 et des bicelles anioniques (identiques à celles présentées dans le chapitre *III-Etude des interactions hydrophobes du faisceau d'hélice R1-3*) indépendamment l'une de l'autre. Il a ainsi été observé une légère augmentation de la taille du fragment ABD2 en présence de bicelles anioniques, mais pas zwitterioniques, de manière un peu similaire à ce qui avait été précédemment observé de façon beaucoup plus nette avec le fragment R1-3.

La simulation basse-résolution du système R11-15 / bicelle anionique prédit une région étendue d'interaction aux lipides incluant le peptide qui a été identifié par click-chemistry/spectrométrie de masse situé au niveau C-terminal de la répétition R12. Elle ne permet cependant pas d'expliquer l'élongation observée par SANS. Toutefois, les résultats de la simulation sont à nuancer au regard du contact très étendu établi entre la dystrophine et la bicelle, et surtout, de la diminution de la longueur maximale du fragment protéique après sa liaison aux lipides, contrairement à ce qui est observé expérimentalement.

Plusieurs hypothèses ont été avancées pour expliquer l'augmentation de longueur de R11-15 en présence de bicelles anioniques mesurée par SANS. Il a été dans un premier temps avancé que les faisceaux d'hélices aux extrémités subissaient une ouverture, comme pour R1-3. L'augmentation de taille de 7 nm que subissait R1-3 en présence de bicelles anioniques représentait 40% de sa longueur initiale (de 17,7 nm à 24,8 nm), compatible avec l'ouverture d'une partie du faisceau d'hélice. R11-15 s'étend de 2,2 nm (de 27,4 nm à 29,6 nm, soit 8 %), une longueur incompatible avec une ouverture du faisceau, ce qui semble confirmé par les mesures faites par dichroïsme circulaire présentées dans notre étude. Une seconde hypothèse est que cette élongation de R11-15 correspond à des changements de conformations des *linkers*. En effet, les modèles à notre disposition montrent une structure avec de nombreux coudes entre les répétitions. Or, l'utilisation du champ de force MARTINI demande, lors de la préparation du système, de définir la structure secondaire de l'objet étudié pour la contraindre durant la simulation. Généralement, la structure secondaire d'un modèle moléculaire est déterminée en utilisant des outils comme STRIDE (Frishman and Argos, 1995) ou DSSP (Kabsch and Sander, 1983). Cependant, cette détermination peut être délicate dans des régions telles que les *linkers*, où l'hélicité n'est que partiellement maintenue ou hautement dynamique. C'est ainsi que dans les premières simulations du complexe R11-15 /

bicelle, le *linker* entre R14 et R15 présentait une interruption locale de l'hélicité. Le *linker* s'est alors replié à 180° en présence d'une bicelle anionique, la répétition R15 étant ramenée le long de R14. Dans les simulations présentées ici, le linker R14-15 a été défini comme entièrement hélical, au risque de limiter la dynamique de cette région au cours de la simulation. Ce type de choix s'est avéré délicat et semble altérer la portée explicative de nos simulations et de leur confrontation aux données expérimentales de SANS.

Globalement, l'étude présentée dans ce chapitre semble montrer une limite à l'approche par fragment que nous utilisons pour étudier le domaine central. En effet, le domaine R1-3 est délimité de chaque part par une charnière (*hinge* H1 et H2 en N-terminal et C-terminal, respectivement), lui conférant une certaine indépendance du reste de la protéine. Dans le cas de l'étude du fragment R11-15, la délimitation de cette construction a plus été guidée par des nécessités techniques (expression, purification des fragments, etc), que pour des raisons structurales. En effet, même si ce sous-domaine semble avoir une signature propre en terme d'interaction (lipides et actine-F), il n'est pas forcément pertinent de l'étudier de manière isolée quand, sur le plan structural, le sous-domaine cohérent correspondrait plutôt au fragment R4-19, délimité à ses extrémités par les deux charnières H2 et H3. C'est ainsi que l'interaction entre l'actine et l'extrémité N-terminale de R11, qui a été observée dans les

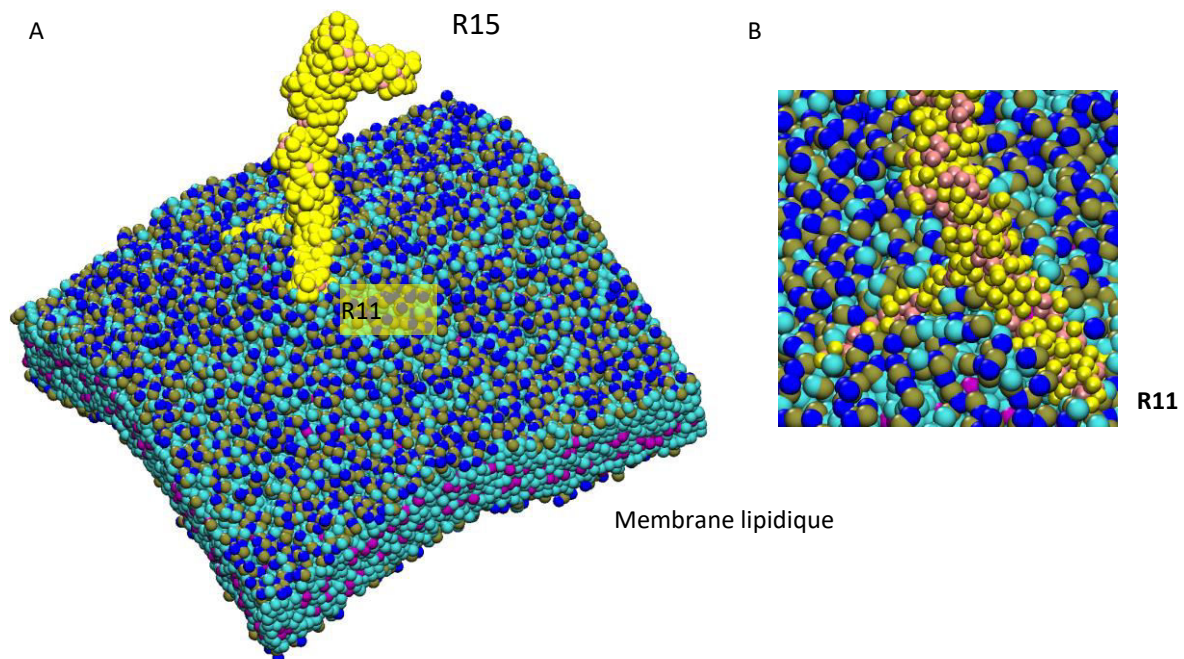


Figure 49 : Résultat de simulation de dynamique moléculaire gros grain avec le champ de force Martini d'un complexe R11-15 / membrane anionique au terme de 8  $\mu$ s de calcul ; A) Vue d'ensemble B) Zoom sur l'ouverture du faisceau de la répétition R11.

approches d'amarrages rigides et par spectrométrie de masse, est probablement un artefact provenant de cette approche par fragment qui a longtemps imposé l'utilisation du sous-domaine R11-15 isolé dans nos expériences et nos simulations.

Nous avons effectué une simulation basse résolution d'un modèle R11-15 avec un modèle de bicouche lipidique anionique, sur une durée de 8 microsecondes (Figure 49). Comme dans nos premières simulations R11-15 / bicelle anionique, on observe un coude très serré entre R14 et R15, résultat de l'absence de correction de la définition de la structure secondaire préalablement à la simulation. De plus, le contact entre la membrane et le fragment de dystrophine s'effectue via la répétition R11, qui n'est pas la zone de contact identifiée par spectrométrie de masse (R12). Nous interprétons ce contact comme le résultat d'un effet de bord, résultant de l'approche par fragment décrite plus tôt. Enfin, on observe l'ouverture du faisceau d'hélice de R11, avec l'écartement de l'hélice A des hélices B et C, incompatible avec les expériences de dichroïsme circulaire présentées dans notre étude.

Il semble que l'unité structurale de R4-19 soit à nuancer par la présence d'un *mini-hinge* entre R15 et R16, qui séparerait cette région en deux sous-domaines fonctionnels, R4-15 et R16-19. Durant l'expression et la purification du grand fragment R11-19, ce *mini-hinge* constitue un site de coupure spontané, qui impose d'utiliser cette protéine rapidement après sa purification. La structure de ce domaine *mini-hinge* a été simulée à la fois par des approches *ab initio* (PEPFOLD (Thévenet et al., 2012)) et par homologie (I-TASSER (Yang and Zhang, 2015))

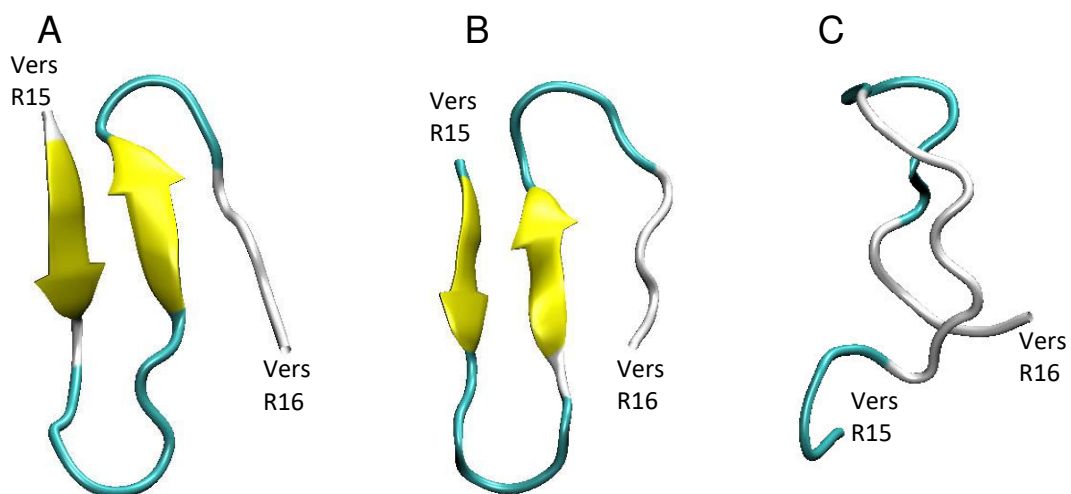


Figure 50 : Modèles de *mini-hinge*, obtenus A) *ab initio* (PEP-FOLD) B) par homologie (PEP-FOLD) C) ou d'après le modèle R11-19 obtenus par ajustement flexible sous contraintes SAXS.

(Figure 50A et B). Ces modèles prédisent une structure en feuillets  $\beta$ , qui n'est pas retrouvée dans R11-19 (Figure 50C).

C'est pour essayer de compenser cette limitation que nous avons entrepris d'étendre le modèle R11-15 à un sous-domaine allant de R4 à R19, notamment via l'obtention de nouveaux modèles par ajustement flexibles pour de nouvelles données SAXS des fragments R8-15 et R11-19. L'étude de ces fragments a de fait permis la modélisation d'un complexe R4-19 / actine-F, qui a servi de base à l'obtention d'un modèle plus complet comprenant un fragment du domaine central de la dystrophine associé à un filament d'actine ainsi qu'à une bicelle (Figure 51 A). De plus, dans le modèle obtenu, l'orientation relative de la bicelle par rapport à l'actine et à la dystrophine permet son remplacement par un modèle de membrane (Figure 51 B). En particulier, cette association est possible par la courbure de la membrane, permettant d'obtenir localement une très bonne superposition entre la bicelle et la membrane.

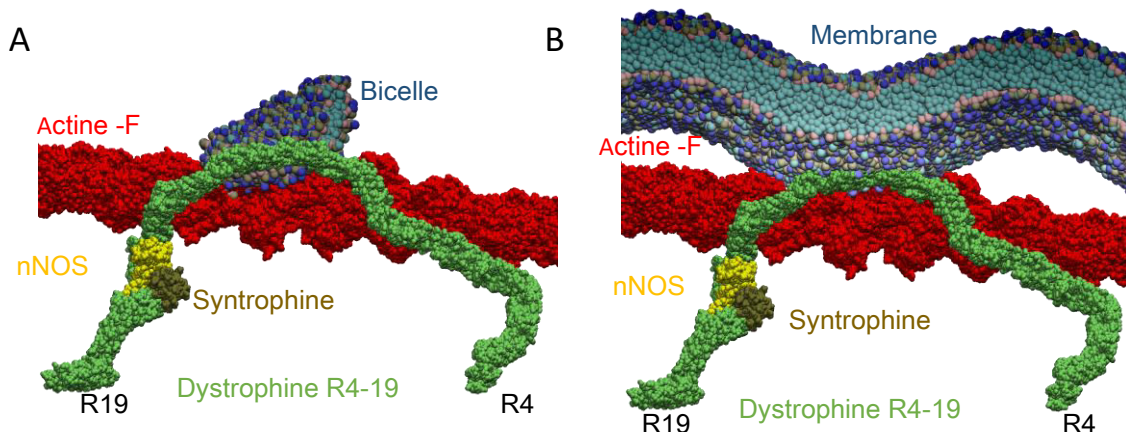


Figure 51 : Modèles préliminaires de complexes (rouge) actine-F (vert) R4-15 (jaune) nNOS-PDZ (marron) syntrophine PDZ (bleu) lipides ; A) Modèle de bicelle anionique ; B) Modèle de bicouche anionique.

Des travaux préliminaires effectués dans l'équipe ont porté sur l'étude de l'interaction entre les domaines PDZ de la nNOS et de la syntrophine, et le fragment R16-17 de la dystrophine. Les domaines PDZ sont les domaines d'interaction avec la dystrophine. Ce modèle de complexe a été obtenu suivant le même protocole d'amarrage rigide que celui mis en place pour la construction du complexe R11-15 / actine-F. Ce modèle a pu être transposé sans modification à notre modèle R4-19, permettant de proposer un concept de complexe

dystrophine R4-19 / actine-F / bicelle ou membrane / nNOS-PDZ / syntrophine-PDZ (Figure 51).



## F. Perspectives

Le muscle est un organe très dynamique, dont les différents composants sont soumis en permanence à des contraintes mécaniques. Pour aller plus loin et étudier l'aspect dynamique et mécanique de la dystrophine au sein de complexes macro-moléculaires, nous envisageons la mise en œuvre de simulations de dynamique moléculaire sous contraintes (voir // *Méthodes*). Brièvement, ces expériences consisteront à effectuer une simulation de dynamique moléculaire du complexe macromoléculaire d'intérêt en appliquant une force sur une partie du système, et en observant comment cette perturbation est transmise aux partenaires. De plus, il est également prévu d'effectuer des expériences qui alternent les phases avec et sans application de contraintes. Ces expériences ont pour objectif d'étudier la dynamique de retour au repos du complexe après une déformation mécanique.

Plusieurs paramètres du protocole de simulation sont à l'étude. En particulier, plusieurs scénarios seront testés, en variant le composant subissant la contrainte : l'actine, la membrane, ou la dystrophine. De plus, de nombreux scénarios sont également envisagés quant à la directionnalité des contraintes. Alors qu'il semble évident que celle-ci seront appliquées dans toutes les directions pour la membrane, et dans la direction de l'axe principal pour les protéines filamenteuses (dystrophine et actine), il sera également nécessaire d'étudier des situations plus « exotiques », avec par exemple des contraintes perpendiculaires à l'axe de l'actine. Ces approches de SMD, pour le moment à l'état de projet, seront essentielles pour la pleine compréhension de la fonction de la dystrophine et du rôle de ses partenaires dans un contexte dynamique.

## V. Discussion générale

La dystrophine est un élément essentiel à l'assemblage de l'échafaudage moléculaire retrouvé au niveau des sarcomères, et qui offre au sarcolemme un soutien mécanique pour résister au stress mécanique des contractions musculaires (Figure 52). Les myopathies de Duchenne et de Becker sont causées respectivement par l'absence totale ou la dysfonction partielle de la dystrophine. La méconnaissance de la structure de la dystrophine, et en particulier de son domaine central, est un frein au développement de thérapies efficaces des DMD et BMD. En particulier, il a été récemment mis en évidence le lien qui existe entre le type de mutation rencontré chez des patients BMD, la protéine tronquée codée et la sévérité du phénotype (Nicolas et al., 2014b). Ainsi la conception rationnelle de gènes codant une dystrophine raccourcie mais fonctionnelle, tenant compte des différentes contraintes imposées comme la limite de taille d'encapsulation par les vecteurs viraux, nécessite la compréhension des bases moléculaires et physico-chimiques qui permettent les interactions entre la dystrophine et ses partenaires.

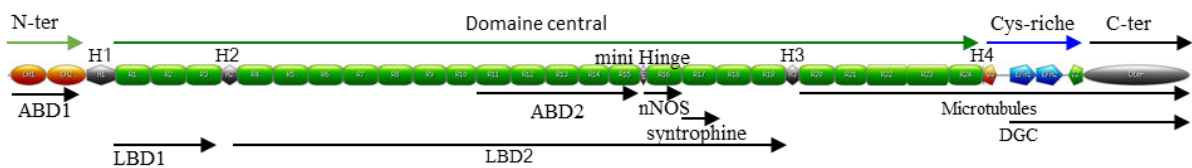


Figure 52: Organisation de la dystrophine et régions d'interactions avec ses partenaires ; H1-H4 : hinges 1 à 4 ; ABD : Domaines de liaison à l'actine ; LBD : Domaines de liaison aux lipides ; nNOS : Domaine de liaison à l'Oxyde nitrique synthase neuronale ; DGC : Domaine de liaison au complexe glycoprotéique associé à la dystrophine.

Cette thèse avait pour objectif de caractériser la structure du domaine central de la dystrophine et ses interactions avec ses partenaires filamenteux du cytosquelette, actine et microtubules. Ce travail a été plus tard étendu à l'interaction avec les lipides membranaires.

Les interactions hydrophobes sont les principales responsables du maintien de la structure en faisceau d'hélice du domaine central de la dystrophine. Dans une première partie, nous avons présenté le travail effectué sur le développement d'une nouvelle approche de détection de ces interactions hydrophobes. Nous l'avons appliqué à la répétition R1 pour étudier l'influence de son contexte (R1 seul, dans R1-2 ou dans R1-3) sur le réseau d'interactions hydrophobes. Nous avons pu ainsi démontrer la plasticité de ces contacts. De plus, des différences ont été observées au niveau des *linkers* entre les répétitions : le coude

entre les répétitions R1 et R2 est associé avec un faible chevauchement des interactions entre les faisceaux de R1 et R2, alors qu'elles forment un maillage dense entre R2 et R3 associé à un enchaînement plus linéaire des répétitions.

Ce travail a été mis en lien avec les travaux récemment publiés de Dos Santos Morais et al. (voir Annexe : Dos Santos Morais *et al*, 2018) sur l'interaction R1-3 / lipides. Par des expériences de SANS et de simulations interactives, il a été montré qu'en présence de membranes anioniques, les hélices A et B de la répétition R1 s'écartaient du faisceau. Cet évènement est associé à la présence de deux verrous dans R1, correspondant à deux acides aminés impliqués dans de nombreuses interactions hydrophobes, et qui semblent maintenir le faisceau fermé en absence de lipides. Cette ouverture pourrait également être assistée par la présence d'un motif senseur de la courbure membranaire ALPS, qui rentrerait en contact avec les lipides avant l'ouverture du faisceau.

Le fragment R11-15 est le seul qui a été identifié comme capable d'interagir à la fois avec l'actine-F et les lipides membranaires. Dans une seconde partie, nous avons cherché à construire un modèle de complexe tripartite actine / dystrophine / bicelle basé sur des données expérimentales. Cette construction s'est effectuée en plusieurs étapes : la modélisation du fragment R11-15, les expériences d'amarrage moléculaire R11-15 / actine associées à des données de spectrométrie de masse, des simulations gros-grains de complexes R11-15 / bicelles anioniques supportées par des expériences de spectrométrie de masse entre protéine et lipides photoactivables, et finalement la construction du modèle de complexe à trois partenaires. Dans le détail, nous avons dans un premier temps associé des données basses résolutions obtenues par SAXS à un protocole d'ajustement flexible pour obtenir un ensemble de modèles du fragment R11-15. Cette diversité est nécessaire pour l'étape suivante, d'amarrages moléculaires rigides avec l'actine : le modèle ne pouvant s'adapter activement au cours de la simulation, la diversité dynamique est simulée par la diversité des modèles statiques utilisés. Les résultats d'amarrages moléculaires ont ensuite été raffinés par l'obtention de cartes de contact 1D entre l'actine et la dystrophine par spectrométrie de masse de complexes réticulés. Cependant, alors que la cartographie identifiait deux régions de la dystrophine interagissant avec l'actine, sur R11 et R15, nos meilleures solutions de complexes théoriques n'en impliquaient qu'une seule. Grâce à des expériences d'amarrages interactifs, dans lesquelles un domaine est fixé à l'actine et l'autre libre, nous avons pu

construire un modèle de complexe R11-15-actine dont les deux domaines identifiés par spectrométrie de masse sont au contact de l'actine. L'étude de l'amarrage entre R11-15 et des membranes anioniques a été effectuée par une simulation gros-grains, permettant d'effectuer la simulation de système sur plusieurs microsecondes. En parallèle, des complexes ont été formés entre le fragment de dystrophine et des bicelles contenant un type de lipide photoactivable, pouvant créer des liaisons covalentes avec les résidus protéiques proches lors d'une exposition aux UV. Ces expériences ont permis d'identifier une zone de contact située au niveau de R12, également retrouvée dans les simulations. Enfin, les deux modèles de complexes ont été associés, les zones d'interactions n'étant pas exclusives. En particulier, alors que la dystrophine s'étend sur l'actine, les acides aminés de la répétition R12 impliqués dans la liaison aux lipides membranaires restent accessibles pour les lipides.

En plus du fragment R11-15, des données expérimentales de SAXS ont été obtenues pour d'autres fragments du domaine central (R4-9, R8-15, R11-19 et R16-19), et ont permis l'obtention de modèles hautes résolutions, également par des approches d'ajustement flexible. Ceux-ci ont été utilisés pour étendre d'une part le modèle R11-15 en solution vers un modèle R4-19 en solution, et d'autre part un modèle R11-15/actine-F vers un modèle R4-19/actine-F. Ce dernier a servi de base à la création d'un modèle de macro-complexe dystrophine / actine / lipides / nNOS / syntrophine.

La dystrophine est une protéine complexe. Il sera ainsi nécessaire d'obtenir des modèles de la protéine complète, seule et en complexes macro-moléculaires, pour pouvoir comprendre la fonction de la protéine dans son ensemble. Des méthodes de CET ont ainsi été utilisées pour explorer le complexe composé d'une part de la région de la dystrophine s'étendant de R20 à l'extrémité C-terminale, et d'autre part aux microtubules. L'absence d'un tel fragment à notre disposition ne nous a pas permis de poursuivre les expériences en ce sens. Cependant, des travaux en cours dans l'équipe sont actuellement menés pour tenter de produire l'isoforme dp140, s'étendant de R20 à l'extrémité C-terminale. Si ce travail se révèle fructueux, il permettra de réitérer les expériences de CET. De plus, le  $K_d$  entre la dystrophine et les microtubules ayant été déterminé de l'ordre du micromolaire (Belanto et al., 2016), les acquisitions pourront s'effectuer dans des conditions optimales.

D'autre part, l'étude de l'interaction avec l'actine pourra être approfondie lorsque des constructions s'étendant de l'extrémité N-terminale à la répétition R15 seront disponibles. Ce

fragment comporte à la fois les domaines ABD1 et ABD2 et présente une constante d'affinité avec l'actine inférieure à celle de chaque domaine pris indépendamment. La formation de complexe pourra être obtenue *in vitro* dans des conditions de concentration en dystrophine adéquates. Il sera donc vraisemblablement possible d'approfondir les études par CET.

Ce travail de thèse est une avancée dans l'étude des myopathies de Duchenne et de Becker, en proposant le plus grand modèle de macro-complexe construit autour de la dystrophine aujourd'hui disponible. Ce dernier est la synthèse des travaux récents menés dans l'équipe, il présente également la synthèse des méthodes qui ont et qui seront utilisées pour l'obtention d'un modèle complet de la dystrophine en complexe avec ses partenaires.

Enfin, nous pensons que les travaux présentés dans cette thèse sont une base sérieuse pour l'exploration de la faisabilité des différentes constructions étudiées dans les différentes approches de thérapies des DMD et BMD. Ainsi la prédiction de la fonction de produit du gène DMD tronqué chez les patients BMD et la prédiction du saut d'exon optimal à mettre en œuvre dans les thérapies doit s'appuyer sur une connaissance approfondie de la structure de la dystrophine et de ses interactions avec ses partenaires.

## Références bibliographiques

Aartsma-Rus, A., Janson, A.A.M., Kaman, W.E., Bremmer-Bout, M., van Ommen, G.-J.B., den Dunnen, J.T., and van Deutekom, J.C.T. (2004). Antisense-induced multiexon skipping for Duchenne muscular dystrophy makes more sense. *Am. J. Hum. Genet.* *74*, 83–92.

Aartsma-Rus, A., Fokkema, I., Verschuuren, J., Ginjaar, I., van Deutekom, J., van Ommen, G.-J., and den Dunnen, J.T. (2009). Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum. Mutat.* *30*, 293–299.

Abe, A., Saeki, K., Yasunaga, T., and Wakabayashi, T. (2000). Acetylation at the N-terminus of actin strengthens weak interaction between actin and myosin. *Biochem. Biophys. Res. Commun.* *268*, 14–19.

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* *1–2*, 19–25.

Adams, M.E., Odom, G.L., Kim, M.J., Chamberlain, J.S., and Froehner, S.C. (2018). Syntrophin binds directly to multiple spectrin-like repeats in dystrophin and mediates binding of nNOS to repeats 16-17. *Hum. Mol. Genet.* *in press*.

Akey, D.L., Malashkevich, V.N., and Kim, P.S. (2001). Buried Polar Residues in Coiled-Coil Interfaces. *Biochemistry* *40*, 6352–6360.

Altmann, S.M., Grünberg, R.G., Lenne, P.-F., Ylänne, J., Raae, A., Herbert, K., Saraste, M., Nilges, M., and Hörber, J.K.H. (2002). Pathways and Intermediates in Forced Unfolding of Spectrin Repeats. *Structure* *10*, 1085–1096.

Alushin, G.M., Lander, G.C., Kellogg, E.H., Zhang, R., Baker, D., and Nogales, E. (2014). High-resolution microtubule structures reveal the structural transitions in  $\alpha\beta$ -tubulin upon GTP hydrolysis. *Cell* *157*, 1117–1129.

Amann, K.J., Renley, B.A., and Ervasti, J.M. (1998). A Cluster of Basic Repeats in the Dystrophin Rod Domain Binds F-actin through an Electrostatic Interaction. *J. Biol. Chem.* *273*, 28419–28423.

Amann, K.J., Guo, A.W.-X., and Ervasti, J.M. (1999). Utrophin Lacks the Rod Domain Actin Binding Activity of Dystrophin. *J. Biol. Chem.* *274*, 35375–35380.

Ameziane-Le Hir, S., Raguénès-Nicol, C., Paboeuf, G., Nicolas, A., Le Rumeur, E., and Vié, V. (2014). Cholesterol favors the anchorage of human dystrophin repeats 16 to 21 in membrane at physiological surface pressure. *Biochim. Biophys. Acta BBA - Biomembr.* *1838*, 1266–1273.

Anderson, J.T., Rogers, R.P., and Jarrett, H.W. (1996). Ca<sup>2+</sup>-calmodulin binds to the carboxyl-terminal domain of dystrophin. *J. Biol. Chem.* *271*, 6605–6610.

Ayalon, G., Davis, J.Q., Scotland, P.B., and Bennett, V. (2008). An ankyrin-based mechanism for functional organization of dystrophin and dystroglycan. *Cell* *135*, 1189–1200.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* *315*, 1709–1712.

Barreto-Ojeda, E., Corradi, V., Gu, R.-X., and Tieleman, D.P. (2018). Coarse-grained molecular dynamics simulations reveal lipid access pathways in P-glycoprotein. *J. Gen. Physiol.* *jgp.201711907*.

Barrientos, G., Sánchez-Aguilera, P., Jaimovich, E., Hidalgo, C., and Llanos, P. (2017). Membrane Cholesterol in Skeletal Muscle: A Novel Player in Excitation-Contraction Coupling and Insulin Resistance. *J. Diabetes Res.* *2017*, 3941898.

Beaugrand, M., Arnold, A.A., Hénin, J., Warschawski, D.E., Williamson, P.T.F., and Marcotte, I. (2014). Lipid concentration and molar ratio boundaries for the use of isotropic bicelles. *Langmuir ACS J. Surf. Colloids* *30*, 6162–6170.

Becker, P.E., and Kiener, F. (1955). Eine neue x-chromosomale Muskeldystrophie. *Arch. Für Psychiatr. Nervenkrankh.* *193*, 427–448.

Belanto, J.J., Mader, T.L., Eckhoff, M.D., Strandjord, D.M., Banks, G.B., Gardner, M.K., Lowe, D.A., and Ervasti, J.M. (2014). Microtubule binding distinguishes dystrophin from utrophin. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 5723–5728.

Belanto, J.J., Olthoff, J.T., Mader, T.L., Chamberlain, C.M., Nelson, D.M., McCourt, P.M., Talsness, D.M., Gundersen, G.G., Lowe, D.A., and Ervasti, J.M. (2016). Independent variability of microtubule perturbations associated with dystrophinopathy. *Hum. Mol. Genet.*

Berthelot, M.P.E., and Jungfleisch, E.C. (1872). Sur les lois qui prbsident au partage d'un corps entre deux dissolvants (experiences). *Ann Chim Phys 4th Ser.* 396–407.

Bhat, H.F., Adams, M.E., and Khanday, F.A. (2013). Syntrophin proteins as Santa Claus: role(s) in cell signal transduction. *Cell. Mol. Life Sci.* *70*, 2533–2554.

Bhosle, R.C., Michele, D.E., Campbell, K.P., Li, Z., and Robson, R.M. (2006). Interactions of intermediate filament protein synemin with dystrophin and utrophin. *Biochem. Biophys. Res. Commun.* *346*, 768–777.

Bladen, C.L., Salgado, D., Monges, S., Foncuberta, M.E., Kekou, K., Kosma, K., Dawkins, H., Lamont, L., Roy, A.J., Chamova, T., et al. (2015). The TREAT-NMD DMD Global Database: analysis of more than 7,000 Duchenne muscular dystrophy mutations. *Hum. Mutat.* *36*, 395–402.

Blake, D.J., Tinsley, J.M., Davies, K.E., Knight, A.E., Winder, S.J., and Kendrick-Jones, J. (1995). Coiled-coil regions in the carboxy-terminal domains of dystrophin and related proteins: potentials for protein-protein interactions. *Trends Biochem. Sci.* *20*, 133–135.

Blake, D.J., Weir, A., Newey, S.E., and Davies, K.E. (2002). Function and Genetics of Dystrophin and Dystrophin-Related Proteins in Muscle. *Physiol. Rev.* *82*, 291–329.

Bloch, R.J., and Pumphlin, D.W. (1992). A model of spectrin as a concertina in the erythrocyte membrane skeleton. *Trends Cell Biol.* *2*, 186–189.

Bodensteiner, J.B., and Engel, A.G. (1978). Intracellular calcium accumulation in Duchenne dystrophy and other myopathies: a study of 567,000 muscle fibers in 114 biopsies. *Neurology* *28*, 439–446.

Bork, P., and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? *Trends Biochem. Sci.* *19*, 531–533.

Bulfield, G., Siller, W.G., Wight, P.A., and Moore, K.J. (1984). X chromosome-linked muscular dystrophy (mdx) in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* *81*, 1189–1192.

Byers, T.J., Brandin, E., Lue, R.A., Winograd, E., and Branton, D. (1992). The complete sequence of *Drosophila* beta-spectrin reveals supra-motifs comprising eight 106-residue segments. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 6187–6191.

Byers, T.J., Lidov, H.G., and Kunkel, L.M. (1993). An alternative dystrophin transcript specific to peripheral nerve. *Nat. Genet.* *4*, 77–81.

Carballido-López, R. (2006). The Bacterial Actin-Like Cytoskeleton. *Microbiol. Mol. Biol. Rev.* *70*, 888–909.

Chang, P., and Stearns, T. (2000). Delta-tubulin and epsilon-tubulin: two new human centrosomal tubulins reveal new aspects of centrosome structure and function. *Nat. Cell Biol.* *2*, 30–35.

Chéron, J.-B., Zacharias, M., Antonczak, S., and Fiorucci, S. (2017). Update of the ATTRACT force field for the prediction of protein–protein binding affinity. *J. Comput. Chem.* *38*, 1887–1890.

Cheung, A.K., Hoggan, M.D., Hauswirth, W.W., and Berns, K.I. (1980). Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. *J. Virol.* *33*, 739–748.

Clark, K.A., McElhinny, A.S., Beckerle, M.C., and Gregorio, C.C. (2002). Striated muscle cytoarchitecture: an intricate web of form and function. *Annu. Rev. Cell Dev. Biol.* *18*, 637–706.

Clarke, M.S., Khakee, R., and McNeil, P.L. (1993). Loss of cytoplasmic basic fibroblast growth factor from physiologically wounded myofibers of normal and dystrophic muscle. *J. Cell Sci.* *106*, 121–133.

Cohen, C., and Hanson, J. (1956). An X-ray diffraction study of F-actin. *Biochim. Biophys. Acta* *21*, 177–178.

Collins, C.A., Olsen, I., Zammit, P.S., Heslop, L., Petrie, A., Partridge, T.A., and Morgan, J.E. (2005). Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche. *Cell* *122*, 289–301.

Colognato, H., and Yurchenco, P.D. (2000). Form and function: the laminin family of heterotrimers. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* *218*, 213–234.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.

Conway, J.F., and Parry, D.A.D. (1991). Three-stranded  $\alpha$ -fibrous proteins: the heptad repeat and its implications for structure. *Int. J. Biol. Macromol.* *13*, 14–16.

Cotter, M.J., and Muruve, D.A. (2005). The induction of inflammation by adenovirus vectors used for gene therapy. *Front. Biosci. J. Virtual Libr.* *10*, 1098–1105.

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. (2017). Lentiviral vectors can be used for full-length dystrophin gene therapy. *Sci. Rep.* *7*.



- Crick, F.H.C. (1953). The packing of  $\alpha$ -helices: simple coiled-coils. *Acta Crystallogr.* 689–697.
- Cruz-Guzmán, O. del R., Rodríguez-Cruz, M., and Escobar Cedillo, R.E. (2015). Systemic Inflammation in Duchenne Muscular Dystrophy: Association with Muscle Function and Nutritional Status. *BioMed Res. Int.* 2015.
- Cui, Y., and Bastien, D.A. (2011). Water transport in human aquaporin-4: Molecular dynamics (MD) simulations. *Biochem. Biophys. Res. Commun.* 412.
- Dalous, J., Burghardt, E., Müller-Taubenberger, A., Bruckert, F., Gerisch, G., and Bretschneider, T. (2008). Reversal of Cell Polarity and Actin-Myosin Cytoskeleton Reorganization under Mechanical and Chemical Stimulation. *Biophys. J.* 94, 1063–1074.
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092.
- Deconinck, A.E., Rafael, J.A., Skinner, J.A., Brown, S.C., Potter, A.C., Metzinger, L., Watt, D.J., Dickson, J.G., Tinsley, J.M., and Davies, K.E. (1997). Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell* 90, 717–727.
- Delalande, O., Czogalla, A., Hubert, J.-F., Sikorski, A., and Rumeur, E.L. (2017). Dystrophin and Spectrin, Two Highly Dissimilar Sisters of the Same Family. In *Fibrous Proteins: Structures and Mechanisms*, (Springer, Cham), pp. 373–403.
- Delalande, O., Molza, A.-E., Dos Santos-Morais, R., Chéron, A., Pollet, É., Ragueneau-Nicol, C., Tascon, C., Giudice, E., Guilbaud, M., Nicolas, A., et al. (2018). Dystrophin's central domain forms a complex filament that becomes disorganized by in-frame deletions. *J. Biol. Chem.*
- Deplazes, E., Davies, J., Bonvin, A.M.J.J., King, G.F., and Mark, A.E. (2016). Combination of Ambiguous and Unambiguous Data in the Restraint-driven Docking of Flexible Peptides with HADDOCK: The Binding of the Spider Toxin PcTx1 to the Acid Sensing Ion Channel (ASIC) 1a. *J. Chem. Inf. Model.* 56, 127–138.
- DeWolf, C., McCauley, P., Sikorski, A.F., Winlove, C.P., Bailey, A.I., Kahana, E., Pinder, J.C., and Gratzer, W.B. (1997). Interaction of dystrophin fragments with model membranes. *Biophys. J.* 72, 2599–2604.
- van Dijk, J., Miro, J., Strub, J.-M., Lacroix, B., van Dorsselaer, A., Edde, B., and Janke, C. (2008). Polyglutamylation is a post-translational modification with a broad range of substrates. *J. Biol. Chem.* 283, 3915–3922.
- Dominguez, R., and Holmes, K.C. (2011). Actin structure and function. *Annu. Rev. Biophys.* 40, 169–186.
- Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: A Protein–Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* 125, 1731–1737.
- Domogatskaya, A., Rodin, S., and Tryggvason, K. (2012). Functional Diversity of Laminins. *Annu. Rev. Cell Dev. Biol.* 28, 523–553.
- Dong, J.Y., Fan, P.D., and Frizzell, R.A. (1996). Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum. Gene Ther.* 7, 2101–2112.
- Dore, K., Stein, I.S., Brock, J.A., Castillo, P.E., Zito, K., and Sjöström, P.J. (2017). Unconventional NMDA Receptor Signaling. *J. Neurosci. Off. J. Soc. Neurosci.* 37, 10800–10807.

Dos Santos Morais, R., Delalande, O., Pérez, J., Mouret, L., Bondon, A., Martel, A., Appavou, M.-S., Le Rumeur, E., Hubert, J.-F., and Combet, S. (2017). Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS. *Langmuir ACS J. Surf. Colloids* *33*, 6572–6580.

Drin, G., Casella, J.-F., Gautier, R., Boehmer, T., Schwartz, T.U., and Antonny, B. (2007). A general amphipathic  $\alpha$ -helical motif for sensing membrane curvature. *Nat. Struct. Mol. Biol.* *14*, 138–146.

D'Souza, V.N., Nguyen, T.M., Morris, G.E., Karges, W., Pillers, D.A., and Ray, P.N. (1995). A novel dystrophin isoform is required for normal retinal electrophysiology. *Hum. Mol. Genet.* *4*, 837–842.

Dubreuil, R.R. (1991). Structure and evolution of the actin crosslinking proteins. *BioEssays* *13*, 219–226.

Duchenne (1868). Recherches sur la paralysie musculaire pseudohypertrophique ou paralysie myo-sclérotique. *Arch. Générales Médecine* *11* 179–209.

Duchesne, L., Gentili, D., Comes-Franchini, M., and Fernig, D.G. (2008). Robust Ligand Shells for Biological Applications of Gold Nanoparticles. *Langmuir* *24*, 13572–13580.

Duff, R.M., Tay, V., Hackman, P., Ravenscroft, G., McLean, C., Kennedy, P., Steinbach, A., Schöffler, W., van der Ven, P.F.M., Fürst, D.O., et al. (2011). Mutations in the N-terminal Actin-Binding Domain of Filamin C Cause a Distal Myopathy. *Am. J. Hum. Genet.* *88*, 729–740.

Dutcher, S.K. (2001). The tubulin fraternity: alpha to eta. *Curr. Opin. Cell Biol.* *13*, 49–54.

Dutcher, S.K., and Trabuco, E.C. (1998). The UNI3 gene is required for assembly of basal bodies of *Chlamydomonas* and encodes delta-tubulin, a new member of the tubulin superfamily. *Mol. Biol. Cell* *9*, 1293–1308.

Ebashi, S., and Ebashi, F. (1964). A New Protein Factor Promoting Contraction Of Actomyosin. *Nature* *203*, 645–646.

Endo, T., Imaizumi, T., Tagawa, T., Shiramoto, M., Ando, S., and Takeshita, A. (1994). Role of nitric oxide in exercise-induced vasodilation of the forearm. *Circulation* *90*, 2886–2890.

England, S.B., Nicholson, L.V., Johnson, M.A., Forrest, S.M., Love, D.R., Zubrzycka-Gaarn, E.E., Bulman, D.E., Harris, J.B., and Davies, K.E. (1990). Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* *343*, 180–182.

Ervasti, J.M. (2013). Structure and Function of the Dystrophin-Glycoprotein Complex (Landes Bioscience).

Ferey, N., Delalande, O., and Baaden, M. (2012). BioSpring: an interactive and multi-resolution software for flexible docking and for mechanical exploration of large biomolecular assemblies. (Inria Rennes - Bretagne Atlantique, Campus universitaire de Beaulieu 35042 Rennes Cedex), pp. 433–434.

Fiehn, W., Peter, J.B., Mead, J.F., and Gan-Elepano, M. (1971). Lipids and fatty acids of sarcolemma, sarcoplasmic reticulum, and mitochondria from rat skeletal muscle. *J. Biol. Chem.* *246*, 5617–5620.

Flores, S.C. (2014). Fast fitting to low resolution density maps: elucidating large-scale motions of the ribosome. *Nucleic Acids Res.* *42*, e9–e9.

Folker, E.S., and Baylies, M.K. (2013). Nuclear positioning in muscle development and disease. *Front. Physiol.* *4*, 363.

Franco, A., and Lansman, J.B. (1990). Calcium entry through stretch-inactivated ion channels in mdx myotubes. *Nature* *344*, 670–673.

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. (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* *50*, 1212–1225.

Frecha, C., Szécsi, J., Cosset, F.-L., and Verhoeven, E. (2008). Strategies for targeting lentiviral vectors. *Curr. Gene Ther.* *8*, 449–460.

Frishman, D., and Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins* *23*, 566–579.

Fuchs, E., and Karakesisoglou, I. (2001). Bridging cytoskeletal intersections. *Genes Dev.* *15*, 1–14.

Gadadhar, S., Bodakuntla, S., Natarajan, K., and Janke, C. (2017). The tubulin code at a glance. *J. Cell Sci.* *130*, 1347–1353.

Gaglio, T., Dionne, M.A., and Compton, D.A. (1997). Mitotic spindle poles are organized by structural and motor proteins in addition to centrosomes. *J. Cell Biol.* *138*, 1055–1066.

Galkin, V.E., Orlova, A., VanLoock, M.S., Rybakova, I.N., Ervasti, J.M., and Egelman, E.H. (2002). The utrophin actin-binding domain binds F-actin in two different modes: implications for the spectrin superfamily of proteins. *J. Cell Biol.* *157*, 243–251.

Ganguly, A., Yang, H., Sharma, R., Patel, K.D., and Cabral, F. (2012). The role of microtubules and their dynamics in cell migration. *J. Biol. Chem.* *287*, 43359–43369.

Gao, Q.Q., and McNally, E.M. (2011). The Dystrophin Complex: Structure, Function, and Implications for Therapy. In *Comprehensive Physiology*, (John Wiley & Sons, Inc.), p.

Geng, C., Narasimhan, S., Rodrigues, J.P.G.L.M., and Bonvin, A.M.J.J. (2017). Information-Driven, Ensemble Flexible Peptide Docking Using HADDOCK. In *Modeling Peptide-Protein Interactions*, (Humana Press, New York, NY), pp. 109–138.

Gentil, C., Leturcq, F., Ben Yaou, R., Kaplan, J.-C., Laforet, P., Pénisson-Besnier, I., Espil-Taris, C., Voit, T., Garcia, L., and Piétri-Rouxel, F. (2012). Variable phenotype of del45-55 Becker patients correlated with nNOS $\mu$  mislocalization and RYR1 hypernitrosylation. *Hum. Mol. Genet.* *21*, 3449–3460.

Ghose, A.K., and Crippen, G.M. (1987). Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. *J. Chem. Inf. Comput. Sci.* *27*, 21–35.

Ghose, A.K., Viswanadhan, V.N., and Wendoloski, J.J. (1998). Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of ALOGP and CLOGP Methods. *J. Phys. Chem. A* *102*, 3762–3772.

Giudice, E., Molza, A.-E., Laurin, Y., Nicolas, A., Le Rumeur, E., and Delalande, O. (2013). Molecular Clues about the Dystrophin–Neuronal Nitric Oxide Synthase Interaction: A Theoretical Approach. *Biochemistry* *52*, 7777–7784.

Górecki, D.C., Monaco, A.P., Derry, J.M., Walker, A.P., Barnard, E.A., and Barnard, P.J. (1992). Expression of four alternative dystrophin transcripts in brain regions regulated by different promoters. *Hum. Mol. Genet.* *1*, 505–510.

Green, S.J., Mellouk, S., Hoffman, S.L., Meltzer, M.S., and Nacy, C.A. (1990). Cellular mechanisms of nonspecific immunity to intracellular infection: Cytokine-induced synthesis of toxic nitrogen oxides from l-arginine by macrophages and hepatocytes. *Immunol. Lett.* *25*, 15–19.

Griffiths, W.J., Abdel-Khalik, J., Yutuc, E., Morgan, A.H., Gilmore, I., Hearn, T., and Wang, Y. (2017). Cholesterolomics: An update. *Anal. Biochem.* *524*, 56–67.

Grounds, M.D., and Torrisi, J. (2004). Anti-TNF $\alpha$  (Remicade) therapy protects dystrophic skeletal muscle from necrosis. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *18*, 676–682.

Grum, V.L., Li, D., MacDonald, R.I., and Mondragón, A. (1999). Structures of Two Repeats of Spectrin Suggest Models of Flexibility. *Cell* *98*, 523–535.

Guesdon, A., Bazile, F., Buey, R.M., Mohan, R., Monier, S., García, R.R., Angevin, M., Heichette, C., Wieneke, R., Tampé, R., et al. (2016). EB1 interacts with outwardly curved and straight regions of the microtubule lattice. *Nat. Cell Biol.* *18*, 1102–1108.

Gussoni, E., Soneoka, Y., Strickland, C.D., Buzney, E.A., Khan, M.K., Flint, A.F., Kunkel, L.M., and Mulligan, R.C. (1999). Dystrophin expression in the mdx mouse restored by stem cell transplantation. *Nature* *401*, 390–394.

Hack, A.A., Cordier, L., Shoturma, D.I., Lam, M.Y., Sweeney, H.L., and McNally, E.M. (1999). Muscle degeneration without mechanical injury in sarcoglycan deficiency. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 10723–10728.

Hainfeld, J.F., and Furuya, F.R. (1992). A 1.4-nm gold cluster covalently attached to antibodies improves immunolabeling. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* *40*, 177–184.

Hainfeld, J.F., and Powell, R.D. (1999). Nanogold Technology: New Frontiers in Gold Labeling. *Appl. Immunohistochem. Mol. Morphol.* *7*, 92.

Halgren, T.A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* *17*, 490–519.

Hammonds, R.G. (1987). Protein sequence of DMD gene is related to actin-binding domain of alpha-actinin. *Cell* *51*, 1.

Hanson, J., and Lowy, J. (1964). The structure of actin filaments and the origin of the axial periodicity in the I-substance of vertebrate striated muscle. *Proc R Soc Lond B* *160*, 449–460.

Harayama, T., and Riezman, H. (2018). Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* *19*, 281–296.

Hastie, E., and Samulski, R.J. (2015). Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success--a personal perspective. *Hum. Gene Ther.* *26*, 257–265.

Havens, M.A., and Hastings, M.L. (2016). Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.* *44*, 6549–6563.

Heier, J.A., Dickinson, D.J., and Kwiatkowski, A.V. (2017). Measuring Protein Binding to F-actin by Co-sedimentation. *J. Vis. Exp. JoVE*.

Heissler, S.M., and Sellers, J.R. (2016). Kinetic Adaptations of Myosins for their Diverse Cellular Functions. *Traffic Cph. Den.* *17*, 839–859.

Hemmings, L., Kuhlman, P., and Critchley, D. (1992). Analysis of the actin-binding domain of alpha-actinin by mutagenesis and demonstration that dystrophin contains a functionally homologous domain. *J. Cell Biol.* *116*, 1369–1380.

Henderson, D.M., Lin, A.Y., Thomas, D.D., and Ervasti, J.M. (2012). The Carboxy-Terminal Third Of Dystrophin Enhances Actin Binding Activity. *J. Mol. Biol.* *416*, 414–424.

Henry, M.D., and Campbell, K.P. (1999). Dystroglycan inside and out. *Curr. Opin. Cell Biol.* *11*, 602–607.

Herman, I.M. (1993). Actin isoforms. *Curr. Opin. Cell Biol.* *5*, 48–55.

Hoffman, E.P., Brown, R.H., and Kunkel, L.M. (1987a). Dystrophin: The protein product of the duchenne muscular dystrophy locus. *Cell* *51*, 919–928.

Hoffman, E.P., Monaco, A.P., Feener, C.C., and Kunkel, L.M. (1987b). Conservation of the Duchenne muscular dystrophy gene in mice and humans. *Science* *238*, 347–350.

Holmes, K.C., Popp, D., Gebhard, W., and Kabsch, W. (1990). Atomic model of the actin filament. *Nature* *347*, 44–49.

Hong, W., and Lev, S. (2014). Tethering the assembly of SNARE complexes. *Trends Cell Biol.* *24*, 35–43.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M.J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* *7*, 634–638.

Huard, J., Bouchard, J.P., Roy, R., Malouin, F., Dansereau, G., Labrecque, C., Albert, N., Richards, C.L., Lemieux, B., and Tremblay, J.P. (1992). Human myoblast transplantation: preliminary results of 4 cases. *Muscle Nerve* *15*, 550–560.

Ikonen, E. (2008). Cellular cholesterol trafficking and compartmentalization. *Nat. Rev. Mol. Cell Biol.* *9*, 125–138.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.

Joshi, H.C., and Cleveland, D.W. (1989). Differential utilization of beta-tubulin isoforms in differentiating neurites. *J. Cell Biol.* *109*, 663–673.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* *22*, 2577–2637.

Kaverina, I., and Straube, A. (2011). Regulation of cell migration by dynamic microtubules. *Semin. Cell Dev. Biol.* *22*, 968–974.

Kharraz, Y., Guerra, J., Pessina, P., Serrano, A.L., and Muñoz-Cánoves, P. (2014). Understanding the Process of Fibrosis in Duchenne Muscular Dystrophy. *BioMed Res. Int.* 2014.

Khurana, T.S., and Davies, K.E. (2003). Pharmacological strategies for muscular dystrophy. *Nat. Rev. Drug Discov.* 2, 379–390.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1156–1160.

Kimura (2010). 468. Lentiviral Vector Mediated Delivery of Full-Length Dystrophin for Gene Therapy of Muscular Dystrophy. *Mol. Ther.* 18, S180–S181.

Kirsch, J., and Betz, H. (1995). The postsynaptic localization of the glycine receptor-associated protein gephyrin is regulated by the cytoskeleton. *J. Neurosci. Off. J. Soc. Neurosci.* 15, 4148–4156.

Klip, A., Volchuk, A., He, L., and Tsakiridis, T. (1996). The glucose transporters of skeletal muscle. *Semin. Cell Dev. Biol.* 7, 229–237.

Koenig, M., and Kunkel, L.M. (1990). Detailed analysis of the repeat domain of dystrophin reveals four potential hinge segments that may confer flexibility. *J. Biol. Chem.* 265, 4560–4566.

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C., and Kunkel, L.M. (1987). Complete cloning of the duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* 50, 509–517.

Koenig, M., Monaco, A.P., and Kunkel, L.M. (1988). The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* 53, 219–228.

Körbelin, J., Sieber, T., Michelfelder, S., Lunding, L., Spies, E., Hunger, A., Alawi, M., Rapti, K., Indenbirken, D., Müller, O.J., et al. (2016). Pulmonary Targeting of Adeno-associated Viral Vectors by Next-generation Sequencing-guided Screening of Random Capsid Displayed Peptide Libraries. *Mol. Ther.* 24, 1050–1061.

Kouranova, E., Forbes, K., Zhao, G., Warren, J., Bartels, A., Wu, Y., and Cui, X. (2016). CRISPRs for Optimal Targeting: Delivery of CRISPR Components as DNA, RNA, and Protein into Cultured Cells and Single-Cell Embryos. *Hum. Gene Ther.* 27, 464–475.

Krag, T.O.B., Bogdanovich, S., Jensen, C.J., Fischer, M.D., Hansen-Schwartz, J., Javazon, E.H., Flake, A.W., Edvinsson, L., and Khurana, T.S. (2004). Heregulin ameliorates the dystrophic phenotype in mdx mice. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13856–13860.

Krieger, E., Koraimann, G., and Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins Struct. Funct. Bioinforma.* 47, 393–402.

Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., and Vriend, G. (2004). Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins Struct. Funct. Bioinforma.* 57, 678–683.

Kuriu, T., Inoue, A., Bito, H., Sobue, K., and Okabe, S. (2006). Differential control of postsynaptic density scaffolds via actin-dependent and -independent mechanisms. *J. Neurosci. Off. J. Soc. Neurosci.* *26*, 7693–7706.

Kusumi, A., Sako, Y., and Yamamoto, M. (1993). Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. *Biophys. J.* *65*, 2021–2040.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* *157*, 105–132.

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. (2009). 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.* *119*, 624–635.

Lai, Y., Zhao, J., Yue, Y., and Duan, D. (2013).  $\alpha 2$  and  $\alpha 3$  helices of dystrophin R16 and R17 frame a microdomain in the  $\alpha 1$  helix of dystrophin R17 for neuronal NOS binding. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 525–530.

Lapidos, K.A., Kakkar, R., and McNally, E.M. (2004). The Dystrophin Glycoprotein Complex: Signaling Strength and Integrity for the Sarcolemma. *Circ. Res.* *94*, 1023–1031.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* *26*, 283–291.

Lau, T.-L., Partridge, A.W., Ginsberg, M.H., and Ulmer, T.S. (2008). Structure of the integrin beta3 transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry* *47*, 4008–4016.

Le Rumeur, E., Fichou, Y., Pottier, S., Gaboriau, F., Rondeau-Mouro, C., Vincent, M., Gallay, J., and Bondon, A. (2003). Interaction of dystrophin rod domain with membrane phospholipids. Evidence of a close proximity between tryptophan residues and lipids. *J. Biol. Chem.* *278*, 5993–6001.

Le Rumeur, E., Pottier, S., Da Costa, G., Metzinger, L., Mouret, L., Rocher, C., Fourage, M., Rondeau-Mouro, C., and Bondon, A. (2007). Binding of the dystrophin second repeat to membrane di-oleyl phospholipids is dependent upon lipid packing. *Biochim. Biophys. Acta* *1768*, 648–654.

Lecine, P., Italiano, J.E., Kim, S.W., Villeval, J.L., and Shivdasani, R.A. (2000). Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. *Blood* *96*, 1366–1373.

Lederfein, D., Levy, Z., Augier, N., Mornet, D., Morris, G., Fuchs, O., Yaffe, D., and Nudel, U. (1992). A 71-kilodalton protein is a major product of the Duchenne muscular dystrophy gene in brain and other nonmuscle tissues. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 5346–5350.

Legardinier, S., Hubert, J.-F., Bihan, O.L., Tascon, C., Rocher, C., Raguénès-Nicol, C., Bondon, A., Hardy, S., and Rumeur, E.L. (2008). Sub-domains of the dystrophin rod domain display contrasting lipid-binding and stability properties. *Biochim. Biophys. Acta BBA - Proteins Proteomics* *1784*, 672–682.

Legardinier, S., Raguénès-Nicol, C., Tascon, C., Rocher, C., Hardy, S., Hubert, J.-F., and Le Rumeur, E. (2009). Mapping of the Lipid-Binding and Stability Properties of the Central Rod Domain of Human Dystrophin. *J. Mol. Biol.* *389*, 546–558.

Legrand, B., Giudice, E., Nicolas, A., Delalande, O., and Le Rumeur, E. (2011). Computational Study of the Human Dystrophin Repeats: Interaction Properties and Molecular Dynamics. *PLoS ONE* *6*, e23819.

Leikin, S., Kozlov, M.M., Fuller, N.L., and Rand, R.P. (1996). Measured effects of diacylglycerol on structural and elastic properties of phospholipid membranes. *Biophys. J.* *71*, 2623–2632.

Leo, A., Hansch, C., and Elkins, D. (1971). Partition coefficients and their uses. *Chem. Rev.* *71*, 525–616.

Leung, C.L., Green, K.J., and Liem, R.K.H. (2002). Plakins: a family of versatile cytolinker proteins. *Trends Cell Biol.* *12*, 37–45.

Leventis, P.A., and Grinstein, S. (2010). The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys.* *39*, 407–427.

Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., et al. (2015). Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Rep.* *4*, 143–154.

Li, W., Asokan, A., Wu, Z., Van Dyke, T., DiPrimio, N., Johnson, J.S., Govindaswamy, L., Agbandje-McKenna, M., Leichtle, S., Eugene Redmond Jr, D., et al. (2008). Engineering and Selection of Shuffled AAV Genomes: A New Strategy for Producing Targeted Biological Nanoparticles. *Mol. Ther.* *16*, 1252–1260.

Liao, G., Nagasaki, T., and Gundersen, G.G. (1995). Low concentrations of nocodazole interfere with fibroblast locomotion without significantly affecting microtubule level: implications for the role of dynamic microtubules in cell locomotion. *J. Cell Sci.* *108 ( Pt 11)*, 3473–3483.

Lidov, H.G., Selig, S., and Kunkel, L.M. (1995). Dp140: a novel 140 kDa CNS transcript from the dystrophin locus. *Hum. Mol. Genet.* *4*, 329–335.

Lim, K.R.Q., Maruyama, R., and Yokota, T. (2017). Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des. Devel. Ther.* *11*, 533–545.

Little, W., University College, L.L.S., and University College, L.L.S. (1853). On the nature and treatment of the deformities of the human frame [electronic resource] : being a course of lectures delivered at the Royal Orthopaedic Hospital in 1843 ; with numerous notes and additions to the present time (London : Longman, Brown, Green, and Longmans).

Liu, N., Xiong, Y., Ren, Y., Zhang, L., He, X., Wang, X., Liu, M., Li, D., Shui, W., and Zhou, J. (2015). Proteomic Profiling and Functional Characterization of Multiple Post-Translational Modifications of Tubulin. *J. Proteome Res.* *14*, 3292–3304.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). The Dynamics of Actin Assembly.



Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2015). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* aad5725.

Lupas, A.N., and Bassler, J. (2017). Coiled Coils – A Model System for the 21st Century. *Trends Biochem. Sci.* 42, 130–140.

Macé, K., Giudice, E., Chat, S., and Gillet, R. (2018). The structure of an elongation factor G-ribosome complex captured in the absence of inhibitors. *Nucleic Acids Res.* 46, 3211–3217.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 102, 3586–3616.

MacKerell, A.D., Banavali, N., and Foloppe, N. (2000). Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56, 257–265.

Maksymiuk, R., Sui, S.F., Gaub, H., and Sackmann, E. (1987). Electrostatic coupling of spectrin dimers to phosphatidylserine containing lipid lamellae. *Biochemistry* 26, 2983–2990.

Mann, C.J., Honeyman, K., Cheng, A.J., Ly, T., Lloyd, F., Fletcher, S., Morgan, J.E., Partridge, T.A., and Wilton, S.D. (2001). Antisense-induced exon skipping and synthesis of dystrophin in the mdx mouse. *Proc. Natl. Acad. Sci. U. S. A.* 98, 42–47.

Marchesi, V.T., and Steers, E. (1968). Selective solubilization of a protein component of the red cell membrane. *Science* 159, 203–204.

Marletta, M.A. (1993). Nitric oxide synthase structure and mechanism. *J. Biol. Chem.* 268, 12231–12234.

Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H. (2007). The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* 111, 7812–7824.

Martens, S., Kozlov, M.M., and McMahon, H.T. (2007). How synaptotagmin promotes membrane fusion. *Science* 316, 1205–1208.

Martín-García, F., Papaleo, E., Gomez-Puertas, P., Boomsma, W., and Lindorff-Larsen, K. (2015). Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS ONE* 10.

Marty, I., and Fauré, J. (2016). Excitation-Contraction Coupling Alterations in Myopathies. *J. Neuromuscul. Dis.* 3, 443–453.

McCloy, G., Moulton, H.M., Iversen, P.L., Fletcher, S., and Wilton, S.D. (2006). Antisense oligonucleotide-induced exon skipping restores dystrophin expression in vitro in a canine model of DMD. *Gene Ther.* 13, 1373–1381.

McGreevy, J.W., Hakim, C.H., McIntosh, M.A., and Duan, D. (2015). Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Dis. Model. Mech.* 8, 195–213.

Mendell, J.R., and Clark, K.R. (2006). Risks, benefits, and consent in the age of gene therapy. *Neurology* 66, 964–965.

Mendell, J.R., Goemans, N., Lowes, L.P., Alfano, L.N., Berry, K., Shao, J., Kaye, E.M., Mercuri, E., and Eteplirsen Study Group and Telethon Foundation DMD Italian Network (2016).

Longitudinal effect of eteplirsen versus historical control on ambulation in Duchenne muscular dystrophy. *Ann. Neurol.* *79*, 257–271.

Meng, J., Counsell, J.R., Reza, M., Laval, S.H., Danos, O., Thrasher, A., Lochmüller, H., Muntoni, F., and Morgan, J.E. (2016). Autologous skeletal muscle derived cells expressing a novel functional dystrophin provide a potential therapy for Duchenne Muscular Dystrophy. *Sci. Rep.* *6*, 19750.

Meryon, E. (1852). On Granular and Fatty Degeneration of the Voluntary Muscles. *Medico-Chir. Trans.* *35*, 73-84.1.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* *29*, 143–148.

Miller, R.G., Sharma, K.R., Pavlath, G.K., Gussoni, E., Mynhier, M., Lanctot, A.M., Greco, C.M., Steinman, L., and Blau, H.M. (1997). Myoblast implantation in Duchenne muscular dystrophy: the San Francisco study. *Muscle Nerve* *20*, 469–478.

Milligan, J.F., Matteucci, M.D., and Martin, J.C. (1993). Current concepts in antisense drug design. *J. Med. Chem.* *36*, 1923–1937.

Mimori-Kiyosue, Y., and Tsukita, S. (2003). “Search-and-capture” of microtubules through plus-end-binding proteins (+TIPs). *J. Biochem. (Tokyo)* *134*, 321–326.

Miralles, F., and Visa, N. (2006). Actin in transcription and transcription regulation. *Curr. Opin. Cell Biol.* *18*, 261–266.

Mirza, A., Sagathevan, M., Sahni, N., Choi, L., and Menhart, N. (2010). A biophysical map of the dystrophin rod. *Biochim. Biophys. Acta BBA - Proteins Proteomics* *1804*, 1796–1809.

Miskew Nichols, B., Aoki, Y., Kuraoka, M., Lee, J.J.A., Takeda, S., and Yokota, T. (2016). Multi-exon Skipping Using Cocktail Antisense Oligonucleotides in the Canine X-linked Muscular Dystrophy. *J. Vis. Exp. JoVE*.

Mitchison, T., and Kirschner, M. (1984). Dynamic instability of microtubule growth. *Nature* *312*, 237–242.

Mokri, B., and Engel, A.G. (1975). Duchenne dystrophy: electron microscopic findings pointing to a basic or early abnormality in the plasma membrane of the muscle fiber. *Neurology* *25*, 1111–1120.

Molza, A.-E., Férey, N., Czjzek, M., Le Rumeur, E., Hubert, J.-F., Tek, A., Laurent, B., Baaden, M., and Delalande, O. (2014). Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly. *Faraday Discuss.* *169*, 45–62.

Molza, A.-E., Mangat, K., Le Rumeur, E., Hubert, J.-F., Menhart, N., and Delalande, O. (2015). Structural Basis of Neuronal Nitric-oxide Synthase Interaction with Dystrophin Repeats 16 and 17. *J. Biol. Chem.* *290*, 29531–29541.

Monaco, A.P., Neve, R.L., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M., and Kunkel, L.M. (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* *323*, 646–650.

Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H., and Kunkel, L.M. (1988). An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* 2, 90–95.

Mongini, T., Ghigo, D., Doriguzzi, C., Bussolino, F., Pescarmona, G., Pollo, B., Schiffer, D., and Bosia, A. (1988). Free cytoplasmic Ca<sup>++</sup> at rest and after cholinergic stimulus is increased in cultured muscle cells from Duchenne muscular dystrophy patients. *Neurology* 38, 476–480.

Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.-J. (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* 4, 819–834.

Munshi, N.V. (2016). CRISPR (Clustered Regularly Interspaced Palindromic Repeat)/Cas9 System: A Revolutionary Disease-Modifying Technology. *Circulation* 134, 777–779.

Murphy, S., and Ohlendieck, K. (2016). The biochemical and mass spectrometric profiling of the dystrophin complexome from skeletal muscle. *Comput. Struct. Biotechnol. J.* 14, 20–27.

Murphy, S., Zweyer, M., Henry, M., Meleady, P., Mundegar, R.R., Swandulla, D., and Ohlendieck, K. (2018). Proteomic analysis of the sarcolemma-enriched fraction from dystrophic mdx-4cv skeletal muscle. *J. Proteomics*.

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.

Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Rivera, R.M.C., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2015). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* aad5143.

Newey, S.E., Benson, M.A., Ponting, C.P., Davies, K.E., and Blake, D.J. (2000). Alternative splicing of dystrobrevin regulates the stoichiometry of syntrophin binding to the dystrophin protein complex. *Curr. Biol. CB* 10, 1295–1298.

Nicolas, A., Delalande, O., Hubert, J.-F., and Le Rumeur, E. (2014a). The spectrin family of proteins: A unique coiled-coil fold for various molecular surface properties. *J. Struct. Biol.* 186, 392–401.

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. (2014b). Becker muscular dystrophy severity is linked to the structure of dystrophin. *Hum. Mol. Genet.*

Nicolas, G., Pedroni, S., Fournier, C., Gautero, H., Craescu, C., Dhermy, D., and Lecomte, M.C. (1998). Spectrin self-association site: characterization and study of beta-spectrin mutations associated with hereditary elliptocytosis. *Biochem. J.* 332, 81–89.

Nishimune, H., and Shigemoto, K. (2018). Practical Anatomy of the Neuromuscular Junction in Health and Disease. *Neurol. Clin.* 36, 231–240.

Norwood, F.L., Sutherland-Smith, A.J., Keep, N.H., and Kendrick-Jones, J. (2000). 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 8, 481–491.

Nudel, U., Zuk, D., Einat, P., Zeelon, E., Levy, Z., Neuman, S., and Yaffe, D. (1989). Duchenne muscular dystrophy gene product is not identical in muscle and brain. *Nature* 337, 76–78.

Nyman, T., Schüler, H., Korenbaum, E., Schutt, C.E., Karlsson, R., and Lindberg, U. (2002). The role of MeH73 in actin polymerization and ATP hydrolysis. *J. Mol. Biol.* *317*, 577–589.

Oakley, C.E., and Oakley, B.R. (1989). Identification of gamma-tubulin, a new member of the tubulin superfamily encoded by mipA gene of *Aspergillus nidulans*. *Nature* *338*, 662–664.

Ogura, Y., Tajrishi, M.M., Sato, S., Hindi, S.M., and Kumar, A. (2014). Therapeutic potential of matrix metalloproteinases in Duchenne muscular dystrophy. *Front. Cell Dev. Biol.* *2*.

Ohta, Y., Akiyama, T., Nishida, E., and Sakai, H. (1987). Protein kinase C and cAMP-dependent protein kinase induce opposite effects on actin polymerizability. *FEBS Lett.* *222*, 305–310.

Otterbein, L.R., Graceffa, P., and Dominguez, R. (2001). The Crystal Structure of Uncomplexed Actin in the ADP State. *Science* *293*, 708–711.

Ousterout, D.G., Kabadi, A.M., Thakore, P.I., Majoros, W.H., Reddy, T.E., and Gersbach, C.A. (2015). Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* *6*, 6244.

Paci, E., and Karplus, M. (2000). Unfolding proteins by external forces and temperature: The importance of topology and energetics. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 6521–6526.

Park, Y., Powell, R.T., Tripathi, D.N., Dere, R., Ho, T.H., Blasius, T.L., Chiang, Y.-C., Davis, I.J., Fahey, C.C., Hacker, K.E., et al. (2016). Dual Chromatin and Cytoskeletal Remodeling by SETD2. *Cell* *166*, 950–962.

Parry, D.A., Dixon, T.W., and Cohen, C. (1992). Analysis of the three-alpha-helix motif in the spectrin superfamily of proteins. *Biophys. J.* *61*, 858–867.

Partridge, T.A. (2004). Stem cell therapies for neuromuscular diseases. *Acta Neurol. Belg.* *104*, 141–147.

Pascual, J., Castresana, J., and Saraste, M. (1997). Evolution of the spectrin repeat. *BioEssays* *19*, 811–817.

Pastoret, C., and Sebillé, A. (1995). mdx mice show progressive weakness and muscle deterioration with age. *J. Neurol. Sci.* *129*, 97–105.

Percival, J.M., Gregorevic, P., Odom, G.L., Banks, G.B., Chamberlain, J.S., and Froehner, S.C. (2007). rAAV6-Microdystrophin Rescues Aberrant Golgi Complex Organization in mdx Skeletal Muscles. *Traffic* *8*, 1424–1439.

Péréon, Y., Mercier, S., and Magot, A. (2015). Physiopathologie de la dystrophie musculaire de Duchenne. *Arch. Pédiatrie* *22*, 12S18-12S23.

Periole, X., and Marrink, S.-J. (2013). The Martini Coarse-Grained Force Field. In *Biomolecular Simulations*, (Humana Press, Totowa, NJ), pp. 533–565.

Peter, B.J., Kent, H.M., Mills, I.G., Vallis, Y., Butler, P.J.G., Evans, P.R., and McMahon, H.T. (2004). BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* *303*, 495–499.

Petry, S., Groen, A.C., Ishihara, K., Mitchison, T.J., and Vale, R.D. (2013). Branching microtubule nucleation in *Xenopus* egg extracts mediated by augmin and TPX2. *Cell* *152*, 768–777.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* *26*, 1781–1802.

Piluso, G., Mirabella, M., Ricci, E., Belsito, A., Abbondanza, C., Servidei, S., Puca, A.A., Tonali, P., Puca, G.A., and Nigro, V. (2000).  $\gamma$ 1- and  $\gamma$ 2-Syntrophins, Two Novel Dystrophin-binding Proteins Localized in Neuronal Cells. *J. Biol. Chem.* *275*, 15851–15860.

Pizon, V., Gerbal, F., Diaz, C.C., and Karsenti, E. (2005). Microtubule-dependent transport and organization of sarcomeric myosin during skeletal muscle differentiation. *EMBO J.* *24*, 3781–3792.

Pollard, T.D., and Borisy, G.G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell* *112*, 453–465.

Pollard, T.D., and Cooper, J.A. (2009). Actin, a Central Player in Cell Shape and Movement. *Science* *326*, 1208–1212.

Pons, F., Nicholson, L.V.B., Robert, A., Voit, T., and Leger, J.J. (1993). Dystrophin and dystrophin-related protein (utrophin) distribution in normal and dystrophin-deficient skeletal muscles. *Neuromuscul. Disord.* *3*, 507–514.

Ponting, C.P., Blake, D.J., Davies, K.E., Kendrick-Jones, J., and Winder, S.J. (1996). ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem. Sci.* *21*, 11–13.

Prat, A.G., Bertorello, A.M., Ausiello, D.A., and Cantiello, H.F. (1993). Activation of epithelial Na<sup>+</sup> channels by protein kinase A requires actin filaments. *Am. J. Physiol.* *265*, C224–233.

Prins, K.W., Humston, J.L., Mehta, A., Tate, V., Ralston, E., and Ervasti, J.M. (2009). Dystrophin is a microtubule-associated protein. *J. Cell Biol.* *186*, 363–369.

PROSITE entry PS00018, P. entry P. (1990). PROSITE entry PS00018.

Prota, A.E., Magiera, M.M., Kuijpers, M., Bargsten, K., Frey, D., Wieser, M., Jaussi, R., Hoogenraad, C.C., Kammerer, R.A., Janke, C., et al. (2013). Structural basis of tubulin tyrosination by tubulin tyrosine ligase. *J. Cell Biol.* *200*, 259–270.

Pyrkov, T.V., Chugunov, A.O., Krylov, N.A., Nolde, D.E., and Efremov, R.G. (2009). PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organization of biomolecular complexes. *Bioinformatics* *25*, 1201–1202.

Raff, E.C., Hoyle, H.D., Popodi, E.M., and Turner, F.R. (2008). Axoneme beta-tubulin sequence determines attachment of outer dynein arms. *Curr. Biol. CB* *18*, 911–914.

Ragusa, R.J., Chow, C.K., and Porter, J.D. (1997). Oxidative stress as a potential pathogenic mechanism in an animal model of Duchenne muscular dystrophy. *Neuromuscul. Disord.* *NMD* *7*, 379–386.

Rapaport, D.C. (1997). Interactive molecular dynamics. *Phys. Stat. Mech. Its Appl.* *240*, 246–254.

Relizani, K., Griffith, G., Echevarría, L., Zarrouki, F., Facchinetti, P., Vaillend, C., Leumann, C., Garcia, L., and Goyenvallé, A. (2017). Efficacy and Safety Profile of Tricyclo-DNA Antisense Oligonucleotides in Duchenne Muscular Dystrophy Mouse Model. *Mol. Ther. Nucleic Acids* *8*, 144–157.

Renault, V., Piron-Hamelin, G., Forestier, C., DiDonna, S., Decary, S., Hentati, F., Saillant, G., Butler-Browne, G.S., and Mouly, V. (2000). Skeletal muscle regeneration and the mitotic clock. *Exp. Gerontol.* *35*, 711–719.

Rentschler, S., Linn, H., Deininger, K., Bedford, M.T., Espanel, X., and Sudol, M. (1999). The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan. *Biol. Chem.* *380*, 431–442.

Richfield, D. (2014). Medical gallery of David Richfield 2014. *WikiJournal Med.* *1*, 9.

Rose, J.A., Hoggan, M.D., and Shatkin, A.J. (1966). Nucleic acid from an adeno-associated virus: chemical and physical studies. *Proc. Natl. Acad. Sci. U. S. A.* *56*, 86–92.

Rougerie, P., Miskolci, V., and Cox, D. (2013). Generation of membrane structures during phagocytosis and chemotaxis of macrophages: role and regulation of the actin cytoskeleton. *Immunol. Rev.* *256*, 222–239.

Rubenstein, P.A. (1990). The functional importance of multiple actin isoforms. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *12*, 309–315.

Ruiz, F., Krzywicka, A., Klotz, C., Keller, A., Cohen, J., Koll, F., Balavoine, G., and Beisson, J. (2000). The SM19 gene, required for duplication of basal bodies in *Paramecium*, encodes a novel tubulin, eta-tubulin. *Curr. Biol. CB* *10*, 1451–1454.

Rybakova, I.N., Amann, K.J., and Ervasti, J.M. (1996). A new model for the interaction of dystrophin with F-actin. *J. Cell Biol.* *135*, 661–672.

Rybakova, I.N., Patel, J.R., and Ervasti, J.M. (2000). The Dystrophin Complex Forms a Mechanically Strong Link between the Sarcolemma and Costameric Actin. *J. Cell Biol.* *150*, 1209–1214.

Sadoulet-Puccio, H.M., and Kunkel, L.M. (1996). Dystrophin and Its Isoforms. *Brain Pathol.* *6*, 25–35.

Sadoulet-Puccio, H.M., Rajala, M., and Kunkel, L.M. (1997). Dystrobrevin and dystrophin: an interaction through coiled-coil motifs. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 12413–12418.

Sakamoto, M., Yuasa, K., Yoshimura, M., Yokota, T., Ikemoto, T., Suzuki, M., Dickson, G., Miyagoe-Suzuki, Y., and Takeda, S. (2002). Micro-dystrophin cDNA ameliorates dystrophic phenotypes when introduced into mdx mice as a transgene. *Biochem. Biophys. Res. Commun.* *293*, 1265–1272.

Saladin, A., Fiorucci, S., Poulain, P., Prévost, C., and Zacharias, M. (2009). PTools: an open-source molecular docking library. *BMC Struct. Biol.* *9*, 27.

Salomon-Ferrer, R., Case, D.A., and Walker, R.C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* *3*, 198–210.

Sampaolesi, M., Blot, S., D'Antona, G., Granger, N., Tonlorenzi, R., Innocenzi, A., Mognol, P., Thibaud, J.-L., Galvez, B.G., Barthélémy, I., et al. (2006). Mesoangioblast stem cells ameliorate muscle function in dystrophic dogs. *Nature* *444*, 574–579.

Sarkis, J., Hubert, J.-F., Legrand, B., Robert, E., Chéron, A., Jardin, J., Hitti, E., Le Rumeur, E., and Vié, V. (2011). Spectrin-like repeats 11-15 of human dystrophin show adaptations to a lipidic environment. *J. Biol. Chem.* *286*, 30481–30491.

Sarkis, J., Vié, V., Winder, S.J., Renault, A., Rumeur, E.L., and Hubert, J.-F. (2013). Resisting sarcolemmal rupture: dystrophin repeats increase membrane-actin stiffness. *FASEB J.* *27*, 359–367.

Schwer, H.D., Lecine, P., Tiwari, S., Italiano, J.E., Hartwig, J.H., and Shivdasani, R.A. (2001). A lineage-restricted and divergent beta-tubulin isoform is essential for the biogenesis, structure and function of blood platelets. *Curr. Biol. CB* *11*, 579–586.

Seo, M.-D., Velamakanni, S., Ishiyama, N., Stathopoulos, P.B., Rossi, A.M., Khan, S.A., Dale, P., Li, C., Ames, J.B., Ikura, M., et al. (2012). Structural and functional conservation of key domains in InsP<sub>3</sub> and ryanodine receptors. *Nature* *483*, 108–112.

Serdaroglu, A., Gücüyener, K., Erdem, S., Köse, G., Tan, E., and Okuyaz, Ç. (2002). Role of Apoptosis in Duchenne’s Muscular Dystrophy. *J. Child Neurol.* *17*, 66–68.

Setny, P., Bahadur, R.P., and Zacharias, M. (2012). Protein-DNA docking with a coarse-grained force field. *BMC Bioinformatics* *13*, 228.

Shams, H., Golji, J., Garakani, K., and Mofrad, M.K. (2016). Dynamic Regulation of  $\alpha$ -Actinin’s Calponin Homology Domains on F-Actin. *Biophys. J.* *110*, 1444–1455.

Shaw, D.E., Grossman, J.P., Bank, J.A., Batson, B., Butts, J.A., Chao, J.C., Deneroff, M.M., Dror, R.O., Even, A., Fenton, C.H., et al. (2014). Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. In *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 41–53.

Shen, W., Wu, B., Zhang, Z., Dou, Y., Rao, Z.-R., Chen, Y.-R., and Duan, S. (2006). Activity-induced rapid synaptic maturation mediated by presynaptic cdc42 signaling. *Neuron* *50*, 401–414.

Shi, Y., Fernandez-Martinez, J., Tjioe, E., Pellarin, R., Kim, S.J., Williams, R., Schneidman-Duhovny, D., Sali, A., Rout, M.P., and Chait, B.T. (2014). Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell. Proteomics MCP* *13*, 2927–2943.

Shunchang, S., Haitao, C., Weidong, C., Jingbo, H., and Yunsheng, P. (2008). Expression of truncated dystrophin cDNAs mediated by a lentiviral vector. *Neurol. India* *56*, 52.

Sicinski, P., Geng, Y., Ryder-Cook, A.S., Barnard, E.A., Darlison, M.G., and Barnard, P.J. (1989). The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science* *244*, 1578–1580.

Singer, S.J., and Nicolson, G.L. (1972). The fluid mosaic model of the structure of cell membranes. *Science* *175*, 720–731.

Singh, S.M., Bandi, S., and Mallela, K.M.G. (2017). The N-Terminal Flanking Region Modulates the Actin Binding Affinity of the Utrophin Tandem Calponin-Homology Domain. *Biochemistry* *56*, 2627–2636.

Song, L.-S., Guatimosim, S., Gómez-Viquez, L., Sobie, E.A., Ziman, A., Hartmann, H., and Lederer, W.J. (2005). Calcium biology of the transverse tubules in heart. *Ann. N. Y. Acad. Sci.* *1047*, 99–111.

Speicher, D.W., and Marchesi, V.T. (1984). Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature* *311*, 177–180.

Spencer, M.J., Montecino-Rodriguez, E., Dorshkind, K., and Tidball, J.G. (2001). Helper (CD4(+)) and cytotoxic (CD8(+)) T cells promote the pathology of dystrophin-deficient muscle. *Clin. Immunol. Orlando Fla* 98, 235–243.

Stahl, B.T., Benekareddy, M., Coulon-Bainier, C., Banfal, A.A., Floor, S.N., Sabo, J.K., Urnes, C., Munares, G.A., Ghosh, A., and Doudna, J.A. (2017). Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. *Nat. Biotechnol.* 35, 431–434.

Stone, M.R., O'Neill, A., Catino, D., and Bloch, R.J. (2005). Specific Interaction of the Actin-binding Domain of Dystrophin with Intermediate Filaments Containing Keratin 19. *Mol. Biol. Cell* 16, 4280–4293.

Sun, H. (1998). COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase Applications Overview with Details on Alkane and Benzene Compounds. *J. Phys. Chem. B* 102, 7338–7364.

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. (2015). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 407–411.

Taylor, W.R. (1986). The classification of amino acid conservation. *J. Theor. Biol.* 119, 205–218.

Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., and Tufféry, P. (2012). PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res.* 40, W288–293.

Tinsley, J.M., and Davies, K.E. (1993). Utrophin: A potential replacement for dystrophin? *Neuromuscul. Disord.* 3, 537–539.

Tinsley, J.M., Blake, D.J., Roche, A., Fairbrother, U., Riss, J., Byth, B.C., Knight, A.E., Kendrick-Jones, J., Suthers, G.K., Love, D.R., et al. (1992). Primary structure of dystrophin-related protein. *Nature* 360, 591–593.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics. *Struct. Lond. Engl.* 1993 16, 673–683.

Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* 31, 455–461.

Truebestein, L., and Leonard, T.A. (2016). Coiled-coils: The long and short of it. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 38, 903–916.

Tufty, R.M., and Kretsinger, R.H. (1975). Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. *Science* 187, 167–169.

Tutdibi, O., Brinkmeier, H., Rüdell, R., and Föhr, K.J. (1999). Increased calcium entry into dystrophin-deficient muscle fibres of MDX and ADR-MDX mice is reduced by ion channel blockers. *J. Physiol.* 515, 859–868.

Ursu, D., Sebille, S., Dietze, B., Freise, D., Flockerzi, V., and Melzer, W. (2001). Excitation-contraction coupling in skeletal muscle of a mouse lacking the dihydropyridine receptor subunit gamma1. *J. Physiol.* 533, 367–377.



Uusitalo, J.J., Ingólfsson, H.I., Akhshi, P., Tieleman, D.P., and Marrink, S.J. (2015). Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* *11*, 3932–3945.

Vanni, S., Vamparys, L., Gautier, R., Drin, G., Etchebest, C., Fuchs, P.F.J., and Antony, B. (2013). Amphipathic lipid packing sensor motifs: Probing bilayer defects with hydrophobic residues. *Biophys. J.* *104*, 575–584.

Vasiliev, J.M., Gelfand, I.M., Domnina, L.V., Ivanova, O.Y., Komm, S.G., and Olshevskaja, L.V. (1970). Effect of colcemid on the locomotory behaviour of fibroblasts. *J. Embryol. Exp. Morphol.* *24*, 625–640.

Vaughan, S., Attwood, T., Navarro, M., Scott, V., McKean, P., and Gull, K. (2000). New tubulins in protozoal parasites. *Curr. Biol. CB* *10*, R258-259.

Vickers, T.A., and Crooke, S.T. (2015). The rates of the major steps in the molecular mechanism of RNase H1-dependent antisense oligonucleotide induced degradation of RNA. *Nucleic Acids Res.* *43*, 8955–8963.

Vié, V., Legardinier, S., Chieze, L., Le Bihan, O., Qin, Y., Sarkis, J., Hubert, J.-F., Renault, A., Desbat, B., and Le Rumeur, E. (2010). 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 BBA - Biomembr.* *1798*, 1503–1511.

Vincent, S.R. (2010). Nitric oxide neurons and neurotransmission. *Prog. Neurobiol.* *90*, 246–255.

Voit, T., Topaloglu, H., Straub, V., Muntoni, F., Deconinck, N., Champion, G., De Kimpe, S.J., Eagle, M., Guglieri, M., Hood, S., et al. (2014). Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *Lancet Neurol.* *13*, 987–996.

Wade, R.H. (2009). On and around microtubules: an overview. *Mol. Biotechnol.* *43*, 177–191.

Walshaw, J., and Woolfson, D.N. (2003). Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J. Struct. Biol.* *144*, 349–361.

Wassenaar, T.A., Pluhackova, K., Böckmann, R.A., Marrink, S.J., and Tieleman, D.P. (2014). Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* *10*, 676–690.

Watanabe, T., Wang, S., Noritake, J., Sato, K., Fukata, M., Takefuji, M., Nakagawa, M., Izumi, N., Akiyama, T., and Kaibuchi, K. (2004). Interaction with IQGAP1 links APC to Rac1, Cdc42, and actin filaments during cell polarization and migration. *Dev. Cell* *7*, 871–883.

Waters, J.C., Mitchison, T.J., Rieder, C.L., and Salmon, E.D. (1996). The kinetochore microtubule minus-end disassembly associated with poleward flux produces a force that can do work. *Mol. Biol. Cell* *7*, 1547–1558.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* *171*, 737–738.

Way, M., Pope, B., Cross, R.A., Kendrick-Jones, J., and Weeds, A.G. (1992). Expression of the N-terminal domain of dystrophin in *E. coli* and demonstration of binding to F-actin. *FEBS Lett.* *301*, 243–245.

Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. *Cell* 100, 157–168.

Winder, S.J., Gibson, T.J., and Kendrick-Jones, J. (1995). Dystrophin and utrophin: the missing links! *FEBS Lett.* 369, 27–33.

Wise, D.O., Krahe, R., and Oakley, B.R. (2000). The gamma-tubulin gene family in humans. *Genomics* 67, 164–170.

Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S.C., and Branton, D. (1993). Crystal structure of the repetitive segments of spectrin. *Science* 262, 2027–2030.

Yáñez-Muñoz, R.J., Balagán, K.S., MacNeil, A., Howe, S.J., Schmidt, M., Smith, A.J., Buch, P., MacLaren, R.E., Anderson, P.N., Barker, S.E., et al. (2006). Effective gene therapy with nonintegrating lentiviral vectors. *Nat. Med.* 12, 348–353.

Yang, J., and Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* 43, W174–W181.

Yao, J., Qi, J., and Chen, G. (2006). Actin-dependent activation of presynaptic silent synapses contributes to long-term synaptic plasticity in developing hippocampal neurons. *J. Neurosci. Off. J. Soc. Neurosci.* 26, 8137–8147.

Yao, X., Grade, S., Wriggers, W., and Rubenstein, P.A. (1999). His(73), often methylated, is an important structural determinant for actin. A mutagenic analysis of HIS(73) of yeast actin. *J. Biol. Chem.* 274, 37443–37449.

Yao, X., Nguyen, V., Wriggers, W., and Rubenstein, P.A. (2002). Regulation of yeast actin behavior by interaction of charged residues across the interdomain cleft. *J. Biol. Chem.* 277, 22875–22882.

Yokota, T., Lu, Q.-L., Partridge, T., Kobayashi, M., Nakamura, A., Takeda, S., and Hoffman, E. (2009). Efficacy of systemic morpholino exon-skipping in Duchenne dystrophy dogs. *Ann. Neurol.* 65, 667–676.

Yokota, T., Duddy, W., Echigoya, Y., and Kolski, H. (2012). Exon skipping for nonsense mutations in Duchenne muscular dystrophy: too many mutations, too few patients? *Expert Opin. Biol. Ther.* 12, 1141–1152.

Yoo, S.K., Lam, P.-Y., Eichelberg, M.R., Zasadil, L., Bement, W.M., and Huttenlocher, A. (2012). The role of microtubules in neutrophil polarity and migration in live zebrafish. *J. Cell Sci.* 125, 5702–5710.

Yoshida, M., and Ozawa, E. (1990). Glycoprotein complex anchoring dystrophin to sarcolemma. *J. Biochem. (Tokyo)* 108, 748–752.

Young, T. (1805). III. An essay on the cohesion of fluids. *Philos. Trans. R. Soc. Lond.* 95, 65–87.

Yucel, N., Chang, A.C., Day, J.W., Rosenthal, N., and Blau, H.M. (2018). Humanizing the mdx mouse model of DMD: the long and the short of it. *Npj Regen. Med.* 3, 4.

Zacharias, M. (2005). ATTRACT: Protein–protein docking in CAPRI using a reduced protein model. *Proteins Struct. Funct. Bioinforma.* 60, 252–256.

Zatz, M., Rapaport, D., Vainzof, M., Passos-Bueno, M.R., Bortolini, E.R., Pavanello, R. de C., and Peres, C.A. (1991). Serum creatine-kinase (CK) and pyruvate-kinase (PK) activities in

Duchenne (DMD) as compared with Becker (BMD) muscular dystrophy. *J. Neurol. Sci.* *102*, 190–196.

Zhang, R., Zhang, C., Zhao, Q., and Li, D. (2013). Spectrin: Structure, function and disease. *Sci. China Life Sci.* *56*, 1076–1085.

Zhao, J., Kodippili, K., Yue, Y., Hakim, C.H., Wasala, L., Pan, X., Zhang, K., Yang, N.N., Duan, D., and Lai, Y. (2016). Dystrophin contains multiple independent membrane-binding domains. *Hum. Mol. Genet.* *25*, 3647–3653.

Zhu, Q., and Asaro, R.J. (2008). Spectrin Folding versus Unfolding Reactions and RBC Membrane Stiffness. *Biophys. J.* *94*, 2529–2545.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastiris, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* *428*, 720–725.





## Annexe : Dos Santos Morais *et al*, 2018

Au cours de la thèse de Raphael Dos-Santos Morais, il a été montré que le fragment R1-3 de la dystrophine subit des changements de conformations importants en présence de bicelles anioniques.

J'ai activement contribué à cette étude, notamment en réalisant les études suivantes :

Production et analyses des simulations de dynamiques moléculaire gros grain de complexes R1-3 / bicelles anioniques ou zwitterioniques

Mise en place des scénarios de simulations de dynamiques moléculaires interactives pour l'ouverture du faisceau d'hélice

Réalisation du protocole de simulations de dynamiques moléculaires interactives pour l'ouverture du faisceau d'hélice

Analyse des résultats des simulations de dynamiques moléculaires interactives

### **Human dystrophin structural changes upon binding to anionic membrane lipids**

R. Dos Santos Morais, O. Delalande, J. Pérez, D. Mias-Lucquin, M. Lagarrigue, A. Martel, AE. Molza, C. Raguénès-Nicol, A. Chéron, T. Chenuel, A. Bondon, MS. Appavou, E. Le Rumeur, S. Combet\*, and JF. Hubert\*.

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

<sup>1</sup>Université de Rennes 1, F-35043 Rennes, France. <sup>2</sup>Institut de Génétique et Développement de Rennes, CNRS UMR 6290, F-35043 Rennes, France. <sup>3</sup>Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, F-91191 Gif-sur-Yvette CEDEX, France. <sup>4</sup>SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP 48, Saint-Aubin, F-91192 Gif-sur-Yvette, France. <sup>5</sup>Inserm U1085, Protim-Plate-forme Protéomique. <sup>6</sup>Institut Laue-Langevin, F-38042 Grenoble, France. <sup>7</sup>CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-35043 Rennes, France. <sup>8</sup>Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany.

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

**Short title: lipid-driven dystrophin 3D changes**

## ABSTRACT

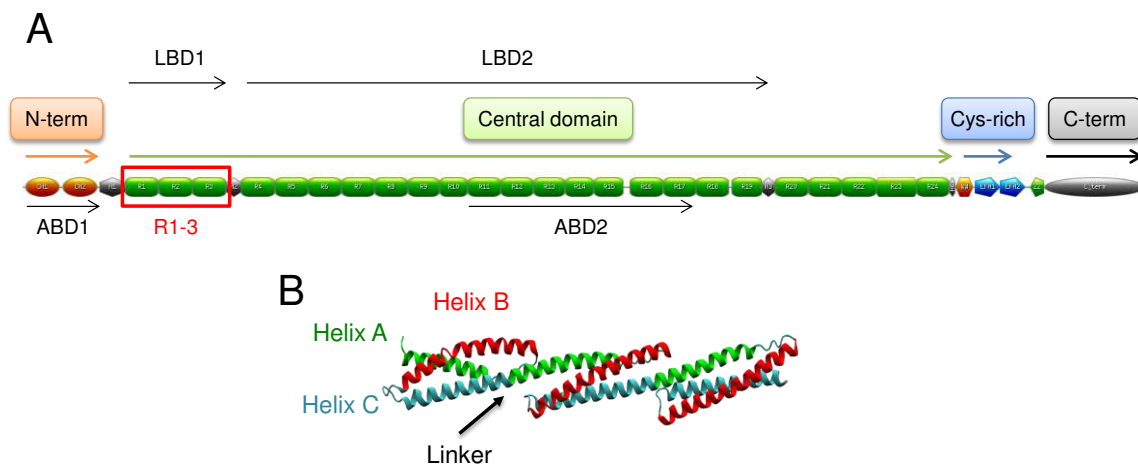
Scaffolding proteins play important roles in supporting the plasma membrane (sarcolemma) of muscle cells. Among them, dystrophin strengthens the sarcolemma through protein-lipid interactions, while its absence due to gene mutations leads to the severe Duchenne muscular dystrophy. Most of the dystrophin protein consists of a central domain made of 24 spectrin-like coiled-coil repeats (R). Using small-angle neutron scattering (SANS) and the contrast variation technique, we specifically probed the structure of the three first consecutive repeats 1 to 3 (R1-3), a part of dystrophin known to physiologically interact with membrane lipids. R1-3 free in solution was compared to its structure adopted in the presence of phospholipid-based bicelles. SANS data for the protein/lipid complexes were obtained with contrast-matched bicelles under various phospholipid compositions to probe the role of electrostatic interactions. When bound to anionic bicelles, large modifications of the protein 3D-structure were detected, as revealed by a significant increase of the protein gyration radius, from  $42 \pm 1$  to  $60 \pm 4$  Å. R1-3/anionic bicelle complexes were further analyzed by coarse-grained molecular dynamics simulations. From these studies, we report an all-atom model of R1-3 that highlights the opening of R1 coiled-coil repeat when bound to the membrane lipids. This model is totally in agreement with SANS and click chemistry/mass spectrometry data. We conclude that the sarcolemma membrane anchoring that occurs during the contraction/elongation process of muscles could be ensured by this coiled-coil opening. Therefore, understanding these structural changes may help in the design of rationalized shortened dystrophins for gene therapy. Finally, our strategy opens up new possibilities for structure determination of peripheral and integral membrane proteins not compatible with different high resolution structural methods.

## INTRODUCTION

Obtaining structural information of membrane proteins is very challenging due to the difficulties associated with their handling (1). Only ~770 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 (2). Furthermore, for membrane proteins not compatible with X-ray crystallography (XRC), NMR, and/or the increasingly used cryo-electron microscopy (3), the challenge could be even trickier. To overcome this, small-angle X-ray or neutron scattering (SAXS or SANS, respectively) are alternative techniques to obtain structural information on biological macromolecules, either free in solution, or bound to various partners (4), such as membrane proteins (5, 6). The SANS-based contrast variation technique (7) allows to specifically obtain low-resolution structures of proteins involved in macromolecular complexes and has been successfully applied to protein-polymer (8), protein-protein (7), protein-DNA/RNA (9, 10), and protein-lipid/detergent complexes (11–13). We highlighted recently that phospholipid-based bicelles can be contrast-matched to be used in SANS experiments as a relevant membrane mimic to probe conformational modifications of peripheral membrane proteins in interaction with phospholipids (14). In the present study, this approach was developed to explore the structure-function relationship of a protein associated with severe genetic diseases, *i.e.* dystrophin (Fig. 1A), a peripheral scaffolding membrane protein essential to protect muscle cell membrane (sarcolemma) against shear stresses (15). Indeed, its absence leads to the severe Duchenne muscular dystrophy (DMD) (16), with an incidence of ~1 in 5,000 male births (17).



Dystrophin consists mainly of a central domain that accounts for 75% of the protein sequence and has the particularity to be made of 24 spectrin-like repeats (denoted R1 to R24) (Fig. 1B). This protein is able to interact directly with the lipid part of the sarcolemma (18, 19). *In vitro* experiments showed that the region from R1 to R19 is able to bind phospholipids (20, 21). Recently, *in vivo* experiments highlighted that R1-3 is exclusively localized at the sarcolemma (19). Combination of these observations led to the conclusion that interactions of the central domain 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 (22). These regions present a high degree of curvature leading to a local lipid packing defect. The influence of this phenomenon on the dystrophin/lipid interactions has been investigated with liposomes of various sizes (18). Our previous works highlighted that the interfacial properties of dystrophin are modulated according to the region of the protein involved (23), the nature of lipids, and the lipid packing (20, 21), which all play key roles in the physiology of muscle cells. Moreover, the R1-3 protein fragment is bordered by two non-structured hinges and is found in several potentially therapeutic shortened dystrophins, making it a relevant protein subdomain (24, 25).



**Fig. 1.** (A) Schematic representation of dystrophin and its four domains, including the central domain composed of 24 spectrin-like repeats (R). The R1-3 protein fragment is framed in red and can interact with membrane phospholipids. LBD: lipid binding domain. ABD: actin binding domain. (B) 3D-structure of three spectrin repeats folded in triple coiled-coil (PDB: 1U4Q) showing the organization of dystrophin central domain. The linker region is the junction between the helix C of one repeat and the helix A of the subsequent one.

By combining SAXS and molecular modeling, we recently proposed solution structures of fragments of the central domain, including the R1-3 fragment (26). We obtained evidence of the conformation adopted by these protein fragments free in solution. However, an exhaustive knowledge of the structure of the dystrophin central domain, either alone, or in interaction with membrane phospholipids must be considered for understanding its crucial scaffolding role in muscle cells and by the way for helping for the design of therapeutic proteins.

In the present study, we probed the 3D-structure of the R1-3 dystrophin lipid-binding fragment, either free in solution, or in interaction with zwitterionic or anionic bicelles, made of

DMPC/DHPC or DMPC/DMPS/DHPC, respectively. PC and PS phospholipids are both well represented in the sarcolemma (27). PC is the most abundant phospholipid in the inner leaflet, representing 45% of the total phospholipids, whereas PE and PS represent 23 and 18%, respectively, and the presence of unsaturated fatty acids contribute to the packing defects that are observed in plasma membranes (27–30). Within the bicelles, the structure of the torus is made of short fatty acids, which reproduce the lipid packing defects of native sarcolemma. These considerations make the phospholipid-based bicelles relevant membrane mimics for both electrostatic and lipid-packing properties (14). Using a combination of SANS, click chemistry/mass spectrometry (MS), coarse-grained molecular dynamics (CG-MD), and interactive CG-MD (CG-IMD) simulations, we show that the tertiary structure of R1-3 is maintained upon binding to zwitterionic phospholipids, whereas, in the presence of anionic phospholipids, R1-3 exhibits significant conformational modifications. We report an all-atom model of R1-3 in interaction with anionic membrane phospholipids, where its structural changes are attributed to an opening of the coiled-coil of repeat 1 (R1).

## MATERIAL AND METHODS

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 “ $q_{eff}$ ” (31, 32), for effective molar ratio, when the proportion of free DHPC is considered and is defined by the equation Eq. 1.

$$q_{eff} = \frac{[LCP]}{[DHPC]_{total} - [DHPC]_{free}} \quad Eq. 1$$

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.* (33), based on the volumes of the phospholipids, was used to estimate the radius and the MW of the bicelles by the equation Eq. 2 (33, 34).

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

where R is the radius of the bicelle,  $r_{\perp}$  is the thickness of the rim (11 Å, the length of a DHPC molecule), and  $\Lambda$  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 (35), 979 (36), and 660 (37) Å<sup>3</sup>, respectively. See reference (14).

**Protein purification.** R1-3 (333 residues, human dystrophin (Uniprot P11532) *GSEVNLD...QISQA*, the residues in italic being from thrombin cleavage site) is designed according to the alignment of Winder *et al.* (38). R1-3 was expressed and purified as previously described with small modifications (21, 39). Briefly, the His-tagged protein was produced in BL21 *E. coli* (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 protein was 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 a molar extinction coefficient at 280 nm of  $59,720 \text{ M}^{-1} \cdot \text{cm}^{-1}$  and a molecular weight of 38.5 kDa (21). 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-glycéro-3-phosphocholine (DMPC), 1,2-dimyristoyl-sn-glycéro-3-phospho-L-serine (DMPS), 1,2-dihexanoyl-sn-glycéro-3-phosphocholine (DHPC), 1,2-dimyristoyl-d54-sn-glycéro-3-phosphocholine-1,1,2,2-d4-N,N,N-trimethyl-d9 (DMPC-d67), 1,2-dimyristoyl-d54-sn-glycéro-3-[phospho-L-serine] (DMPS-d54), and 1,2-dihexanoyl-d22-sn-glycéro-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<sub>2</sub>O, 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<sub>2</sub>, 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 versions, denoted HZB/DZB and HAB/DAB, respectively. Stock solutions were diluted in TNE or d-TNE depending on the experiments.

**Intrinsic tryptophan (Trp) fluorescence.** Fluorescence measurements were carried out on a Fluorolog spectrofluorometer (Jobin-Yvon). Tryptophan fluorescence emission spectra were recorded in low volume quartz Hellma cells (120  $\mu\text{L}$ ) 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  $\mu\text{M}$  and the total lipid concentration was 50 mM (~150  $\mu\text{M}$  of bicellar objects). The measurement temperature was 26°C.

**Microscale thermophoresis (MST).** Bicelles were labeled with 1  $\mu\text{M}$  of 1,2-Dioléoyl-sn-glycéro-3-phospho-L-serine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl) for a final concentration of 15 mM of lipids (~50  $\mu\text{M}$  of bicellar objects) at  $q_{\text{eff}} = 1.3$  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 R1-3 from 338 to 4.5  $\mu\text{M}$ . The 16 solutions were then loaded into Monolith NT premium capillaries (NanoTemper) and thermophoresis was measured with a Monolith NT.115 instrument (NanoTemper). 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 (NanoTemper). From the fit of the MST data, the concentration of the binding site ranges from 50 to 100  $\mu\text{M}$ . 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  $q_{\text{eff}}$ . Indeed, working at 15 mM of lipids tends to be close to the  $[\text{DHPC}]_{\text{free}}$  that could drastically affect  $q_{\text{eff}}$  even for a small misestimate of the total lipid concentration. The measurement temperature was 26°C.

**Far-UV circular dichroism (CD) analysis.** Far-UV CD spectra of the proteins were recorded on a Jasco J-815 spectropolarimeter 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(20, 40). The protein concentration was 20  $\mu\text{M}$  and the total lipid concentration was 50 mM (~150  $\mu\text{M}$  of bicellar objects). The measurement temperature was 26°C.

**Small-angle neutron scattering (SANS) experiments.** Preparatory works were carried out at KWS-2 instrument (MLZ, Garching) while final investigations were done at PACE (LLB, Saclay), D22 (ILL, Grenoble), and 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<sup>-1</sup> for the largest one, where  $Q = \frac{4\pi \sin \theta}{\lambda}$  is the momentum transfer,  $\lambda$  is the wavelength, and  $2\theta$  is the scattering angle. All measurements were performed in 1 mm thickness Hellma QS cells, and the intensities obtained are in absolute units (cm<sup>-1</sup>). 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. The protein concentration was 103  $\mu\text{M}$  (4.2 g/L) and the measurement temperature was 22°C.

**SANS data analysis.** SANS data were analyzed with the ATSAS suite (41) 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 with the following equation Eq. 3 (42):

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

The apparent molecular weight (MW, in g·mol<sup>-1</sup>) of the protein was estimated with the following equation Eq. 4 (43):

$$MW = \frac{I(0) \cdot N_A}{c \cdot (\Delta\rho V_P)^2} \quad \text{Eq. 4}$$

where  $\Delta\rho$  (cm<sup>-2</sup>) is the contrast (*i.e.*, the difference in the neutron scattering length densities, SLD) between R1-3 and d-TNE buffer,  $V_P$  is the partial specific volume (cm<sup>3</sup>·g<sup>-1</sup>) of the protein

fragments,  $c$  is the protein concentration ( $\text{g}\cdot\text{cm}^{-3}$ ), and  $N_A$  is the Avogadro number ( $\text{mol}^{-1}$ ). The  $V_P$  was determined from the density of the protein according to its amino acid sequence (<http://psldc.isis.rl.ac.uk/Psldc/>). Kratky and Porod-Debye SANS plots confirmed the well-folded and compact states of R1-3 (Fig. S1). Moreover, the calculated apparent MWs (~45 kDa) in the presence of both types of bicelles exclude the possibility of the dimerization upon lipid binding. Fitting procedures were performed using the SASview v3.1.0 software (<http://www.sasview.org/>). Bicelle data were fitted with the “core-shell cylinder” form factor model (44). The GNOM module from the ATSAS suite was used to determine the pair-distribution function  $P(r)$ , the maximum distance  $D_{\text{max}}$ , and the  $R_g$  in real space (denoted  $R_g^{\text{real}}$ ) from  $P(r)$ . Twenty *ab initio* models were generated with the DAMMIF software considering the data up to  $Q = 0.25 \text{ \AA}^{-1}$ . The models obtained were aligned and averaged with the DAMAVER software. The presented *ab initio* models correspond to the DAMMIF model with the smallest normalized spatial discrepancy (NSD) surrounded by the corresponding DAMAVER model.

**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 (45) (Fig. S2A in the Supporting Material). The modified bicelles were mixed with R1-3 to reach a final concentration of 50 mM of lipids and 15  $\mu\text{M}$  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 (Fig. S2B), 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 (46). The peptide extract was loaded (10  $\mu\text{L}$ ) on a nano-HPLC system (LC Packings Ultimate 3000, Dionex) equipped with a trapping precolumn (5 mm x 300  $\mu\text{m}$  id, 300  $\text{\AA}$  pore size, Pepmap C18, 5  $\mu\text{m}$ , ThermoScientific) and an analytical column (15 cm x 75  $\mu\text{m}$  id, 300  $\text{\AA}$  pore size, Pepmap C18, 5  $\mu\text{m}$ , ThermoScientific). Reversed-phase separation was performed with the same gradient as described previously (47). 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 (47). 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 ( $\text{C}_{39}\text{H}_{72}\text{NO}_8\text{P}$ , 713.4996 Da) was manually defined in the Mascot configuration editor with a neutral loss of phosphocholine

(C<sub>5</sub>H<sub>14</sub>NO<sub>4</sub>P, 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) (48). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (49) partner repository with the dataset identifier PXD007716 and 10.6019/PXD007716 (username: reviewer99373@ebi.ac.uk, password: gKQNG7SE).

**Coarse-grained molecular dynamics (CG-MD) simulations.** CG-MD simulations were run using the GROMACS 5.0 program (50) with the Martini coarse-grained force field (51, 52). Two protein/lipid systems were simulated, *i.e.*, the R1-3 dystrophin fragment with, either hydrogenated zwitterionic bicelles (HZB), or hydrogenated anionic bicelles (HAB). The R1-3 model is from <http://www.sasbdb.org/data/SASDB53/>. In a manner similar to a previously improved protocol (53), bicelles ( $q_{\text{eff}} = 1.3$ ) were built through the generation of a DMPC/DHPC or DMPC/DMPS/DHPC bilayer with the insane tool (54) by defining the number of lipids within the IBM (33). This step was followed by a second step consisting of the enhancement of the size of the cubic simulation boxes up to 250 Å<sup>3</sup> 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 (55). Theoretical SANS curves were generated with defect parameters (except 50 harmonics) from 0.008 to 0.38 Å<sup>-1</sup>, with the CRYSON program from the ATSAS suite. The theoretical curves were smeared using two “resolution files” 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 (56). Two different scenarios of coiled-coil opening were tested through five independent simulations for each. 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

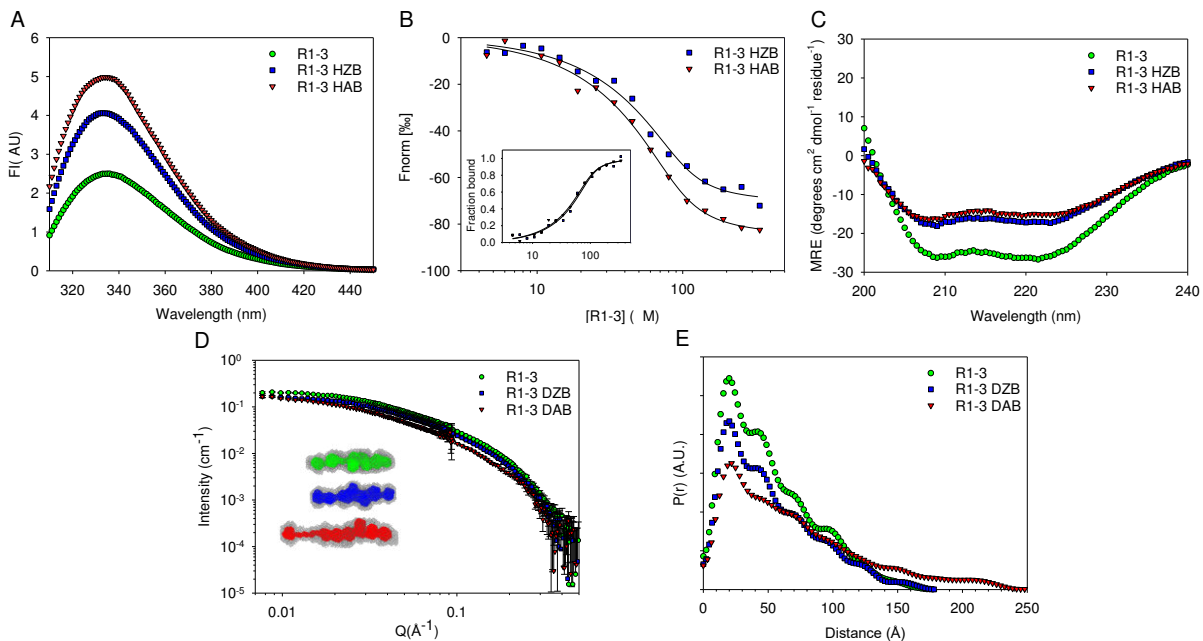
### Protein and bicelle characterization

R1-3 was obtained with a high degree of purity, as assessed by SDS-PAGE analysis (Fig. S3A in the Supporting Material). The molecular weight (MW) of R1-3 was determined by high-pressure size-exclusion chromatography coupled with multi-angle light scattering (HPSEC-MALS). The obtained MW of 38.2 ± 0.4 kDa is in line with the expected values of 38.5 kDa, assessing the monomeric state of R1-3 (Fig. S3B). The quality control of the bicelles was assessed by NMR, HPSEC-dynamic light scattering, HPSEC-MALS, and SANS. All data,

presented in Figs. S3B-E and S4A (in the Supporting Material), allowed to conclude that both R1-3 and bicelles exhibit the expected sizes and monodispersity and therefore can be further used for the characterization of protein/lipid interactions.

### Protein structure changes upon binding to bicelles

The protein/bicelle interactions were highlighted by tryptophan (Trp) intrinsic fluorescence variation, attributed to a modification in the Trp environment due to lipid binding. As expected, Trp intrinsic fluorescence of R1-3 in the presence of HZB or HAB increases significantly (Fig. 2A). To go deeper into the characterization of the protein/bicelle interactions, microscale thermophoresis (MST) was used to determine the dissociation constant ( $K_d$ ) of the complexes, which are in the range of 10  $\mu$ M for both R1-3/HZB and R1-3/HAB (Fig. 2B). Circular dichroism (CD) probed the potential secondary structure modifications of the protein during its interactions with both types of bicelles (Fig. 2C).



**Fig. 2.** Characterization of the protein/lipid interactions of the R1-3 dystrophin fragment. (A) Tryptophan fluorescence intensity (FI) of R1-3 in the absence (green circles) or in the presence of zwitterionic (HZB, blue squares) or anionic (HAB, red triangles) bicelles. (B) Microscale thermophoresis data: the dissociation constant ( $K_d$ ) was determined to be approximately 10  $\mu$ M for protein/bicelle complex. Inset: the same MST data but using the “fraction bound *versus* protein concentration” representation. (C) Far-UV CD spectra with the same legend as for FI. CD spectra highlight that the  $\alpha$ -helical folding of R1-3 is maintained upon lipid binding. (D) SANS intensities measured for R1-3 alone in solution (green circles) or in interaction with contrast-matched deuterated zwitterionic (DZB, blue squares) or anionic (DAB, red triangles) bicelles. Inset: *ab initio* shapes of (from top to bottom) R1-3 free in solution (green) or in interaction with DZB (blue) or DAB (red), corresponding to the DAMMIF models obtained with the smallest normalized spatial discrepancy and surrounded by the corresponding DAMAVER model (gray). (E) Pair-distribution function  $P(r)$  analysis with the same color code

showing an increase of  $D_{\max}$  of R1-3 when the protein fragment is in interaction with anionic bicelles.

The CD spectra of R1-3 alone are typical of an overall  $\alpha$ -helical folding with the presence of two minima at 222 and 208 nm. These data confirm that R1-3 is properly folded (18, 20). For R1-3 in the presence of HZB or HAB, although the  $\alpha$ -helical folding is globally preserved during the interaction with both types of bicelles, a partial loss of helicity occurs. The 3D-structure of R1-3 was first analyzed by SANS in the absence of bicelles (Fig. 2D). The measured MW ( $45 \pm 9$  kDa) and radius of gyration ( $R_g$ ,  $42 \pm 1$  Å) are in good agreement with a monomeric state of the protein and consistent with HPSEC-MALS data. The pair-distance distribution  $P(r)$  is typical of an elongated protein with a peak observed at approximately 20 Å and a smooth fall to  $D_{\max}$  at 177 Å (Fig. 2E). The  $R_g^{\text{real}}$  value of 43 Å is consistent with the above Guinier analysis. We recorded SANS data of R1-3 at exactly the same concentration as for the protein free in solution, but in the presence of zwitterionic or anionic bicelles in deuterated version (DZB and DAB, respectively), which can be contrast-matched in 100% d-TNE buffer (14, 57) (Fig. S4B, inset, in the Supporting Material). The disk-shaped morphology of bicelles is supported by the fit of the SANS data obtained at 42% d-TNE buffer and this morphology of the bicelles is preserved in R1-3/DZB and R1-3/DAB complexes (Fig. S3B-C), highlighting that R1-3 is not located inside the bilayer. This result discards the unlikely scenario, where R1-3 would be located inside the bilayer part of the bicelles, allowing also the click reaction to occur.

For R1-3/DZB complex, the SANS signal is different from that of the protein alone but  $R_g$  ( $41 \pm 1$  Å) and  $D_{\max}$  (178 Å) remain constant (Fig. 2, D-E). We conclude that, either the tertiary structure of R1-3 is maintained in the presence of zwitterionic bicelles, or a change, if any, is not large enough to be detectable by SANS. Surprisingly, for R1-3 in interaction with anionic bicelles,  $R_g$  ( $60 \pm 4$  Å) and  $D_{\max}$  (248 Å) are significantly larger. Such increases indicate a large and significant elongation of the tertiary structure of R1-3 when bound to anionic bicelles. The generated *ab initio* models support this conclusion, showing more extended low-resolution envelopes of R1-3 when bound to DAB, compared to the envelopes found for the protein free in solution or in interaction with DZB (Figs. 2D, inset, and 2E). The structural parameters and the shape-model fitting results are summarized in Table S1 in the Supporting Material, and have been deposited at the SASBDB database (58, 59), with the accession codes of SASDDJ9, SASDDK9, and SASDDL9 for R1-3 alone or in interaction with zwitterionic or anionic phospholipid bicelles, respectively.

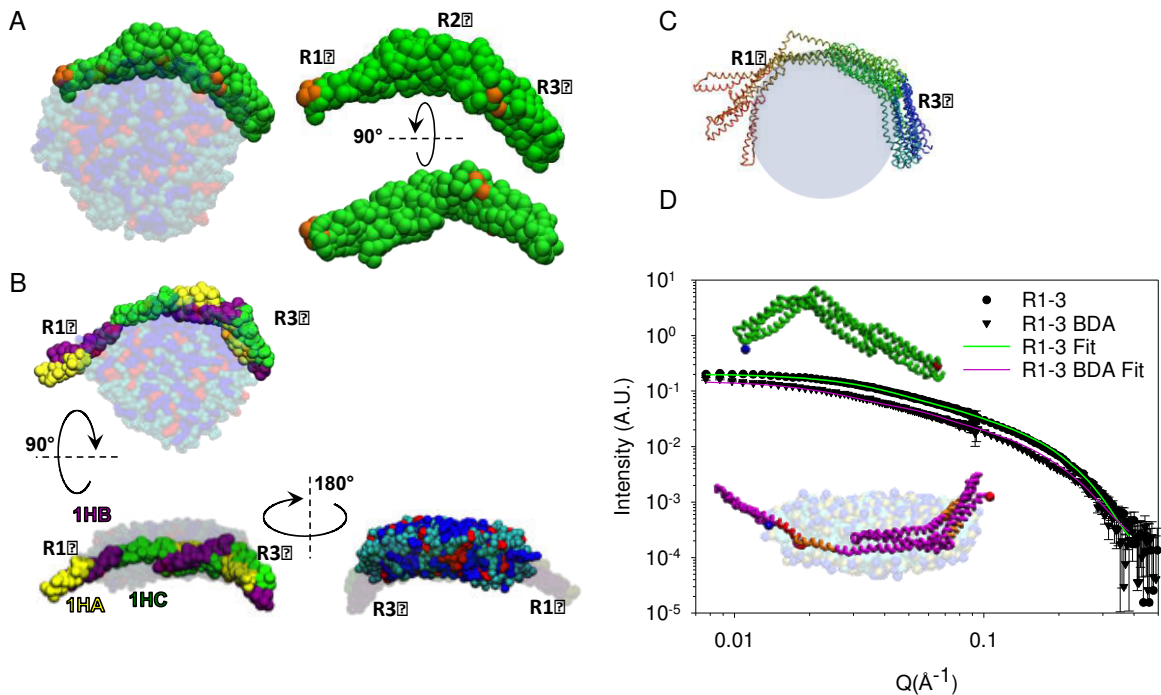
### **Opening of repeat 1 when R1-3 is bound to anionic bicelles**

The interactions of R1-3 with both types of bicelles were more deeply studied by CG-MD experiments. The radial distribution function of the lipids indicates that all simulations report stable self-organized bicelles after 20 to 30 ns of CG-MD, a long time before the 440 and 320 ns recruitment of the protein fragment observed for R1-3/HZB and R1-3/HAB complexes, respectively (Fig. S5, A-B, in the Supporting Material). R1-3/HZB and R1-3/HAB CG models, obtained after 20  $\mu$ s 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 (Fig. S5, C-D). Surprisingly, in both simulations, R1-3 is spontaneously recruited to the bicelle



surface at its torus part (Fig. 3A), which presents a high level of lipid packing defect, whatever the nature of the bicelle. To check these *in silico* results experimentally, we labeled the bicelles with a bifunctional phospholipid (45) (pacFA PC) (Fig. S2A in the Supporting Material). The cross-link reaction should occur only if R1-3 is close to this diazirin group, *i.e.* if R1-3 is located at the bicelle torus, as suggested by CG-MD data. Three peptides of R1-3 linked to the phospholipid were identified by MS (Table S2 and Figs. S2, B-C, and S6, A-C, in the Supporting Material), corresponding to the protein/lipid interaction areas suggested by CG-MD located in the R1 and R3 repeats (Fig. S2D) for both types of bicelles. Transmission electron microscopy analysis of the R1-3/HAB complex highlights that R1-3 is found at the torus of the bicelles (Fig. S6D, in the Supporting Material). All together, 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.

The CG-MD dataset for R1-3/HAB was chosen for further investigation by CG-IMD to understand such conformational modification. An opening of the coiled-coil could explain the  $R_g$  and  $D_{max}$  increase of R1-3 obtained in the corresponding SANS experiments, as well as the extended *ab initio* envelope. 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 on the R1-3/HAB system. Modeling an extended protein, in line with the SANS data, was achieved through two CG-IMD scenarios (Fig. S7A-B and Video S1, in the Supporting Material). Five all-atom models were generated (Fig. 3, B-C) and compared with the experimental data. Among them, the model having the best overall agreement with MS contact mapping (Fig. S7, C-E), a suitable  $R_g$  (59 Å), and the best fit of the SANS data (lowest  $\chi^2$  value) was considered the most reliable structural model for R1-3 interacting with an anionic bicelle (Fig. 3D), namely opening the 1HA-1HB helices of the R1 repeat. This opening may be necessary for the membrane anchoring of dystrophin into the curved parts (22) of the sarcolemma to firmly support the membrane of muscle cells during the muscle elongation/contraction process (refer to the new all-atom R1-3 model accessible with SASDDL9 code).



**Fig. 3.** (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 MS. Lipids are colored in blue for DMPC, red for DMPS, and cyan for DHPC. (B) Final structural model obtained by CG-IMD 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 contrast-matched 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 C-termini, respectively. The peptides bearing the pacFA PC are colored in orange and the putative motif (QGRVGNILQLGSKLIGTG) is colored in red.

Several peripheral proteins recognize lipid packing defects by adsorbing preferentially to positively curved membranes through the insertion of amphipathic sequences (60). Interestingly, based on sequence similarity, a putative amphipathic lipid packing sensor (ALPS) motif was found in 1HB helix of R1 repeat of human dystrophin (61) that we found to open during the interaction with anionic bicelles. This putative ALPS motif is always in contact with the bicelle (Fig. S8, in the Supporting Material). In line with the literature (60), the preference of the ALPS motif of dystrophin for the torus part of the bicelle is likely related to hydrophobic interaction, rather than electrostatic one, since the contact frequencies with tails are higher compared to heads (Fig. S8). Finally, it has been suggested that in helical ALPS motifs, hydrophobic residues, once inserted, stabilize large lipid packing defects by filling them, the molecular details of the interaction between these motifs and biological membranes remaining however elusive (60). In the present case, it is reasonable to hypothesize that such a behavior could be involved in the native dystrophin/lipid interaction. Nevertheless, understanding how the

specific amino acid composition of 1HB helix of R1 repeat might recognize and stabilize the lipid-packing defects at the torus of the bicelle requires additional work.

## CONCLUSION

We characterized the protein/lipid interactions of R1-3 protein fragment from 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 highlight by MST that R1-3 interacts with zwitterionic bicelles and even more with anionic bicelles. By using the contrast-matched bicelle tool (14) in a SANS approach, we demonstrate that R1-3 undergoes large conformational modifications when bound to bicelles containing anionic phospholipids. We used click-chemistry and mass spectrometry coupled with *in silico* CG-MD to determine a reliable mapping of the R1-3/lipid interactions. Finally, using the innovative CG-IMD method, we propose a relevant new all-atom model of R1-3 in interaction with anionic bicelles, providing evidence that the R1 first repeat of the dystrophin central domain undergoes an opening of its coiled-coil structure. This opening may be necessary for the membrane anchoring of dystrophin into the sarcolemma, to firmly support the plasma membrane of muscle cells during the muscle elongation/contraction process. Therefore, our work may contribute to the design of rationalized mini-dystrophins as a gene therapy for the treatment of DMD 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.

## AUTHOR INFORMATION

### Corresponding Authors

\*sophie.combet@cea.fr

\*jfhubert@univ-rennes1.fr

### Present Address

### Author Contribution

RDSM, OD, JP, ELR, SC, and JFH designed the study ; RDSM, AC, TC prepared the samples ; RDSM, OD, DML, ML, AM, AEM, CRN, AB, MSA, and SC acquired the data ; RDSM, OD, DML, ML, SC, and JFH analyzed and/or interpreted the data ; RDSM, OD, SC, and JFH wrote the manuscript ; all authors reviewed the manuscript and approved the final version.

### Note

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank AFM-Telethon, the Conseil Régional de Bretagne, LLB, and SOLEIL for funding. LLB, ILL, and JCNS neutron facilities are acknowledged for beamtime on PACE, D22, KWS-1, and KWS-2, as well as the GENCI program. We thank C. Tascon (CPR, UMR-CNRS 6290), A. Burel (MRic platform), and NMR PRISM platform and the Spectroscopies-DCTP core

facility from BIOSIT at the University of Rennes 1, as well as G. David and B. Pineau (SOLEIL).

## SUPPORTING CITATIONS

### REFERENCES

1. Seddon, A.M., P. Curnow, and P.J. Booth. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta.* 1666: 105–117.
2. Hendrickson, W.A. 2016. Atomic-level analysis of membrane-protein structure. *Nat. Struct. Mol. Biol.* 23: 464–467.
3. Fernandez-Leiro, R., and S.H.W. Scheres. 2016. Unravelling biological macromolecules with cryo-electron microscopy. *Nature.* 537: 339–346.
4. Petoukhov, M. V, and D.I. Svergun. 2007. Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Curr. Opin. Struct. Biol.* 17: 562–571.
5. Pérez, J., and A. Koutsioubas. 2015. Memprot: A program to model the detergent corona around a membrane protein based on SEC-SAXS data. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71: 86–93.
6. Koutsioubas, A. 2017. Low-Resolution Structure of Detergent-Solubilized Membrane Proteins from Small-Angle Scattering Data. *Biophys. J.* 113: 2373–2382.
7. Heller, W.T. 2010. Small-angle neutron scattering and contrast variation: a powerful combination for studying biological structures. *Acta Crystallogr. Sect. D.* 66: 1213–1217.
8. Le Cœur, C., S. Combet, G. Carrot, P. Busch, J. Teixeira, and S. Longeville. 2015. Conformation of the Poly(ethylene Glycol) Chains in DiPEGylated Hemoglobin Specifically Probed by SANS: Correlation with PEG Length and in Vivo Efficiency. *Langmuir.* 31: 8402–8410.
9. Gabel, F. 2015. Small-Angle Neutron Scattering for Structural Biology of Protein-RNA Complexes. *Methods Enzymol.* 558: 391–415.
10. Sonntag, M., P.K.A. Jagtap, B. Simon, M.S. Appavou, A. Geerlof, R. Stehle, F. Gabel, J. Hennig, and M. Sattler. 2017. Segmental, Domain-Selective Perdeuteration and Small-Angle Neutron Scattering for Structural Analysis of Multi-Domain Proteins. *Angew. Chemie - Int. Ed.* 56: 9322–9325.
11. Breyton, C., F. Gabel, M. Lethier, A. Flayhan, G. Durand, J.-M. Jault, C. Juillan-Binard, L. Imbert, M. Moulin, S. Ravaud, M. Härtlein, and C. Ebel. 2013. Small angle neutron scattering for the study of solubilised membrane proteins. *Eur. Phys. J. E. Soft Matter.* 36: 71.
12. Naing, S.-H., R.C. Oliver, K.L. Weiss, V.S. Urban, and R.L. Lieberman. 2018. Solution Structure of an Intramembrane Aspartyl Protease via Small Angle Neutron Scattering. *Biophys. J.* 114: 602–608.
13. Midtgaard, S.R., T.A. Darwish, M.C. Pedersen, P. Huda, A.H. Larsen, G.V. Jensen, S.A.R. Kynde, N. Skar-Gislinge, A.J.Z. Nielsen, C. Olesen, M. Blaise, J.J. Dorosz, T.S. Thorsen, R. Venskutonytė, C. Krintel, J. V. Møller, H. Frielinghaus, E.P. Gilbert, A. Martel, J.S. Kastrup, P.E. Jensen, P. Nissen, and L. Arleth. 2018. Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering. *FEBS J.* 285: 357–371.
14. Dos Santos Morais, R., O. Delalande, J. Pérez, L. Mouret, A. Bondon, A. Martel, M.-S.M.-S. Appavou, E. Le Rumeur, J.-F.J.-F. Hubert, and S. Combet. 2017. Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS. *Langmuir.* 33: 6572–6580.
15. Le Rumeur, E., S.J. Winder, and J.-F.F. Hubert. 2010. Dystrophin: More than just the sum of its parts. *Biochim. Biophys. Acta - Proteins Proteomics.* 1804: 1713–1722.

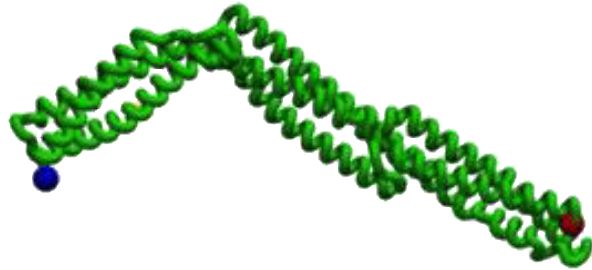
16. Monaco, A.P., R.L. Neve, C. Colletti-Feener, C.J. Bertelson, D.M. Kurnit, and L.M. Kunkel. 1986. Isolation of candidate cDNAs for portions of the Duchenne of muscular dystrophy gene. *Nature*. 323: 646–650.
17. Stark, A.E. 2015. Determinants of the incidence of Duchenne muscular dystrophy. *Ann Transl Med*. 3: 287.
18. Legardinier, S., J.F. Hubert, O.L. Bihan, C. Tascon, C. Rocher, C. Raguénès-Nicol, A. Bondon, S. Hardy, and E.L. Rumeur. 2008. Sub-domains of the dystrophin rod domain display contrasting lipid-binding and stability properties. *Biochim. Biophys. Acta - Proteins Proteomics*. 1784: 672–682.
19. Zhao, J., K. Kodippili, Y. Yue, C.H. Hakim, L. Wasala, X. Pan, K. Zhang, N.N. Yang, D. Duan, and Y. Lai. 2016. Dystrophin contains multiple independent membrane-binding domains. *Hum. Mol. Genet*. 25: 3647–3653.
20. Legardinier, S., C. Raguénès-Nicol, C. Tascon, C. Rocher, S. Hardy, J.-F. Hubert, and E. Le Rumeur. 2009. Mapping of the lipid-binding and stability properties of the central rod domain of human dystrophin. *J. Mol. Biol*. 389: 546–558.
21. Vié, V., S. Legardinier, L. Chieze, O. Le Bihan, Y. Qin, J. Sarkis, J.-F. Hubert, A. Renault, B. Desbat, and E. Le Rumeur. 2010. 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*. 1798: 1503–1511.
22. García-Pelagio, K.P., R.J. Bloch, A. Ortega, and H. González-Serratos. 2011. Biomechanics of the sarcolemma and costameres in single skeletal muscle fibers from normal and dystrophin-null mice. *J. Muscle Res. Cell Motil*. 31: 323–336.
23. Legrand, B., E. Giudice, A. Nicolas, O. Delalande, and E. Le Rumeur. 2011. Computational study of the human dystrophin repeats: interaction properties and molecular dynamics. *PLoS One*. 6: e23819.
24. Chamberlain, J.R., and J.S. Chamberlain. 2017. Progress toward Gene Therapy for Duchenne Muscular Dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther*. 25: 1125–1131.
25. Le Guiner, C., L. Servais, M. Montus, T. Larcher, B. Fraysse, S. Moullec, M. Allais, V. François, M. Dutilleul, A. Malerba, T. Koo, J.L. Thibaut, B. Matot, M. Devaux, J. Le Duff, J.Y. Deschamps, I. Barthelemy, S. Blot, I. Testault, K. Wahbi, S. Ederhy, S. Martin, P. Veron, C. Georger, T. Athanasopoulos, C. Masurier, F. Mingozzi, P. Carlier, B. Gjata, J.Y. Hogrel, O. Adjali, F. Mavilio, T. Voit, P. Moullier, and G. Dickson. 2017. Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat. Commun*. 8.
26. Delalande, O., A.E. Molza, R. Dos Santos Morais, A. Chéron, É. Pollet, C. Raguénès-Nicol, C. Tascon, E. Giudice, M. Guilbaud, A. Nicolas, A. Bondon, F. Leturcq, N. Férey, M. Baaden, J. Perez, P. Roblin, F. Piétri-Rouxel, J.F. Hubert, M. Czjzek, and E. Le Rumeur. 2018. Dystrophin's central domain forms a complex filament that becomes disorganized by in-frame deletions. *J. Biol. Chem*. 293: 6637–6646.
27. Fiehn, W., J. Peter, J.F. Mead, and M. Gan-Elepano. 1971. Lipids and fatty acids of sarcolemma, sarcoplasmic reticulum, and mitochondria from rat skeletal muscle. *J. Biol. Chem*. 246: 5617–5620.
28. Leventis, R., and J.R. Silvius. 2010. Quantitative experimental assessment of macromolecular crowding effects at membrane surfaces. *Biophys. J*. 99: 2125–2133.
29. Nagatomo, T., M. Sasaki, and T. Konishi. 1984. Differences in lipid composition and fluidity of cardiac sarcolemma prepared from newborn and adult rabbits. *Biochem. Med*. 32: 122–131.
30. Leventis, P.A., and S. Grinstein. 2010. The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys*. 39: 407–427.
31. Beaugrand, M., A.A. Arnold, J. Hénin, D.E. Warschawski, P.T.F. Williamson, and I.

- Marcotte. 2014. Lipid concentration and molar ratio boundaries for the use of isotropic bicelles. *Langmuir ACS J. surfaces colloids.* 30: 6162–6170.
32. Glover, K.J., J.A. Whiles, G. Wu, N. Yu, R. Deems, J.O. Struppe, R.E. Stark, E.A. Komives, and R.R. Vold. 2001. Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophys. J.* 81: 2163–2171.
  33. Triba, M.N., D.E. Warschawski, and P.F. Devaux. 2005. Reinvestigation by phosphorus NMR of lipid distribution in bicelles. *Biophys. J.* 88: 1887–1901.
  34. Mineev, K.S., K.D. Nadezhdin, S.A. Goncharuk, and A.S. Arseniev. 2016. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir ACS J. surfaces colloids.* 32: 6624–6637.
  35. Petrache, H.I., S. Tristram-Nagle, and J.F. Nagle. 1998. Fluid phase structure of EPC and DMPC bilayers. *Chem. Phys. Lipids.* 95: 83–94.
  36. Petrache, H.I., S. Tristram-Nagle, K. Gawrisch, D. Harries, V.A. Parsegian, and J.F. Nagle. 2004. Structure and Fluctuations of Charged Phosphatidylserine Bilayers in the Absence of Salt. *Biophys. J.* 86: 1574–1586.
  37. Small, D.M. 1986. *The Physical Chemistry of Lipids: From Alkanes to Phospholipids.* Plenum Press. p. 672.
  38. Winder, S.J., T.J. Gibson, and J. Kendrick-Jones. 1995. Dystrophin and utrophin: the missing links! *FEBS Lett.* 369: 27–33.
  39. Sarkis, J., J.-F.F. Hubert, B. Legrand, E. Robert, A. Chéron, J. Jardin, E. Hitti, E. Le Rumeur, and V. Vié. 2011. Spectrin-like repeats 11-15 of human dystrophin show adaptations to a lipidic environment. *J. Biol. Chem.* 286: 30481–30491.
  40. Mehboob, S., B.H. Luo, B.M. Patel, and L.W. Fung. 2001. alpha beta Spectrin coiled coil association at the tetramerization site. *Biochemistry.* 40: 12457–12464.
  41. Franke, D., M. V Petoukhov, P. V Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, and D.I. Svergun. 2017. ATSAS 2.8 : a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* 50: 1212–1225.
  42. Glatter, O., and O. Kratky. 1982. *Small Angle X-ray Scattering.* Academic Press.
  43. Jeffries, C.M., M.A. Graewert, C.E. Blanchet, D.B. Langley, A.E. Whitten, and D.I. Svergun. 2016. Preparing monodisperse macromolecular samples for successful biological small-Angle X-ray and neutron-scattering experiments. *Nat. Protoc.* 11: 2122–2153.
  44. Kline, S.R. 2006. Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Crystallogr.* 39: 895–900.
  45. Haberkant, P., R. Raijmakers, M. Wildwater, T. Sachsenheimer, B. Brügger, K. Maeda, M. Houweling, A.C. Gavin, C. Schultz, G. Van Meer, A.J.R. Heck, and J.C.M. Holthuis. 2013. In vivo profiling and visualization of cellular protein-lipid interactions using bifunctional fatty acids. *Angew. Chemie - Int. Ed.* 52: 4033–4038.
  46. Lavigne, R., E. Becker, Y. Liu, B. Evrard, A. Lardenois, M. Primig, and C. Pineau. 2012. Direct Iterative Protein Profiling (DIPP) - an Innovative Method for Large-scale Protein Detection Applied to Budding Yeast Mitosis. *Mol. Cell. Proteomics.* 11: M111.012682.
  47. Jumeau, F., E. Com, L. Lane, P. Duek, M. Lagarrigue, R. Lavigne, L. Guillot, K. Rondel, A. Gateau, N. Melaine, B. Guével, N. Sergeant, V. Mitchell, and C. Pineau. 2015. Human spermatozoa as a model for detecting missing proteins in the context of the chromosome-centric human proteome project. *J. Proteome Res.* 14: 3606–3620.
  48. Carapito, C., L. Lane, M. Benama, A. Opsomer, E. Mouton-Barbosa, L. Garrigues, A. Gonzalez de Peredo, A. Burel, C. Bruley, A. Gateau, D. Bouyssié, M. Jaquinod, S. Cianferani, O. Burlet-Schiltz, A. Van Dorsselaer, J. Garin, and Y. Vandenbrouck. 2015. Computational and Mass-Spectrometry-Based Workflow for the Discovery and

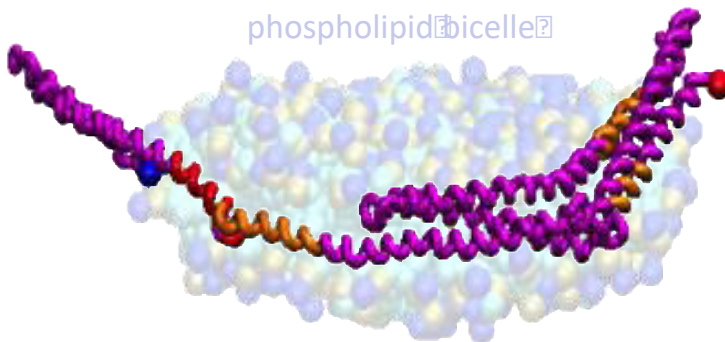
- Validation of Missing Human Proteins: Application to Chromosomes 2 and 14. *J. Proteome Res.* 14: 3621–3634.
49. Vizcaíno, J.A., A. Csordas, N. del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.-W. Xu, R. Wang, and H. Hermjakob. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44: D447–456.
  50. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 4: 435–447.
  51. Marrink, S.J., H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. de Vries. 2007. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B.* 111: 7812–7824.
  52. Monticelli, L., S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and S.-J. Marrink. 2008. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* 4: 819–834.
  53. Vácha, R., and D. Frenkel. 2014. Stability of bicelles: a simulation study. *Langmuir ACS J. surfaces colloids.* 30: 4229–4235.
  54. Wassenaar, T.A., H.I. Ingólfsson, R.A. Böckmann, D.P. Tieleman, and S.J. Marrink. 2015. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* 11: 2144–2155.
  55. Wassenaar, T.A., K. Pluhackova, R.A. Böckmann, S.J. Marrink, and D.P. Tieleman. 2014. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* 10: 676–690.
  56. Delalande, O., N. Férey, G. Grasseau, and M. Baaden. 2009. Complex molecular assemblies at hand via interactive simulations. *J. Comput. Chem.* 30: 2375–2387.
  57. Struppe, J., J.A. Whiles, and R.R. Vold. 2000. Acidic phospholipid bicelles: a versatile model membrane system. *Biophys. J.* 78: 281–289.
  58. Trehwella, J., A.P. Duff, D. Durand, F. Gabel, J.M. Guss, W.A. Hendrickson, G.L. Hura, D.A. Jacques, N.M. Kirby, A.H. Kwan, J. Pérez, L. Pollack, T.M. Ryan, A. Sali, D. Schneidman-Duhovny, T. Schwede, D.I. Svergun, M. Sugiyama, J.A. Tainer, P. Vachette, J. Westbrook, and A.E. Whitten. 2017. 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: An update. *Acta Crystallogr. Sect. D Struct. Biol.* 73: 710–728.
  59. Valentini, E., A.G. Kikhney, G. Previtali, C.M. Jeffries, and D.I. Svergun. 2015. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* 43: D357–D363.
  60. Vanni, S., L. Vamparys, R. Gautier, G. Drin, C. Etchebest, P.F.J. Fuchs, and B. Antonny. 2013. Amphipathic lipid packing sensor motifs: Probing bilayer defects with hydrophobic residues. *Biophys. J.* 104: 575–584.
  61. Drin, G., J.F. Casella, R. Gautier, T. Boehmer, T.U. Schwartz, and B. Antonny. 2007. A general amphipathic  $\alpha$ -helical motif for sensing membrane curvature. *Nat. Struct. Mol. Biol.* 14: 138–146.

## SYNOPSIS

SANS-based model of R1-3 dystrophin fragment in solution



contrast-matched anionic phospholipid bicelle



SANS-based model of R1-3 dystrophin fragment bound to bicelle



## Supporting Material

### Human dystrophin structural changes upon binding to anionic membrane lipids

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

<sup>1</sup>Université de Rennes 1, F-35043 Rennes, France. <sup>2</sup>Institut de Génétique et Développement de Rennes, CNRS UMR 6290, F-35043 Rennes, France. <sup>3</sup>Laboratoire Léon-Brillouin, UMR 12 CEA-CNRS, Université Paris-Saclay, CEA-Saclay, F-91191 Gif-sur-Yvette CEDEX, France. <sup>4</sup>SWING Beamline, Synchrotron SOLEIL, L'Orme des Merisiers, BP 48, Saint-Aubin, F-91192 Gif-sur-Yvette, France. <sup>5</sup>Inserm U1085, Protim-Plate-forme Protéomique. <sup>6</sup>Institut Laue-Langevin, F-38042 Grenoble, France. <sup>7</sup>CNRS 6226, Institut des Sciences Chimiques de Rennes, PRISM, F-35043 Rennes, France. <sup>8</sup>Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, D-85748 Garching, Germany.

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

## METHODS

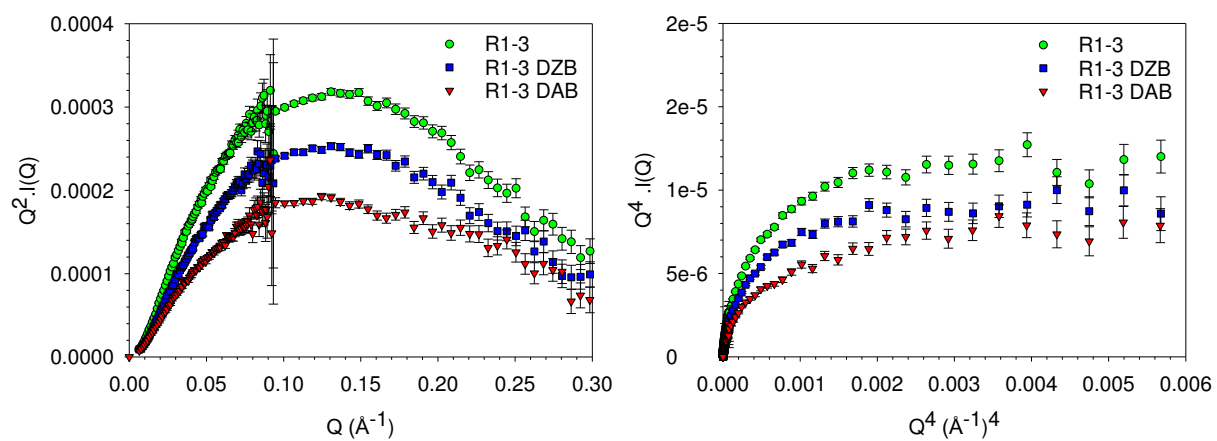
**<sup>31</sup>P NMR spectroscopy.** <sup>31</sup>P NMR spectra of the HZB and HAB were recorded on a Bruker spectrometer Avance III HD 500 equipped with a 5-mm BBO probe operating at 202.46 MHz. Data were recorded in d-TNE buffer at a concentration of 240 mM of lipids, with  $q = 1$  in 3 mm capillary (200  $\mu$ L). 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<sub>3</sub>PO<sub>4</sub> was used as an external standard for the <sup>31</sup>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 Technologies) 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 protein fragment and 0.16 mL/g for the phospholipids (refractive index increment  $dn/dc$  for proteins, polymers SLS. Available at: <http://www.materials-talks.com/blog/2013/06/18/refractive-index-increment-dn-dc-values/>). The SEC column (5  $\mu$ m, 4.6 mm x 300 mm, Agilent) was equilibrated with TNE buffer supplemented with 6 mM of DHPC, corresponding to the [DHPC]<sub>free</sub>, also known as the critical bicellar concentration (Beaugrand et al., 2014) to avoid bicelle deformation under diluted conditions (Dos Santos Morais et al., 2017b). The flow rate was 0.3 mL/min and data were processed with the Astra software v6.1 (Wyatt Technologies).

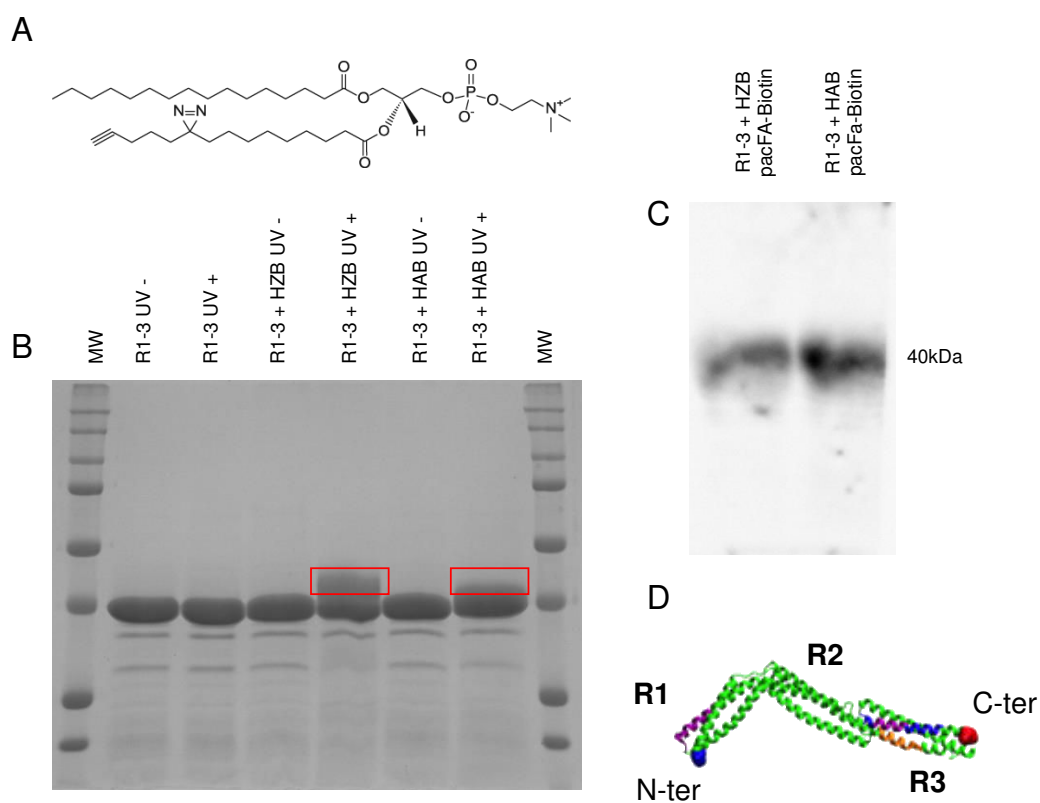
**High-pressure size-exclusion chromatography coupled with dynamic light scattering (HPSEC-DLS).** HPSEC-DLS were recorded with another HPLC system (Agilent) coupled with a light scattering detector (DAWN HELEOS II, Wyatt Technologies) in the same conditions as for HPSEC-MALS.

**Western blotting.** The protein fragment of the protein/lipid complex was precipitated with CHCl<sub>3</sub>/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<sub>4</sub> 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).

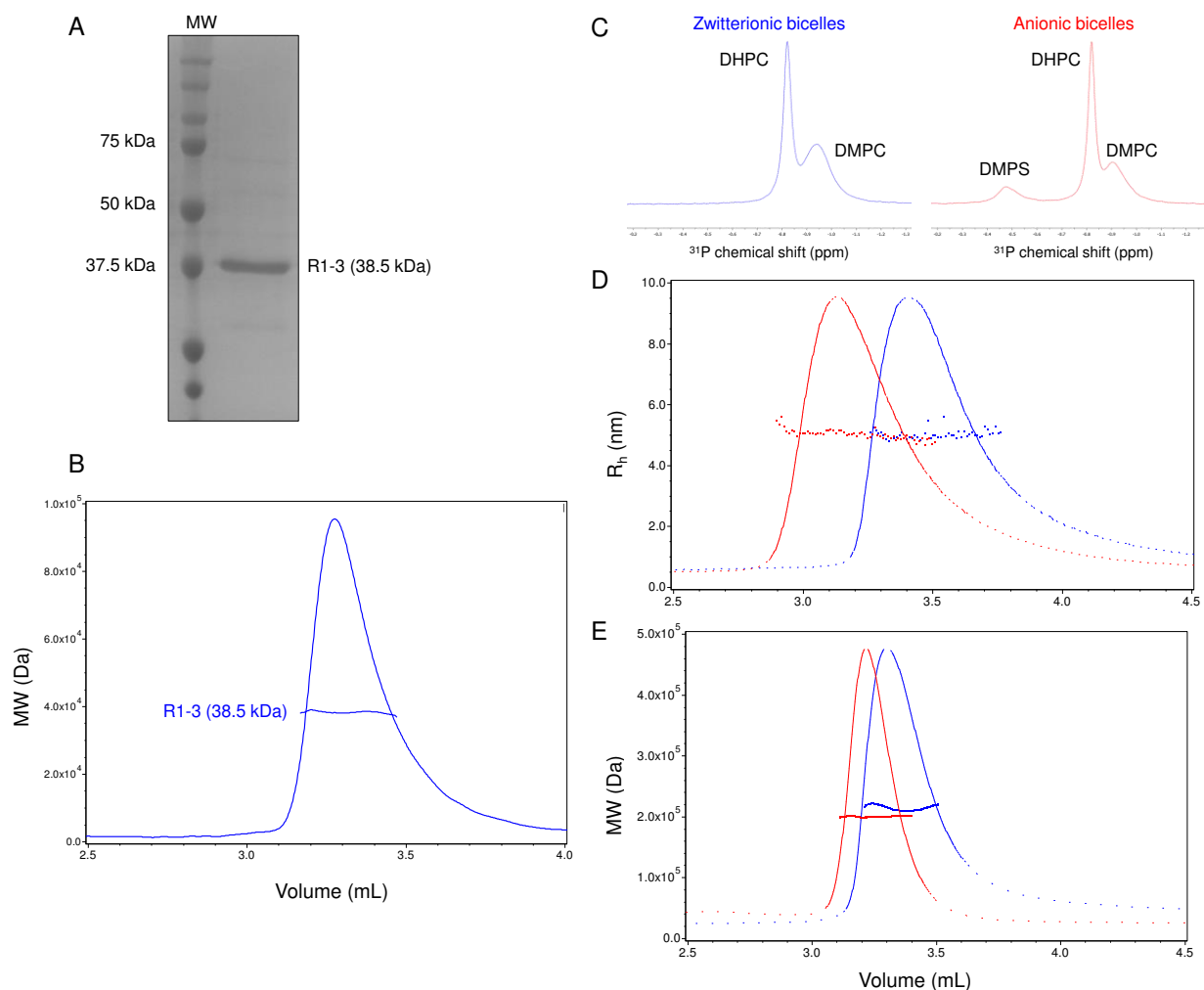
**Gold nanoparticle labeling, negative staining, and transmission electron microscopy (TEM).** The R1-3 protein fragment was labeled with gold nanoparticle (1.4 nm monomaleimido Nanogold, Nanoprobes) *via* cysteines according to the manufacturer's instructions. The excess of free nanoparticles was removed 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  $\mu$ M of conjugate and 15mM of lipids. The sample was deposited 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).



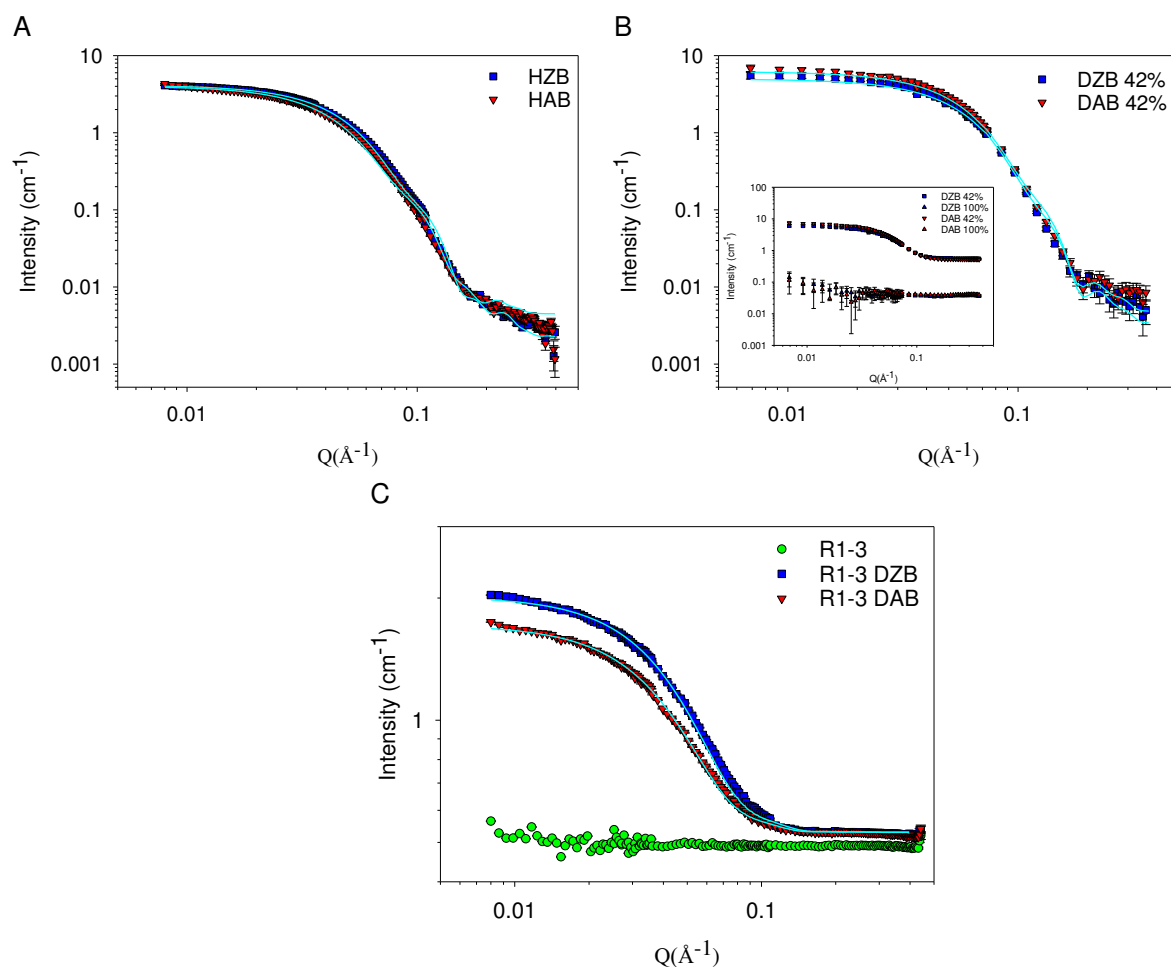
**Fig. S1.** (*left*) Kratky and (*right*) Porod-Debye plots of R1-3 alone (green circles) or in interaction with DZB (blue squares) or DAB (red triangles) bicelles, showing that the proteins are in a good compaction state.



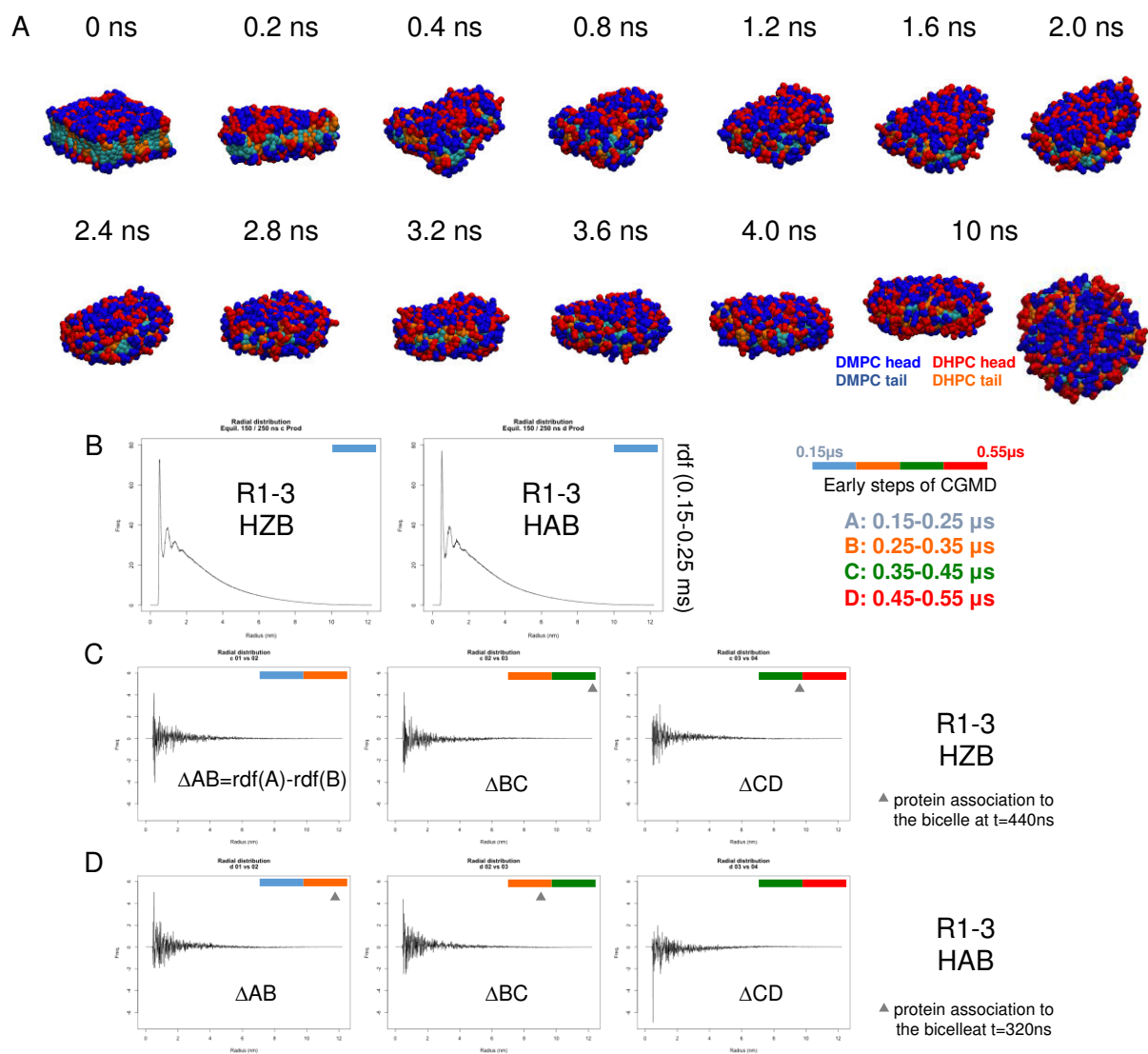
**Fig. S2.** (A) Chemical structure of the bifunctional phospholipid (pacFA PC) bearing a photoactivable diazirine group and a clickable alkyne group on the tail. (B) SDS-PAGE analysis of R1-3 in the presence of bicelles labeled with pacFA PC, with or without UVA exposure. The gel was revealed by Coomassie 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<sub>3</sub> revealed with a streptavidin-HRP fusion protein. (D) R1-3 model (green) with the identified protein/lipid areas after binding with HZB (blue), HAB (orange), or both (magenta).



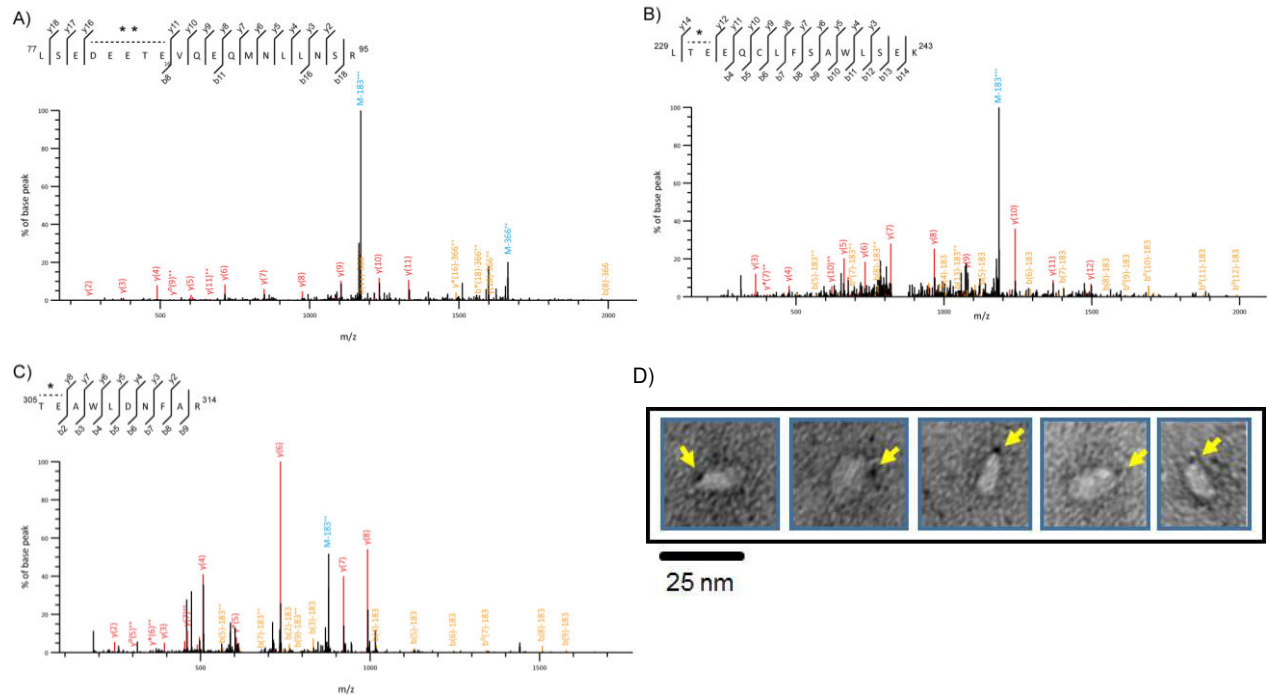
**Fig. S3.** (A) and (B) Quality control of the protein fragment. (A) SDS-PAGE analysis revealed by Coomassie Blue staining. (B) HPSEC-MALS chromatogram of the R1-3 protein fragment showing a MW of  $38.2 \pm 0.4$  kDa. SDS-PAGE and SEC-MALS chromatograms demonstrate the purity, the monodispersity, as well as the monomeric state of the protein fragment. (C), (D), and (E) Quality control of the zwitterionic (blue) and anionic (red) bicelles. (C) <sup>31</sup>P NMR spectra assess the proper organization in bicelles with a  $q$  ratio equal to 1. (D) HPSEC-DLS chromatograms highlight the monodispersity of the bicelles with a constant  $R_h$  of  $4.9 \pm 0.1$  and  $5.0 \pm 0.1$  nm for the zwitterionic and anionic bicelles, respectively. (E) HPSEC-MALS chromatograms show a constant MW of approximately 200 kDa for both types of bicelles all along the peak, in agreement with the aggregation number predicted by the IBM. The void and total volume of the SEC column are 2.0 and 4.1 mL, respectively.



**Fig. S4.** (A) Structural characterization of the HZB (blue squares) and HAB (red triangles) bicelles at 15 mM of lipids ( $q_{eff} = 1.3$ ), 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. (B). SANS intensities for deuterated zwitterionic (DZB) (blue squares) and anionic (DAB) (red triangles) bicelles used in the present study at 50 mM of lipids ( $q_{eff} = 1.3$ ) in 42% (the classical protein match-point) d-TNE buffer. SANS data are fitted with the “core-shell cylinder” model. *Inset*: non-subtracted SANS data in 42% (squares) and 100% (triangles) d-TNE buffer. Both types of bicelles are contrast-matched in 100% d-TNE buffer. (C) SANS data for DZB (blue squares) or DAB (red triangles) in the presence of contrast-matched R1-3 compared to contrast-matched R1-3 alone (green circles) in 42% d-TNE buffer. 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. Therefore, the regular disk-shaped morphology of the bicelles is preserved in R1-3/DZB and R1-3/DAB complexes, highlighting that R1-3 is not located inside the bilayer. This result discards the unlikely scenario, where R1-3 would be located inside the bilayer part of the bicelles allowing also the click reaction to occur (see Fig. S2).

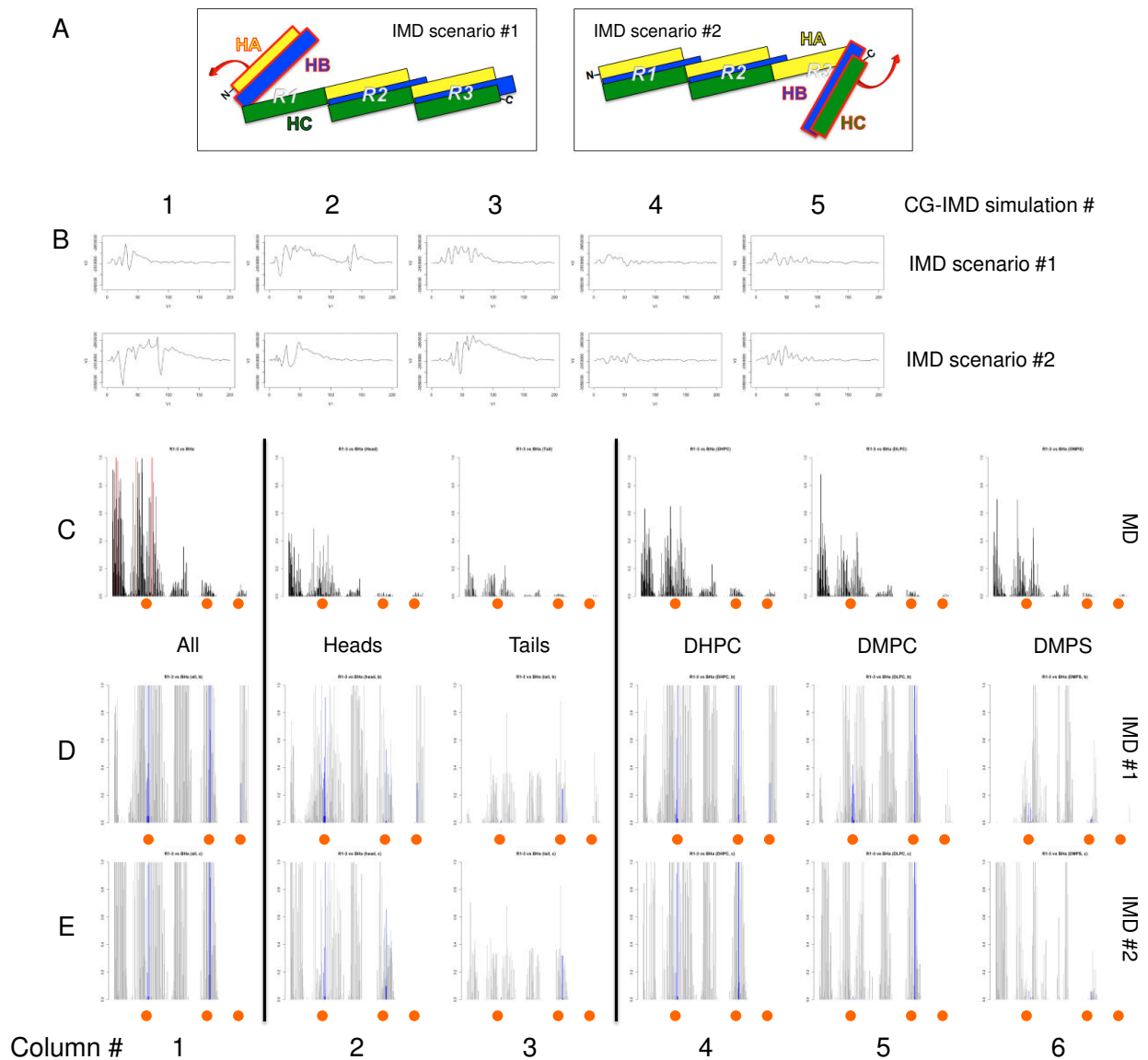


**Fig. S5.** (A) Equilibration of zwitterionic bicelle (HZB) spontaneously formed after 10 ns of CG-MD simulation, from a randomly generated phospholipid bilayer composed of DMPC and DHPC ( $q = 1$ ), by using insane.py script provided for the GROMACS software. Phospholipid head and tail are shown, respectively, in blue and cyan for DMPC and in red and orange for DHPC. (B) Radial distribution function (rdf) of lipids in the early steps of the CG-MD trajectory, including system states before and after R1-3 recruitment to HZB or HAB. Plots of rdf for the 0.15-0.25  $\mu$ s-range of the trajectory show a well-stabilized bicelle organization before the recruitment of R1-3. (C and D) 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 gray triangle) with the bicelle indicate that only a light deformation of the bicelle is observed upon association with the protein.



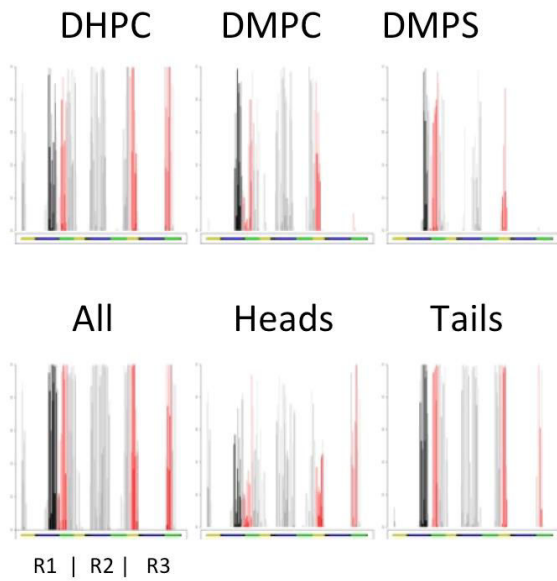
**Fig. S6.** (A, B, and C) Annotated MS/MS spectra corresponding to the three peptides identified with a pacFA PC modification: (A) R1-3 [77-95], (B) R1-3 [229-243], and (C) R1-3 [305-314]. The presence of the phospholipid modification is indicated by an asterisk on the peptide sequence. Dashed lines indicate the amino acids that can bear the phospholipid modification according to MS/MS spectra. (D) 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.





**Fig. S7.** (A) Two 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 two helices of a repeat (red frames and arrows) opening, either 1HA-1HB in R1, or 3HB-3HC in R3, respectively. (B) Fluctuations of the system total energy followed throughout the five independent CG-IMD simulations computed for scenarios #1 and #2. 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. (C, D, and E) Frequencies of protein/lipid contacts measured during classical CG-MD before (C) and after (D and E) 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 #1 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 Fig. 3D).



**Fig. S8.** Contact frequency measurements all along the CG molecular simulation, performed from the best pose obtained by CG-IMD (and after validation by CRYSON calculation), between R1-3 dystrophin fragment (primary sequence on the  $x$ -axis) and phospholipids, by type (DHPC, DMPC, DMPS) or topology (all, heads, tails), and measured all along the CG molecular dynamics simulation trajectory following CG-IMD. Black, red, and gray bars indicate the frequencies involving, respectively, the putative ALPS motif of R1-3, the residues identified by cross-linking/MS/MS experiments using pacFA PC lipids (see the main text of the article), and the rest of the sequence.

These measurements show that this putative ALPS motif is always in contact with the anionic bicelle, as shown by the protein-lipid tails graph. Protein interaction with the torus part of the bicelle is well depicted through the plot of contact frequencies with DHPC *versus* DMPC/DMPS, even if no clear specificity of the anionic bicelle can be evidenced from this cartography. In line with the literature (Vanni et al., 2013), the preference of the ALPS motif of dystrophin for the torus part of the bicelle is likely related to hydrophobic interaction, rather than electrostatic one, since the contact frequencies with the tails are higher compared to the heads.

**Table S1.** Sample details, structural parameters, shape-model fitting results, and atomistic modelling of R1-3 dystrophin fragment, alone or in interaction with zwitterionic or anionic phospholipid bicelles.

<i>Sample details</i>			
Sample name	<b>R1-3 dystrophin fragment</b>		
Organism	<i>Homo sapiens</i>		
UniProt sequence ID (residues in construct)	P11532 (338-668)		
Molar extinction coefficient ( $M^{-1} \text{ cm}^{-1}$ )	59,720		
Source	Recombinant protein produced in BL21 <i>E. coli</i> (DE3)		
$v$ from chemical composition ( $\text{cm}^3 \text{ g}^{-1}$ )	0.725		
Protein SLD ( $10^{-6} \text{ \AA}^{-2}$ )	3.11		
$M$ from chemical composition (kDa)	38.5		
Concentration ( $\text{mg mL}^{-1}$ )	4.2		
<i>Structural parameters</i>	<b>R1-3</b>	<b>R1-3/ zwitterionic bicelles</b>	<b>R1-3/ anionic bicelles</b>
<b>Guinier analysis:</b>			
$I(0)$ ( $\text{cm}^{-1}$ )	0.214	0.170	0.175
$R_g$ ( $\text{\AA}$ )	$42 \pm 1$	$41 \pm 1$	$62 \pm 4$
$Q_{min}$ ( $\text{\AA}^{-1}$ )	0.008	0.008	0.008
$QR_g$ range	0.32-0.97 (16)	0.32-0.97 (16)	0.47-0.99 (13)
Coefficient of correlation	0.97	0.97	0.94
$M$ from $I(0)$ (kDa) (ratio to predicted value)	56 (0.69)	45 (0.86)	46 (0.84)
<b><math>P(r)</math> analysis:</b>			
$I(0)$ ( $\text{cm}^{-1}$ )	0.215	0.171	0.168

$R_g^{real}$ (Å)	43	44	61
$D_{max}$ (Å)	177	178	248
$Q$ -range (Å <sup>-1</sup> )	0.0087-0.2506	0.0087-0.2506	0.0119-0.2505
$\chi^2$ (total estimate from GNOM)	0.666	0.668	0.645
$M$ from $I(0)$ (kDa) (ratio to predicted value)	56 (0.69)	45 (0.86)	44 (0.88)
<b>Shape model-fitting results</b>			
<b>DAMMIF</b> (default parameters, 20 calculations):			
$Q$ -range for fitting	0.0087-0.2506	0.0087-0.2506	0.0119-0.2505
Symmetry	P1	P1	P1
NSD (Standard deviation)	1.270	1.049	1.078
$\chi^2$ range	0.698-0.722	0.585-0.595	0.682-0.720
Resolution from SASRES (Å)	$34 \pm 3$	$27 \pm 2$	$34 \pm 3$
<b>Atomistic modelling</b>			
SASBDB accession codes	SASDDJ9	SASDDK9	SASDDL9
<b>CRYSON:</b>		not determined	
Fitting $Q$ -range (Å <sup>-1</sup> )	0.008-0.38		0.008-0.38
$\chi^2$	6.60		6.97
Predicted $R_g$ (Å)	40.55		58.95

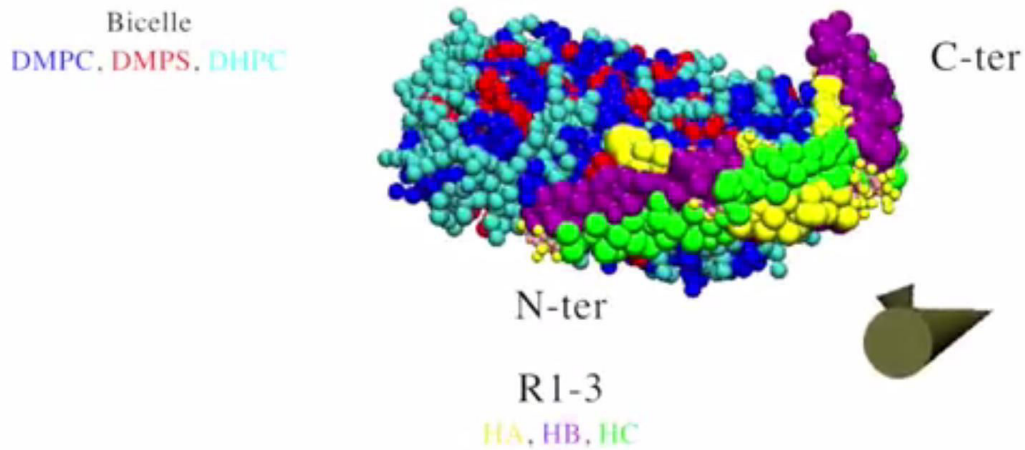
**Table S2.** List of peptides identified by LC-MS/MS for the R1-3 dystrophin fragment. Peptides linked to the phospholipid (pacFA PC) are indicated in bold (corresponding to the annotated MS/MS spectra presented in Fig. S6).

Protein	Score	Sequence	Number of peptides							
R1-3 fragment of dystrophin	2176.8	80.18 %	45							
Peptide sequence	Modifications	Ion Score	Sta rt	Sto p	Calculat ed Mass	Exp. m/z	z	Erro r	Missed Cleava ge	Ran k
<b>LSEDEETE VQEQM NLLNSR</b>	<b>pacFA PC (D4-E8); pacFA PC (D4-E8)</b>	<b>27.6777</b>	<b>95</b>	<b>3690.03</b>	<b>923.54</b>	<b>3.14</b>	<b>0</b>	<b>1</b>	<b>1</b>	
<b>LSEDEETE VQEQM NLLNSR</b>	<b>pacFA PC (E3); pacFA PC (E8)</b>	<b>22.877</b>	<b>95</b>	<b>3690.03</b>	<b>923.54</b>	<b>2.26</b>	<b>0</b>	<b>1</b>	<b>1</b>	
<b>LSEDEETE VQEQM NLLNSR</b>	<b>pacFA PC (E3); pacFA PC (T7)</b>	<b>20.8477</b>	<b>95</b>	<b>3690.03</b>	<b>923.54</b>	<b>2.6</b>	<b>0</b>	<b>1</b>	<b>1</b>	
<b>LTEEQCLFSAWLSEK</b>	<b>pacFA PC (T2 or E3);</b>	<b>23.69229</b>	<b>243</b>	<b>2553.37</b>	<b>852.13</b>	<b>1.04</b>	<b>0</b>	<b>1</b>	<b>1</b>	
<b>TEAWLDNFAR</b>	<b>pacFA PC (T1 or E2)</b>	<b>18.22305</b>	<b>314</b>	<b>1935.08</b>	<b>646.03</b>	<b>0.88</b>	<b>0</b>	<b>1</b>	<b>1</b>	
GSEVNLDR		53.751	8	888.43	445.22	1.2	0	1	1	
YQTALEEVL SWLLSAEDTLQAQGEISN		35.499	41	3676.84	920.24	1.42	0	1	1	
DQFHTHEGYMMDLTAHQGR		60.0242	60	2272.97	758.63	-1.620	1	1	1	
DQFHTHEGYMMDLTAHQGR	Oxidation (M10)	63.7542	60	2288.97	764.3	0.9	0	1	1	
DQFHTHEGYMMDLTAHQGR	Oxidation (M11)	43.442	60	2288.97	763.93	-3.810	1	1	1	
DQFHTHEGYMMDLTAHQGR	Oxidation (M10); Oxidation (M11)	22.9842	60	2304.96	769.33	-0.080	1	1	1	
VGNILQLGSK		61.4761	70	1027.6	514.82	-0.070	1	1	1	
LSEDEETE VQEQM NLLNSR		78.9677	95	2263.03	1132.2	-0.5	0	1	1	
LSEDEETE VQEQM NLLNSR	Oxidation (M13)	55.3777	95	2279.02	1140.2	2.26	0	1	1	
LSEDEETE VQEQM NLLNSRWECLR	Oxidation (M13); Carbamidomethyl	17.7977	100	3023.36	1008.3	-0.941	1	1	1	
VLMDLQNQK		49.46113	121	1087.57	544.72	-2.6	0	1	1	
VLMDLQNQK	Oxidation (M3)	62.12113	121	1103.56	552.72	2.23	0	1	1	
LKELNDWLTK		48.47122	131	1258.69	630.32	-0.711	1	1	1	
ELNDWLTK		37.74124	131	1017.51	509.72	2.04	0	1	1	
KMEEEPLGPDLEDLK		66	138	1521741.84	871.92	-0.661	1	1	1	
KMEEEPLGPDLEDLK	Oxidation (M2)	17.94138	152	1757.84	879.92	-0.331	1	1	1	
KMEEEPLGPDLEDLKR		95.63138	153	1897.95	633.63	-1.482	1	1	1	
KMEEEPLGPDLEDLKR	Oxidation (M2)	53.18138	153	1913.94	638.93	-3.082	1	1	1	
MEEEPLGPDLEDLK		61.63139	152	1613.75	807.82	1.11	0	1	1	

MEEEEPLGPDLEDLK	Oxidation (M1)	47.93139	1521629.74	815.8	2	1.22	0	1
MEEEEPLGPDLEDLKR		51.84139	1531769.85	885.9	2	0.94	1	1
MEEEEPLGPDLEDLKR	Oxidation (M1)	55.01139	1531785.85	893.9	2	-0.211		1
VLQEDLEQEQR		82.91160	1711484.75	743.3	2	-1.410		1
VNSLTHMVVVVDESSGDHATAALEEQ		66.64172	1992978.47	993.8	3	0.79	0	1
VNSLTHMVVVVDESSGDHATAALEEQ	Oxidation (M7)	45.54172	1992994.46	749.6	4	1.06	0	1
WTEDRWVLLQDILLK		35.99211	2251927.06	964.5	2	0.7	1	1
WVLLQDILLK		56.59216	2251239.76	620.8	2	1.07	0	1
LTEEQCLFSAWLSEK	Carbamidomethyl (C6)	111.0229	2431839.87	920.9	2	2.26	0	1
LTEEQCLFSAWLSEKEDAVNK	Carbamidomethyl (C6)	51.65229	2492496.18	833.0	3	0.75	1	1
IHTTGFK		29.52250	256802.43	402.2	2	-0.220		1
IHTTGFKDQNEMLSSLQK		48.43250	2672076.03	1039.	2	-1.021		1
IHTTGFKDQNEMLSSLQK	Oxidation (M12)	55.47250	2672092.03	1047.	2	4.47	1	1
DQNEMLSSLQK		61.23257	2671291.61	646.8	2	0.94	0	1
DQNEMLSSLQK	Oxidation (M5)	54.15257	2671307.6	654.8	2	2.92	0	1
LAVLKADLEK		38.76268	2771098.66	550.3	2	-5.031		1
LYSLKQDLLSTLK		38.8	2852971520.88	507.9	3	-2.6	1	1
QDLLSTLK		35.41290	297916.52	459.2	2	0.12	0	1
TEAWLDNFAR		45.13305	3141221.58	611.8	2	0.58	0	1
CWDNLVQK	Carbamidomethyl (C1)	40.58315	3221061.5	531.7	2	-1.060		1
CWDNLVQKLEK	Carbamidomethyl (C1)	25.83315	3251431.72	478.2	3	-1.891		1

**Video S1.** A movie showing the CG-IMD process. The phospholipids are represented in blue (DMPC), red (DMPS) and cyan (DHPC). The  $\alpha$ -helices of R1-3 are represented in yellow (HA), purple (HB) and green (HC).

Supplementary Movie for Dos-Santos Morais et al.



[https://www.youtube.com/watch?v=Aa0jvL3IBpQ&feature=em-share\\_video\\_user](https://www.youtube.com/watch?v=Aa0jvL3IBpQ&feature=em-share_video_user)



**SUPPORTING REFERENCES**

1. Beaugrand, M., A.A. Arnold, J. Hénin, D.E. Warschawski, P.T.F. Williamson, and I. Marcotte. 2014. Lipid concentration and molar ratio boundaries for the use of isotropic bicelles. *Langmuir ACS J. surfaces colloids.* 30: 6162–6170.
2. Dos Santos Morais, R., O. Delalande, J. Pérez, L. Mouret, A. Bondon, A. Martel, M.-S.M.-S. Appavou, E. Le Rumeur, J.-F.J.-F. Hubert, and S. Combet. 2017. Contrast-Matched Isotropic Bicelles: A Versatile Tool to Specifically Probe the Solution Structure of Peripheral Membrane Proteins Using SANS. *Langmuir.* 33: 6572–6580.
3. Vanni, S., L. Vamparys, R. Gautier, G. Drin, C. Etchebest, P.F.J. Fuchs, and B. Antony. 2013. Amphipathic lipid packing sensor motifs: Probing bilayer defects with hydrophobic residues. *Biophys. J.* 104: 575–584.



**Titre :** Dynamique et mécanique de complexes dystrophine-actine-lipides membranaires

**Mots clés :** Dystrophine, myopathies, modélisation moléculaire, interaction protéine-protéine, interaction protéine-lipides

**Résumé :** La dystrophine est une protéine filamenteuse qui contribue à la structuration des cellules musculaires en créant un lien entre le cytosquelette et le sarcolemme. Avec l'actine du cytosquelette et les lipides membranaires, la dystrophine représente l'un des éléments d'un complexe macromoléculaire, localisé sous la membrane plasmique, dont le rôle est la prévention des dommages qui pourraient être induits à force de contractions-relâchements répétés. De tels dommages, notamment des ruptures du sarcolemme, sont observés chez des personnes atteintes de myopathies de Duchenne (DMD) et de Becker (BMD), des maladies causées par des mutations qui altèrent l'expression ou la fonction de la dystrophine. Ces myopathies sont actuellement incurables et une connaissance approfondie de la relation structure-fonction de la dystrophine et

de ses interactions avec ses partenaires s'avère absolument nécessaire à la mise au point de nouvelles stratégies de thérapies géniques. Cette protéine se compose de quatre domaines fonctionnels, dont un domaine central filamentaire, constitué de 24 répétitions successives d'un même motif structural, un faisceau de trois hélices alpha ou « coiled-coil ». Or, il a été récemment montré que la structure de ce domaine central n'est pas strictement linéaire et que certaines régions inter-répétitions (linker) forment des coudes, délimitant ainsi des sous-domaines d'interaction spécifiques. Cette thèse a pour objectif de comprendre l'origine de cette diversité de conformations inter-répétition dans un domaine structurellement homogène, et d'explorer comment elle permet à certaines régions de se différencier afin d'interagir avec l'actine et/ou les lipides membranaires.

**Title :** Dynamics et mechanics of dystrophin complexes with actin and mambrane lipids

**Keywords :** Dystrophin, myopathies, molecular modeling, protein-protein interaction, protein-lipid interaction

**Abstract :** Dystrophin is a filamentous protein involved in muscular cells structure which links the cytoskeleton to the sarcolemma. Together with cytoskeletal actin and membrane lipids, dystrophin is a part of a macromolecular complex, located under the sarcolemma, which prevents damages induced during repeated muscle contractions and relaxations. Such damages, including sarcolemma disruption, are found in people with Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD), diseases caused by mutations altering dystrophin expression or function. There is currently no treatment to cure these myopathies, and a deep understanding of the structure-function of the dystrophin and its interactions with its partners is necessary to the development of gene therapy strategies.

Structurally, this protein is composed of four functional domains, including a long filamentous central domain, composed of 24 successive coiled-coil repeats. It was recently showed that the central domain is not rod shaped and some inter-repeats regions (linker) are kinked, delimiting specific interaction sub-domains.

This thesis aims to bring knowledge about the basis of the conformational diversity of linkers in a structurally homogenous domain in human dystrophin. We explore how dystrophin delimits some regions that interact with f-actin and/or membrane lipids.