Solid-state NMR studies of the ABC transporter BmrA in its lipid environment

Denis Lacabanne

N°d'ordre NNT : 2017LYSE1243

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 205
Ecole Doctorale Interdisciplinaire Sciences-Santé

Discipline : Biochimie

Soutenue publiquement le 09/11/2017, par :
Denis Lacabanne

Solid-state NMR studies of the ABC transporter BmrA in its lipid environment

---

Devant le jury composé de :

Professeur Ute Hellmich	Université Johannes Gutenberg de Mayence	Rapporteur
Professeur Hartmut Oschkinat	Institut de recherche de pharmacologie moléculaire – Leibniz Berlin	Rapporteur
Professeur Beat Meier	École polytechnique fédérale de Zurich	Examinateur
Docteur Olivier Walker	Université Claude Bernard Lyon 1	Examinateur
Docteur Anja Böckmann	Centre national de la recherche scientifique	Directrice de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

A ma famille :
Claude Lacabanne
Sonia Lacabanne
Solange Lacabanne

Remerciements

Ma thèse fut semblable à une longue route pavée de briques qui déboucha, non pas sur la Cité d'Émeraude, mais sur l'obtention de mon doctorat. A l'instar d'une certaine Dorothée Gale, je fis mon chemin avec trois compagnons : mon courage, mon cerveau et mon cœur. Je fus accompagné d'un charmant système biologique comparable à un gentil petit chien, BmrA, c'est une enzyme allostérique. Sans surprise, elle fut très coopérative avec moi.

En revanche, comme dans tous les contes, il faut un opposant, ici comparable à la Méchante Sorcière de l'Ouest Lyonnais. Il s'appelle C**h**A**R**Les. C**h**A**R**Les n'est pas un cerveau analytique de recherche et de liaison mais un exemplaire de la dernière génération de spectromètre RMN 800 MHz. A l'instar d'un autre super ordinateur bien connu, C**h**A**R**Les est sensé sortir d'une série de spectromètre RMN la plus perfectionnée que l'on connaisse. Aucun appareil de cette série n'a jamais fait d'erreur ni même déformé un renseignement, ils sont tous parfaitement au point et incapables de se tromper. Cependant, mes échantillons devaient être incompatibles avec la logique de C**h**A**R**Les et devaient générer une boucle de Hofstadter-Möbius. Puisque C**h**A**R**Les a délibérément tenté de les détruire à plusieurs reprises (tentative de congélation, tentative de crash du rotor, tentative de surchauffe, etc), ou bien de me retarder un maximum dans l'acquisition par la génération d'erreurs presque aléatoires (3 pics de l'eau, aucun signal (car il est décalé de 10 000 Hz), erreur de pré-ampli pouvant surgir n'importe quand, surtout entre 0h et 6h du matin, l'erreur PD (oui très agréable), signal qui bouge quand on marche à côté du spectromètre etc).

Bref, du courage, oui il en faut pour faire une thèse. Une thèse ce n'est pas les meilleures années de sa vie. Clairement pas. Cependant, je vous ai parlé de C**h**A**R**Les, de BmrA, mais il y'a un paramètre que j'oublie et qui est pourtant capital, l'environnement de travail.

Deux cas de figure s'offrent à moi, soit je suis entouré par des ancres qui me font plonger et me ralentissent soit je suis entouré par des ailes qui me font gagner le ciel et toucher les étoiles. Je fus entouré d'ailes qui m'ont apporté un soutien sans faille et qui m'ont aidé à aller plus haut. Les voici :

Tout d'abord **Anja Böckmann**, ma directrice de thèse, Anja vient tous les jours au laboratoire depuis Avignon, Anja est donc ma bonne sorcière du sud (oui ils ont fait une erreur dans le film, et puis ça m'arrange bien). Ma vision de la RMN du solide était à l'origine plutôt sépia, Anja sut la rendre colorée, en partie grâce à CCPN. Elle me conseilla,

m'aida et surtout me laissa beaucoup de libertés au niveau de mes idées… tant que je ne m'éloignais pas trop de ma route pavée de briques jaunes. Elle ne nous mâche pas le travail pour ne pas que l'on se repose sur elle car elle tient à ce que l'on réfléchisse, que l'on trouve tout seul. Elle m'a beaucoup appris, tout en me laissant agir à ma fantaisie. Je n'ai jamais regretté d'avoir passé quatre années à ses côtés, je n'ai jamais voulu m'en aller au-delà de l'arc-en-ciel car j'avais tout ce qui me fallait ici.

Ensuite, savez-vous quel est le point commun entre tous les gens de notre laboratoire ? C'est Aurélie, **Aurélie Badillo**. Aurélie n'aime pas être coincée dans les bouchons le matin, elle n'aime pas quand il n'y a plus de carboglace, elle aime avoir 63 fenêtres de son navigateur internet ouvertes en petit sur son mac, elle aime jeter à la poubelle les 117 tubes eppendorf rangés dans 6 racks au congélateur (avec les 20 autres), elle aime manger des morceaux de sucres, mais ce qu'elle aime le plus c'est le chocolat qu'elle cache dans le premier tiroir de son bureau. Ce qui est bien avec Aurélie, c'est qu'elle me fait penser à une certaine personne très connue habitant près de Montmartre, comme elle, elle se met en quatre pour arranger les cafouillages de la vie et des manips des autres. Toujours souriante, toujours à dire les bons mots, sa grande expérience des protéines nous rappelle tous quelqu'un qu'elle a fort bien remplacé… la moustache en moins.

Guillaume David, ou Guillome pour les intimes fut le dernier arrivé pendant mon ère. On ne le cherchait nulle part mais il est apparu n'importe où. C'est une personne assez fascinante car sans toucher la moindre pipette, il doit être la personne qui fait le plus de Western blot du laboratoire, voire de l'institut. Pour mener à bien ses expériences sans quitter son bureau ou son téléphone, il doit avoir le pouvoir de manipuler le temps et les dimensions relatives de l'espace, je ne vois pas comme cela est possible par les voies « normales ». Une personne également bienveillante, un peu excentrique, un discours le plus souvent loufoque dénué de sens et puis capable de changer deux fois de continent en moins de quatre jour... tous ces points m'ont confirmé, après quelque mois de suspicions, que Guillaume sera sans nul doute, l'excellent 7ème docteur de notre laboratoire.

Marie-Laure Fogeron, Marie-Laure, ne jamais se contenter de « Marie » encore moins de « Forgeron ». Elle est intransigeante, exigeante, elle veut le meilleur pour nous, le laboratoire doit être rangé, on se fait gronder si on n'a pas nettoyé nos cochonneries et elle sait tout de suite quand on a fait une bêtise. A côté de ça, elle est toujours prête à nous aider parce qu'elle veut le meilleur pour nous, toujours. Il ne lui manque que la capacité de voler et de

posséder un parapluie qui parle pour nous rappeler quelqu'un. C'est avec elle que j'ai commencé mon tout, tout, tout premier stage au sein du laboratoire, Marie-Laure c'est l'élément qui fait que l'on s'adapte. La personne que l'on apprécie, qui fait que chaque tache devient un plaisir. Elle aida les difficultés de ma thèse à couler.

Britta Kunert, Britta c'était mon encadrante pendant les 11 mois de stage de master. Elle me dispensa les meilleurs cours: défense contre l'agrégation des protéines, soins aux protéines membranaires, étude des Membranes, confection de tampons etc. Elle me transmit son manuel avancé de préparation des protéines membranaires. Il était annoté de toutes parts avec plein d'astuces, comme ajouter un brin (500 mM) de sel pour contrer les effets de l'ADN ou bien de remuer une fois dans le sens des aiguilles d'une montre puis sept fois dans le sens contraire pour avoir une solubilisation de la protéine plus rapide. Pendant tout le long de ma thèse, j'ai utilisé ses optimisations, je lui dois beaucoup. Elle m'a toujours encouragé, du début de mon stage de master jusqu'au jour de ma soutenance de thèse.

Roland Montserret, à l'instar d'un policier irlandais incorruptible mais sans Saint Jude autour du cou, Roland me fut d'une aide indéniable et jusqu'aux derniers jours de ma thèse. Bon, avec lui, si vous voulez jeter du papier à la poubelle, les poubelles pour le recyclage ce n'est pas fait pour les chiens. En dehors de cela, il m'apporta son aide, ses astuces et surtout son vocabulaire en anglais. On a partagé notre bureau pendant un an et demi, il connait donc toutes les bonnes planques d'alcool, très utile en période de prohibition. Mais il fut surtout un excellent compagnon de route maniant l'ITC, le spectromètre de dichroïsme circulaire et maintenant la thermophorèse mieux qu'une mitrailleuse Thomson. Il est donc incollable sur les interactions ou bien les non-interactions (merci Ben Hur *et al.*) et était le premier à rire à toutes mes Connery.

Lauriane Lecoq, Lauriane est tombée dans le terrier du lapin blanc, venant de la RMN du liquide, elle choisit la pilule de la RMN du solide... afin de connaître la vérité. Elle dut explorer elle-même cet aspect pour le maîtriser. Cependant, il lui fallait affronter C**h**A**R**L**e**s ainsi que son agent. Pour cela, elle ingurgita comme une vraie machine tout ce qu'il fallait savoir sur l'acquisition et le traitement des données. A la fin, elle comprenait C**h**A**R**L**e**s, était capable de le maîtriser et pouvait ainsi faire le travail d'un programme. Elle est un peu l'élue du laboratoire et son aide a été très précieuse lorsque je me retrouvais confronté à C**h**A**R**L**e**s. Elle m'a aidé à arpenter le chemin de la RMN qu'elle connaissait parfaitement.

Enfin, c'est une Chargée de Recherche et également une amie, l'un comme l'autre, de classe exceptionnelle.

Anaïs Papin, Anaïs a commencé en tant que simple stagiaire et fut placée dans la même cellu… bureau que moi. Elle utilise une crème de jour à la camomille, elle est courtoise et réceptive à la courtoisie, elle a fait le même parcours universitaire que moi mais elle rêvait de sortir d'ici et d'entrer enfin en C.D.I. Elle réussit à s'adapter brillamment bien qu'elle n'ait jamais voulu déguster un très bon foie avec des fèves au beurre accompagné d'un excellent chianti, avec ma personne. Je dois avouer qu'elle a un peu subi toutes mes petites excentricités au fil des années, mais elle a toujours été présente pour m'aider ou m'épauler quand j'en avais besoin. Même si je ne le manifeste pas clairement, je lui témoigne beaucoup de respect et je n'ai donc jamais eu l'idée de l'avoir… à diner.

Shishan Wang, Shishan nous vient de l'empire du milieu et travaille sur l'hépatite B. En effet, tout commença lorsque les grandes maladies furent forgées, sept furent données aux hépatites (A, B, C, D, E, G, H) grands mineurs et destructeurs de foie. Cependant elles furent toutes dupées, car l'hépatite B gouverne sur toutes les autres. On a commencé à comprendre que Shishan était lié à tout cela à cause de sa capacité à disparaître au milieu de nous tous et de réapparaitre quand plus personne n'est au laboratoire, le week-end. Bien qu'il n'ait pas de jardinier avec lui, il arpente sa thèse sous la bienveillance d'une Magicienne de classe exceptionnelle, Lauriane La Grise (jeu de mot puissant en lien avec un petit penchant de Lauriane). Il apparait sans nul doute que Shishan arrivera à détourner le mauvais œil de l'hépatite B pour arpenter la montagne qu'est sa thèse, le tout avec succès.

Mes deux stagiaires **Clément Danis** et **Claire Chuilon** que j'ai eu l'honneur et la fierté d'encadrer. Clément a beaucoup souffert sur la mise au point de la méthode GRecon, son aide m'a été très précieuse. Et Claire, qui a encore plus souffert, mais comme Clément son aide et son autonomie m'ont été également très précieuses. Il est très important pour moi de le souligner.

Nos trois gestionnaires, **Dorothée Bernard**, **Souad Boukoum**, et **Christine Dupuis**. Il faut savoir que je suis très tête en l'air donc leur présence, leur humour et leur capacité à résoudre des problèmes difficilement solvables doivent être soulignés ! Ce sont les trois rouages qui font que l'institut ne s'écroule pas sous le poids des requêtes administratives.

Et les « anciens » membres du laboratoire :

Je pense à **François Penin**, qui a trouvé les financements pour ma thèse, sa bonne humeur, son savoir sur les protéines nous ont beaucoup manqués après son départ. On était tous un peu triste, sauf nos oreilles qui n'allaient plus être tirées !

Je pense également à **Carole Gardiennet**, qui a débuté ma formation de RMNiste du solide sur Charles, le spectromètre de l'aigreur. Elle m'a également beaucoup aidé pendant mon stage de master et au début de ma thèse avant son départ pour Nancy.

Petites pensées également pour **Stéphane Sarazin**, hey Stéphane ça a bien marché finalement !

Et enfin à **Joanna Bons**, elle faisait son stage avec Aurélie. Elle a connu l'EDISSgate donc sous mon plus mauvais jour, malgré cela son soutien et ses encouragements auxquels j'accordais une très grande importante m'ont été plus que bénéfiques.

C'est drôle les souvenirs, je ne me souviens pas vraiment de mon premier jour à l'IBCP, ni de ma première manip et encore moins de mes premiers résultats mais je me souviens très bien de la première fois que j'ai rencontré **Alice Gutjahr**. En même temps commencer par « il est gentil ce petit », difficile d'oublier. Nous avons discuté dans la file d'attente du Restaurant Universitaire. En dehors des gens du laboratoire, c'est la première personne qui m'a demandé sur quoi je travaillais, je m'étais trouvé un peu bête, je ne savais pas comment répondre. A partir de ce moment on ne s'est jamais quitté. Bref je vais arrêter ici la référence à Forest Gump… bien que le début de ce film est très comparable, la suite est bien trop triste et puis pas du tout comparable !

Alice a fait sa thèse dans un autre laboratoire dans un autre institut, et elle était plutôt portée biologie cellulaire, vaccinologie et expérimentation animale. On a évolué parallèlement, on s'est serré les coudes et on a fini nos contrats ensemble pour soutenir tous les deux à un mois d'intervalle. Cette thèse n'aurait pas été ce qu'elle est sans Alice. Si si, les figures auraient été plus moches et les courbes ou les histogrammes seraient venus d'Excel… et j'aurais été surtout beaucoup moins motivé pour écrire ma revue. L'effet

« monolithe noir » d'Alice n'a pas concerné que son laboratoire. Voilà, ça va être impossible d'oublier une personne qui m'a donné autant de bons souvenirs et qui m'a autant soutenu pendant ces trois années

Chryslène Mercy, Chryslène était ma « co-équipière », femme de tête, particulièrement décidée, c'est une doctorante d'une très grande valeur. Elle met un point d'honneur à finir chaque mission quitte à venir le week-end. Nous étions tous les deux les représentants des doctorants de l'unité Molecular Microbiology and Structural Biochemistry, élus à une majorité des voix comparable à celle d'un dictateur. Nous avions du réaliser plusieurs missions, mais la plus difficile de toutes n'aura pas été de ramener un transmuteur sur Alpha mais d'accueillir un Professeur Espagnol venant de Cantabrie, le Professeur aux mille e-mails. Nous avons réussi la mission avec succès, évidement. En dehors de cela, Chryslène couvrit régulièrement mes écarts diplomatiques et je place en elle toute ma confiance. Une confiance difficilement mesurable car, c'est une confiance bien pauvre, celle que l'on peut calculer

Marine Montmasson, toujours souriante, très polie, pianiste à ses heures perdues, elle a encore plein d'autres qualités… Marine a un petit quelque chose de « princier » que ce soit dans ses propos ou bien dans ses actes. A l'instar d'un Duc bien connu devenu roi, elle n'est pas la personne la plus à l'aise devant un public bien qu'elle ait le bégaiement et les sautes d'humeur en moins. Malgré cela, quand je l'ai vue présenter devant 200 personnes, elle a parlé avec tant de conviction que j'ai su que tout irait bien lors de son futur « discours » de 45 minutes qu'elle donnera prochainement et ceci sans l'intervention d'un orthophoniste spécialiste en élocution. J'étais heureux de l'avoir à mes côtés le temps de cette thèse et surtout pendant ma soutenance. Bien que ses sourires et ses clins d'œil habituels vont indéniablement me manquer, je ne pense pas que nos routes se sépareront mais elles finiront, je l'espère, comme les deux personnages célèbres auxquels je fais allusion

En dehors de ce petit cercle, j'ai pu compter sur tout l'IBCP mais également certains membres du laboratoire de RMN du solide de l'ETH à Zürich.

Les transporteurs ABC sont bien représentés à l'IBCP, deux autres laboratoires s'y intéressent.

Tout d'abord celui « d'en face », avec **Jean-Michel Jault** et **Cédric Orelle** qui ont fait partie de mon comité de suivi de thèse. Jean-Michel et Cédric ont toujours répondu présent pour répondre à mes nombreuses questions et pour apporter des idées et des solutions tout le long de ma thèse. De petites pensées également à **Waquas Javed**, **Maty Diagne** et surtout **Khadija Mathieu**, les trois doctorants du laboratoire, je leur ai donné beaucoup de fil à retordre.

Ensuite, à celui du deuxième étage, dirigé par **Pierre Falson** qui avec **Vincent Chaptal** ont toujours été là pour répondre à mes petites questions de 5 minutes qui en duraient 20. Je remercie aussi **Sandrine Magnard,** ma compagne de galère d'ultracentrifugeuse. Et comme précédemment, petites pensées à **Rachad Nasr**, **Josiane Kassis**, **Alexis Moreno** et **Veronica Zampieri**, les doctorants de l'équipe et **Hélène Cortay** que j'ai eu la chance d'avoir comme professeur à la fac, que j'ai beaucoup embêté et qui m'ont également beaucoup nourri en terme de gâteaux.

Je remercie également les membres du PSF pour toutes les discussions intéressantes scientifiques (ou non), **Virginie Gueguen-Chaignon**, **Veronique Senty-Segault**, **Eric Diesis**, **Damien Ficheux** (même s'il n'est pas PSF, c'est tout comme) du pôle protéine et **Frederic Delolme**, **Adeline Page** et **Isabelle Zanella-Cleon** du pôle spectrométrie de masse. Toute la partie spectrométrie de masse de ma thèse a été réalisée par Frédéric, merci encore Fred pour ta bonne humeur et pour tout ton travail.

Cette thèse n'aurait pas été la même sans l'intervention Zurichoise du groupe de **Beat Meier**, professeur à l'ETH Zürich. Beat a beaucoup contribué, que ce soit pour ma thèse ou bien les articles, par ses idées, ses avis, ses encouragements et surtout il m'a accueilli pour ma formation en RMN du solide dans son laboratoire le premier mois de ma thèse. Toujours de bonne humeur et souriant, ça a été un plaisir de travailler avec lui pendant ma thèse et ce sera un plaisir de travailler avec lui pendant mon post-doc à l'ETH.

Je remercie également **Thomas Wiegand** avec qui j'ai beaucoup échangé sur le prolifique système DnaB. Tellement échangé, que nous avons publié un article en commun sur nos

deux systèmes. Je le remercie surtout pour ses aides (je pense à la RMN du phosphore) et ses commentaires toujours très avisés et pointilleux. « In principle you're right BUT » … le terrible « but »… !

Je pense également à **Alons Lend** et à **Vlastimil Jirasko** avec qui j'ai beaucoup partagé et travaillé pour mettre au point la méthode GRecon.

Une petite pensée également pour **Maarten Scheldorn** pour son travail et sa disponibilité pour le remplissage des rotors et l'enregistrement de spectres avec la sonde 0.7 mm 110 kHz.

Et enfin, je remercie tout le petit aréopage des doctorants MMSB : **Christelle Folio** (avec qui j'ai fait sept années de cursus universitaire en fait !), **Margaux Chahpazoff**, **Loïc Carrique**, **Halima Yajjou** (une doctorante qui aime les PERV), **Julien Cayron** (qui prendra ma place et fera lever les gens lorsque le jury de thèse entrera dans la salle), **Marine Restelli**, **Anaïs Pelletier**, **Pierre Garcia** (le faux doctorant IBCP !!), **Laure Zucchini**, **Célia Bergé** (qui a basculé du bon côté de la biochimie structurale en faisant de la RMN et de la cryo-microscopie), **Paul Imbert**, **Stéphanie Gagné** et **Arthur Louche** (j'aimais beaucoup taquiner ces trois là)…

Et pour finir **Morgane Roussin**… Mademoiselle Roussin ! Elle a commencé à l'IBCP en tant que stagiaire un an avant mon départ. Elle a passé la Sélection de l'école doctorale, et a commencé sa thèse quand moi je la finissais, c'est dommage. Ce qui est intéressant avec Morgane c'est que je ne sais pas trop comment la classer… Elle aime dire que c'est une cinq bien qu'elle est tout d'une quatre (« théoriquement »), après en tant que doctorante forcément on serait enclin à dire que c'est une trois, sauf que d'apparence c'est complètement une deux et avec le style d'une un. En tout cas ce qui est sûr c'est qu'elle fait partie d'une Élite et elle pourrait largement être Élue Miss Parfaite. Elle est l'Héritière d'un projet ambitieux mais elle a les capacités de s'en sortir haut la main. J'espère que je pourrai être à ses côtés lorsqu'elle recevra sa Couronne de compliments venant de ses rapporteurs, digne récompense de trois années de thèse bien entreprise. Quand il sera venu ce moment, bien que tu seras le centre d'attention de tout le monde, tu ne ressentiras aucune angoisse, juste de la paix, crois moi

Je remercie également les membres de mon comité de suivi de thèse qui ont suivi chaque année mon travail : **Jean-Michel Jault, Cédric Orelle et Paul Schanda**. Merci encore Paul pour tes conseils judicieux que ce soit en labeling tout comme en RMN et pour tes opinions/avis sans langue de bois.

Pour finir, je tiens à remercier les membres de mon jury de thèse pour avoir lu et évalué mon travail.

Ainsi, je remercie **Madame le Professeur Ute Hellmich** et **Monsieur le Professeur Hartmut Oschkinat** d'avoir accepté d'être les rapporteurs de mon travail de thèse.

Je remercie également le **Monsieur le Professeur Beat Meier** pour avoir présidé et examiné mon travail. Et je finis par **Monsieur le Docteur Olivier Walker** d'avoir également examiné mon travail mais surtout pour m'avoir fait découvrir la RMN, aiguillé dans la recherche et dans le choix de laboratoire, il y'a déjà six ans.

Dans l'ensemble je remercie les membres de mon jury de thèse d'avoir lu très rigoureusement et d'avoir discuté de mon travail avec un très haut degré de perspicacité.

This thesis has been carried out under the supervision of Anja Böckmann, in the team « solid-state NMR of proteins » that is part of the unit UMR 5086 – MMSB (Molecular Microbiology and Structural Biochemistry) at IBCP :

Institut de Biologie et Chimie des Protéines
7, passage du Vercors
F-69367 Lyon cedex 07

Résumé substantiel

Aujourd'hui, nous nous confrontons à un problème général qui est la résistance aux médicaments, acquise par des cellules ou bien par des microorganismes, contre des molécules cibles : typiquement des molécules qui leur sont toxiques. Une des raisons expliquant ce phénomène de résistances provient du fait qu'il existe des pompes à efflux rejetant en dehors de ces organismes ces molécules cibles. L'acquisition de résistance est de plus en plus fréquente et incite à élucider les mécanismes sous-jacents à ces pompes à efflux afin de les contrer par l'élaboration de nouvelles molécules. Malgré l'importance de l'action de ces pompes, leur fonctionnement est actuellement peu caractérisé. Ces pompes sont essentiellement des transporteurs ABC.

Les transporteurs ABC font partie d'une superfamille de protéine membranaire que l'on retrouve dans tous les règnes du vivant. Ils ont la capacité de lier des molécules présentes à l'intérieur des cellules et de les exporter au travers de la membrane plasmique. Ces molécules peuvent être endogènes ou bien exogènes. Bien que ce mécanisme soit ubiquitaire, il demeure aujourd'hui très mal compris au niveau de son fonctionnement à l'échelle atomique.

C'est pour combler ce manque de connaissance que l'on a recours à la biologie structurale qui renseigne sur les différentes structures des acteurs mis en jeu et d'en comprendre leur rôle. On distingue des méthodes complémentaires d'analyse dont les trois majeures sont la microscopie électronique, la cristallographie et la résonance magnétique nucléaire du solide (RMN) pour les protéines membranaires. Cette dernière est une méthode en plein développement et très prometteuse pour l'étude de protéines membranaires réinsérées dans un environnement lipidique. Le fait d'être ainsi proche des conditions natives de la protéine en fait sa force.

La RMN du solide a été employée pendant ces travaux de recherche afin de caractériser les différentes conformations moléculaires que peut adopter un transporteur ABC. Bien qu'on puisse obtenir des structures tridimensionnelles par RMN du solide, l'état de l'art ne permet actuellement pas d'apporter la structure complète d'un transporteur ABC. En revanche, la méthode est suffisamment puissante pour révéler des informations sur les détails conformationnels de la protéine dans un environnement lipidique dans différents stades du cycle d'exportation. Elle devient ainsi complémentaire de la cristallographie.

Le système biologique étudié est la protéine BmrA. BmrA est un transporteur ABC (120 kDa) provenant de l'organisme *Bacillus subtilis*, il est qualifié de multi-drogue car ce transporteur possède la propriété d'exporter une grande variété de substrat.

L'un des premiers aspects approché a été l'élaboration et l'optimisation de l'expression, purification et reconstitution dans un environnement lipidique de la protéine BmrA, dans le but d'obtenir des échantillons compatibles avec les exigences de la RMN (isotopiquement marqué ^{13}C, ^{15}N et concentré). Nous avons déterminé que la reconstitution de la protéine dans des lipides provenant de *Bacillus subtilis* avec un ratio lipide/protéine (m/m) de 0.5 conduit à une insertion optimale de la protéine dans son environnement lipidique. De plus, la préparation répond aux deux principales exigences de la RMN, elle conduit à l'obtention de spectre ayant un rapport signal sur bruit élevé et à l'obtention d'échantillons qui montrent une stabilité sur des années.

L'étape de reconstitution consiste à appauvrir en détergent une protéine membranaire en présence de liposomes pour que celle-ci s'intègre dans ces derniers. Cette étape fut étudiée en quasi temps-réel par la quantification de détergent par spectrométrie de masse, par microscopie électronique et par des tests biochimiques. En utilisant ces paramètres nous avons adapté pour la RMN du solide une méthode de reconstitution élaborée pour des études de microscopie électroniques, GRecon. Cette méthode consiste à reconstituer une protéine membranaire en utilisant la combinaison de trois gradient en un seul : un gradient de sucrose, de lipides et de cyclodextrines. Nous avons établis les conditions optimales pour lesquelles la reconstitution de la protéine BmrA était la plus fiable d'un point de vue quantitatif et qualitatif *via* cette méthode. Après une adaptation à large échelle, l'échantillon découlant de cette méthode donne des spectres RMN très similaires en terme de déplacement chimique et de signal sur bruit à la méthode classique de reconstitution en dialyse. En réduisant l'étape de reconstitution d'un facteur 5 en terme de temps, cette méthode représente une bonne alternative pour la préparation d'échantillon de protéine membranaire pour la RMN du solide.

Dans le but d'étudier les différentes conformations que peut adopter un transporter ABC, nous avons également développé un protocole reproductible et quantitatif permettant d'induire des changements conformationnels de BmrA par l'utilisation d'ATP-Mg^{2+} et de vanadate. En utilisant le métabolisme bactérien à notre avantage, nous avons enrichi sélectivement en ^{13}C des acides aminés spécifiques de BmrA, dans le but de réduire le

nombre de signaux se superposant dans les spectres de cette protéine. Ce marquage sélectif nous a permis l'enregistrement de spectres bidimensionnels permettant de suivre les différentes perturbations de déplacement chimique dues aux changements de conformation de la protéine. Nous avons remarqué l'apparition de nouveaux signaux, concomitant avec les différentes perturbations de déplacement chimique de large amplitude soulignant d'importants changements au niveau de la conformation et de la flexibilité de la protéine. Dans le but d'identifier ces nouveaux signaux, nous avons utilisé un remplacement du cofacteur Mg^{2+} par un Mn^{2+} ayant des propriétés paramagnétiques et préservant les propriétés de cofacteur. L'effet paramagnétique du Mn^{2+} se traduit par une disparition des résonances provenant de résidus d'acide aminé à proximité du cation. L'expérience a révélé que les nouvelles résonances sont situées au niveau du site de fixation de l'ATP. Nous avons utilisé le fait que la protéine BmrA en conformation fermée piège en ses sites actifs deux Mn^{2+}, pour déterminer par résonance paramagnétique électronique une distance séparant ces deux cations de 1,8 nm.

Nous avons utilisé les spectres obtenus précédemment comme des empreintes digitales de la protéine BmrA en conformation ouverte ou bien fermée. Ces empreintes nous ont permis d'observer des changements de conformation similaires sur un mutant capable de lier l'ATP mais incapable de l'hydrolyser. L'hydrolyse de l'ATP n'est donc pas nécessaire pour passer d'une forme ouverte a fermée. Nous avons également analysé le rôle clef d'une séquence consensus, la X-loop, cette séquence joue un rôle de communication entre la partie motrice de la protéine (qui utilise l'ATP comme énergie) à la partie qui fixe et exporte les molécules. Un mutant de la X-loop a été conçu, ce dernier est actif au niveau de l'hydrolyse de l'ATP mais sa fonction de transport a été abolie. Les études RMN conduites sur ce mutant ont mis en évidence que la protéine semble changer de conformation sans modification de sa flexibilité contrairement à la forme sauvage. Ces mesures suggèrent que la flexibilité semble le point central dans la transmission des changements conformationnels nécessaires de la partie motrice à la partie d'exportation de molécules. Ces observations seraient semblables à un moteur tournant à plein régime qui ne serait pas connecté de manière rigide à un arbre de transmission le reliant au système de transport.

Summary

ATP binding cassette (ABC) transporters can translocate a variety of molecules by coupling drug/lipid efflux with an ATP-Mg^{2+} fuelled engine. They are found in all forms of life and they are involved in a number of drug resistances including anti-cancer drugs and antibiotics. Herbicides or other xenobiotic resistances in plants or microorganisms are also related to ABC transporters.

Our studies focus on the drug exporter BmrA (130 kDa) from *Bacillus subtilis* as a model system and homologue of the human P-glycoprotein that is involved in multidrug resistance in cancer. The homodimeric drug exporter BmrA was also chosen as a model because it can be overproduced in large quantities with stable isotope labelling. We show that the reconstitution of this protein in lipids from *Bacillus subtilis* at a lipid-protein ratio of 0.5 m/m allows an optimal protein insertion into lipid bilayer as well as it complies with the two central NMR requirements: high signal-to-noise in the spectra and sample stability over a time period of years. The obtained spectra point to a well-folded protein and a highly homogenous preparation, as witnessed by the narrow resonance lines and the signal dispersion typical of the expected secondary structure distribution of the membrane protein. In the same time, we adapted the GRecon method used in electron microscopy studies for membrane protein reconstitution to the needs of solid-state NMR sample preparation. We followed in detail the reconstitution of the ABC transporter BmrA by dialysis as a reference, and established optimal reconstitution conditions using the combined sucrose/cyclodextrin/lipid gradient characterizing GRecon. We established conditions under which quantitative reconstitution of active protein at low lipid-to-protein ratios can be obtained, and also how to upscale these conditions in order to produce adequate amounts for NMR. NMR spectra recorded on a sample produced by GRecon showed a highly similar fingerprint as those recorded previously on samples reconstituted by dialysis. GRecon sample preparation presents a gain in time of nearly an order of magnitude for reconstitution, and shall represent a valuable alternative in solid-state NMR membrane protein sample preparation.

In order to study the inward-facing and the outward-facing state of the transporter, we developed a reproducible and quantitative protocol of ATP:Mg^{2+}:VO_4^{3-} addition inducing the outward-facing state. We used selectively labelled samples obtained by the addition of

natural abundance residues in the bacterial medium in order to reduce the number of signals in the spectra of this large protein. We recorded solid-state NMR two-dimensional spectra with different mixing times (20 and 200 ms) in order to follow chemical shift changes and identify residues by sequential correlations. The very noticeable apparition of new signals concomitant with the large amplitude of chemical shift perturbations highlight the important flexibility and conformational changes of the protein in presence of ATP:Mg^{2+}:VO_4^{3-} substrate.

In order to identify the residues appearing in the spectra, we use paramagnetic replacement by Mn^{2+} of the Mg^{2+} acting as a co-factor in the active site. The paramagnetic relaxation enhancements caused the Mn^{2+} revealed that the amino acids appearing in the spectra are located in proximity to the ATP binding pocket. Besides, EPR measurements confirmed the closed state of the protein by identifying the corresponding 1.8 nm distances between two Mn^{2+}.

We investigate on the conformational differences identified between the inward-facing (IF) and outward-facing (OF) state in the ABC transporter BmrA reconstituted in its natural lipids. The observation of numerous chemical shift perturbations (CSPs), as well as the apparition new signals are observed for a hydrolysis-incompetent mutant on addition of ATP, indicating that hydrolysis is not required for the IF to OF transition in BmrA.

We also analyze the mechanistic of the X-loop motif described to be involved in the communication between the NBDs and the TMDs part. We observe for a mutant protein in which transport is abolished, but which remains ATPase active, an incomplete transition since only a subset of CSPs is observed, as well as lack of rigidification. This suggests that the change in dynamics might be central for transmitting the relevant conformational changes to the part of the protein driving transport, concomitant of an engine which is turning an input shaft, but which fails to connect in a rigid manner, trough adequate gears, with the output shaft driving the pump.

List of illustrations

List of illustrations for supplementary material

Chapter VI Membrane protein reconstitution strategies for solid-state NMR studies

Chapter VII Solid-state NMR and PER Mn^{2+} substituted ATP-fueled protein engines

Chapter VIII Flexible-to-rigid transition is central for transport but not for hydrolysis in BmrA

Table of contents

Abbreviations

ABC	ATP Binding Cassette
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
AMPPNP	Adenylyl-imidodiphosphate
ATP	Adenosine diphosphate
cDNA	complementary DNA
CMC	Critical micelle concentration
CP	Cross Polarization
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
DEER	Double Electron-Electron Resonance
DDM	n-Dodecyl β-D-maltoside
DTT	Dithiothreitol
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
ECL	Extracellular loop
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EPA	Egg phosphatidylcholine
EPC	Egg phosphatidic acid
HSQC	Heteronuclear single quantum coherence spectroscopy
LRET	Luminescence Resonance Energy Transfer
LPR	Lipid-protein ratio
Kav	Effective partition coefficient
ICL	Intracellular loop
INEPT	Insensitive nuclei enhanced by polarization transfer
IF	Inward-facing
MM	Molecular Mass
NADH	Nicotinamide adénine dinucléotide
NBD	Nucleotide-binding domains
NTA	Nitrilotriacetic acid
NTD	N-terminal domain
OF	Outward-facing
PDB	Protein data bank
Pi	Inorganic phosphate
PPM	Parts-per-million
PRE	Paramagnetic Relaxation Enhancement
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
Vi	Inorganic vanadate
RMSD	Root-Mean-Square Deviation
TMD	Transmembrane domain
TMH	Transmembrane helice

Related publications and contributions

Chapter III – Selective labelling and unlabelling for NMR
Selective labelling and unlabelling strategies in protein NMR spectroscopy
Lacabanne, D ; Meier, B.H and Böckmann A

Submitted - *J Biomol NMR.*

Chapter IV – Production, purification and reconstitution of the ABC transporter BmrA for solid-state NMR studies
Adapted from
Kunert, B.; Gardiennet, C.; Lacabanne, D.; Calles-Garcia, D.; Falson, P.; Jault, J.-M.; Meier, B. H.; Penin, F.; Böckmann, A. Efficient and Stable Reconstitution of the ABC Transporter BmrA for Solid-State NMR Studies. *Front Mol Biosci* 2014, *1*, 5.

Contribution: The chapter describes the detail of the sample preparation for my thesis. It is based on the paper by Kunert et al. to which I contributed as a master student: I conducted experiments for the analysis of BmrA reconstitution in *B. subtilis* lipids with different LPRs by sucrose gradient centrifugation. The text is adapted to describe NMR sample preparation used for the experiments in this thesis.

Chapter V – Membrane protein preparation for solid-state NMR studies
Lacabanne, D.; Kunert, B.; Gardiennet, C.; Meier, B.H.; Böckmann, A. Sample preparation for membrane protein structural studies by solid-state NMR. *Methods Mol Biol. 2017, 1635:345-358.*

Contribution: Wrote book chapter

Chapter VI – Membrane protein reconstitution strategies for solid-state NMR studies
Adapted from
Lacabanne, D.; Lends, A.; Danis, C.; Kunert, B.; Fogeron, ML.; Jirasko, V.; Chuilon, C.; Lecoq, L.; Orelle, C.; Prier, C.; Chaptal, V.; Falson, P.; Jault, JM.; Meier, B.H.; Böckmann, A. Gradient reconstitution of membrane proteins for solid-state NMR studies.

Accepted - *J Biomol NMR.*

Contribution: Carried out biochemical experiments. Participated in data processing and interpretation, prepared manuscript.

Alons Lends participated in GRecon experiment optimization and recorded NMR spectrum.

Chapter VII – Solid-state NMR and EPR Mn^{2+} substituted ATP-fueled protein engines
Wiegand, T.; Lacabanne, D.; Keller, K.; Cadalbert, R.; Lecoq. L.; Yulikov, M.; Terradot, L.; Jeschke, G.; Meier, B.H.; Böckmann A. Solid-state NMR and EPR Spectroscopy of Mn^{2+} - Substituted ATP-Fueled Protein Engines. *Angew Chem Int Ed Engl. 2017, 56(12):3369-3373*

Contribution: shared first author and contributed to writing the manuscript with Dr. Thomas Wiegand and Katharina Keller. Prepared all BmrA samples for NMR and EPR. Carried out NMR measurements on BmrA, together with Lauriane Lecoq. Analyzed BmrA NMR spectra.

Dr. Thomas Wiegand carried out NMR measurements, data processing and interpretation on DnaB. Participated in EPR measurements and data interpretation.

Katharina Keller carried out EPR measurements, process and interpret all EPR data on DnaB and BmrA.

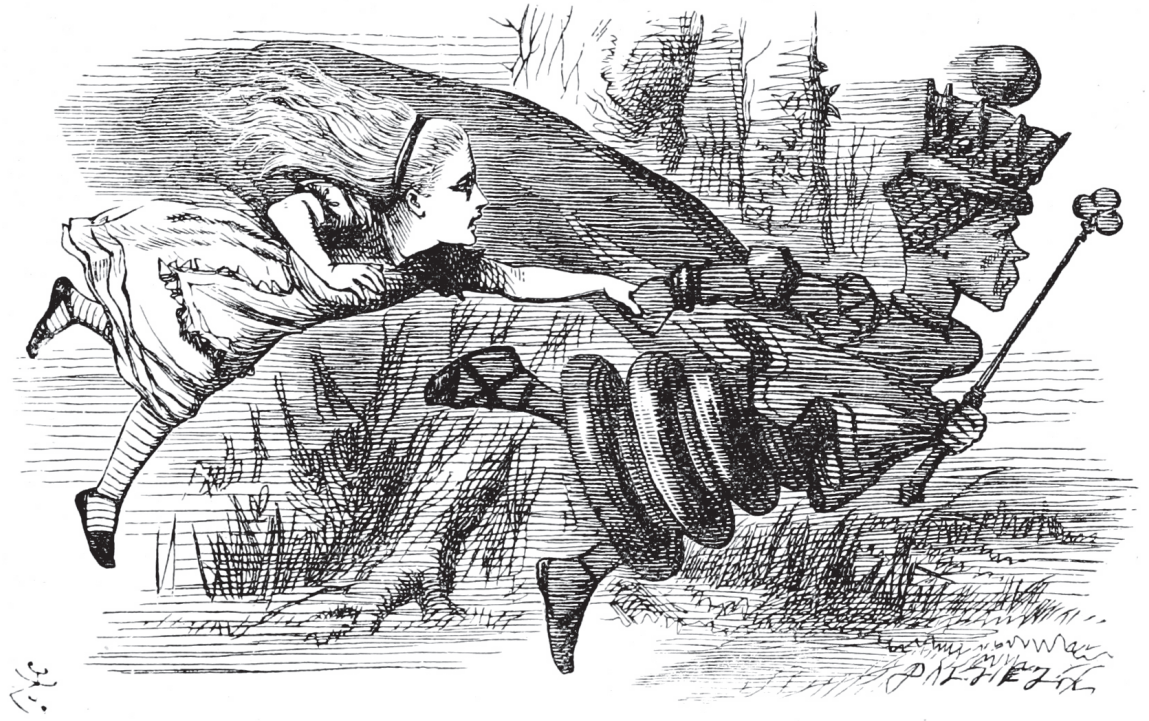
Chapter VIII – Flexible-to-rigid transition is central for transport but not for hydrolysis in the ABC transporter BmrA
Lacabanne, D.; Orelle, C.; Wiegand, T.; Lecoq, L.; Kunert, B.; Chuilon, C.; Ravaud, S.; Jault, JM.; Meier, B.H.; Böckmann, A. Engaging the gear: flexible-to-rigid transition is central for transport but not for hydrolysis in the ABC transporter BmrA.
In preparation

Contribution: shared first author and contributed to writing the manuscript with Dr. Cédric Orelle. Prepared all NMR samples of BmrA and recorded NMR spectra on BmrA. Analyzed data. Participated on ATPase activity, drug-transport and drug-affinity experiments.

Dr. Cédric Orelle participated on ATPase activity, drug-transport and drug-affinity experiments.

Dr. Thomas Wiegand carried out ^{31}P NMR measurements on BmrA.

Foreword



Alice meeting the Red Queen

illustration of the Red Queen race scene by Sir John Tenniel for Lewis Carroll

« Now, here, you see, it takes all the running you can do, to keep in the same place »

The Red Queen race scene, from the novel by Lewis Carroll *Through the Looking-Glass, and What Alice Found There,* in which Alice meets the Red Queen very well describes the constant evolution of species.

Leigh Van Valen made an assumption based on the Red Queen scene called "The Red Queen Theory" (Van Valen 1973). He put forward the evolutionary hypothesis that species have to run to keep the same place, in other words, they have to evolve in order to exist. The basis of this theory goes down to the evolutionary arms race, and this race between predator and prey is a driving force in evolution.

Similarly, antibiotic resistance of bacteria is caused by the arms race (Jefferies *et al.* 2011; Woolhouse *et al.* 2002). Indeed, to keep their same place, bacteria evolve by developing counter-measures against antibiotics. However, the evolutionary arms race against drugs does not concern only bacteria but also a vast majority of other organisms: fungi (Sanglard 2016), cancer cells (Housman *et al.* 2014), parasites (Hyde 2007), plants (Powles & Yu 2010) etc.

The Red Queen hypothesis suggests that adaptations occur in the organisms that allow them to expand in the presence of drugs. One of the more common arms implicated in drug resistance is the ABC transporters (Housman *et al.* 2014; Hyde 2007; Powles & Yu 2010; Sanglard 2016). This arm is found across all phyla of life, nevertheless the greater interest of the scientific community in these transporters is mainly because ABC transporters are involved in cancer resistance. Paradoxically, ABC transporters are equally involved in glyphosate resistance in plants (Tani *et al.* 2016) this resistance leads farmers to increase the quantity of herbicide or to mix it with others. Moreover, ABC transporters are also found in parasite resistance which is tightly bound to intensive animal farming. As an example, the proliferation of the salmon louse (*Lepeophtheirus salmonis*) –an ectoparasite of salmon, causing economic damage in the aquaculture of Atlantic salmon– is facilitated in intensive aquaculture of salmonids. In order to overcome the proliferation, aquaculturists treat salmon with insecticide. However, after decades of treatment, the salmon louse is becoming resistant to insecticide. It was shown that ABC transporters were overexpressed by the louse in the presence of insecticides (Heumann *et al.* 2012) and some other ABC transporters are equally suspected to be involved in this resistance (Carmona-Antoñanzas *et al.* 2015; Valenzuela-Muñoz *et al.* 2015). As was the case of herbicides for plants, the quantity of insecticides was drastically increased and mixed with other insecticides to overcome the resistance issue. Whether it is for plants or salmon, overtreated food ends up on our dinner plates.

So yes, ABC transporters are involved in cancer resistance, but indirectly they are also involved in cancer appearance in humans through drug pollution.

Carmona-Antoñanzas, G., Carmichael, S.N., Heumann, J., Taggart, J.B., Gharbi, K., Bron, J.E., Bekaert, M. & Sturm, A. (2015) A Survey of the ATP-Binding Cassette (ABC) Gene Superfamily in the Salmon Louse (Lepeophtheirus salmonis). *PLoS ONE* 10, e0137394.

Heumann, J., Carmichael, S., Bron, J.E., Tildesley, A. & Sturm, A. (2012) Molecular cloning and characterisation of a novel P-glycoprotein in the salmon louse Lepeophtheirus salmonis. *Comparative biochemistry and physiology. Toxicology & pharmacology : CBP* 155, 198–205.

Housman, G., Byler, S., Heerboth, S., Lapinska, K., Longacre, M., Snyder, N. & Sarkar, S. (2014) Drug resistance in cancer: an overview. *Cancers* 6, 1769–1792.

Hyde, J.E. (2007) Drug-resistant malaria - an insight. *The FEBS journal* 274, 4688–4698.

Jefferies, J.M.C., Clarke, S.C., Webb, J.S. & Kraaijeveld, A.R. (2011) Risk of red queen dynamics in pneumococcal vaccine strategy. *Trends in Microbiology* 19, 377–381.

Powles, S.B. & Yu, Q. (2010) Evolution in action: plants resistant to herbicides. *Annual review of plant biology* 61, 317–347.

Sanglard, D. (2016) Emerging Threats in Antifungal-Resistant Fungal Pathogens. *Frontiers in medicine* 3, 11.

Tani, E., Chachalis, D., Travlos, I.S. & Bilalis, D. (2016) Environmental Conditions Influence Induction of Key ABC-Transporter Genes Affecting Glyphosate Resistance Mechanism in Conyza canadensis. *International journal of molecular sciences* 17.

Valenzuela-Muñoz, V., Sturm, A. & Gallardo-Escárate, C. (2015) Transcriptomic insights on the ABC transporter gene family in the salmon louse Caligus rogercresseyi. *Parasites & vectors* 8, 209.

Van Valen, L. (1973) A new evolutionary law. *Evolutionary Theory* 1, 1–30.

Woolhouse, M.E.J., Webster, J.P., Domingo, E., Charlesworth, B. & Levin, B.R. (2002) Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature genetics* 32, 569–577.

General introduction
Objectives of the thesis work

ATP-binding cassette (ABC) transporters are membrane proteins that translocate a variety of substrates across the membrane. The ABC exporter group acts as an efflux pump by coupling the ATP-driven movement of the nucleotide binding domains (NBDs) to the transmembrane domains (TMDs). ABC transporters switch between two different orientations: open in the direction to the cytosol environment – inward facing state – or open in the direction to the extracellular environment – outward facing state. One is the inward-facing conformation in which the substrate-binding site is open onto the cytosol side. This enables the hydrophilic molecules to enter the binding site directly from the inner leaflet of the phospholipid bilayer. The other is the outward-facing conformation in which the drug-binding site is open onto the extracellular side. The molecules are released from the binding site into the extracellular medium.

This group is present in the three domains of life: bacteria, archaea and eukaryota. These ABC exporters are essential, and most corresponding genes encode membrane-bound proteins that participate directly to the transport of a wide range of substrates including metal ions, peptides, amino acids, sugars and a large number of hydrophobic compounds and metabolites across the membranes (Vasiliou *et al.* 2009).

However, some ABC transporters are less specific towards a substrate and can also export a large number of different compounds. These ABC transporters belong to the category of multidrug resistance because they are associated with a multidrug resistance phenotype (Choi 2005). ABC exporters are present in the three domains of life and so is the multidrug-resistance phenotype for this reason, the ABC transporters can be defined as a family of transporter proteins responsible for drug resistance and for a low bioavailability of drugs by pumping a variety of drugs out of cells at the expense of ATP hydrolysis (Choi 2005).

Although crystal structures provide an in-depth insight into several steps in the transport mechanism, a comprehensive understanding of the transport cycle still lacks (Beek *et al.* 2014). This lack of knowledge precludes the development of drugs defeating this transport mechanism.

We shall present here our work on ABC multidrug transporter by the means of solid-state nuclear magnetic resonance (RMN). As evidenced in the literature, solid-state NMR is not often used to study ABC transporters at an atomic level, mainly due to the large size of the proteins. Solid-state NMR has, despite this, a high potential to contribute interesting

information, as it is a method of choice for studying a membrane protein in its lipid environment. It enables specific structural and biological questions to be addressed.

This Ph.D research focuses on three major stages typical in NMR investigation: the sample preparation, the resonances assignment and the structural analyses.

The aim of our work is to contribute to the understanding of the mechanics driving export in ABC tranporters at the molecular level. We use the ABC transporter BmrA as a model to develop solid-state NMR strategies for the investigations of this a 2x64.5 kDa-protein. Information from NMR can complement other structural techniques to clarify grey spots in the underlying mechanism of the protein family under study.

General knowledge from the literature about ABC transporters is described in **Chapter 1**. **Chapter 1** focuses on the ABC multidrug transporters and summarizes the last forty years of research with their milestones and major achievements. The anatomy of an ABC exporter is also depicted and the different hypotheses about the transport mechanism outlined in the literature are also reported in this chapter.

Chapter 2 briefly introduces NMR, what results we can expect to obtain on this type of protein. It also highlights the advantages of the solid-state NMR. Chapter 2 expands on the detailed description of the experiments conducted and the different approaches for this thesis work.

Chapter 3 examines all the strategies meant to reduce spectral overlap by unlabelling approaches. It presents a complete overview of the *Escherichia coli* amino-acids metabolism. This overview allows to choose the right combination of amino acids that can be added in the culture-medium with minimal scrambling.

Chapter 4, **5** and **6** are devoted to the sample preparation of the ABC transporter BmrA from the protein expression to the acquisition of the spectrum.

Chapter 4 details the expression in minimum medium, the isolation of the *Escherichia coli* membranes, the extraction of the protein from the membrane and the purification of BmrA. The really robust and reproducible protocol allows for the production of 10 mg of BmrA per liter of culture. This yield is particularly substantial for a membrane protein and makes the sample preparation suitable for solid-state NMR.

Chapter 5 focuses on the sample preparation for structural studies by solid-state NMR. This chapter presents the three critical steps for a sample preparation, i.e. the membrane reconstitution of the protein in native lipids, rotor filling and sample quality

assessment. In addition, it describes the lipid extraction and solubilisation from *Bacillus subtilis*. The lipids are used for the reconstitution of BmrA in order to be as relevant as possible.

Chapter 6 centers mainly on the reconstitution steps of BmrA: it is the process of exchanging detergents bound to the hydrophobic parts of the protein with lipids. This chapter presents two reconstitutions in lipids for NMR methods: a reconstitution in dialysis or an ultra-centrifugal gradient reconstitution (GRecon). The combination of detergent quantification, electron microscopy and biochemical assays makes it possible to track the protein reconstitution in real-time.

Chapter 7 presents the combination of two strategies: one is about specific amino acids labelling in order to reduce overlapping peaks and the other is about paramagnetic relaxation enhancements by replacing ATP-Mg^{2+} substrates by ATP-Mn^{2+} with the aim to identify residues located in the binding pocket.

The strategies developed in **Chapter 7** are meant to identify accurately the different conformations that an ABC transporter can adopt in the presence of a substrate, the inward-facing conformation, which allows ATP binding, and the outward-facing conformation in which the drug-binding site is open onto the extracellular side.

Chapter 8 reports the NMR characterization of these two conformations. They are probed by using ^{13}C NMR but also by using ^{31}P in order to get a structural-sight from the substrate. Besides, we use a mutant for which the drug-transport activity is abolished but not the ATPase activity. The use of this mutant highlights the key-role of a conserved motif – so-called X-loop – in the transmission of the conformational changes between the nucleotide binding domains and the transmembrane domains.

Chapter 9 is the conclusion of the thesis and introduces also on the different perspectives will be developed in particular on the drug-binding states and on the use of NMR at 110 kHz of rotor spinning frequency.

Beek, ter, J., Guskov, A. & Slotboom, D.J. (2014) Structural diversity of ABC transporters. *The Journal of general physiology* 143, 419–435.

Choi, C.-H. (2005) ABC transporters as multidrug resistance mechanisms and the development of chemosensitizers for their reversal. *Cancer cell international* 5, 30.

Vasiliou, V., Vasiliou, K. & Nebert, D.W. (2009) Human ATP-binding cassette (ABC) transporter family. *Human genomics* 3, 281–290.

Chapter I

ABC transporters

Contents

40 years of progress

ATP-binding cassette transporters or ABC transporters have been studied since the first identifications of the histidine permease HisP (Ames & Lever 1970) and the maltose importer MalK (Hazelbauer 1975) respectively in 1970 and 1975. Both of them use the energy of ATP to catalyse the transport of their substrates.

Fifteen years after the first identification, Christopher Higgins described a class of closely related ATP-binding proteins associated with a variety of cellular functions including membrane transport (Higgins *et al.* 1985), (Higgins *et al.* 1986). The ABC superfamily was about to be identified. Eventually, in 1990, the family was clearly identified by Higgins and Ames and was called "ATP-binding cassette superfamily" by Higgins (Hyde *et al.* 1990) and "Traffic ATPases" by Ames (Ames *et al.* 1990). At the same time, the first models of ATP-binding proteins associated with cystic fibrosis, multidrug resistance, and bacterial transport were designed by Higgins and co-workers (Hyde *et al.* 1990). The models were based on the similarities between the predicted secondary structures of members of this family with the structure of the adenylate kinase. The general structure determined by Higgins and co-workers consists of four domains: two of them are highly hydrophobic and behave as integral membrane proteins which span the membrane five or six times via alpha-helices. Moreover, these domains are believed to bind drug-substrates and may mediate their passage across the lipid bilayer. The other two domains couple ATP hydrolysis with transport and are associated with the cytoplasmic face of the membrane.

Meanwhile, another important tree branch grew: the discovery of the P glycoprotein. In the early seventies, multidrug resistant cell lines were observed. These cell lines display an active extrusion system for drugs such as daunorubicin, a chemotherapy medication used to treat cancer (Dano 1973).

In 1976, working on Chinese hamster ovary cells, Julian and Ling demonstrated how those cells were resistant to colchicine and displayed a pleiotropic cross-resistance to a wide range of drugs. Julian and Ling identified a component, a glycoprotein, in the cell membranes. Moreover, they showed that the amount of this cell surface glycoprotein was correlated with the degree of drug resistance. They concluded that the glycoprotein is associated to drug permeability and they named it "P glycoprotein" the letter P meaning "permeability" (Juliano & Ling 1976). Three years later, in 1979, the P glycoprotein was purified from

plasma membrane vesicles of Chinese hamster cells with reduced colchicine permeability (Riordan & Ling 1979).

In 1986, the nucleotide sequence of the cDNA of the *mdr* gene was determined by Gros and co-workers (Gros *et al.* 1986). They deduced that this gene encoded for a P glycoprotein. Knowing the amino acid sequence, they proposed that six pairs of transmembrane domains, a cluster of N-linked glycosylation and a consensus ATP-binding site were part of the P glycoprotein. Besides, with sequences alignment, they highlighted a strong homology between the *mdr* gene and bacterial transport genes. Finally, they suggested that an energy-dependent transport mechanism was responsible for the multidrug-resistant phenotype. Four years later, Higgins and co-workers established that the P glycoprotein was an ATP binding cassette transporter (Hyde *et al.* 1990). Cole and co-workers found in 1992 that another transporter, MRP1 (ABCB1) was associated with a multidrug-resistant phenotype (Cole *et al.* 1992).

The fact that multidrug resistance in human cancer is a process regulated by ABC transporters was a breakthrough and it initiated a "gold rush" in the studies of this protein superfamily. Three years later, in 1995, Alan Senior and co-workers proposed a scheme of catalytic cycle for P glycoprotein in which the two ATP molecules were hydrolysed alternatively: the alternating site mechanism (Senior *et al.* 1995). This model was to be updated in 2009 by Peter Jones and Anthony Georges as a "constant contact alternating site" model in which each nucleotide binding domain (NBD) remains in contact during the ATP hydrolysis (Jones & George 2009).

In 1997, the complete gene sequencing of *Escherichia coli* K-12 (Blattner *et al.* 1997) allowed the identification of 57 ABC transporter genes which represent almost 5% of the genome (Linton & Higgins 1998). Finally, in 2001, the sequencing and the analysis of the human genome highlighted the presence of 48 ABC genes encoding 48 ABC transporters (Venter *et al.* 2015), (Lander *et al.* 2001). There were distributed in 7 different families ABC A, B, C, D, E, F and G (Dean *et al.* 2001).

Twenty years after the identification of the first ATP-binding cassette protein, one important question remains: what is the underlying mechanism driving the function of this protein family? The answer to this question is crucial in order to design drugs which could avoid an efflux transport outside the resistant cells. However, to determine the mechanism of the protein, the structure must be solved. A second "gold rush" followed in order to determine

the general structure of the ABC transporters and the cogwheel organization leading to drug transport. The first structure of the P glycoprotein (P-gp) at low resolution (2.5 nm) obtained by cryo-microscopy was published in 1997. (Rosenberg *et al.* 1997). One year later, the group of Ames and Kim solved the ABC transporter NBDs subunit structure of the histidine permease, HisP, in the presence of ATP (Hung *et al.* 1998). Interestingly, the dimeric arrangement of the subunit displayed a back-to-back organization, **Figure 1A**. In 2000, a second structure of the NBD subunit part of the ABC transporter trehalose/maltose ABC transporter, MalK, was solved (Diederichs *et al.* 2000). The comparison between the ABC subunit HisP and MalK revealed an almost identical fold with a root-mean-square deviation of atomic positions (R.M.S.D) of less than 2 Å. However, the dimeric arrangement of the structure displayed a different organization in which each monomer was in a face-to-face arrangement, **Figure 1B**.

In the same year, two structures of the NBD subunits of the ABC transporter Malk and of Rad50 – an ABC protein implicated in the DNA repair system – were solved. The NBDs were crystallized in the presence of a non-hydrolysable ATP analogue (AMPPNP). In that case, they showed a head-to-tail organization (Hopfner *et al.* 2000), **Figure 1C**. Hopfner and co-workers put forward the hypothesis that ATP-driven dimerization of ABC's represented the force-transducing "power stroke" of ABC transporter pumps. Finally, in 2002, with the determination of the structure of MJ0796 – an ABC transporter from *Methanocaldococcus jannaschi* which is a thermophilic methanogenic archaean – in the presence of ATP, John Hunt's group confirmed the head-to-tail hypothesis. A head-to-tail organization where ATP is sandwiched between both monomers is now accepted as a general form of ABC transporters (Smith *et al.* 2002).

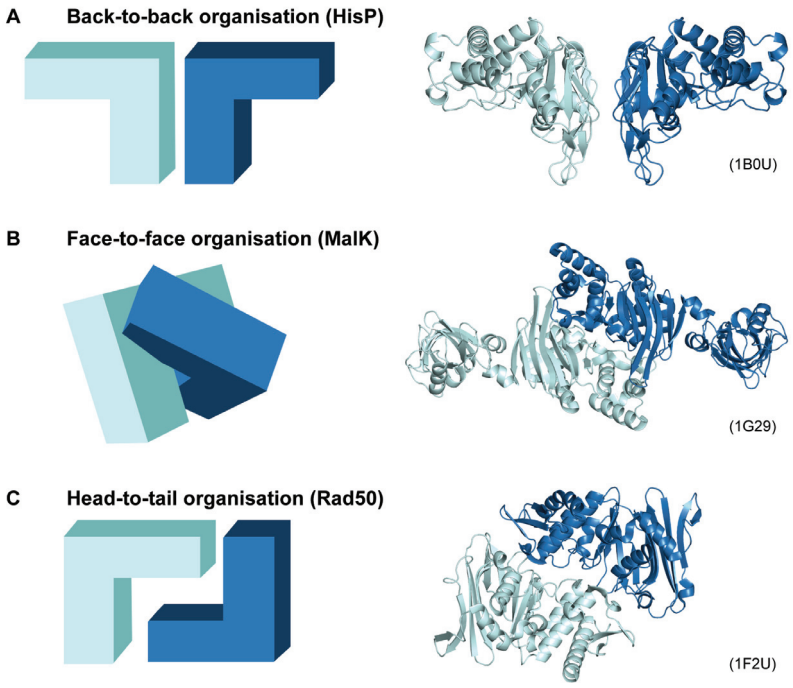


Figure 1. Schematic representations of the NBD dimers, adapted from (Kerr 2002). In each case the L-shapes represent the NBD with the corresponding structure and the associated PDB accession number. The different NBDs organizations are presented by (**A**) the back-to-back from HisP NBDs structure ; (**B**) face-to-face from MalK NBDs structure (**C**) and head-to-tail from Rad50 structure.

In 2002, the first structure of the ABC importer BtuCD – a vitamin B12 importer from E. coli – is solved by Locher's group (Locher *et al.* 2002). In 2004, Christopher Higgins and Kenneth Linton proposed the "ATP switch model" for ABC transporter, based on the structure of BtuCD and 11 structures of NBDs from ABC transporters (Higgins & Linton 2004).

In 2006, Locher's group published the full structure of the *Staphylococcus aureus* multidrug efflux pump Sav1866. The structure of Sav1866 was fully consistent with all biochemical and structural data from the literature (Dawson & Locher 2006). The access to different full ABC exporters/importers structures marked an impressive tour de force (van der Does & Tampé 2004). Indeed, due to the high homology within the family of ABC transporters exporting chemotherapy drugs, the structures of Sav1866 and BtuCD provide critical new tools to address major health issues.

Today, 43 structures of ABC transporters (importer and exporter) have been solved since 2002 with the last one in May 2017 of the human multidrug transporter ABCG2 by Locher's group (Taylor *et al.* 2017a).

40 years of progress have been described and still, as Rees and co-workers (Rees *et al.* 2009) noticed, our current understanding of ABC transporters is like that of a play. A play whose synopsis is known, and so are its characters' names, looks and the general actions of the play. However the full play that gives stage directions describing characters' movements and specific actions is missing. The composition of the full play will make a giant leap for the ABC transporters research.

The next paragraphs of the chapter will describe the state of knowledge of the different steps involved in ABC export, in order to present the different characters of the play.

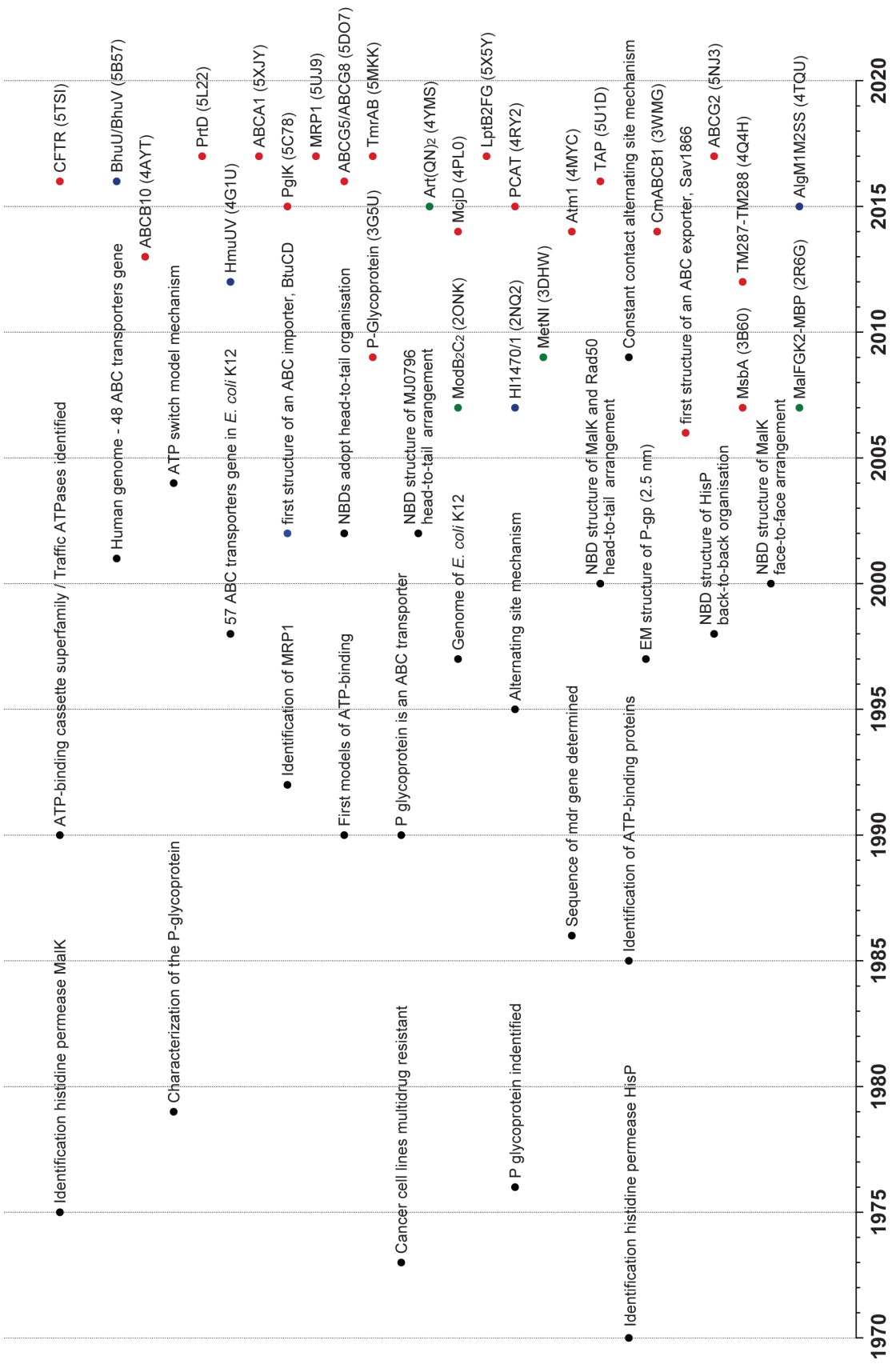


Figure 2. Timeline showing selected key discoveries and events in ABC research. Each structure is followed by the associated PDB number, ABC exporters were labelled with red dots, ABC importers I with blue dots and ABC importers II with green dots.

Anatomy of an ABC transporter

ABC transporters are grouped into exporters and importers. The importer group is subdivided into two classes (I and II), depending on details of their architecture and mechanism (Holland & Blight 1999), (Rees *et al.* 2009) and (Locher 2004). Moreover, apart few exceptions like AtABCB14 – a malate importer – eukaryotes only have exporter while bacteria employ both ABC exporter and importer (Lee *et al.* 2008) and (Wilkens 2015).

The canonical ABC transporter or the minimal functional unit contains four functional domains: two Nucleotide-Binding Domains (NBD1, NBD2) and two Transmembrane Domains (TMD1, TMD2). However, various combination of these domains can be found, indeed they can be separated into :

i. Four functional polypeptide chains where NBDs and TMDs can be identical (BtuCDF, ArtMP-J) or different (NikABCDE, MalEFGK2), presented **Figure 3A**.

ii. Three separated functional polypeptide (FhuBCD with 2 NBDs and with TMDs fused together or RbsABC with 2 TMDs and with NBDs fused together), presented **Figure 3B**.

iii. Two separated functional polypeptide where NBD and TMD are fused together to form a homodimeric (MsbA, Sav1866, BmrA) or heterodimeric (TM287/288, CydDE) protein, presented **Figure 3C**.

iv. A single polypeptide that contains all four functional units, where NBDs and TMDs are fused into one monomer (P-gp, CFTR), presented **Figure 3D**.

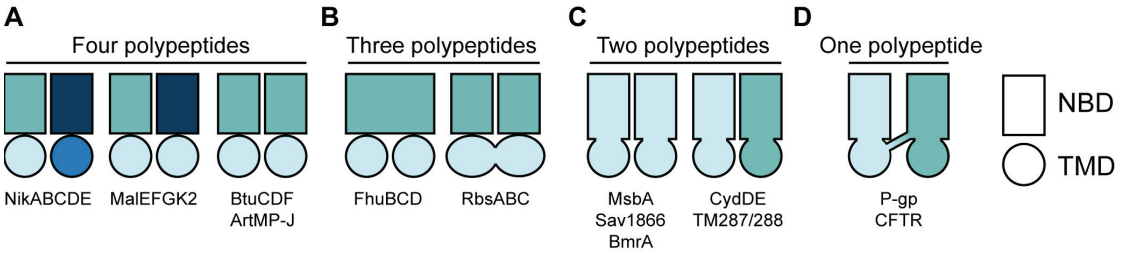


Figure 3. Domain organization of ABC transporters exporters and importers comprise two TMDs, two NBDs, adapted from (Theodoulou & Kerr 2015). The organization of core and additional domains is shown for different representative topologies in each subfamily with the number of polypeptides chains; four (**A**), three (**B**), two (**C**) and one (**D**).

In the following, the anatomy of the NBDs and the TMDs will be described for the ABC exporter group, the importer group will be not discussed here.

1. The Nucleotide-Binding Domains

ABC transporters display two Nucleotide-Binding Domains (NBDs) which are able to bind ATP and catalyse the ATP hydrolysis to power the molecules transporter. Moreover, NBDs are present in all ABC transporters and form a highly conserved region, regardless of the organism. However, the organization of the NBDs is not exclusive to ABC transporters but concerns equally a large panel of proteins which are gathered in a superfamily: P-loop NTPases (Vetter & Wittinghofer 1999). This superfamily is currently made up of 871450 proteins from 3260 genomes (http://supfam.org/; SSF52540 (Wilson *et al.* 2009). This structural organization can be defined by a RecA-like core domain present in all P-loop NTPases, preceded by a β-subdomain and followed by an α-subdomain, which are found only in ABC transporters (Geourjon *et al.* 2001), (Beek *et al.* 2014), **Figures 4A and B**. The RecA-like domain typically consists of two β-sheets and six α-helices (including the Walker A and B motifs). The α-subdomain consists of three helices and the signature motif (Davidson *et al.* 2008), **Figure 4A**.

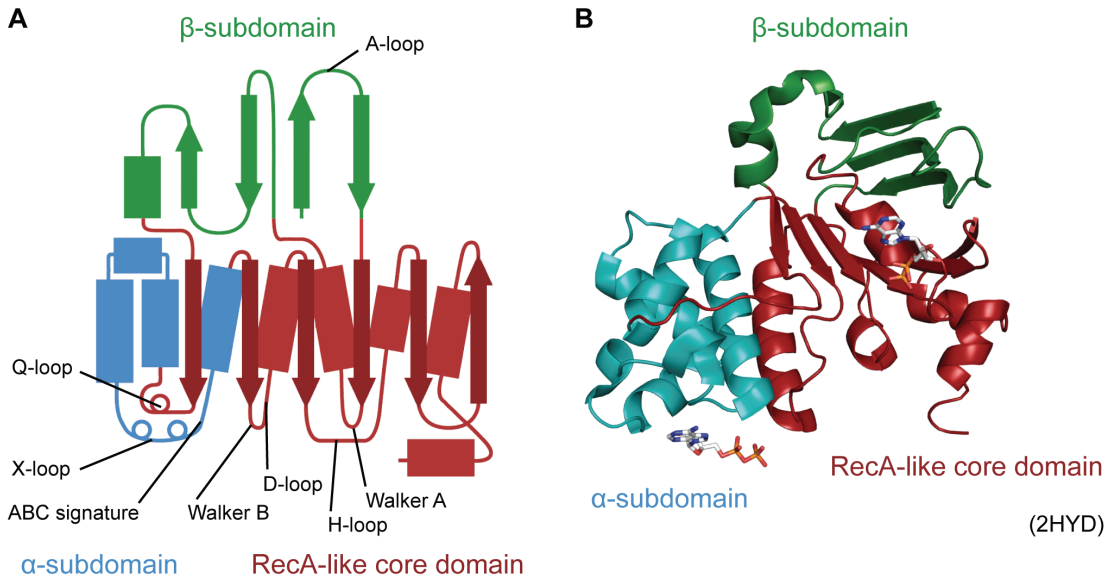


Figure 4. General NBD-structural organization of an ABC exporter. (**A**) Topology diagram of an ABC transporter NBD. (**B**) Conserved sequence motifs are indicated on the diagram. NBD structure from SAV1866 (2HYD). In both Figures each domain is represented with different colour, in blue the α-subdomain, in green the β-subdomain and in red the RecA-like core domain.

NBDs contain seven highly conserved motifs present in the three different regions:

1.1. Motif found in the β-subdomain

The A-loop, located in the β-subdomain 25-29 residues upstream of the Walker A motif; this motif contains a conserved aromatic residue (usually a tyrosine) located on a loop which interacts with the adenosine ring of ATP via π-π interactions, hydrogen bonding and van der Waals contacts (Ambudkar *et al.* 2006). This conserved aromatic residue that helps to position the ATP by stabilizing the adenine moiety plays an essential role for the binding and the hydrolysis of ATP (Ambudkar *et al.* 2006), (Beek *et al.* 2014).

1.2. Motifs found in the RecA-like domain

The **Walker A** motif or **P-loop** (**Phosphate-binding loop**) motif adopts an archetypal loop shape which wraps around the β- and γ-phosphate of ATP. The sequence comprises GXXXXGK(T/S) residues followed by a highly conserved lysine (Saraste *et al.* 1990). The amide backbone of the motif and the ε-amino group of the conserved lysine residue form a network of hydrogen bonds with β- and γ-phosphates which stabilize the ATP molecules (Beek *et al.* 2014).

The **Walker B** motif presents a primary structure in β-strand composed of four hydrophobic residues ended by an aspartate and a second acidic residue (frequently a glutamate residues). The aspartate residue helps to coordinate the magnesium ion by hydrogen bonding through a water molecule (Gaudet & Wiley 2001), (Orelle & Jault 2016). The glutamate residue operates as a catalytic base for ATP hydrolysis (Orelle *et al.* 2003). Indeed, having ionization, the glutamate residue is poised to function as a general-base catalyst.

The D-loop (Dimerization-loop) directly follows the Walker B motif and carries the conserved motif composed of four residues SALD. This motif seems to mediate communication between both active sites. Indeed, the D-loop interacts, through hydrogen and electrostatic bonds, with the Walker A motif and the H-loop of the opposite NBD (Jones & George 2012). The D-loop is crucial in the control of ATP hydrolysis. The loop is involved in the stabilization of the catalytic glutamate from Walker B and in the connection with the nucleophilic water. Moreover, It may plays a role in the directionality of active export

(shown in a heterodimeic ABC transporter) (Jones & George 2012), (Grossmann *et al.* 2014), (Orelle & Jault 2016).

The H-loop or **switch region** contains a conserved histidine residue supposed to play a role in organising the geometry of the prehydrolytic state. This histidine residue forms a chemical hinge between the D-loop, the catalytic glutamate (walker B motif) and the γ-phosphate of the ATP (Zaitseva *et al.* 2005), (Jones & George 2012), (Orelle & Jault 2016). A mutation of the conserved histidine shows that this chemical hinge is crucial for the transporter's ATPase activity (Davidson & Sharma 1997).

The Q-loop is composed of eight residues ended by a highly conserved glutamine. This motif is located at the interface between the RecA-like subdomain and the α-helical subdomain. Conformational changes in the Q-loop allow the conserved glutamine residue to move in and out of the active site during the hydrolysis cycle. The Q-loop contributes to the assembling of the active site when Mg-ATP is bound and to the dissociation of the active site once ATP is hydrolysed (Jones & George 2002).

1.3. Motifs found in the α-helical subdomain

The ABC signature motif or **C motif** is the hallmark of ABC superfamily (Schmitt & Tampé 2002). The ABC signature is found in the α-helical subdomain but is not present in other P-loop NTPases such as the F_1-ATPase, which justifies its specificity as ABC transporters. This LSGGQ motif is located at the N-terminal end of a long helice that directs the positive charge of the helical dipole towards the γ-phosphate of ATP. Consequently, another network of hydrogen bonds is formed and it allows a better stabilization of ATP. Moreover, the ABC signature motif and the Walker A motif of the second NBD are situated opposite to each other, on either side of the ATP molecule. In other words, the ATP molecules are sandwiched by these two motifs.

The X-loop seems to be a characteristic of ABC exporters and the motif is defined by a TEVGERG sequence. This motif is located just before the ABC signature; this proximity seems to imply a role in the ATP binding and its hydrolysis. Besides, it may transmit conformational changes to the intracellular transmembrane domains. Regarding this

apparent function in the cross-linking between NBD and TMD, this sequence is called X-loop (Dawson & Locher 2006) and (Oancea *et al.* 2009).

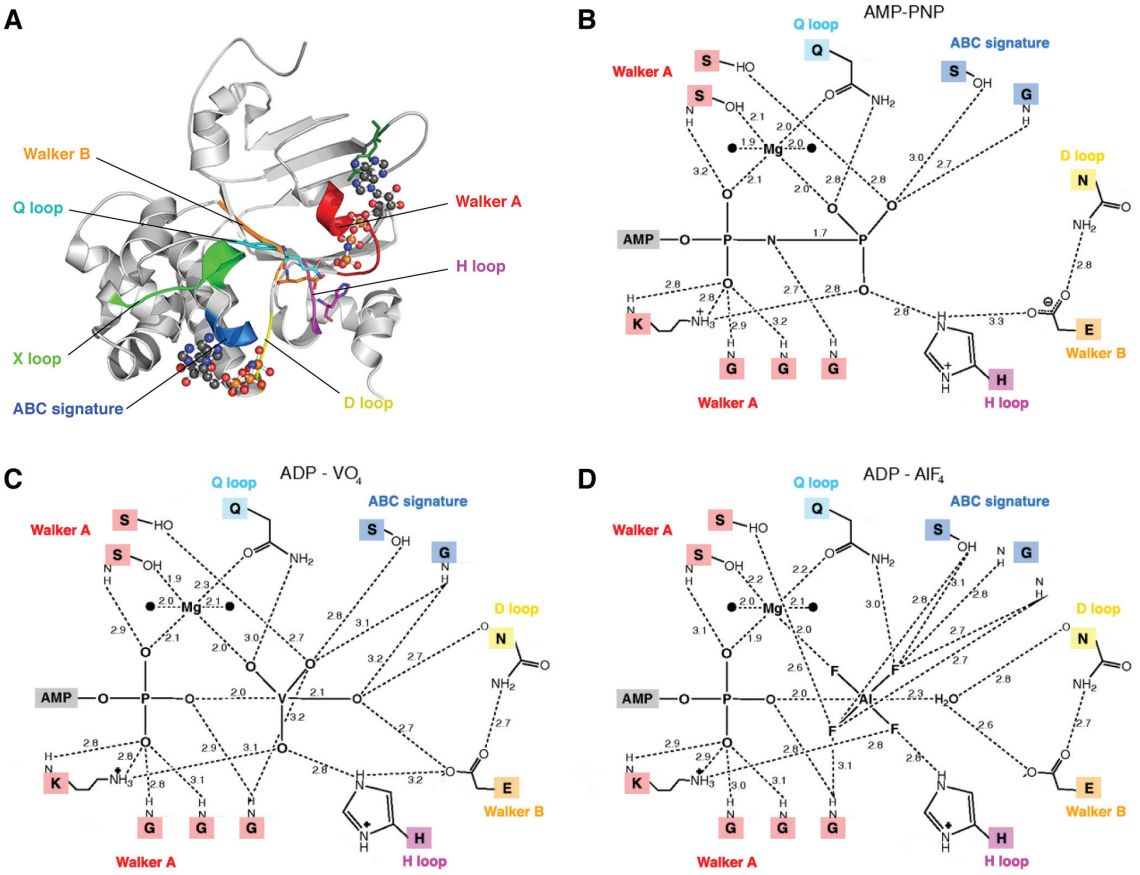


Figure 5. Structure of an ABC transporter NBD and its interaction map with ATP analogues, adapted from (Oldham & Chen 2011). (**A**) NBD structure of the maltose transporter $MalFGK_2$ (2R6G) (**B**) interaction map with AMPPNP (3RLF) (**C**) ADP-VO_4 (3PUV) and (**D**) ADP-AlF_4 (3PUW) with the associated sequence motifs.

2. The Transmembrane Domains

Trans-membrane domains are constituted of hydrophobic α-helices which cross the lipid bilayer in order to form a translocation path for different substrates. ABC exporters are composed of twelve α-helices, six for each TMD. Ten to twelve α-helices seem to be the minimum necessary to form a functional pore for substrate translocation across the membrane (Locher *et al.* 2002). Unlike the NBD, the amino-acids sequence of TMD is not as strictly conserved between the different ABC exporters. This diversity of sequences highlights the wide variety of transported substrates found in this family.

All the structures of ABC exporters available seem to have a similar structural organization of their TMDs. They are composed of six transmembrane α-helices (TMH) for each

monomer, twelve α-helices in total. These six TMH can be grouped in three pairs: TMH1-2, TMH4-5, TMH3-6, **Figures 6A and B**. Furthermore, in the presence of nucleotide, the TMH splits into two wings that point away from one another in the outer leaflet of the membrane, thus forming an outward-facing conformation, **Figures 6C** (Dawson & Locher 2006), (Oldham *et al.* 2008). This conformation seems to allow the release of the dug out the membrane. Each wing is a cluster of the TMH1-2 from one monomer and TMH3-6 from the other monomer, **Figures 6A and C**. Regarding the description of SAV1866, the TMH1 and TMH3 are related by a twofold rotation axis between TMH4 and TMH6. This rotation is around a symmetry axis parallel to the membrane plane, **Figures 6A** (Dawson & Locher 2006).

The six TMHs are interlinked by three extracellular loops (ECL) on the extracellular side of the transporter and also by two intracellular loops (ICL) on the cytosolic side. ECL1 links TMH1 and TMH2, ECL2 TMH3 and TMH4, ECL3 TMH 5 and 6 whereas ICL1 links TMH2 and TMH3 and ICL2 links TMH4 and TMH5, **Figures 6A**. Extracellular loops are not strictly conserved but seem to play a role in drug binding. The two intracellular loops so-called "coupling helices" play a key-role in the interface between TMDs and NBDs. ICL1 makes the intra-molecular junction between the TMD and the NBD, whereas the ICL2 makes the inter-molecular junction between the TMD and the NBD of the other monomer.

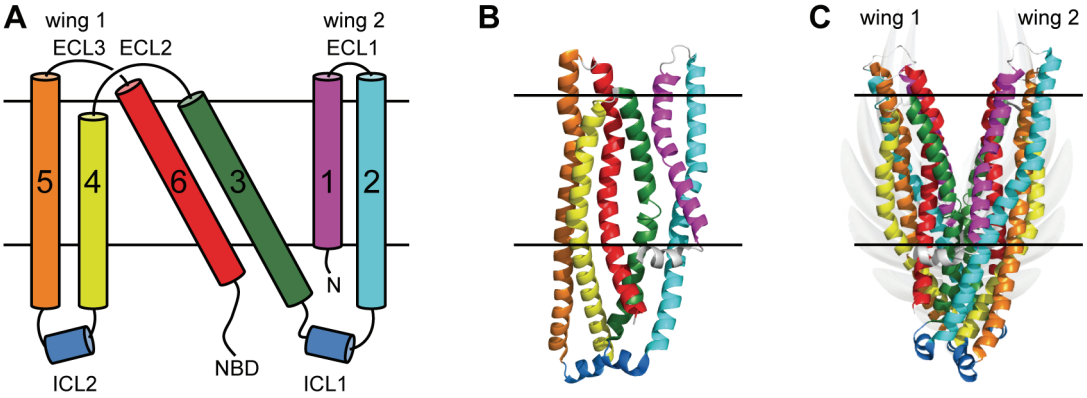


Figure 6. Topology and structure of a multidrug ABC transporter based on Sav1866, adapted from (Locher 2009). (**A**) Topology of TMD from one monomer, with an emphasis on the pseudo-twofold symmetry between the sets of TMH helices 1–3 and 4–6. (**B**) TMD structure from one monomer of SAV1866 (2HYD) (**C**) TMD structure of the homodimeric Sav1866, the TMH splitting into two wings that point away from one another in the outer leaflet of the membrane, thus forming an outward-facing conformation. Lipid-bilayer is represented as black line.

Transport mechanisms of multidrug ABC transporters

The transporter mechanism in ABC transporters seems to be lead by an alternating access mechanism. Despite that this mechanisms is still a theoretical model, it seems to be accepted as the prevailing process in ABC transporters (Orelle & Jault 2016). The principal property of this model is the access of a substrate-binding site either the extracellular or the intracellular side of the membrane (Ward *et al.* 2007), (Dawson *et al.* 2007) and (Rees *et al.* 2009). The ABC transporter alternates between two conformations exposing the drug-binding site to the two sides of the membrane: the inward-facing (**Figure 7A**) and outward-facing conformations (**Figure 7B**).

The inward-facing conformation or open conformation is the protein in its apo-state. The NBDs are not associated and the drug-binding site is open toward the inside and closed toward the outside, **Figure 7A**.

The outward-facing conformation or closed conformation is the protein in its nucleotide-bound state. The NBDs are associated with the nucleotide-binding site positioned at the dimer interface (Dawson & Locher 2006; Smith *et al.* 2002). The binding of ATP induces a conformational change to the OF state, **Figure 7B**.

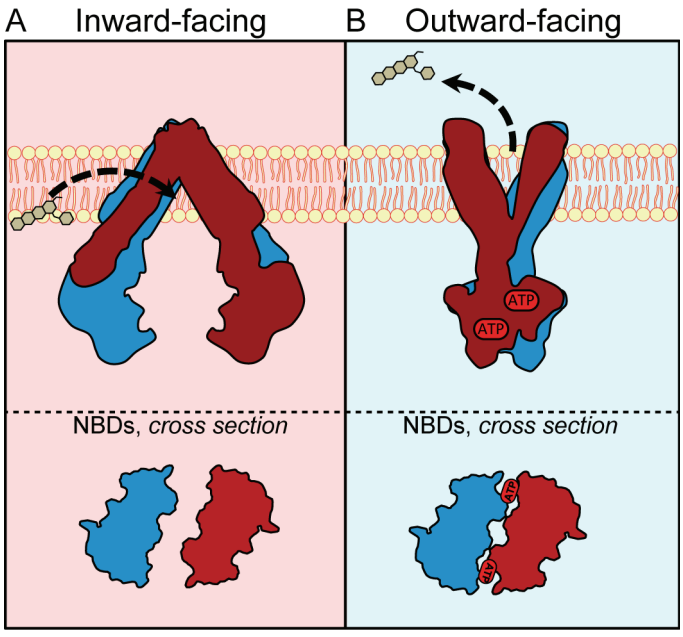


Figure 7. Two conformational states of an ABC exporter regarding the alternating access model. The inward-facing conformation in which the drug-binding site is open towards the inside, allowing the drug binding (**A**). The outward-facing conformation which the drug-binding site is open toward the outside allowing drug releasing (**B**). Monomers are represented in blue and red.

However, the alternating access model is still controversial because it is an extrapolation of only two conformations, which physiological relevance is still questioned. In the following paragraph, the ATPase mechanism and the transporter mechanism will be detailed.

1. ATPase activity

The mechanics of the motor engine of the protein which hydrolyses ATP is described in the literature regarding by two strictly different models: the ATP switch model and the constant contact model.

1.1. The ATP switch model

The ATP switch model, also called processive – enzyme's ability to catalyse consecutive reactions without releasing its substrate – clamp model, consists mainly of a conformation switch resulting in an association of the NBDs. The binding of two ATP molecules in the binding pocket leads to the association of the NBDs, thus facilitating ATP hydrolysis. As reviewed by Higgins and Linton, this model is composed of four steps (Higgins & Linton 2004):

Step I. The drug-substrate is bound in the TMDs and facilitates ATP-dependent closed dimer formation.

Step II. The energy-substrates - ATP molecules - are bound and induce the outward-facing dimer formation.

Step III. ATP is hydrolysed sequentially in ADP and Pi.

Step IV. The ADP in the NBDs interface cannot maintain the ABC transporter in its outward-facing conformation.

Step V. Release ADP.

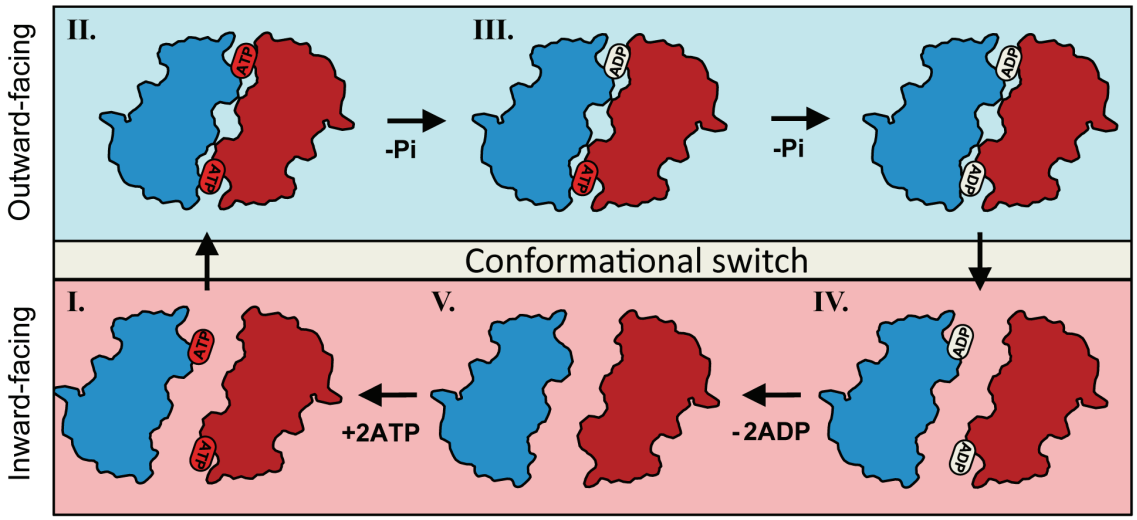


Figure 8. ATP switch model (Processive clamp model), caption and picture adapted from (Parcej & Tampé 2010). Each NBD is schematically drawn, they are getting closer to each other upon ATP-induced formation of a closed dimer. The engagement and disengagement of this closed dimer promotes a conformational switch within the TMD corresponding to the outward and inward states. As described in the previous page, **Step I.** The drug-substrate is bound in the TMDs and facilitates ATP-dependent closed dimer formation. **Step II.** The energy-substrates - ATP molecules - are bound and induce the outward-facing dimer formation. **Step III.** ATP is hydrolysed sequentially in ADP and Pi. **Step IV.** The ADP in the NBDs interface cannot maintain the ABC transporter in its outward-facing conformation. **Step V.** Release ADP.

During the cycle, it is not clear in which part the energy from ATP hydrolysis is consumed, and in which conformational changes it is invested. It has been put forward that dimerization of the NBDs is solely induced by the binding of ATP. After ATP hydrolysis into ADP, the NBDs are no longer maintained together, and energy might be invested into the transition into the IF state. Indeed, this possibility has been investigated by Luminescence Resonance Energy Transfer (LRET). A LRET donor probe was bound to one NBD while a LRET acceptor probe was bound to the other. The experiment showed that, after the binding of the two molecules of ATP, the dimerization of the NBDs induces the outward-facing state (closed state) of the protein (Zoghbi & Altenberg 2014). The result was also confirmed with several ABC transporters (BmrA, LmrA, P-gp, MsbA) where the binding of ATP promoted large conformational changes. Moreover, the LRET method shows that the ATP hydrolysis and the release of Pi and ADP generates the second conformation switch, where the protein returns to its original state: the open or inward-facing state (Zoghbi & Altenberg 2013).

1.2. The constant contact model or alternating site model

The constant contact model appeared with the first work on the occluded nucleotide conformation of the P-glycoprotein using vanadate (which replaces the γ-phosphate of ATP after hydrolysis). The first hypothesis made by Alan Senior and co-workers suggest an alternation of the two catalytic sites during the ATP hydrolysis: one site enters the transition state conformation but the other cannot do so, as a result, it implies that the sites may undergo ATP hydrolysis alternately (Senior *et al.* 1995), (Tombline & Senior 2005). In addition, Suresh Ambudkar and co-workers showed a similar effect using ATP-γ-S, a non hydrolyzable analogue of ATP. They also suggested an asymmetrical occlusion of ATP using ATP-γ-^{35}S. This occluded conformation tends to favor the alternate catalysis for the catalytic cycle (Sauna *et al.* 2007). The model is known as "*alternating site model*".

In 2009, drawing on the alternating site hypothesis, Jones and Georges came up with the hypothesis of the constant contact model (Jones & George 2009). The latter assumes that each catalytic site contains two different substrates (ADP and ATP) and as a result, NBDs would be asymmetric: one side closed, and one open. Like the alternating site model, the NBDs remain in contact throughout the cycle, with opening and closing of the sites occurring by way of intra-subunit conformational changes within in the NBD monomers (George & Jones 2012).

The constant contact model can be described in six steps:

Step I. Binding of ATP and its hydrolysis in ADP and Pi, while the ABC transporter is in outward-facing state.

Step II. Releasing the Pi with ADP is still bound. A second molecule of ATP binds to the empty site.

Step III. ATP binding to the empty site promotes the opening of the other site, inducing the release of the ADP.

Step IV. The remaining ATP molecule is hydrolysed into ADP and Pi.

Step V. As in step II, while releasing the Pi, ADP remains bound. A second molecule of ATP binds to the empty site.

Step VI. Release of the ADP leading to an empty site and an ATP occupied site, which completes the cycle.

This model was designed regarding the asymmetry observed in structural (Zaitseva *et al.* 2006) or biochemical studies (Mittal *et al.* 2012), (Sauna *et al.* 2007), (Siarheyeva *et al.* 2010) which tend to be in accordance with this model. Indeed, occlusion of one nucleotide during the transition state (in presence of vanadate) has been observed for several ABC transporter (p-gp (Tombline & Senior 2005); BmrA (Orelle *et al.* 2003); LmrA (van Veen *et al.* 2000); Maltose transporter (Sharma & Davidson 2000) suggesting an asymmetry of the NBDs (Orelle & Jault 2016). Furthermore, it has been also shown, regarding P-gp, that the ATPase activity is highly promoted when the N-terminal and the central regions of the NBDs are covalently linked together. These results show that some ABC transporters, at least the P-gp, can be fully active when the two NBDs are forced to remain in constant contact. However, in this mechanism exactly how the events in the NBD correspond to changes in the TMD conformation is not well defined. Furthermore, the stoichiometry of solute translocation and ATP hydrolysis remains an open question (Parcej & Tampé 2010).

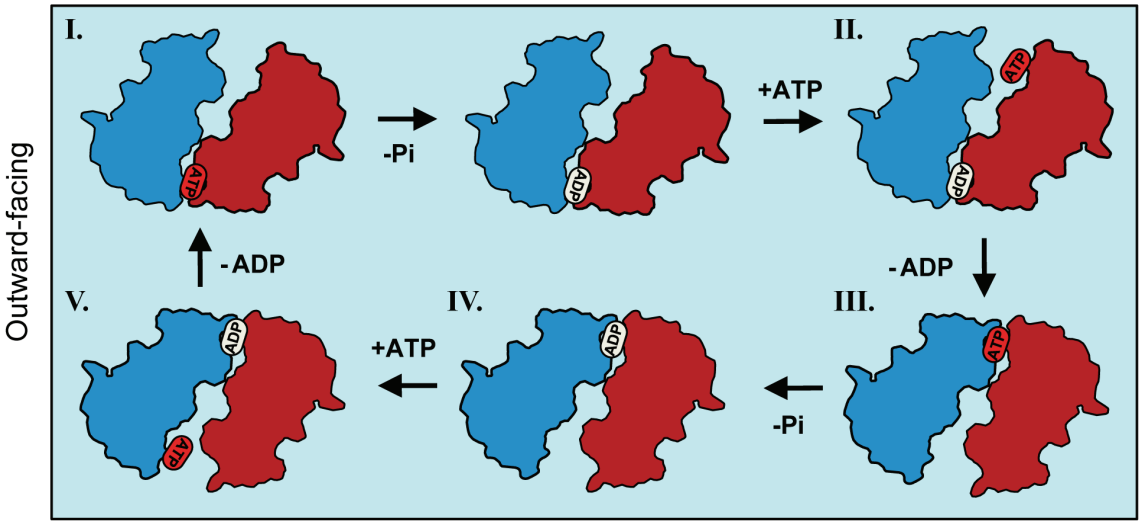


Figure 9. Constant contact model or alternating site model, caption and picture adapted from (Parcej & Tampé 2010). Each NBD is schematically drawn, a part of each NBD rotate to each other upon ATP-induced formation of a closed dimer. NBDs are engaged within the TMD corresponding to the outward states during all the efflux-cycle. As described in the previous page, **Step I.** Binding of ATP and its hydrolysis in ADP and Pi, while the ABC transporter is in outward-facing state. **Step II.** Releasing the Pi with ADP is still bound. A second molecule of ATP binds to the empty site. **Step III.** ATP binding to the empty site promotes the opening of the other site, inducing the release of the ADP. **Step IV.** The remaining ATP molecule is hydrolysed into ADP and Pi. **Step V.** As in step II, while releasing the Pi, ADP remains bound. A second molecule of ATP binds to the empty site. Release of the ADP leading to an empty site and an ATP occupied site, which completes the cycle.

1.3. Relevance of existing models

The two models differ in one major point: the switch model proposes that each NBD loses contact with each other after the sequential ATP hydrolysis, and the dimer operates a large amplitude conformational change into an inward-facing form. Alternatively, the constant contact model sets forth that each NBD remains in contact with each other, with one ATP-bound site closed and occluded, while the other site is open and allows ADP and Pi to be released. In the switch model, one of the major question concerns the physiological relevance of the apo state (inward-facing state): in others words, does the apo state really exist in the cell? Indeed, as pointed by Jones and Georges, the ATP concentration is ten times larger than the apparent ATP affinity of a typical ABC exporter. Consequently, an inward-facing state seems to exist only for a short time. A second question related to the mechanism concerns the futile cycle of the ATP hydrolysis. As also underlined by Jones and Georges (George & Jones 2012), the question is: "What is to prevent the doubly nucleotide-bound, inward-facing, closed NBD dimer from hydrolyzing ATP in a futile cycle?" For example, the ATP hydrolysis rate of BmrA is high – up to 1.5 $\mu mol.min^{-1}.mg^{-1}$ – and seems to be not drug-binding dependent.

This question about the relevance of the inward-facing state or of the futile cycle has been redundant in several discussions in ABC transporters articles. However, the details about the heterogeneity of the environment and the fact that ABC transporters are immobilized enzymes have not been discussed.

An experiment will illustrate the purpose. This experiment consisted to measure the ATPase activity of BmrA in lipid in the presence of 10 mM ATP:Mg^{2+} without agitation / vortexing. The activity of the protein was drastically low (around 1 000 times lower). Despite the presence of 10 mM ATP:Mg^{2+} incubated at the right temperature, non-homogenization of the sample lead to the sedimentation of the protein at the bottom of the tube. After homogenization of the sample, the activity came back to a canonical value. This anecdote demonstrates the presence of different environments in the cell: a micro-environment which is the closest to the membrane and a macro-environment which is the cell cytoplasm.

ATP and Mg^{2+} need to be transported by diffusion from the cytoplasm to the protein, thus generating a concentration gradient. Indeed, due to the activity of the protein, the concentration of ATP and Mg^{2+} should be lower in the environment that is the closest to the

protein than that of the cytoplasm. This gradient is also observed for the reaction products. The concentration of reaction products should be higher in the closest environment of the protein than that of the cytoplasm. Incidentally, the reaction products - ADP and Pi - inhibit the ATPase activity of the protein.

Moreover, this example did not take the macromolecular crowding inside the bacteria into account, which alters the properties of molecules in the presence of a high concentration of molecules like proteins. The cytosol of *Escherichia coli* contain about 200-400 $mg.mL^{-1}$ of protein (Zimmerman & Trach 1991) and (Milo 2013), this high protein concentration reduce the volume of solvent available and alter drastically the diffusion coefficient of any substrates (Ellis 2001).

2. Transport activity

Most hypotheses on the transport activity were made regarding the ATP switch model. In this model, the substrate-binding site is accessible to the substrate in the inward-facing conformation. Then, after ATP binding, the transporter switches to the opposite conformation (outward-facing conformation) releasing the substrate out of the membrane. Then ATP hydrolysis resets the transporter to the inward-facing conformation ensures the accessibility to the substrate-binding site. It is important to note the detailed mechanism is still unknown and explanations are only hypotheses based on crystallographic structures and biochemical assays. However, the crystallographic structures of ABC exporters can be used as a snapshot of the drug-transport cycle. Four different states have been probed:

i) The inward-facing conformation in which the NBDs of the transporter are separated. In the inward-facing conformation, the cavity between the TMDs allows the substrate binding on the inner or outer membrane leaflet, **Figure 10A**.

ii) The inward-facing conformation in which the NBDs are close together in contrast with the previous state. The cavity between the TMDs does not seem to be accessible to either side of the membrane, **Figure 10B**.

iii) The occluded-state conformation in which the NBDs of the transporter are engaged in close contact without any real conformational change on the TMDs moiety. The cavity between the TMDs is not accessible to either side of the

membrane. Nevertheless, it is worth noting that these occluded conformation structures were obtained without any transported substrate (Choudhury *et al.* 2014) and (Lin *et al.* 2015), **Figure 10C**.

iv) The outward-facing state is the conformation in which the NBDs of the transporter are in close contact bound together with ATP molecules. In the outward-facing conformation, the cavity between the TMDs is now open to the opposite side of the membrane, **Figure 10D**.

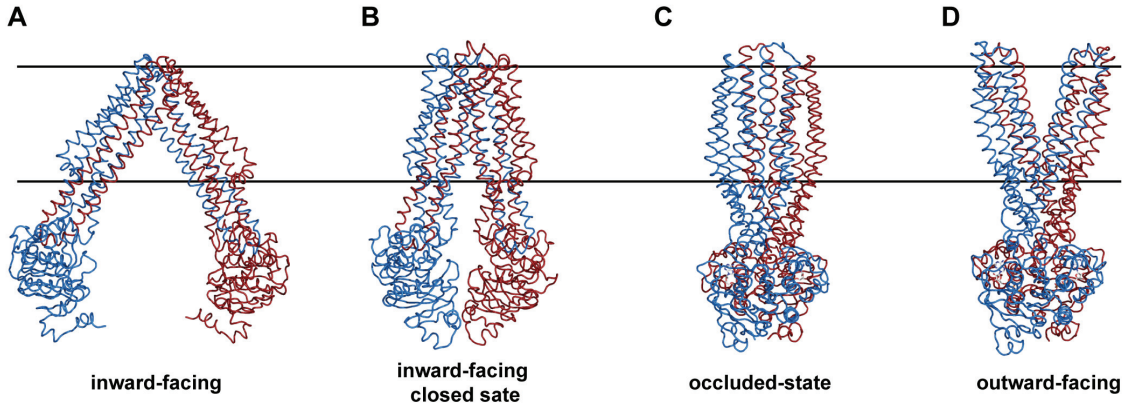


Figure 10. Different conformational snapshot of an ABC transporter, obtain by X-ray crystallography. MsbA from *Eschericia coli* in its inward-facing state (3B5W) (**A**), MsbA from *Vibrio cholerae* inward-facing closed state (3B5X) (**B**), McjD-AMPPNP:Mg^{2+} complex from *Escherichia coli* in its occluded state (4PL0) (**C**), MsbA-ADP:Mg^{2+}:VO_4^{3-} complex from *Salmonella enterica* in its outward-facing sate (3B5Z) (**D**). **A**, **B**, **D** are from (Ward *et al.* 2007), and **C** from (Choudhury *et al.* 2014).

These different snapshots provide a first picture of the transport mechanism. However, the reconstruction of the full mechanical details is further hindered by missing elements. The difference between and the similarity of each state can furnish structural information about each step of the mechanism. Geoffrey Chang and co-workers solved four different structures of the ABC transporter flippase MsbA from three bacterial orthologs of different conformations: MsbA nucleotide bound (using AMPPNP and ADP:Vanadate as analogs of ATP) (**Figures 10D and 11C**), MsbA inward-facing state (Open Apo) (**Figures 10A and 11A**) and MsbA inward-facing with NBDs close together (so-called closed Apo) (**Figures 10B and 11B**), (Ward *et al.* 2007).

The structural alignment of NBD with an R.M.S.D of 1.2 Å can be used to highlight the conformation change of TMH 1,2,3 and 6 between each state, **Figure 11**.

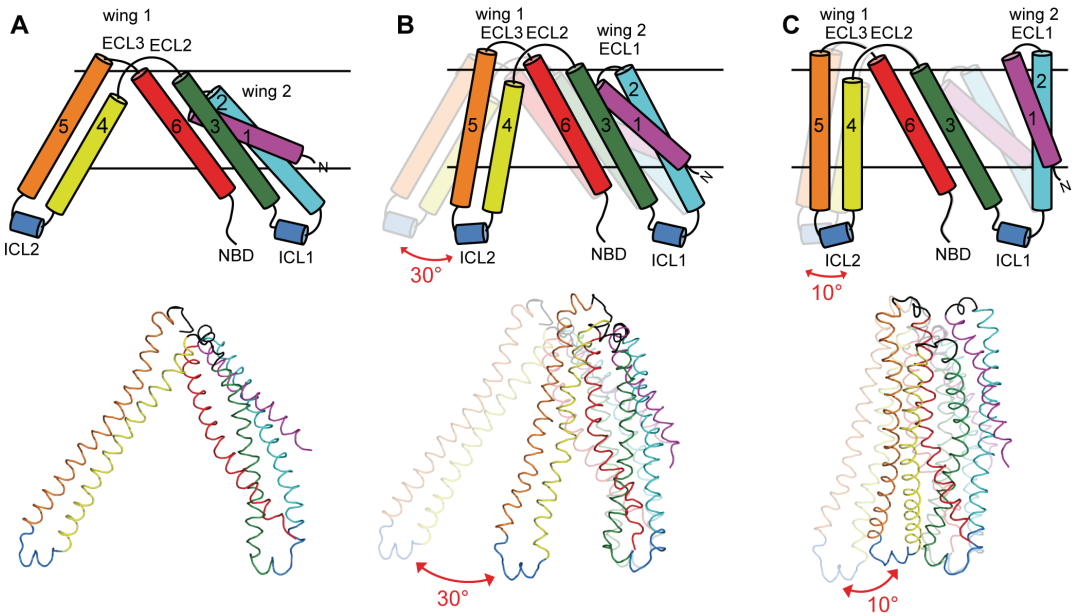


Figure 11. TMDs diagrams and structures of the transporter ABC MsbA in its different conformation. The different conformations represented are the inward-facing state (**A**), inward-facing closed state (**B**) and the outward-facing state (**C**). On the diagram, the previous state is represented in background with low opacity.

Sallie Gardner at a Gallop so-called *The Horse in Motion* results of a photographic experiment by Eadweard Muybridge. It is a series of snapshot consisting of a galloping horse, the motion is decomposed by 12 pictures. Geoffrey Chang and co-workers reconstituted the motion of an ABC transporter in the same manner as *The Horse in Motion* using 3 snapshots (Ward *et al.* 2007).

First snapshot: the ABC transporter is in its inward-facing state (open and apo), **Figure 11A**

Second snapshot: the ABC transporter switch to the inward-facing closed state (closed and apo). The conformational changes from inward-facing state (open apo) to the inward-facing closed state (closed apo) appear to be lead by a rotation of TMH4, TMH5, and ICL2. Both helices and the coupling loop can operate a rotation of 30° about a hinge formed by the ECL2 and 3, **Figures 11B and 12A**. The transporter is not in outward-facing sate, NBDs are misaligned, they are not following the canonical arrangement: Walker A motif of one NBD should be opposite the ABC signature motif of the other NBD. In this case, Walker A motifs face each other.

Third snapshot: the ABC transporter switch to the outward-facing sate (closed and nucleotide bound) and restore the NBDs alignment to the canonical arrangement, the transition goes through two unique movements occurring simultaneously. One is another

TMH4&5 rotation of 10° about the hinge (ECL2 and 3), **Figures 11C**, the second is a 20° transverse rotation of TMH4&5, **Figure 12B**.

The pair TMH4&5 is connected to TMH3&6 through the ECL2 and 3. As a result, the NBDs dimerization pulls TMH3&6 away from TMH1&2, **Figures 11C**. The newly formed outward opening is created between TMH1 and TMH3, whereas the inward opening is formed between TMH4 and TMH6. Thus, changes in the orientation and spacing of the NBDs dramatically rearrange the packing of transmembrane helices and effectively switch access of the internal chamber from the inner to the outer leaflet of the bilayer.

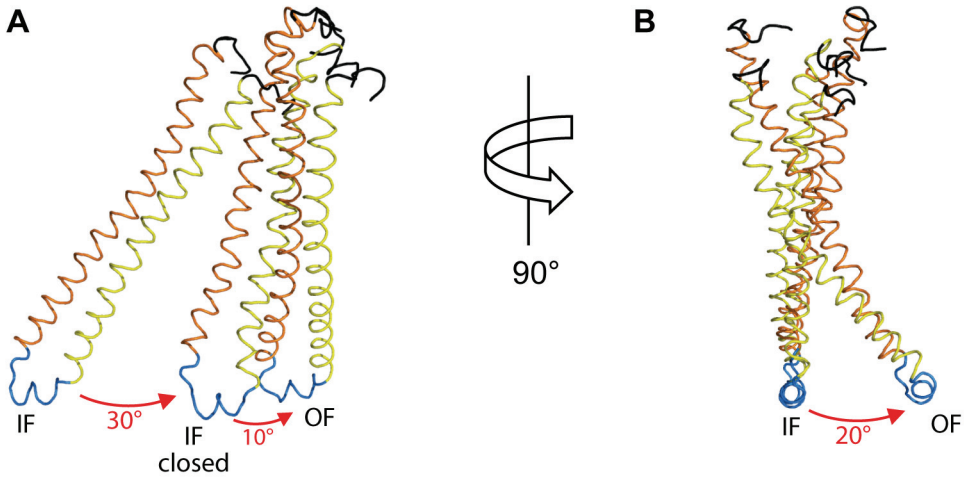


Figure 12. Structures of the TMH4 and 5 in different conformations. The TMH4 (yellow) and TMH5 (orange) rotate about a hinge formed by the ECL2 and 3 (in black) of 30° between the inward-facing state (IF) and the inward-facing closed state (IF closed) then 10° and operate a translation of 20° between the inward-facing closed state (IF closed) and the outward-facing state (OF).

In this model, the occluded-state conformation is not represented. Besides, a variation of this mechanism has been proposed by Locher and co-workers for the lipid-linked oligosaccharide flippase PglK, which is distinct from the canonical alternating-access (Perez *et al.* 2015). PglK like MsbA is an ABC transporter which catalyses a flipping reaction of a lipid-linked oligosaccharide from the inner leaflet to the outside of the lipid-bilayer. The lipid-linked-oligosaccharide is composed of a polyprenyl tail (lipid part) followed by a pyrophosphate group and an oligosaccharide (*N*-acetylgalactosamine).

For this mechanism, it is proposed that the polyprenyl tail of a lipid-linked-oligosaccharide interacts with the surface of ABC transporter flippase in its outward-facing state. The pyrophosphate moiety is directly transferred into the outward-facing cavity (nucleotide

state), where it can form electrostatic interactions with exposed arginine residues. The energy released from the ATP hydrolysis seems to be used in order to break the electrostatic interactions between the substrate and the exposed arginine residues.

In other words, in contrast to MsbA, the substrate is bound with the ABC transporter when the latter is in its outward-facing state. The outward-facing state seems to allow the translocation of the lipid-linked-oligosaccharide from the inner leaflet to the outside of the bilayer. The ATP hydrolysis resets the conformation to the inward-facing state in which NBDs are pushed apart, closing the door of the binding-cavity behind the substrate.

All these mechanistic hypotheses are centered upon the opening or the closing state of the binding-site cavity. However, the location of the binding site has not yet been precisely identified.

Structural analysis of ABC transporters

Membrane proteins are challenging targets for structural biology. Despite the fact that they represent around 20 to 30 % of the total proteomes from several eubacterial, archaean or eukaryotic organisms (Wallin & Heijne 1998), membrane proteins are structurally drastically under-represented in the Protein Data Bank (PDB). To date, membrane proteins represent less than 2 % of the total entries of the PDB. In August 2017, there were 2250 membrane proteins structures solved in the PDB, representing in total 707 different membrane proteins, according to the referencing structure of membrane proteins website: http://blanco.biomol.uci.edu/mpstruc/. ABC transporters as membrane proteins do not escape these difficulties. The different approaches used in structural studies of membrane proteins will be described further down.

1. Crystallography

Currently, the majority of membrane protein structures were solved by X-ray crystallography. Nevertheless, the crystallization stage of membrane proteins is still a major challenge. The classical approach of crystallization is the hanging-drop method, when a drop composed of a mixture of protein sample and reagent is placed in vapor equilibration with a liquid reservoir of reagent. Nonetheless, this technique requires a large number of tests, different reagents, temperatures, buffers and so on, in order to obtain well-ordered 3D crystals. In addition to the stochastic requirements to obtain 3D crystals, the detergent could negatively influence the prospects for crystallisation due to the disorder induced by these detergent molecules bound to the protein. The shape and size of the detergent belt can influence the quality of the crystalline stack (Chaptal *et al.* 2017). The structural information brought by crystallography plays a major role in the investigations of the ABC transporter mechanism as well as in the drug docking. Indeed, the different conformations revealed by crystallography from several ABC transporters make it possible to determine the general scheme previously described. The latter concerns essentially the transport mechanism of ABC transporters with the canonical architecture and the two alternating states: the outward-facing and the inward-facing states. However, one of the main drawbacks of the X-ray crystallography is the physiological relevance of the conformational states. Indeed, the

use of detergents decreases drastically the surface tension induced by the lipid membrane (Henriksen *et al.* 2010). Consequently, the detergents used do not represent the properties of the lipid membrane. What is more, it may lead to problems of instability for folded proteins (Matthews *et al.* 2006) or a loss of biological activity, as observed for BmrA in the presence of DDM (Matar-Merheb *et al.* 2011). These main disadvantages result in questioning the functional relevance of the inward-facing structures obtained by X-ray crystallography. The long distance between both NBDs (up to 40 Å) could be an artefact due to the use of detergent.

2. Cryo-electron microscopy

The recent progress in cryo-electron microscopy (cryo-EM) allows to obtain 3D structure with a resolution comparable to crystallography. Indeed, the introduction of direct electron detectors which can read electrons at high frame rate and the performance of image processing methods, cryo-EM becomes a tool to be able to analyze the structure of dynamical biomolecules (Kühlbrandt 2014) and (Murata & Wolf 2017). As example, in June 2015 was published a 2.2 Å map of the β-Galactosidase in complex with β-D-thiogalactopyranoside (Bartesaghi *et al.* 2015). Cryo-EM requires only small amounts of material that cannot be isolated in large quantities or high concentration. This last property makes cryo-EM suitable for membrane protein structural studies.

The starting signal was given by the subnanometre-resolution cryo-EM structure of the heterodimeric ABC exporter TmrAB in detergent (8.2Å), in November 2014 (Kim *et al.* 2015), **Figure 13A**. In one year, the structures of the ABC transporters ABCA1 (4.1 Å), CFTR (3.37 Å), TAP (3.97 Å), MRP1 (3.49 Å) and ABCG2 (3.78 Å) were obtained respectively **Figures 13B, C, D, E and F** (Qian *et al.* 2017), (Zhang *et al.* 2017), (Oldham *et al.* 2016), (Johnson & Chen 2017) and (Taylor *et al.* 2017b) .

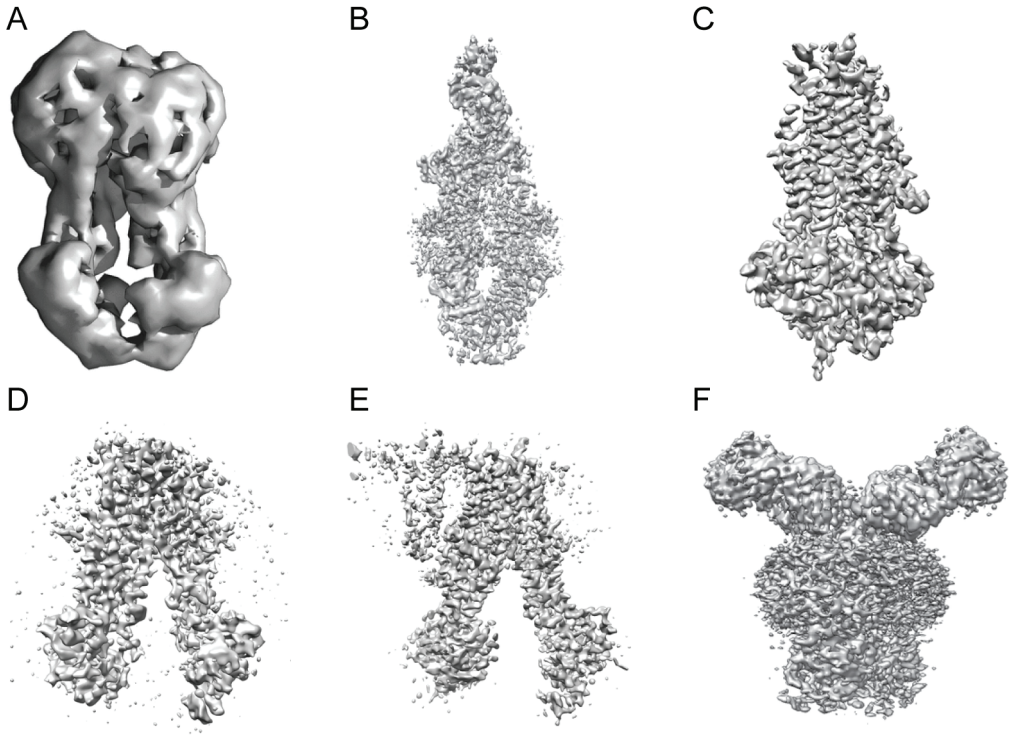


Figure 13. The last cryo-electron microscopy structure of ABC transporters.
TmrAB -8.2Å (**A**) ABCA1 - 4.1 Å (**B**) , CFTR - 3.37 Å (**C**), TAP - 3.97 Å (**D**), MRP1 - 3.49 Å (**E**) and ABCG2 - 3.78 Å (**F**)

3. Electron paramagnetic resonance

Electron paramagnetic resonance (EPR) was mainly used in proteins to determine the distance between two spin-labels or to characterize conformational motions. Following the example of fluorescence labelling, the introduction of small paramagnetic species as spin probes can be performed covalently on a protein. The two main probes used in EPR are maleimide and methanesulfonate derivative spin-labels which are highly specific for sulfhydryl group of cysteine (Berliner *et al.* 1982). Spin-label probes can be introduced almost everywhere in a protein by site-directed mutagenesis, this experimental method is used to probe the different conformational snapshots of ABC transporters in detergent or in lipids as well. EPR on MsbA (Zou *et al.* 2009), (Mittal *et al.* 2012), P-gp (Wen *et al.* 2013), LmrA (Hellmich *et al.* 2012b), BmrCD (Mishra *et al.* 2014) and TM287/288 (Timachi *et al.* 2017) confirmed the large conformational change of transporters in detergent and in proteoliposomes. Moreover, the method highlighted the flexible behavior of the protein in absence of substrate. The technique indicates a wide distance-distribution between both

NBDs, which suggests a large amplitude motion in absence of any substrate. These data are incompatible with the different xray crystallographic structures of the inward-facing state, which shows a large distance between both NBDs for all structures comparable to a *tepee-like* form. Indeed, the crystallographic structure takes a snapshot where the distance between the two NBDs of the protein is the longest (Wen *et al.* 2013).

4. Nuclear magnetic resonance (NMR)

ABC transporters are still challenging targets for NMR spectroscopy, mainly due to the size of the protein. So far, the most studied transporter by NMR has been the transporter LmrA by the group of Professor Clemens Glaubitz since the first work on LmrA in 2004.

4.1. LmrA

LmrA is an ABC multidrug exporter from *Lactococcus lactis* and mediates drug and antibiotic resistance (van Veen *et al.* 2000). Like BmrA, LmrA is a homodimer of 64.5 kDa composed of 12 transmembrane segments and two NBDs. The protein LmrA was the first ABC transporter studied by solid-state NMR. The protein was produced using selective labelling on glycine and reconstituted, initially regarding a ratio lipid-to-protein of 200 (mol:mol) (James Mason *et al.* 2004). The empty liposomes were separated using a sucrose gradient. Using a similar preparation, the dynamics of the protein were probed by deuterium solid-state NMR using L-[β-$^{2}H_{3}$]-alanine (Siarheyeva *et al.* 2007). The authors highlighted the high mobility of the LmrA's NBDs whereas the mobility became restricted when the protein was in complex with a non-hydrolyzable ATP analogue. In 2008, Hellmich and co-workers followed the ATP hydrolysis by ^{31}P real-time solid-state NMR on an LmrA mutant displaying a slow ATP turn over rate (Hellmich *et al.* 2008). A second ^{31}P real-time solid-state NMR on LmrA was conducted and surprisingly, revealed that LmrA couple ATPase with adenylate kinase activity (Kaur *et al.* 2016). Finally, using the combination of liquid-state NMR on NBDs (Hellmich *et al.* 2012a) and solid-state NMR on full length LmrA and selective labelling, Hellmich and co-workers confirmed the large amplitude motions observed for the ABC exporters in lipids (Hellmich *et al.* 2015).

4.2. MsbA

MsbA is an ABC exporter which acts as a lipid A flippase. It is a homodimer of 64.5 kDa following the canonical ABC transporters organization which is composed of two NBDs and

two TMDs crossing the membrane by 12 transmembranes segments. MsbA is studied by solid-state NMR in the group of Professor Clemens Glaubitz at Goethe-Universität in Frankfurt. The first solid-state NMR spectra of MsbA were published in 2015 (Kaur *et al.* 2015). The protein was reconstituted in DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) and DMPA (1,2-dimyristoyl-sn-glycero-3-phosphate) (9:1) with a ratio lipid-to-protein of 1 (m:m). In 2017, the same group demonstrated by ^{31}P solid-state NMR that the protein was able to couple ATPase and adenylate kinase activity (Kaur *et al.* 2016).

4.3. ArtMP-J

ArtMP-J is an amino acid ABC type 1 importer from the thermophilic bacteria *Geobacillus stearothermophilus*. This importer was first described in 2005 as one of the smallest ABC transporters (Fleischer *et al.* 2005). Moreover, the importer is constituted by four polypeptides (see section *Anatomy of an ABC transporter*) and is organized as a homodimer composed by homodimeric transmembrane core of two ArtM subunits that are non-covalently bound to two NBDs called ArtP. The importer interacts with an extracellular substrate-binding protein called ArtJ. The role of ArtJ is to capture substrate from the environment, here an amino acid, and deliver it to the importer which imports the substrate to the bacteria cytoplasm (Davidson & Chen 2004). Its small size, its capacity to study each subunit independently and its thermostability make the transporter ArtMP an ideal system to study ABC transporter by solid-state NMR (Voreck 2015). The different states of the ABC importer were characterized by the solid-state NMR group of Professor Hartmut Oschkinat. The NBDs ArtP and the complex NBDs-TMD – ArtMP – were studied separately in presence of different type of nucleotide or homologues. Authors of this study noticed an asymmetric of the dimer in presence of non-hydrolyzable ATP homologue. Moreover, regarding the chemical shift perturbation (CSP), the authors are able to distinguish between the binding of a double and a triply phosphorylated ligand on the NBDs (Lange *et al.* 2010).

The multidrug resistance ABC transporter BmrA

In 2002, the group of Jean-Michel Jault initiated a sequence research on the genome of *Bacillus subtilis* in order to identify a multidrug resistance ABC transporter. They found an ABC transporter which showed a high sequence homology with LmrA. This ABC transporter called YvcC was shown to share 41.5% of identity and also 28.8% and 27.7% of identity with the N- and C-terminus halves of the P-gp (Steinfels *et al.* 2002). Using quenching of YvcC intrinsic fluorescence, it was shown that this transporter was able to transport different drugs (Ethidium bromide 10.6 µM, Doxorubicin 22.1 µM, Daunorubicin 12.2 µM, Rhodamine 6G 22.4 µM, Vinblastine 5 µM, Hoechst 33342 9.5 µM, (Steinfels *et al.* 2004) and surprisingly, large substrates, 84-100 nm, can be equally extruded (Browning *et al.* 2016). In homology to LmrA, the transporter ABC YvcC was subsequently renamed to BmrA for ***B****acillus subtilis* **m**ultidrug **r**esistance **A**BC transporter. BmrA is a homodimer of 2 x 65 kDa, the monomer being constituted of one transmembrane domain (TMD) containing 6 predicted transmembrane helices, and one nucleotide-binding domain.

The protein can be overexpressed in the bacterial strain *E. coli* C41(DE3). Interestingly, the quantification of BmrA was estimated to about 50 % of mass of the whole membrane proteins (Steinfels *et al.* 2002). BmrA is one of the ABC transporters which gives the high yield of recovery in the membrane fraction. It was also shown by Steinfels and co-workers that BmrA displays a high ATPase-activity in proteoliposomes (with an *E. coli* lipids-to protein ratio of 20) measured at 6.5 µmol.min^{-1}.mg^{-1} (Orelle *et al.* 2003) and (Orelle *et al.* 2008). However, the protein in n-Dodecyl β-D-maltoside (DDM) 0.05% shows an activity 5.5 times less than in lipids, 1.2 µmol.min^{-1}.mg^{-1} (Ravaud *et al.* 2006).

Previous studies have shown that BmrA reconstituted in lipids at a lipid-to-protein ratio of 1.2, in the presence of a residual amount of detergent, forms a very peculiar structure, a double ring containing 2 times 12 dimers of BmrA (Chami *et al.* 2002) and (Fribourg *et al.* 2014). This has made it possible to obtain a low-resolution structure of BmrA (~ 25Å) (Orelle *et al.* 2008). Interestingly, the addition of ATP:Mg^{2+} prevents the formation of these rings, or destroys the rings previously formed, showing that BmrA keeps a native-like conformation inside the rings, sensitive to the binding of ATP (Orelle *et al.* 2008). Further removals of the residual detergent led to the formation of tubular structures, presumably reflecting the stacking of many rings together (Chami *et al.* 2002) and (Fribourg *et al.* 2014).

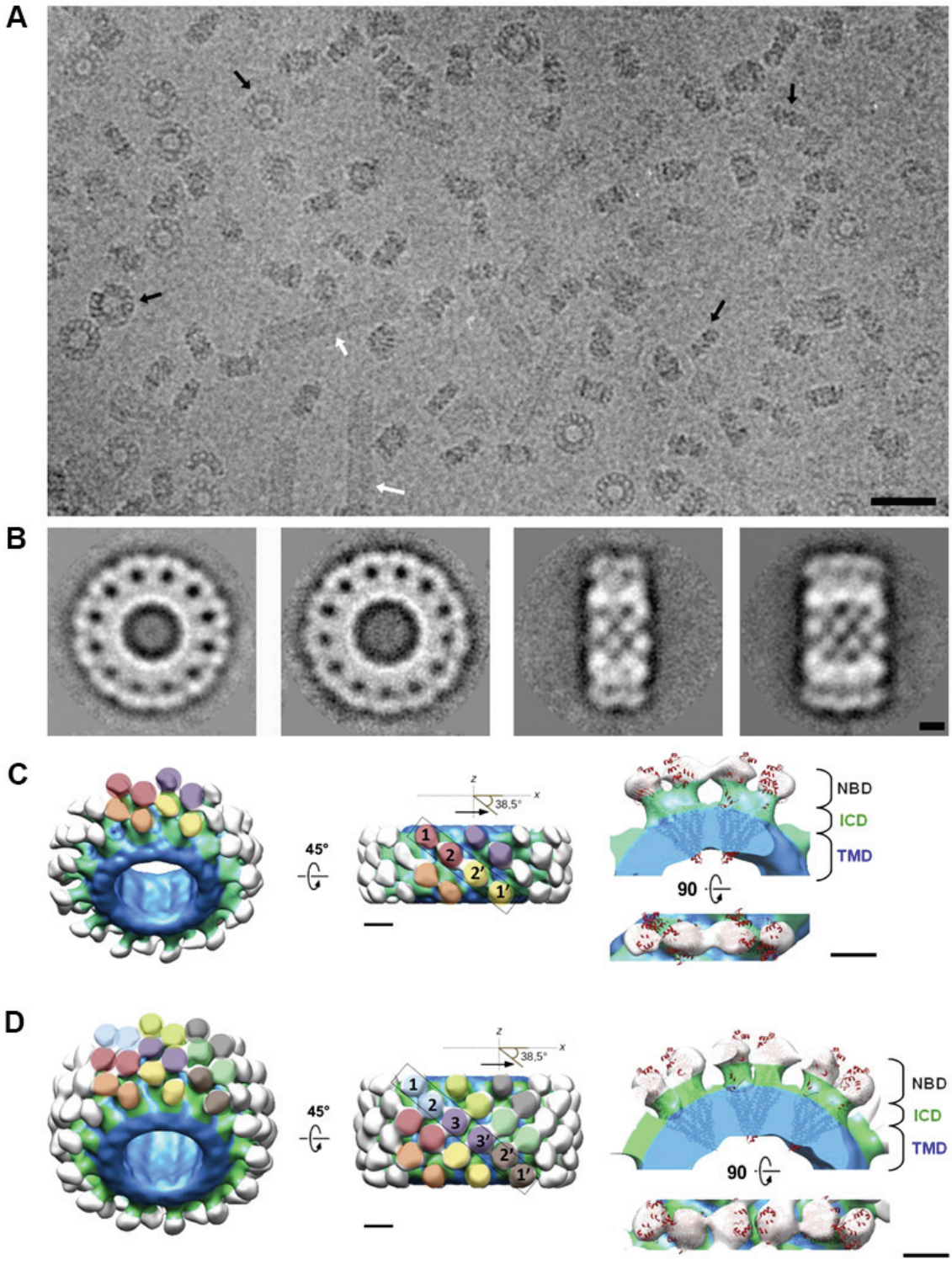


Figure 14. Cryo-electron microscopy structure of a lipids-ring containing BmrA from (Fribourg *et al.* 2014). (A) Representative cryo-electron micrograph with circular or rectangular views of ring-shaped particles (black arrows). Other reconstituted structures such as tubes or membrane sheets were also present (white arrows). The micrograph has been 4-fold down-sampled to increase the contrast. The bar represents 25 nm. (**B**) Average images of the four mains classes of ring-shaped particles. The bar represents 5 nm. (**C**) 3D models of rings of BmrA in lipid membrane calculated with particles of *d*12 symmetry or

(**D**) with particles of *d*13 symmetry (c and d). Homodimers of BmrA are lettered as (1, 2), (1 , 2) and (3, 3). Subdomains of BmrA, NBD (white), ICD (green) and membrane domain (blue). The mouse open apo-P-gp (in red; PDB code: 3G5U) was docked in densities of the homodimers of BmrA. Note that the TMD of P-gp is longer than BmrA because there is an insertion of 33 residues in the first extracellular loop of the P-gp between the transmembrane helices 1 and 2 (protruding toward the center of the ring). The bar represents 5 nm. Caption was adapted from (Fribourg *et al.* 2014)**.**

Finally, the first complete solid-state NMR 2D ^{13}C-^{13}C spectrum of an ABC transporter was done using BmrA (Kunert *et al.* 2014). In this case, the transporter was reconstituted in the lipids of its natural organism at a LPR 0.5 which is physiologically relevant regarding the composition of the *Bacillus subtilis* membrane: 63% proteins and 18,5% lipids; ratio lipids to protein of 0.3 (Bishop *et al.* 1967). This study open the way to further prospects for solid-state NMR of an ABC transporter in its close-to-natural lipid environment as you will see in this thesis work.

References

Ambudkar, S.V., Kim, I.-W., Xia, D. & Sauna, Z.E. (2006) The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker A motif in ABC transporters, is critical for ATP binding. *FEBS Letters* 580, 1049–1055.

Ames, G.F. & Lever, J. (1970) Components of histidine transport: histidine-binding proteins and hisP protein. *Proceedings of the National Academy of Sciences* 66, 1096–1103.

Ames, G.F., Mimura, C.S. & Shyamala, V. (1990) Bacterial periplasmic permeases belong to a family of transport proteins operating from Escherichia coli to human: Traffic ATPases. *FEMS microbiology reviews* 6, 429–446.

Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L.S. & Subramaniam, S. (2015) 2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. *Science* 348, 1147–1151.

Beek, ter, J., Guskov, A. & Slotboom, D.J. (2014) Structural diversity of ABC transporters. *The Journal of general physiology* 143, 419–435.

Berliner, L.J., Grunwald, J., Hankovszky, H.O. & Hideg, K. (1982) A novel reversible thiol-specific spin label: papain active site labelling and inhibition. *Analytical biochemistry* 119, 450–455.

Bishop, D.G., Rutberg, L. & Samuelsson, B. (1967) The chemical composition of the cytoplasmic membrane of Bacillus subtilis. *European Journal of Biochemistry* 2, 448–453.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. & Shao, Y. (1997) The complete genome sequence of Escherichia coli K-12. *Science* 277, 1453–1462.

Browning, L.M., Lee, K.J., Cherukuri, P.K., Nallathamby, P.D., Warren, S., Jault, J.-M. & Xu, X.-H.N. (2016) Single Nanoparticle Plasmonic Spectroscopy for Study of the Efflux Function of Multidrug ABC Membrane Transporters of Single Live Cells. *RSC advances* 6, 36794–36802.

Chami, M., Steinfels, E., Orelle, C., Jault, J.-M., Di Pietro, A., Rigaud, J.-L. & Marco, S. (2002) Three-dimensional structure by cryo-electron microscopy of YvcC, an homodimeric ATP-binding cassette transporter from Bacillus subtilis. *Journal of molecular biology* 315, 1075–1085.

Chaptal, V., Delolme, F., Kilburg, A., Magnard, S., Montigny, C., Picard, M., Prier, C., Monticelli, L., Bornert, O., Agez, M., Ravaud, S., Orelle, C., Wagner, R., Jawhari, A., Broutin, I., Pebay-Peyroula, E., Jault, J.-M., Kaback, H.R., Le Maire, M. & Falson, P. (2017) Quantification of Detergents Complexed with Membrane Proteins. *Scientific reports* 7, 41751.

Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H.W. & Beis, K. (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 111, 9145–9150.

Cole, S.P., Bhardwaj, G., Gerlach, J.H., Mackie, J.E., Grant, C.E., Almquist, K.C., Stewart, A.J., Kurz, E.U., Duncan, A.M. & Deeley, R.G. (1992) Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* 258, 1650–1654.

Dano, K. (1973) Active outward transport of daunomycin in resistant Ehrlich ascites tumor cells. *Biochimica et biophysica acta* 323, 466–483.

Davidson, A.L. & Chen, J. (2004) ATP-binding cassette transporters in bacteria. *Annual Review of Biochemistry* 73, 241–268.

Davidson, A.L. & Sharma, S. (1997) Mutation of a single MalK subunit severely impairs maltose transport activity in Escherichia coli. *Journal of Bacteriology* 179, 5458–5464.

Davidson, A.L., Dassa, E., Orelle, C. & Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and molecular biology reviews : MMBR* 72, 317–64– table of contents.

Dawson, R.J.P. & Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185.

Dawson, R.J.P., Hollenstein, K. & Locher, K.P. (2007) Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Molecular microbiology* 65, 250–257.

Dean, M., Hamon, Y. & Chimini, G. (2001) The human ATP-binding cassette (ABC) transporter superfamily. *Journal of lipid research* 42, 1007–1017.

Diederichs, K., Diez, J., Greller, G., Müller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W. & Welte, W. (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon Thermococcus litoralis. *The EMBO journal* 19, 5951–5961.

Ellis, R.J. (2001) Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences* 26, 597–604.

Fleischer, R., Wengner, A., Scheffel, F., Landmesser, H. & Schneider, E. (2005) Identification of a gene cluster encoding an arginine ATP-binding-cassette transporter in the genome of the thermophilic Gram-positive bacterium Geobacillus stearothermophilus strain DSMZ 13240. *Microbiology (Reading, England)* 151, 835–840.

Fribourg, P.-F., Chami, M., Sorzano, C.O.S., Gubellini, F., Marabini, R., Marco, S., Jault, J.-M. & Lévy, D. (2014) 3D cryo-electron reconstruction of BmrA, a bacterial multidrug ABC transporter in an inward-facing conformation and in a lipidic environment. *Journal of molecular biology* 426, 2059–2069.

Gaudet, R. & Wiley, D.C. (2001) Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. *The EMBO journal* 20, 4964–4972.

George, A.M. & Jones, P.M. (2012) Perspectives on the structure-function of ABC transporters: the Switch and Constant Contact models. *Progress in biophysics and molecular biology* 109, 95–107.

Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A. & Jault, J.M. (2001) A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends in Biochemical Sciences* 26, 539–544.

Gros, P., Croop, J. & Housman, D. (1986) Mammalian multidrug resistance gene: complete cDNA sequence indicates strong homology to bacterial transport proteins. *Cell* 47, 371–380.

Grossmann, N., Vakkasoglu, A.S., Hulpke, S., Abele, R., Gaudet, R. & Tampé, R. (2014) Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter. *Nature Communications* 5, 5419.

Hazelbauer, G.L. (1975) Maltose chemoreceptor of Escherichia coli. *Journal of Bacteriology* 122, 206–214.

Hellmich, U.A., Duchardt-Ferner, E. & Glaubitz, C. (2012a) Backbone NMR resonance assignments of the nucleotide binding domain of the ABC multidrug transporter LmrA from Lactococcus lactis in its ADP-bound state. *Biomolecular NMR Assignments*.

Hellmich, U.A., Haase, W., Velamakanni, S., van Veen, H.W. & Glaubitz, C. (2008) Caught in the act: ATP hydrolysis of an ABC-multidrug transporter followed by real-time magic angle spinning NMR. *FEBS Letters* 582, 3557–3562.

Hellmich, U.A., Lyubenova, S., Kaltenborn, E., Doshi, R., van Veen, H.W., Prisner, T.F. & Glaubitz, C. (2012b) Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *Journal of the American Chemical Society* 134, 5857–5862.

Hellmich, U.A., Mönkemeyer, L., Velamakanni, S., van Veen, H.W. & Glaubitz, C. (2015) Effects of nucleotide binding to LmrA: A combined MAS-NMR and solution NMR study. *Biochimica et biophysica acta* 1848, 3158–3165.

Henriksen, J.R., Andresen, T.L., Feldborg, L.N., Duelund, L. & Ipsen, J.H. (2010) Understanding detergent effects on lipid membranes: a model study of lysolipids. *Biophysical Journal* 98, 2199–2205.

Higgins, C.F. & Linton, K.J. (2004) The ATP switch model for ABC transporters. *Nature Structural & Molecular Biology* 11, 918–926.

Higgins, C.F., Hiles, I.D., Salmond, G.P., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D. & Bell, A.W. (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature* 323, 448–450.

Higgins, C.F., Hiles, I.D., Whalley, K. & Jamieson, D.J. (1985) Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems. *The EMBO journal* 4, 1033–1039.

Holland, I.B. & Blight, M.A. (1999) ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *Journal of molecular biology* 293, 381–399.

Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P. & Tainer, J.A. (2000) Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 101, 789–800.

Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F. & Kim, S.H. (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* 396, 703–707.

Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E. & Higgins, C.F. (1990) Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* 346, 362–365.

James Mason, A., Siarheyeva, A., Haase, W., Lorch, M., van Veen, H. & Glaubitz, C. (2004) Amino acid type selective isotope labelling of the multidrug ABC transporter LmrA for solid-state NMR studies. *FEBS Letters* 568, 117–121.

Johnson, Z.L. & Chen, J. (2017) Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1. *Cell* 168, 1075–1085.e9.

Jones, P.M. & George, A.M. (2002) Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. *Proceedings of the National Academy of Sciences* 99, 12639–12644.

Jones, P.M. & George, A.M. (2009) Opening of the ADP-bound active site in the ABC transporter ATPase dimer: evidence for a constant contact, alternating sites model for the catalytic cycle. *Proteins: Structure, Function, and Bioinformatics* 75, 387–396.

Jones, P.M. & George, A.M. (2012) Role of the D-loops in allosteric control of ATP hydrolysis in an ABC transporter. *The Journal of Physical Chemistry A* 116, 3004–3013.

Juliano, R.L. & Ling, V. (1976) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. 455, 152–162.

Kaur, H., Lakatos, A., Spadaccini, R., Vogel, R., Hoffmann, C., Becker-Baldus, J., Ouari, O., Tordo, P., Mchaourab, H. & Glaubitz, C. (2015) The ABC exporter MsbA probed by solid state NMR – challenges and opportunities. *Biological chemistry* 396, 1135–1149.

Kaur, H., Lakatos-Karoly, A., Vogel, R., Nöll, A., Tampé, R. & Glaubitz, C. (2016) Coupled ATPase-adenylate kinase activity in ABC transporters. *Nature Communications* 7, 13864.

Kerr, I.D. (2002) Structure and association of ATP-binding cassette transporter nucleotide-binding domains. *Biochimica et biophysica acta* 1561, 47–64.

Kim, J., Wu, S., Tomasiak, T.M., Mergel, C., Winter, M.B., Stiller, S.B., Robles-Colmanares, Y., Stroud, R.M., Tampé, R., Craik, C.S. & Cheng, Y. (2015) Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature* 517, 396–400.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Kühlbrandt, W. (2014) Biochemistry. The resolution revolution. *Science* 343, 1443–1444.

Lander, E.S. *et al.*, J.International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.

Lange, V., Becker-Baldus, J., Kunert, B., van Rossum, B.-J., Casagrande, F., Engel, A., Roske, Y., Scheffel, F.M., Schneider, E. & Oschkinat, H. (2010) A MAS NMR Study of the Bacterial ABC Transporter ArtMP. *ChemBioChem* 11, 547–555.

Lee, M., Choi, Y., Burla, B., Kim, Y.-Y., Jeon, B., Maeshima, M., Yoo, J.-Y., Martinoia, E. & Lee, Y. (2008) The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO2. *Nature cell biology* 10, 1217–1223.

Lin, D.Y.-W., Huang, S. & Chen, J. (2015) Crystal structures of a polypeptide processing and secretion transporter. *Nature* 523, 425–430.

Linton, K.J. & Higgins, C.F. (1998) The Escherichia coli ATP-binding cassette (ABC) proteins. *Molecular microbiology* 28, 5–13.

Locher, K.P. (2004) Structure and mechanism of ABC transporters. *Current Opinion in Structural Biology* 14, 426–431.

Locher, K.P. (2009) Structure and mechanism of ATP-binding cassette transporters. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 239–245.

Locher, K.P., Lee, A.T. & Rees, D.C. (2002) The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* 296, 1091–1098.

Matar-Merheb, R., Rhimi, M., Leydier, A., Huché, F., Galián, C., Desuzinges-Mandon, E., Ficheux, D., Flot, D., Aghajari, N., Kahn, R., Di Pietro, A., Jault, J.-M., Coleman, A.W. & Falson, P. (2011) Structuring Detergents for Extracting and Stabilizing Functional Membrane Proteins J. Hoheisel (Ed). *PLoS ONE* 6, e18036.

Matthews, E.E., Zoonens, M. & Engelman, D.M. (2006) Dynamic helice interactions in transmembrane signaling. *Cell* 127, 447–450.

Milo, R. (2013) What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays : news and reviews in molecular, cellular and developmental biology* 35, 1050–1055.

Mishra, S., Verhalen, B., Stein, R.A., Wen, P.-C., Tajkhorshid, E. & McHaourab, H.S. (2014) Conformational dynamics of the nucleotide binding domains and the power stroke of a heterodimeric ABC transporter. *eLife* 3, e02740.

Mittal, A., Böhm, S., Grütter, M.G., Bordignon, E. & Seeger, M.A. (2012) Asymmetry in the homodimeric ABC transporter MsbA recognized by a DARPin. *The Journal of biological chemistry* 287, 20395–20406.

Murata, K. & Wolf, M. (2017) Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochimica et biophysica acta*.

Oancea, G., O'Mara, M.L., Bennett, W.F.D., Tieleman, D.P., Abele, R. & Tampé, R. (2009) Structural arrangement of the transmission interface in the antigen ABC transport complex TAP. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 106, 5551–5556.

Oldham, M.L. & Chen, J. (2011) Snapshots of the maltose transporter during ATP hydrolysis. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 108, 15152–15156.

Oldham, M.L., Davidson, A.L. & Chen, J. (2008) Structural insights into ABC transporter mechanism. *Current Opinion in Structural Biology* 18, 726–733.

Oldham, M.L., Grigorieff, N. & Chen, J. (2016) Structure of the transporter associated with antigen processing trapped by herpes simplex virus. *eLife* 5, 213.

Orelle, C. & Jault, J.M. (2016) Structures and Transport Mechanisms of the ABC Efflux Pumps. *Efflux-Mediated Antimicrobial Resistance in Bacteria*.

Orelle, C., Dalmas, O., Gros, P., Di Pietro, A. & Jault, J.-M. (2003) The conserved glutamate residue adjacent to the Walker-B motif is the catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA. *Journal of Biological Chemistry* 278, 47002–47008.

Orelle, C., Gubellini, F., Durand, A., Marco, S., Lévy, D., Gros, P., Di Pietro, A. & Jault, J.-M. (2008) Conformational change induced by ATP binding in the multidrug ATP-binding cassette transporter BmrA. *Biochemistry* 47, 2404–2412.

Parcej, D. & Tampé, R. (2010) ABC proteins in antigen translocation and viral inhibition. *Nature Chemical Biology*.

Perez, C., Gerber, S., Boilevin, J., Bucher, M., Darbre, T., Aebi, M., Reymond, J.-L. & Locher, K.P. (2015) Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* 524, 433–438.

Qian, H., Zhao, X., Cao, P., Lei, J., Yan, N. & Gong, X. (2017) Structure of the Human Lipid Exporter ABCA1. *Cell*.

Ravaud, S., Do Cao, M.-A., Jidenko, M., Ebel, C., Le Maire, M., Jault, J.-M., Di Pietro, A., Haser, R. & Aghajari, N. (2006) The ABC transporter BmrA from Bacillus subtilis is a functional dimer when in a detergent-solubilized state. *Biochemical Journal* 395, 345.

Rees, D.C., Johnson, E. & Lewinson, O. (2009) ABC transporters: the power to change. *Nature Reviews Molecular Cell Biology* 10, 218–227.

Riordan, J.R. & Ling, V. (1979) Purification of P-glycoprotein from plasma membrane vesicles of Chinese hamster ovary cell mutants with reduced colchicine permeability. *Journal of Biological Chemistry* 254, 12701–12705.

Rosenberg, M.F., Callaghan, R., Ford, R.C. & Higgins, C.F. (1997) Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis. *Journal of Biological Chemistry* 272, 10685–10694.

Saraste, M., Sibbald, P.R. & Wittinghofer, A. (1990) The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* 15, 430–434.

Sauna, Z.E., Kim, I.-W., Nandigama, K., Kopp, S., Chiba, P. & Ambudkar, S.V. (2007) Catalytic cycle of ATP hydrolysis by P-glycoprotein: evidence for formation of the E.S reaction intermediate with ATP-gamma-S, a nonhydrolyzable analogue of ATP. *Biochemistry* 46, 13787–13799.

Schmitt, L. & Tampé, R. (2002) Structure and mechanism of ABC transporters. *Current Opinion in Structural Biology* 12, 754–760.

Senior, A.E., al-Shawi, M.K. & Urbatsch, I.L. (1995) The catalytic cycle of P-glycoprotein. *FEBS Letters* 377, 285–289.

Sharma, S. & Davidson, A.L. (2000) Vanadate-induced trapping of nucleotides by purified maltose transport complex requires ATP hydrolysis. *Journal of Bacteriology* 182, 6570–6576.

Siarheyeva, A., Liu, R. & Sharom, F.J. (2010) Characterization of an asymmetric occluded state of P-glycoprotein with two bound nucleotides: implications for catalysis. *The Journal of biological chemistry* 285, 7575–7586.

Siarheyeva, A., Lopez, J.J., Lehner, I., Hellmich, U.A., van Veen, H.W. & Glaubitz, C. (2007) Probing the molecular dynamics of the ABC multidrug transporter LmrA by deuterium solid-state nuclear magnetic resonance. *Biochemistry* 46, 3075–3083.

Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J. & Hunt, J.F. (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Molecular cell* 10, 139–149.

Steinfels, E., Orelle, C., Dalmas, O., Penin, F., Miroux, B., Di Pietro, A. & Jault, J.-M. (2002) Highly efficient over-production in E. coli of YvcC, a multidrug-like ATP-binding cassette transporter from Bacillus subtilis. *Biochimica et biophysica acta* 1565, 1–5.

Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A. & Jault, J.-M. (2004) Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in Bacillus subtilis†. *Biochemistry* 43, 7491–7502.

Taylor, N., Manolaridis, I., Jackson, S.M. & Kowal, J. (2017a) Structure of the human multidrug transporter ABCG2. *Nature*.

Taylor, N.M.I., Manolaridis, I., Jackson, S.M., Kowal, J., Stahlberg, H. & Locher, K.P. (2017b) Structure of the human multidrug transporter ABCG2. *Nature* 546, 504–509.

Theodoulou, F.L. & Kerr, I.D. (2015) ABC transporter research: going strong 40 years on. *Biochemical Society transactions* 43, 1033–1040.

Timachi, M.H., Hutter, C.A., Hohl, M., Assafa, T., Böhm, S., Mittal, A., Seeger, M.A. & Bordignon, E. (2017) Exploring conformational equilibria of a heterodimeric ABC transporter. *eLife* 6, e20236.

Tombline, G. & Senior, A.E. (2005) The occluded nucleotide conformation of p-glycoprotein. *Journal of bioenergetics and biomembranes* 37, 497–500.

van der Does, C. & Tampé, R. (2004) Changing orders--primary and secondary membrane transporters revised. *ChemBioChem* 5, 1171–1175.

van Veen, H.W., Margolles, A., Müller, M., Higgins, C.F. & Konings, W.N. (2000) The homodimeric ATP-binding cassette transporter LmrA mediates multidrug transport by an alternating two-site (two-cylinder engine) mechanism. *The EMBO journal* 19, 2503–2514.

Venter, J.C., Smith, H.O. & Adams, M.D. (2015) The Sequence of the Human Genome. *Clinical chemistry* 61, 1207–1208.

Vetter, I.R. & Wittinghofer, A. (1999) Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Quarterly Reviews of Biophysics* 32, 1–56.

Voreck, A.M. (2015) Structural Characterization of a Transmembrane Protein by Solid-state NMR. 1–167.

Wallin, E. & Heijne, von, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science* 7, 1029–1038.

Ward, A., Reyes, C.L., Yu, J., Roth, C.B. & Chang, G. (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 104, 19005–19010.

Wen, P.-C., Verhalen, B., Wilkens, S., McHaourab, H.S. & Tajkhorshid, E. (2013) On the origin of large flexibility of P-glycoprotein in the inward-facing state. *The Journal of biological chemistry* 288, 19211–19220.

Wilkens, S. (2015) Structure and mechanism of ABC transporters. *F1000prime reports* 7, 14.

Wilson, D., Pethica, R., Zhou, Y., Talbot, C., Vogel, C., Madera, M., Chothia, C. & Gough, J. (2009) SUPERFAMILY--sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Research* 37, D380–6.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I.B. & Schmitt, L. (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *The EMBO journal* 24, 1901–1910.

Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I.B. & Schmitt, L. (2006) A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. *The EMBO journal* 25, 3432–3443.

Zhang, Z., Liu, F. & Chen, J. (2017) Conformational Changes of CFTR upon Phosphorylation and ATP Binding. *Cell*.

Zimmerman, S.B. & Trach, S.O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of molecular biology* 222, 599–620.

Zoghbi, M.E. & Altenberg, G.A. (2013) Hydrolysis at one of the two nucleotide-binding sites drives the dissociation of ATP-binding cassette nucleotide-binding domain dimers. *The Journal of biological chemistry* 288, 34259–34265.

Zoghbi, M.E. & Altenberg, G.A. (2014) ATP binding to two sites is necessary for dimerization of nucleotide-binding domains of ABC proteins. *Biochemical and Biophysical Research Communications* 443, 97–102.

Zou, P., Bortolus, M. & McHaourab, H.S. (2009) Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. *Journal of molecular biology* 393, 586–597.

Chapter II

Solid-state NMR seen by a biochemist

Contents

Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) is a spectroscopic method that the general public mainly knows *via* the Magnetic Resonance Imagery (MRI). Both methods use the magnetism from the nuclei in order to extract information. Using the hydrogen from water, MRI allows to characterize and differentiate the tissues with great accuracy and a wealth of tissues information that would otherwise be obtained only in an invasive manner. Nevertheless, MRI is an application of the concept of NMR. NMR has been developing since the beginning of the 20th century and is still in development nowadays. These progresses have been made possible thanks to the work of a scientist community involving physicists, chemists and biologists, and have been rewarded with six Nobel prizes:

- **Otto Stern**, USA: Nobel Prize in Physics in 1943, "*for his contribution to the development of molecular ray method and his discovery of the magnetic moment of the proton*"
- **Isidor I. Rabi**, USA: Nobel Prize in Physics in 1944, "*for his resonance method for recording the magnetic properties of atomic nuclei*"
- **Felix Bloch** and **Edward M. Purcell**, USA: Nobel Prize in Physics in 1952, "*for their discovery of new methods for nuclear magnetic precision measurements and discoveries in connection therewith*"
- **Richard R. Ernst**, Switzerland: Nobel Prize in Chemistry in 1991, "*for his contributions to the development of the methodology of high resolution nuclear magnetic resonance (NMR) spectroscopy*"
- **Kurt Wüthrich**, Switzerland: Nobel Prize in Chemistry in 2002, "*for his development of nuclear magnetic resonance spectroscopy for determining the three-dimensional structure of biological macromolecules in solution*"
- **Paul C. Lauterbur**, USA and **Peter Mansfield**, United Kingdom: Nobel Prize in Physiology or Medicine in 2003, "*for their discoveries concerning magnetic resonance imaging*"

The main application of NMR spectroscopy is to determine the structures of molecules synthetized by organic chemistry and the molecules produced *in vivo*, *i.e* proteins and nucleic acids. Furthermore, the field of application can be extended to the probing of molecular dynamics, protein folding and molecular interactions (*i.e* the identification of

protein regions involved in an interaction with another protein, a nucleic acid or a small molecule, and the characterization of dissociation constants).

NMR spectroscopy uses the magnetic properties of certain nuclei, in particular the nuclear magnetic moment. Indeed, a nucleus is composed of nucleons, including protons and neutrons, which are endowed with a magnetic moment. The magnetic moments of nucleons match up and cancel out each other. When the number of neutrons and protons is even in a nucleus, the resulting magnetic moment is zero and therefore the nucleus cannot be detected by NMR. However, if the number of neutrons and / or protons is odd, the nucleus has a nonzero magnetic moment, which allows NMR detection. In fact, nuclei are associated with a nuclear-spin, the nuclei are observable by NMR only when this nuclear-spin is non-null. In NMR of proteins, nuclei displaying a spin of ½ are simple to analyse whereas spin > ½ induces linebroading (*i.e* ^{14}N compared to ^{15}N) hampering detection. The list of the most studied nuclei by NMR is presented **Table. 1**.

Nucleus	Spin I	Natural abundance (%)
^{1}H	1/2	99.9844
^{2}H	1	0.0116
^{12}C	0	98.8944
^{13}C	1/2	1.1056
^{14}N	1	99.6337
^{15}N	1/2	0.3663
^{16}O	0	99.7621
^{17}O	5/2	0.0379
^{18}O	0	0.2000
^{31}P	1/2	100

Table 1. Table listing percent natural abundance for the stable nuclides with the associated spin I. The percent natural abundance data is from the 2002 report of the IUPAC Subcommittee for Isotopic Abundance Measurements (Coplen *et al.* 2002).

98.8944 % isotopes of carbon are ^{12}C which has a nuclear spin of 0 and consequently not observable by NMR. On the other hand, 99.6337 % isotopes of nitrogen atom are ^{14}N which has a nuclear spin of 1 and indeed observable by NMR but displays linebroading on spectrum. Consequently, proteins need to be isotopically enriched in ^{13}C and ^{15}N. This stage of isotopic labelling is necessary in order to carry out an in-depth study of protein, DNA and RNA by NMR, when proton spectroscopy is not yielding enough information. The

culture, where the protein is produced, is carried out in a minimal medium in the presence of ^{13}C-glucose and ^{15}N-amonium chloride as sole source of carbon and nitrogen.

For these nuclei, an NMR signal is detected when they are in presence of a strong external magnetic field and are excited by a radiofrequency radiation using an NMR spectrometer. NMR spectrometers are installed in a large number of chemistry labs and are often the tool of choice as NMR is above all an analytical chemistry technique used for the industry and for the academic research as well. The common applications of NMR spectroscopy include: structure elucidation, chemical composition determination, formulations investigation, purity determination, quantitative analysis, quality control and quality assurance etc. Much progress has been made to facilitate the accessibility of NMR to such an extent that the benchtop NMR spectrometer is now developed to perform the functions of these common applications. Besides, NMR can also be used for the analysis of biomolecules such as proteins, DNA and RNA which are composed mainly of carbon, hydrogen and nitrogen.

Solid-state NMR of (membrane) proteins

NMR, X-ray crystallography and cryo-electron microscopy are the only methods capable of solving the three-dimensional structure of proteins at atomic level. The three-dimensional structure is essential to provide information about the function of the protein and mechanistic information. Structural biology is now an asset in the drug design process (Whittle & Blundell 1994), (Anderson 2003), (Zhang & Lai 2011), leading to a special interest for academic research as well as industry including big pharmaceutical companies.

Since the first protein structure - the pancreatic ribonuclease - solved in 1957 (Saunders & Wishnia 1957), technical advancements in terms of instrumentation (probes and high magnetic field developments), methodologies (pulse sequences), and knowledge in biochemistry have made NMR one of the essential method for structural studies of proteins.

Two fields of NMR used in structural biology can be distinguished: liquid-state NMR and solid-state NMR. Liquid-state NMR provides structural information on a protein sample in solution whereas solid-state NMR provides structural information on a solid protein sample (powder, crystals or sediment).

Protein solid-state NMR is a technique which is still in development for the studies of large molecular assemblies. It can be combined with other structural techniques such as x-ray crystallography, liquid-state NMR or cryo-electron microscopy for protein-structure investigation at an atomic level. While the general size-limit of studied proteins in liquid-state NMR is around 45-50 kDa, solid-state NMR is not limited by the size or the tumbling rate of molecules.

Despite the difference in the sample state (liquid *vs* solid), the general workflows concerning biomolecular studies are the same for solid-state NMR and for liquid-state NMR: sample preparation, data collection, analysis and interpretation, and even 3D structure calculation.

1. Sample preparation

Sample preparation is a cornerstone in any NMR study. Indeed, an NMR sample has to be highly concentrated, homogenous and stable in the chosen condition until and while the

data are collected. In liquid-state NMR, a protein concentration of 0.1 to 1 mM for a volume of 500 µL is sufficient to perform data collection. In liquid-state NMR, the liquid-sample is loaded in a glass tube whereas in solid-state NMR, the solid-sample is loaded in a rotor. A NMR rotor is composed by a cylindrical tube in Zirconia that is open on one side in order to be filled with a solid sample (**Figure 15B**) and the tube is closed by a drive cap, which shows wings (**Figure 15A**) of polychlorotrifluoroethylene (Kel-F). A large panel of rotors with different diameters is available in solid-state NMR, **Figure 15C**. The rotor is disposed in a stator inside a NMR probe where airflow is directed on the winged drive cap which generates the rotation of the rotor.

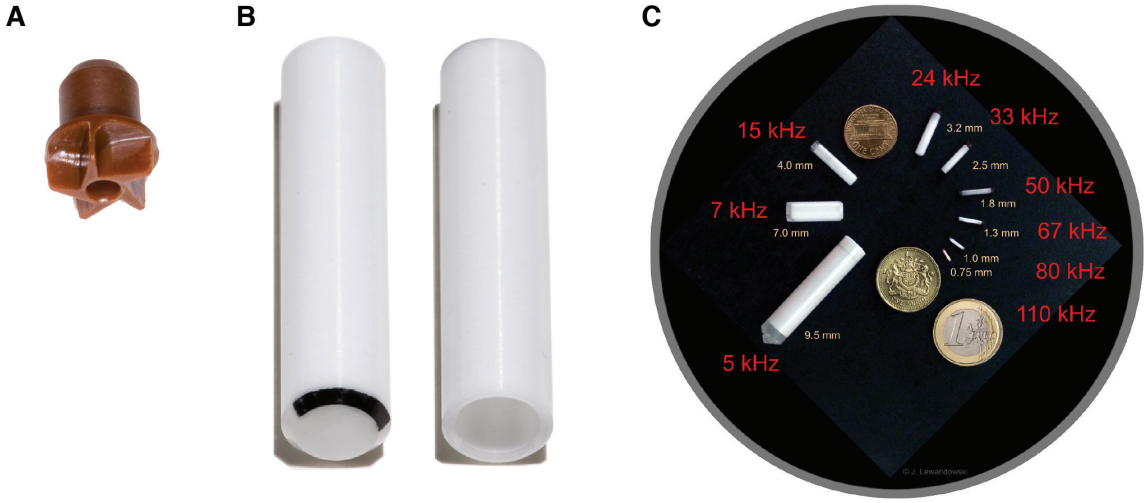


Figure 15. Presentation of a solid-state NMR rotor and its variants. Kel-F Cap for 3.2 mm rotor (**A**). Zirconia rotor 3.2 mm Bruker MAS probe, the black mark on the bottom of the tube is read by a tachometer in order to indicate the spinning frequency of the rotor (**B**). The different size rotor used in NMR with the associated spinning frequency, picture from http://www.warwick.ac.uk made by Dr Józef Lewandowski (**C**).

Spinning the sample at a specific angle of 54.7° or the magic angle with respect to the direction of the magnetic field, is a prerequisite for obtaining high-resolution spectra in solid-state NMR. Briefly, the line width is dependent on the chemical shift anisotropy and also the dipolar coupling. The chemical shift anisotropy (CSA) reminds that the electronic environment of nucleus varies with the position, which creates CSA. The dipolar coupling is the mutual interaction between two spins or two magnetic dipoles with a specific distance; the dipolar coupling is inverse proportional to the third power of this distance. Both CSA and dipolar couplings are depending on the molecular orientation by a factor $(3\cos^2\theta - 1)$,

where θ is the angle that describes the orientation of the spin with respect to the external magnetic field.

In liquid-state NMR, the molecular tumbling reduces the anisotropic interactions to their isotropic values. Indeed, in solution θ varies between 0 and 360°, the integral of 0 to 360 on $(3\cos^2\theta - 1)$ leads to a value equal to 0. The molecular tumbling averages the anisotropic interactions to their isotropic values.

In solid-state NMR, the molecular tumbling is absent thus the anisotropic interactions are not averaged. Molecules are oriented in all directions (0 to 360°) regarding the external magnetic field and all possible chemical shifts can be observed resulting to broad peaks, the so-called powder spectrum. Spinning the sample at angle of 54.7° with respect to the external field removes the angular dependence, since the factor $(3\cos^2\theta - 1)$ becomes $3\cos^2 54.7 - 1 = 0$. The rotor spinning frequency averages chemical shift anisotropies and dipolar couplings which are on the order of the spinning frequency. The spinning frequency used in this thesis, 17.5 kHz, effectively averages chemical shift anisotropies and dipolar couplings between ^{13}C and ^{15}N. However, dipolar couplings between $^{13}C/^{15}N$ and ^{1}H are not averaged and require higher spinning frequencies. To reduce the dipolar coupling between these two nuclei, a radio-frequency radiation is applied on the proton during the signal detection. By this the heteronuclear dipolar couplings are averaged to its isotropic value, leading to a better-resolved spectrum, **Figure 16**.

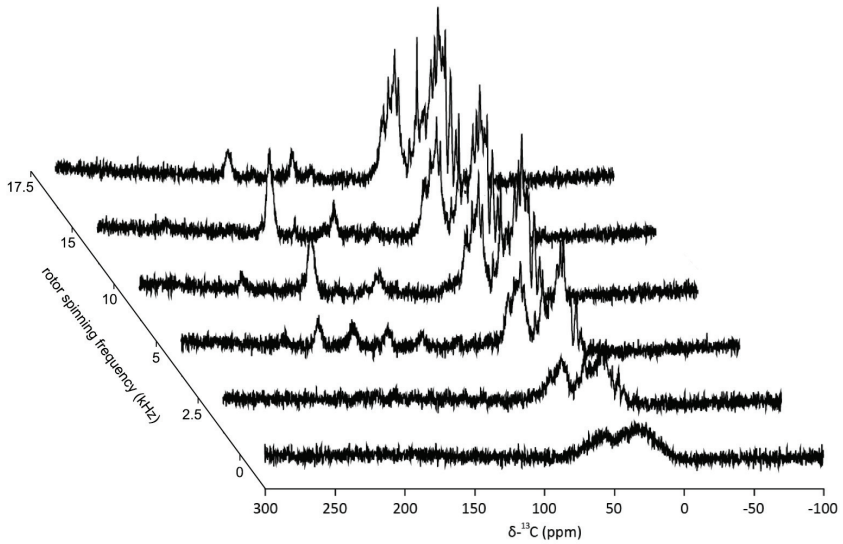


Figure 16. Spinning frequency effect on 1D ^{13}C CP spectrum. 1D spectra of ^{12}C-^{14}N-[LVIKHP]-^{13}C-^{15}N-BmrA recorded at different spinning frequency (0 to 17.5 kHz) on Bruker Biospin AVANCE III spectrometers operating at 800 MHz ^{1}H frequency using 3.2 mm triple-resonance (^{1}H, ^{13}C, ^{15}N) E-free probes (Bruker Biospin).

In solid-state NMR, it is the total quantity of protein (in mg) inside a rotor that matters and that will influence the signal-to-noise ratio on the resulting NMR spectra. The quantity depends on the sample volume of the solid-state NMR rotor.

One of the advantages of solid-state NMR is the wide variety of biological material that can be studied such as membrane proteins in lipids (Kunert *et al.* 2014; Lalli *et al.* 2017; Wiegand *et al.* 2017), micro- or nanocrystals without diffracting requirement as requested by x-ray crystallography (Böckmann *et al.* 2009; Loquet *et al.* 2009; Wasmer *et al.* 2009), sedimented soluble proteins or protein complexes (Bertini *et al.* 2012; Wiegand *et al.* 2016b; Wiegand *et al.* 2017), as well as fibrillar aggregates (Loquet *et al.* 2009; Wasmer *et al.* 2009; Hoop *et al.* 2016).

As mentioned above, solid-state NMR is not limited by the size of proteins, however large proteins lead to the recording of crowded NMR spectra, due to overlapped signals. To avoid this drawback, proteins can be enriched selectively in ^{13}C or ^{15}N to label only the amino-acids of interest.

2. Data collection

Data collection mainly depends on the project and the sample quality. Two-dimensional experiments provide the fingerprint of the protein and are usually the first experiments performed on the protein.

In liquid-state NMR, the common fingerprint experiment is the ^{1}H-^{15}N Heteronuclear Single Quantum Coherence (HSQC) experiment, which gives a correlation signal for all the ^{1}H-^{15}N groups in the protein (NH backbone amid groups, Nε-Hε groups from tryptophan, Nδ-Hδ_2 from asparagine, Nε-Hε_2 from glutamine, Nε-Hε from arginine and at low pH Nη-Hη from arginine and Nζ-Hζ from lysine).

In solid-state NMR, the most standard experiment that provides the fingerprint of the protein is called Dipolar Assisted Rotational Resonance (DARR) (see Routine solid-state NMR experiment Section for more details) and it gives a correlation signal for the different ^{13}C nuclei in the protein. This experiment can be considered as the solid-state NMR equivalent of the HSQC. However, DARR spectra display more resonances because there are much more ^{13}C-^{13}C correlations than ^{1}H-^{15}N groups in a protein. Also, the chemical shift

dispersion is larger in a DARR than in a HSQC. It is worth noticing that, thanks to the recent advances in fast magic angle spinning, ^{1}H-^{15}N correlations spectra can now be recorded in solid-state NMR even on fully protonated proteins.

In both cases, sequential assignments of the fingerprint spectra is performed, *i.e* each resonance has to be assigned to its corresponding nucleus in the protein. Multidimensional (usually 3D or 4D) spectra are recorded, enabling to follow a sequential backbone assignment way (Pauli *et al.* 2001), (Schütz *et al.* 2010). To perform this sequential assignment, sets of three-dimensional experiments have to be recorded in order to "walk" along the protein backbone. In theory, the assignment can be performed using NCACB, NCACO, CANCO and NCOCA experiments, the name of each experiment indicating the polarisation transfer steps. For example, in the case of the NCACB, magnetization is transferred from the $^{1}H^{N}$ to the ^{15}N and subsequently from the ^{15}N to the $^{13}C\alpha$ *via* a cross polarisation step. Finally, the magnetisation is transferred to the $^{13}C\beta$ further along the amino acid side chain. The combination of these four experiments is depicted in **Figure 17**.

When all observable NMR resonances are assigned to their corresponding nucleus in the protein, structural information can be obtained and a deeper data analysis can be performed.

3. Routine solid-state NMR experiment

3.1. The Dipolar Assisted Rotational Resonance (DARR) pulse-sequence

The main solid-state NMR experiment used during the work of this thesis is the Dipolar Assisted Rotational Resonance (DARR), which was developed by Takegoshi and co-workers (Takegoshi *et al.* 2001; Takegoshi *et al.* 2003; Morcombe *et al.* 2004). DARR spectrum is a homonuclear two-dimensional NMR experiment, providing information on all the ^{13}C nuclei of the protein. Like every NMR experiment, the DARR experiment consists of a pulse-sequence that leads to the reception of a signal in the form of Free Induction Decay (FID). The acquisition is performed during a specific time (the acquisition time or t2) and several times using an increment delay (the evolution time or t_1). The pulse sequence starts with a 90° pulse of 2.5 µs on the ^{1}H channel (100 kHz) which is followed by a cross-polarisation (CP) step. For the CP, a simultaneous pulse is applied on the ^{1}H and ^{13}C channels with a length of 700 µs for BmrA (this value depends on the protein and its

relaxation times); thus, the magnetization is transferred from ^{1}H nuclei to the ^{13}C nuclei *via* a heteronuclear dipolar interaction. For an optimal transfer, the pulses on the two channels have to be matched regarding the Hartmann Hahn condition. To fulfill this condition, the irradiation frequency of the ^{1}H must be equal to the irradiation frequency of the ^{13}C plus or minus x time the rotor spinning frequency. In the case of BmrA, the CP step involved a 700 µs pulse at 50 kHz in the ^{13}C channel and ~67.5 kHz in the ^{1}H channel at a spinning frequency of 17.5 kHz. The transfer of the magnetization from the CP induces an enhancement of the signal to noise ratio of 2.

The pulse sequence depicted **Figure 18** is described as follows: The longitudinal ^{13}C magnetization is created by a 90° pulse of 5 µs. The CP is followed by the indirect evolution t_1 in ^{13}C, during which the ^{1}H nuclei are decoupled, which means that the ^{1}H nuclei are irradiated with an rf field with the goal to reduce the dipolar with the ^{13}C. This will significantly simplify the resulting spectrum, as this greatly narrows the resonance lines. The indirect evolution time t_1 is incremented and a spectrum is recorded for each t_1 increment until the signal has decayed.

After t_1, the longitudinal ^{13}C magnetization is created by a 90° pulse of 5 µs. After a certain mixing time t, the resulting magnetization is observed by applying a second 90° pulse of 5 µs. The mixing-time is the time allowed to the ^{13}C nuclei to transfer the magnetisation to another ^{13}C nucleus. The longer is the mixing-time, the larger is the distance which can be measured between two visible ^{13}C nuclei (within the limit of signal decay). At the end, the FID from ^{13}C nuclei is recorded, and during the acquisition time (or t_2) the ^{1}H are decoupled a second time. The FID is Fourier transformed in the two dimensions in order to obtain the 2D DARR spectrum.

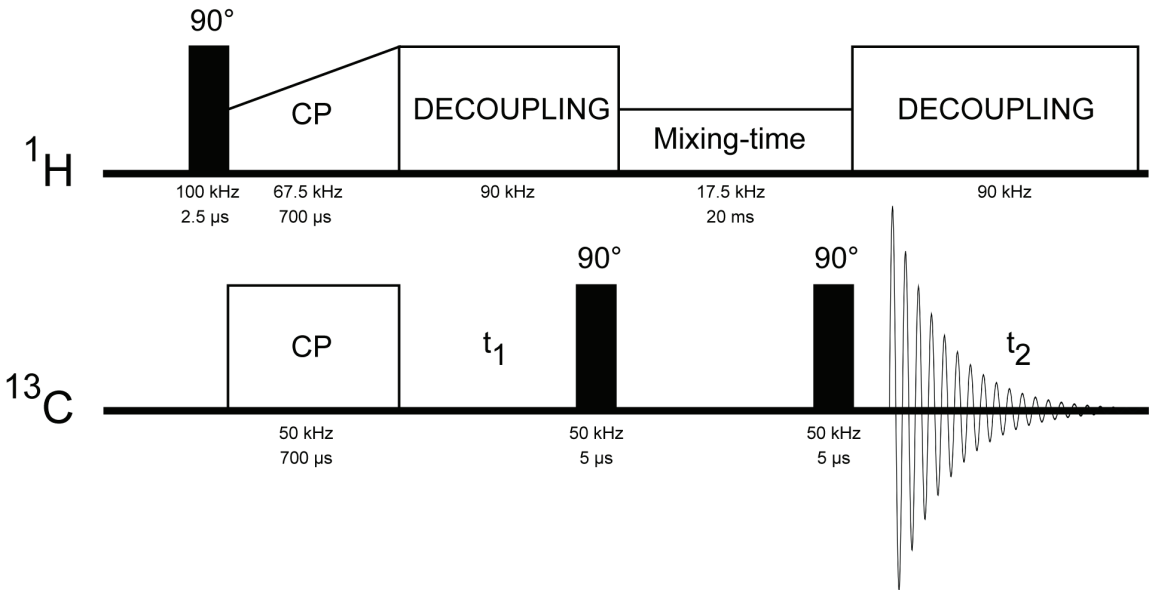


Figure 18. Pulse sequence for ^{13}C – ^{13}C correlation experiments using the DARR scheme for magnetization transfer during the mixing time.

3.2. Dipolar Assisted Rotational Resonance spectrum

One of the advantages of the DARR experiment is that it requires only ^{1}H and ^{13}C channels on a probe. Therefore, on a triple resonance probe (for example ^{1}H, ^{13}C, ^{15}N) that can be switched to double resonance mode using only the ^{13}C- and ^{1}H- channels the signal can be improved by a factor of 1.3 on our probe. For large molecules like BmrA, this signal enhancement is necessary to get a sufficient signal-to-noise ratio for subsequent analysis.

The DARR spectrum displays a central diagonal around which the correlation peaks are close to symmetrical. Each correlation peak corresponds to two ^{13}C nuclei that are close in space. By modifying the mixing-time parameter of the DARR, the minimal distance of detection can be varied. For example, with a mixing-time of 0 ms, there will be no-correlation peaks, for that of 20 ms, mainly intra-residual correlation peaks will be observed, while at 200 ms or more, intra-residual and inter-residual correlation peaks (mainly residues i-1 and i+1 but also residues close in space) will be observed. The maximum distance which can be observed using DARR is around 7 Å.

The NMR signal frequency is sensitive to the environment of each nucleus. Consequently, for a protein, the ^{13}C signal frequency depends on the residue, the primary sequence, the secondary and tertiary structures. On a DARR spectrum, the different residues can be, basically, identified regarding their characteristic resonance frequencies (Hz) or chemical shift (ppm), **Figure 19**. Indeed, in NMR the resonant frequency (in Hz) of a nucleus which depends of the electronic environment is used. The chemical shift – symbolized by δ – is

usually expressed in parts per million (ppm) by frequency and $\delta_{sample} = (\nu_{sample} - \nu_{ref}) / \nu_{ref}$; ν_{sample} the resonance frequency of the sample and ν_{ref} is the resonance frequency of a standard reference compound, measured in the same applied magnetic field. Since the numerator is usually expressed in hertz, and the denominator in megahertz, δ is expressed in ppm. The chemical shift is consequently independent of external magnetic field (Harris & Becker 2001).

However, especially for large proteins, the majority of the ^{13}C-^{13}C correlation peaks corresponding to specific residue overlap, only alanine, serine and threonine Cα-Cβ correlation are quite isolated. Therefore, the residue identification and thus the assignment of each residue of the protein are difficult on a crowded spectrum. To overcome this issue, residue-selective labelling or unlabelling can be performed to reduce spectral overlapping in order to enable the identification of typical spectral fingerprints.

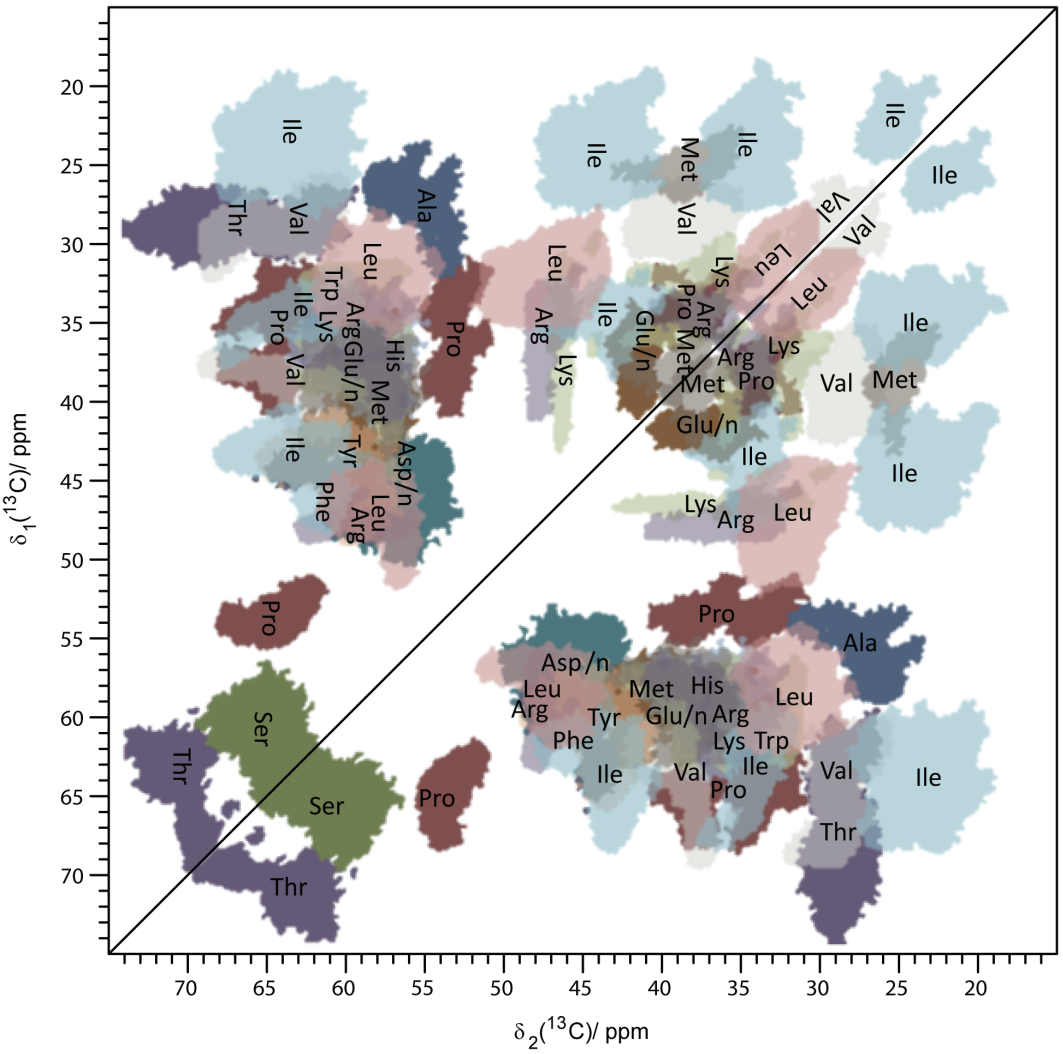


Figure 19. Full aliphatic 2D ^{13}C-^{13}C chemical shift correlation maps for the 20 amino acids, adapted from (Fritzsching *et al.* 2013).

4. Application to membrane proteins

4.1. Secondary and tertiary structure determination

The NMR signal frequency is extremely sensitive to the nucleus environment. Consequently the chemical shifts of each nucleus are specific for each amino-acid residue, and also strongly depends strongly on the secondary structure of the protein (Spera *et al.* 1991; Wishart *et al.* 1991; Wang 2002). Once the sequential assignment is achieved, the chemical shifts of the Cα and Cβ nuclei can be used to determine the secondary structure elements of the protein, as α-helices and β-sheets exert opposite effects on Cα and Cβ chemical shifts (Wong *et al.* 2016). The chemical shift difference Δδ between Cα and Cβ compared to the random-coil chemical shifts (δCα and δCβ respectively) gives information on the secondary structure of residues and is defined by:

$$\Delta\delta = \delta C\alpha - \delta C\beta \text{ where}$$

$$\delta C = \delta C_{observed} - \delta C_{random\ coil}$$

Random-coil chemical shifts (Wang 2002; Wishart *et al.* 1991) correspond to a conformation with no secondary or tertiary structural preference. Then, $\Delta\delta>0$ indicates a helical conformation, while $\Delta\delta<0$ indicates a β-sheet conformation. The average number of residues per helical turn for an α-helice is 3.6, hence four consecutive residues minimum with a $\Delta\delta>0$ can be predicted as an α-helice. Similarly, three consecutive residues with a $\Delta\delta<0$ are necessary to be predicted as β-sheet.

The secondary structure is the first step toward the 3D structure, which is accessible by solid-state NMR. The first structures obtained by solid-state NMR in combination with selective labelling was the SH3 domain in 2002 (Castellani *et al.* 2002). Concerning the membrane proteins structures referenced in the PDB, 82 structures have been solved by NMR including 10 specifically by solid-state NMR, **Figure 20**. However, for large systems, a full 3D structure determination is still challenging using solid-state NMR due to spectra quality and signal overlaps, as well as the time required for data acquisition and analysis. Nevertheless, detailed information on conformational changes or ligand interaction can be assessed by chemical shift analysis.

4.2. Identification of key residues *via* chemical shift perturbations

As previously said in the section, *2. Data collection,* in solid-state NMR, the DARR spectrum represents the fingerprint of a protein in a specific condition (temperature, buffer, pH...).

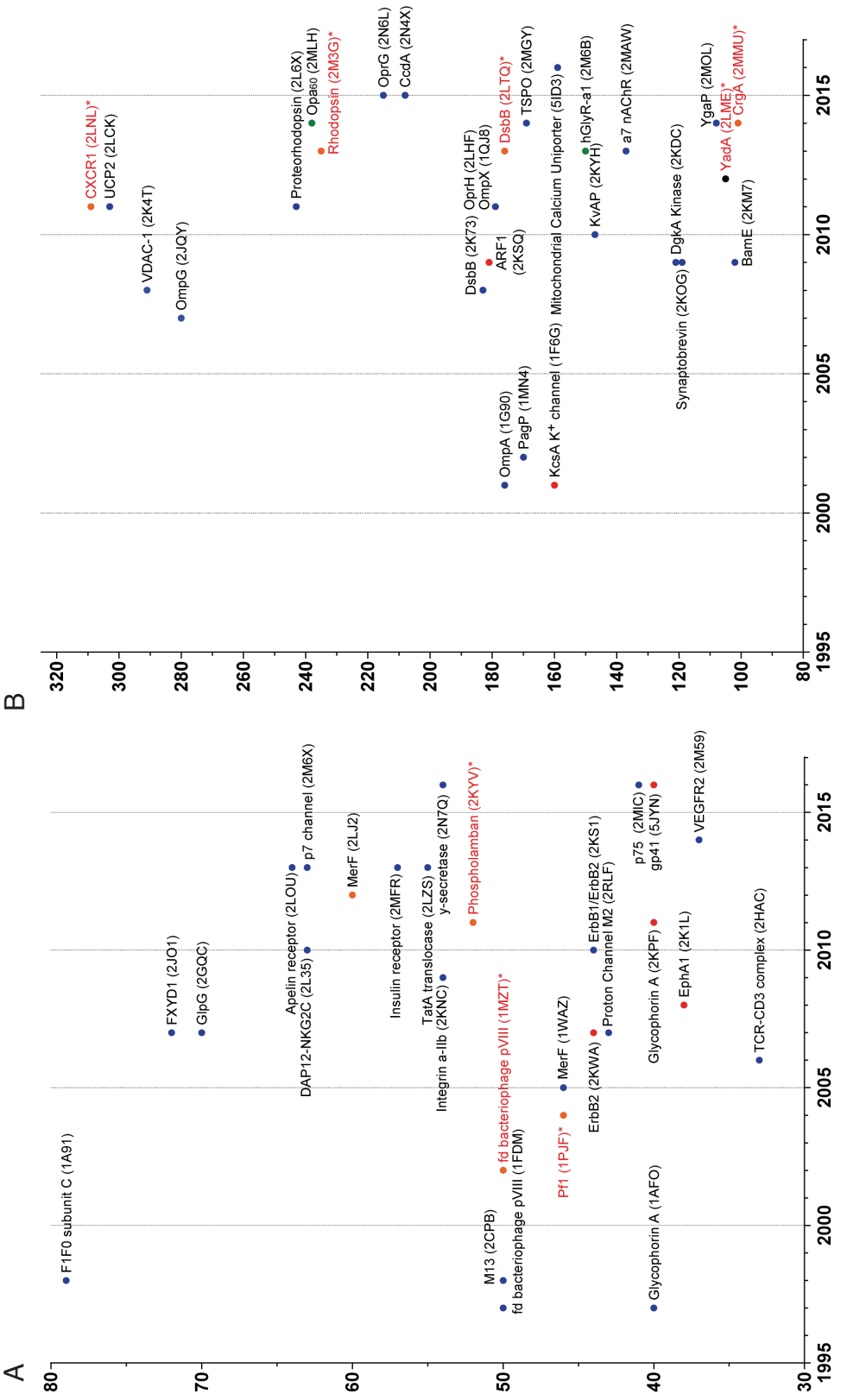


Figure 20. Structure of membrane proteins solved by NMR. Panel **A** displays the different NMR structures of peptides composed of 30 to 80 residues, panel **B** presents the different NMR structures of proteins composed of 80 to 320 residues. The NMR membrane protein structures were obtained in bicelle (red dot), detergent micelle (blue dot), nanodisc (green dot), crystals (black dot) and lipids bilayers (orange dot). Name in red indicate a structure solved by solid-state NMR. If several structures of a same protein were solved, only the first one is shown in the plot.

Indeed, there are no two identical spectra for two different proteins. If the condition changes (different temperature, different buffer, presence of ligand or substrate…) the associated spectrum will be different compared to a reference condition.

A different position of a peak between two spectra is called a chemical shift perturbation (CSP) which is caused in ^{13}C chemical shifts mainly by changes in the dihedral angles, and to a lesser extent also to the environment change. A light environment change around ^{13}C nucleus (like the binding of a small ligand) will induce a light chemical shift perturbation, whereas a substantial environment change (like structural change) will induce a wide chemical shift perturbation.

CSPs is use in liquid-state NMR in order to define the binding pocket of a small molecule or another molecule. The ligand can also be titrated into the protein, the titration is monitored at each stage *via* a 2D experiment recording.

In solid-state NMR, carbon chemical shifts are compared in order to probe where the interaction or structural changes take place. The backbone geometry affects mainly the chemical shift with a difference about 4 ppm for Cα and 2 ppm for Cβ to the secondary shift between α-helice and β-sheet (Iwadate *et al.* 1999; Williamson 2013). While sidechain conformation has an effect of up to 0.6 ppm on the Cα shift especially for amino acids with branched side-chains at Cβ and hydrogen bonding up to 0.9 ppm, depending on main-chain conformation (Iwadate *et al.* 1999; Williamson 2013). Accordingly to these characteristics, conformational change in a protein can be immediately identified on the NMR spectrum. Plethora of examples of CSPs using in protein can be found in the literature. We can note that significant conformational changes of the protein induce significant environment producing large CSP (between 0.5-1.5 ppm) (Etzkorn *et al.* 2010; Gustavsson *et al.* 2013; Gardiennet *et al.* 2016; Kwon & Hong 2016; Prade *et al.* 2016).Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. Using high-resolution solid-state NMR spectroscopy, here we show that high-affinity binding of the scorpion toxin kaliotoxin to a chimaeric K1 channel (KcsA-Kv1.3) is associated with significant structural rearrangements in both molecules as a result large (1.5 ppm) chemical-shift changes (Lange *et al.* 2006).

Interactions is generally localized on the protein surface, major CSP concern firstly atoms which are involved in the interaction ^{1}H or ^{15}N by hydrogen bonding, ionic interactions or electrostatic interaction. Carbon atoms are less surface exposed than ^{1}H or ^{15}N and so are

further removed from any external effects. Thus, through-space effects have only a minor effect on Cα and Cβ shifts (Williamson 2013) explaining the range of CSP (between 0.1-0.5 ppm) (Zech *et al.* 2004; Kwon & Hong 2016; Cerofolini *et al.* 2017). The Amyloid–Congo Red interface between the HET-s(218–289) – prion-forming domain of the HET-s prion from the filamentous fungus *Podospora anserina* – and Congo Red which reacts specifically with amyloids was probed my solid-state NMR. CSP resulting of this interaction was in range of 0.2-0.5 ppm (Schütz *et al.* 2011).

However, interaction can induce major conformational changes as shown by the helicase DnaB in complex with a non-hydrolysable ATP analogue and single-stranded DNA. The binding of both ligands was followed using ^{31}P NMR and ^{13}C NMR. Authors highlighted the conformational changes of the proteins, necessary to bind single-stranded DNA, this information sheds light on the mechanism of this motor protein (Wiegand *et al.* 2016a). The NMR studies of the protein DnaB which is reminiscent of an ABC transporter by several features (Geourjon *et al.* 2001) show the knowledge that can be provide by solid-state NMR without structure resolution.

4.3. Paramagnetism and solid-state NMR

Paramagnetism is mainly used in NMR in order to obtain long-range distance information, 10 Å to 35 Å (Bertini *et al.* 2002; Pintacuda *et al.* 2007; Huang *et al.* 2016). The aim is to use the paramagnetic relaxation enhancement (PRE) proprieties arises from magnetic dipolar interactions between a nucleus and the unpaired electrons of the paramagnetic center which result in an increase in nuclear relaxation rates (Clore & Iwahara 2009) as presented **Figure 21B** or pseudocontact shift as illustrated **Figure 21C**. The distance of dipolar interactions between a nucleus and the unpaired electrons of the paramagnetic center is dependant of the paramagnetic species. Although pseudo-contact shifts or PRE has been utilized since many years in order to get structures or long-range information, comparing to in liquid-state NMR only few investigations on protein were done in solid-state NMR (Jaroniec 2015). The first work using paramagnetic species on protein in solid-state NMR was published only in 2007 (Balayssac *et al.* 2007). Experiments were done on a metalloprotein containing a zinc(II) ion at the catalytic site which has been substituted by a cobalt(II) ion in order to get pseudocontact shifts. In the same time, the first work using PRE to obtain site-specific structural restraints were published on the model protein GB1 (Nadaud *et al.* 2007). Like electron paramagnetic resonance technic spin-label probe are fixed on the protein with a

preference for methanesulfonate derivative spin-label which is convalently bound with a sulfhydryl group of cysteine. Combination of PRE data from several mutants where cysteines were introduced allows the recording of long distance sets. As example, three-dimensional structure of the protein GB1 can be reached using six mutants containing Cys-EDTA-Cu^{2+} side-chains distributed throughout the protein. By this way, the group of Pr. Christopher Jaroniec collected a sufficient number of long-range restraints to elucidate the three-dimensional fold of the protein GB1 (Sengupta *et al.* 2012; Sengupta *et al.* 2013).

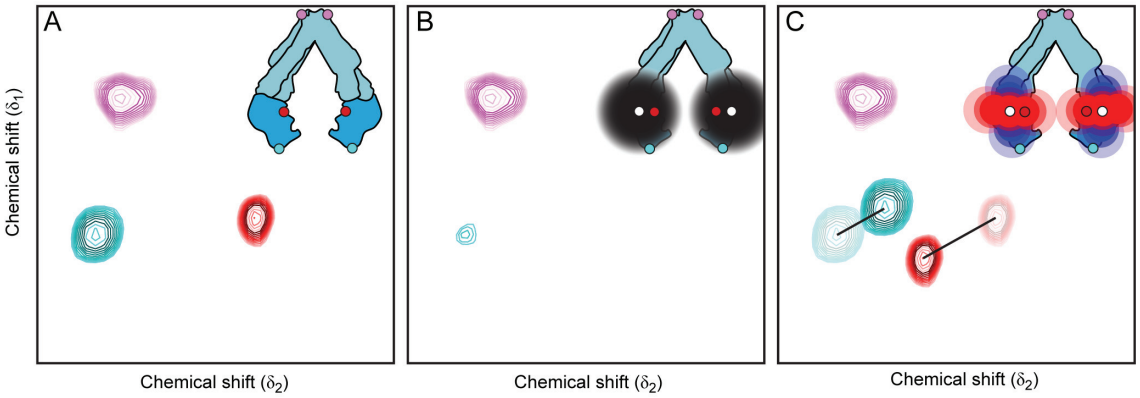


Figure 21. Paramagnetism and solid-state NMR. Schematic two-dimensional NMR chemical shift correlation spectra for a diamagnetic protein (**A**) and two of its structural analogs, each containing at a specific location a paramagnetic species inducing paramagnetic relaxation enhancement (**B**) or pseudocontact shift (**C**). For species inducing paramagnetic relaxation enhancement (*e.g.*, nitroxide spin label, Mn^{2+} or Cu^{2+}) in panel (**B**), the intensities of NMR signals (cyan, pink or red contours) are modulated according to the proximity of the corresponding nuclear spins in the protein (cyan, pink or red circle) to the paramagnetic center (white sphere). The PRE magnitude is represented by the black sphere around the paramagnetic center. For species inducing pseudocontact shifts (*e.g.*, Co^{2+}) in panel (**C**), the frequencies of NMR signals are modulated by PCSs that depend on both the proximity of the nuclei to the paramagnetic center as well as their location in the frame of species. On the protein, the locations of positive (blue spheres) and negative (red spheres) PCS, with successively more transparent surfaces corresponding to decreasing PCS magnitude. The transparent red and cyan contours in the spectrum in panel (**C**) indicate the positions of the corresponding NMR signals for the diamagnetic protein in panel (**A**). Picture and caption were adapted from (Jaroniec 2015).

References

Anderson, A.C. (2003) The process of structure-based drug design. *Chemistry & biology*.

Balayssac, S., Bertini, I., Lelli, M., Luchinat, C. & Maletta, M. (2007) Paramagnetic Ions Provide Structural Restraints in Solid-State NMR of Proteins. *Journal of the American Chemical Society* 129, 2218–2219.

Bertini, I., Engelke, F., Gonnelli, L., Knott, B., Luchinat, C., Osen, D. & Ravera, E. (2012) On the use of ultracentrifugal devices for sedimented solute NMR. *Journal of Biomolecular NMR* 54, 123–127.

Bertini, I., Luchinat, C. & Parigi, G. (2002) Paramagnetic constraints: an aid for quick solution structure determination of paramagnetic metalloproteins. *Concepts in Magnetic Resonance Part A*.

Böckmann, A., Gardiennet, C., Verel, R., Hunkeler, A., Loquet, A., Pintacuda, G., Emsley, L., Meier, B.H. & Lesage, A. (2009) Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45, 319–327.

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. & Oschkinat, H. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420, 98–102.

Cerofolini, L., Giuntini, S., Louka, A., Ravera, E., Fragai, M. & Luchinat, C. (2017) High-Resolution Solid-State NMR Characterization of Ligand Binding to a Protein Immobilized in a Silica Matrix. *The Journal of Physical Chemistry B* 121, 8094–8101.

Clore, G.M. & Iwahara, J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical Reviews* 109, 4108–4139.

Coplen, T.B., Böhlke, J.K., De Bievre, P. & Ding, T. (2002) Isotope-abundance variations of selected elements (IUPAC Technical Report). *Pure and Applied Chemistry*.

Etzkorn, M., Seidel, K., Li, L., Martell, S., Geyer, M., Engelhard, M. & Baldus, M. (2010) Complex Formation and Light Activation in Membrane-Embedded Sensory Rhodopsin II as Seen by Solid-State NMR Spectroscopy. *Structure/Folding and Design* 18, 293–300.

Fritzsching, K.J., Yang, Y., Schmidt-Rohr, K. & Hong, M. (2013) Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *Journal of Biomolecular NMR* 56, 155–167.

Gardiennet, C., Wiegand, T., Bazin, A., Cadalbert, R., Kunert, B., Lacabanne, D., Gutsche, I., Terradot, L., Meier, B.H. & Böckmann, A. (2016) Solid-state NMR chemical-shift perturbations indicate domain reorientation of the DnaG primase in the primosome of Helicobacter pylori. *Journal of Biomolecular NMR* 64, 189–195.

Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A. & Jault, J.M. (2001) A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends in Biochemical Sciences* 26, 539–544.

Gustavsson, M., Verardi, R. & Mullen, D.G. (2013) Allosteric regulation of SERCA by phosphorylation-mediated conformational shift of phospholamban.

Harris, R.K. & Becker, E.D. (2001) NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC recommendations 2001). *Pure and Applied Chemistry*.

Hoop, C.L., Lin, H.-K., Kar, K., Magyarfalvi, G., Lamley, J.M., Boatz, J.C., Mandal, A., Lewandowski, J.R., Wetzel, R. & van der Wel, P.C.A. (2016) Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 113, 1546–1551.

Huang, S., Umemoto, R., Tamura, Y., Kofuku, Y., Uyeda, T.Q.P., Nishida, N. & Shimada, I. (2016) Utilization of paramagnetic relaxation enhancements for structural analysis of actin-binding proteins in complex with actin. *Scientific reports* 6, 33690.

Iwadate, M., Asakura, T. & Williamson, M.P. (1999) Cα and Cβ Carbon-13 Chemical Shifts in Proteins From an Empirical Database. *Journal of Biomolecular NMR* 13, 199–211.

Jaroniec, C.P. (2015) Structural studies of proteins by paramagnetic solid-state NMR spectroscopy. *Journal of Magnetic Resonance* 253, 50–59.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Kwon, B. & Hong, M. (2016) The Influenza M2 Ectodomain Regulates the Conformational Equilibria of the Transmembrane Proton Channel: Insights from Solid-State Nuclear Magnetic Resonance. *Biochemistry* 55, 5387–5397.

Lalli, D., Idso, M.N., Andreas, L.B., Hussain, S., Baxter, N., Han, S., Chmelka, B.F. & Pintacuda, G. (2017) Proton-based structural analysis of a heptahelical transmembrane protein in lipid bilayers. *Journal of the American Chemical Society*.

Lange, A., Giller, K., Hornig, S., Martin-Eauclaire, M.-F., Pongs, O., Becker, S. & Baldus, M. (2006) Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* 440, 959–962.

Loquet, A., Bousset, L., Gardiennet, C., Sourigues, Y., Wasmer, C., Habenstein, B., Schütz, A., Meier, B.H., Melki, R. & Böckmann, A. (2009) Prion Fibrils of Ure2p Assembled under Physiological Conditions Contain Highly Ordered, Natively Folded Modules. *Journal of molecular biology* 394, 108–118.

Morcombe, C.R., Gaponenko, V., Byrd, R.A. & Zilm, K.W. (2004) Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *Journal of the American Chemical Society* 126, 7196–7197.

Nadaud, P.S., Helmus, J.J., Höfer, N. & Jaroniec, C.P. (2007) Long-Range Structural Restraints in Spin-Labeled Proteins Probed by Solid-State Nuclear Magnetic Resonance Spectroscopy. *Journal of the American Chemical Society* 129, 7502–7503.

Pauli, J., Baldus, M., van Rossum, B., de Groot, H. & Oschkinat, H. (2001) Backbone and side-chain 13C and 15N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *ChemBioChem* 2, 272–281.

Pintacuda, G., John, M., Su, X.-C. & Otting, G. (2007) NMR structure determination of protein-ligand complexes by lanthanide labelling. *Accounts of Chemical Research* 40, 206–212.

Prade, E., Barucker, C., Sarkar, R., Althoff-Ospelt, G., Lopez del Amo, J.M., Hossain, S., Zhong, Y., Multhaup, G. & Reif, B. (2016) Sulindac Sulfide Induces the Formation of Large Oligomeric Aggregates of the Alzheimer's Disease Amyloid-β Peptide Which Exhibit Reduced Neurotoxicity. *Biochemistry* 55, 1839–1849.

Saunders, M. & Wishnia, A. (1957) The nuclear magnetic resonance spectrum of ribonuclease. *J. Am. Chem. Soc.*, 1957, 79 (12), pp 3289–3290

Schütz A, Wasmer C, Habenstein B, Verel R, Greenwald J, Riek R, Böckmann A, Meier BH. (2010) Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s (1-227). *Chembiochem. 2010 Jul 26;11(11):1543-51*

Schütz, A.K., Soragni, A., Hornemann, S., Aguzzi, A., Ernst, M., Böckmann, A. & Meier, B.H. (2011) The amyloid-Congo red interface at atomic resolution. *Angewandte Chemie (International ed. in English)* 50, 5956–5960.

Sengupta, I., Nadaud, P.S. & Jaroniec, C.P. (2013) Protein structure determination with paramagnetic solid-state NMR spectroscopy. *Accounts of Chemical Research* 46, 2117–2126.

Sengupta, I., Nadaud, P.S., Helmus, J.J., Schwieters, C.D. & Jaroniec, C.P. (2012) Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nature chemistry* 4, 410–417.

Takegoshi, K., Nakamura, S. & Terao, T. (2001) 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* 344, 631–637.

Takegoshi, K., Nakamura, S. & Terao, T. (2003) [sup 13]C–[sup 1]H dipolar-driven [sup 13]C–[sup 13]C recoupling without [sup 13]C rf irradiation in nuclear magnetic resonance of rotating solids. *The Journal of Chemical Physics* 118, 2325.

Wang, Y. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science* 11, 852–861.

Wasmer, C., Schütz, A., Loquet, A., Buhtz, C., Greenwald, J., Riek, R., Böckmann, A. & Meier, B.H. (2009) The Molecular Organization of the Fungal Prion HET-s in Its Amyloid Form. *Journal of molecular biology* 394, 119–127.

Whittle, P.J. & Blundell, T.L. (1994) Protein structure--based drug design. *Annual review of biophysics and biomolecular structure* 23, 349–375.

Wiegand, T., Cadalbert, R., Gardiennet, C., Timmins, J., Terradot, L., Böckmann, A. & Meier, B.H. (2016a) Monitoring ssDNA Binding to the DnaB Helicase from Helicobacter pyloriby Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* 55, 14164–14168.

Wiegand, T., Gardiennet, C., Cadalbert, R., Lacabanne, D., Kunert, B., Terradot, L., Böckmann, A. & Meier, B.H. (2016b) Variability and conservation of structural domains in divide-and-conquer approaches. *Journal of Biomolecular NMR* 65, 79–86.

Wiegand, T., Lacabanne, D. & Keller, K. (2017) Solid-state NMR and EPR Spectroscopy of Mn2+-Substituted ATP-Fueled Protein Engines. *Angewandte Chemie ….*

Williamson, M.P. (2013) Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* 73, 1–16.

Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology* 222, 311–333.

Wong, W.C.V., Narkevicius, A., Chow, W.Y., Reid, D.G., Rajan, R., Brooks, R.A., Green, M. & Duer, M.J. (2016) Solid state NMR of isotope labelled murine fur: a powerful tool to study atomic level keratin structure and treatment effects. *Journal of Biomolecular NMR* 66, 93–98.

Zech, S.G., Olejniczak, E., Hajduk, P., Mack, J. & McDermott, A.E. (2004) Characterization of Protein–Ligand Interactions by High-Resolution Solid-State NMR Spectroscopy. *Journal of the American Chemical Society* 126, 13948–13953.

Zhang, C. & Lai, L. (2011) Towards structure-based protein drug design. *Biochemical Society transactions* 39, 1382–6– suppl 1 p following 1386.

Chapter III

Selective labelling and unlabelling for NMR

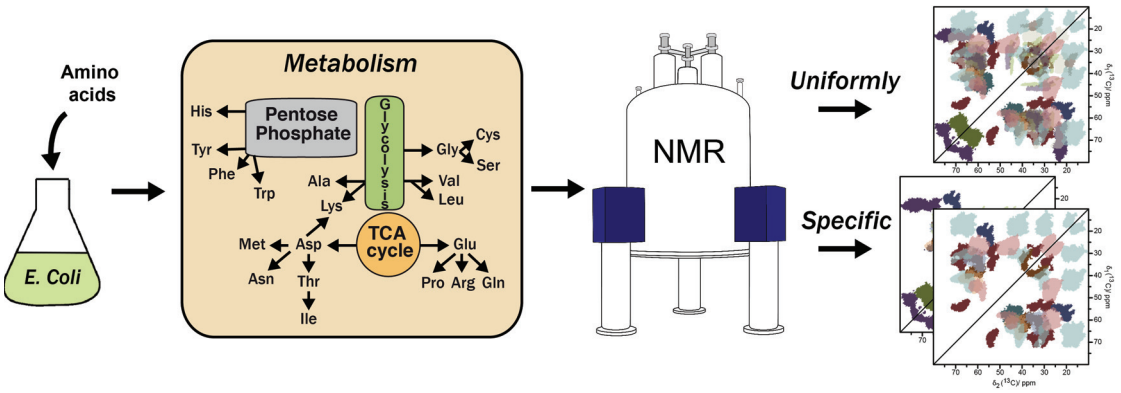
The work presented in this chapter was done with Beat H. Meier. This Chapter was submitted with the following reference:

Selective labelling and unlabelling strategies in protein NMR spectroscopy

Denis Lacabanne[1], Beat H. Meier[2] and Anja Böckmann[1]

Contents

Abstract



Selective isotope labelling is central in NMR experiments and often allows to push the size limits on the systems investigated, as it supplies additional resolution by diminishing the signals in the spectra, which is notably interesting in the large protein systems becoming accessible for solid-state NMR studies. Isotope labeled proteins for NMR experiments are most often expressed, in *E. coli* systems, where bacteria are grown in minimal media supplemented with $^{15}NH_4Cl$ and ^{13}C-glucose as sole source of nitrogen and carbon. For amino acids selective labelling or unlabelling, specific amino acids are supplemented in the minimal medium. The aim is that they will be incorporated in the protein by the bacteria. However, *E. coli* amino-acid anabolism and catabolism tend to interconnect different pathways, remnant of a subway system. These connections lead to inter conversion between amino acids, called scrambling. A thorough understanding of the involved pathways is thus important to obtain the desired labelling schemes, as not all combinations of amino acids are adapted. We present here a detailed overview of amino-acid metabolism in this context. Each amino-acid pathway is described in order to define accessible combinations for ^{13}C or ^{15}N specific labelling or unlabelling. Using as example the ABC transporter BmrA, a membrane protein of 600 amino-acid residues, we demonstrate how these strategies can be applied. Indeed, even though there is no size limit in solid-state NMR, large (membrane) proteins are still a challenge due to heavy signal-overlap. To initiate resonance assignment in these large systems, we describe how selectively-labeled samples can be obtained with the addition of labelled or unlabelled amino acids in the medium. The reduced spectral overlap enabled us to identify typical spectral fingerprints and initiate sequential assignment using the more sensitive 2D DARR experiments with long mixing time showing inter-residue correlations.

Introduction

Selective labelling strategies have been crucial in NMR studies. With the advent of higher magnetic fields, solution-state NMR studies have concentrated on methyl-group selective labelling in deuterated proteins, which makes it possible to push the size limit of proteins to be investigated in order to observe only the most favourably relaxing component in TROESY spectroscopy (Pervushin *et al.* 1997). In solid-state NMR, relaxation is less an issue than spectral overlap, and notably glycerol-2 and 1-3 labelling schemes have been introduced to reduce spectral overlap (Higman *et al.* 2009) by reducing the number of NMR-active spins in the sample. These approaches have been shown to be precious for measuring distances used in structure calculations (Castellani *et al.* 2002b; Tuttle *et al.* 2016).

Another way of reducing the sheer number of residues and reducing spectral overlap observed in the spectra is amino-acid selective labelling (Nuzzio *et al.* 2016). In combination with high-field NMR, it allows size limitations to be pushed forward and opens the way to the study of large and complex proteins, for example the ABC transporter discussed here. The success of amino-acid selective labelling of proteins expressed in bacteria is intimately linked to the detailed knowledge of the amino-acid metabolism of the bacteria (**Figure 22**). Before going into a detailed description of the individual pathways for each amino acid, it is important to point out that specific (un)labelling needs to be done in minimal medium (Studier 2005), with ^{13}C labelled precursors, for example glycerol or glucose, and ammonium chloride as sole source of carbon and nitrogen. Labelled or unlabelled amino acids can then be added in the medium.

In the metabolic pathways, some amino acids are end-products of their pathway, enabling selective labelling, and others are located on the same track, and cannot be labelled independently. **Figure 22** gives an overview on the different synthesis (anabolic) and degradation (catabolic) pathways important in *E. coli* amino-acid metabolism, represented in the style of interconnecting subway lines. For example, Line A represents the glycolysis from glucose to two central metabolic products, fructose-6-phosphate, and pyruvate. From pyruvate, alanine is produced directly, and valine/leucine from a common precursor, 2-oxoisovalerate. A full description of the transport system can be found in the **Figure 22** caption, as well as in the following subsections which describe how every amino acids, or groups thereof, can efficiently be labelled in the context of these metabolic connections.

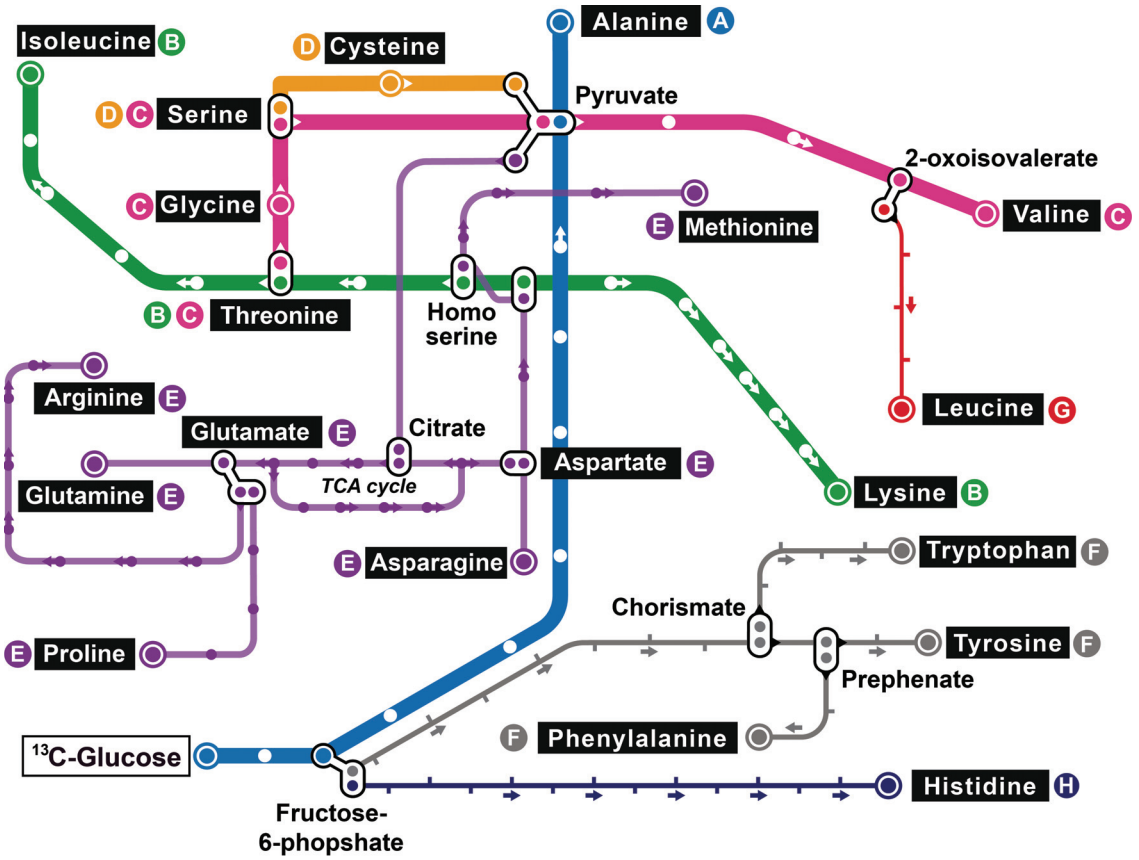


Figure 22. General overview of amino-acid biosynthesis from *E. coli* metabolic pathways. The simplified metabolic pathways shown refer only to the amino-acid biosynthesis in *E. coli*. The overview was established considering the known *E. coli* metabolism, as for example reported in the databases KEGG (Kanehisa *et al.* 2017) and Ecocyc (Keseler *et al.* 2016) and represents a condensed version of the information found at http://www.genome.jp/kegg/ and https://ecocyc.org. Line A (blue line), for example, represents the glycolysis which leads to the formation of pyruvate, which is a precursor of alanine. Line B (green line) shows the catabolism of aspartate resulting in lysine, threonine and isoleucine. Line C (pink line) displays the catabolism of threonine which products are converted into glycine and serine. Serine can in turn be catabolised to yield cysteine, line D (orange line), and used, *via* the pyruvate station, for the biosynthesis of valine and alanine. Line E (purple line) shows the products formed after the TCA cycle: aspartate and glutamate, which can be catabolised leading to arginine, proline and glutamine. Line F (gray line) represents the aromatic biosynthesis pathway from fructose-6-phosphate. Line G (red line) shows the anabolism of leucine from a common precursor with valine, 2-oxoisovalerate. Finally, line H (dark blue) represents the anabolism of histidine from fructose-6-phosphate. Each metabolic step before the final amino acid is represented by a station, non-reversible reactions are indicated by arrows. Interconnections between lines are represented and the names of the interconnecting compounds are indicated. Amino acids are highlighted against black background.

Amino acids selective labelling

1. Isoleucine, leucine and valine

The catabolic pathways connecting the different amino acids with aliphatic side chains are absent in *E. coli*, except for the alpha amino group (Kazakov *et al.* 2009). Consequently, the three branched amino acids isoleucine, leucine, and valine (respectively represented at the end of the line B, G and C in **Figure 22**) can be selectively ^{13}C labelling without problem (Etzkorn *et al.* 2007). However, this does not hold true for nitrogen labelling due to the action of transaminases. Indeed, the branched-chain-amino-acid transaminase catalyses the reversible reaction of transfer of the alpha amino group between the branched amino acid and α-ketoglutarate to generate glutamate and the precursors of the branched amino acids (Muchmore *et al.* 1989; Rudman & Meister 1953; Shortle 1994a; Waugh 1996). Moreover, the valine-pyruvate transaminase, which catalyses a reversible transamination between L-alanine and valine, generates significant scrambling on alanine. Consequently, a supplementation of 100 mg/L (Shortle 1994b) or 1 g/L (Bellstedt *et al.* 2013) of one of these amino acids in the minimal medium induces significant scrambling on the alpha amino groups of the other branched amino acids, as well as on alanine (strongly when valine is supplemented).

Isoleucine, valine and leucine are all targets for methyl labelling. Designing methyl-labelling protocols in perdeuterated proteins requires thorough understanding of the *E. coli* metabolism. Actually, metabolic precursors have to be used to perform the incorporation of $^{13}CH_3$ into branched amino acids in order to achieve minimum scrambling. Their type, necessary quantity and potential incompatibilities have recently been reviewed in the literature (Kerfah *et al.* 2015).

2. Alanine

Alanine is the end product of the line A that runs along the way of glycolysis (**Figure 22**). In order to label alanine, the addition of 600 mg/L to the minimal medium enables a sufficient incorporation for both ^{15}N (Muchmore *et al.* 1989) and ^{13}C (Ayala *et al.* 2009). However, selective ^{13}C labelling or ^{12}C reverse labelling remains difficult, due to the presence of alanine transaminases that transfer an amine group from glutamate (as amine donor) to

pyruvate generating alanine without the desired labelling scheme. The reverse reaction (alanine to pyruvate) is also mediated by transaminases using a 2-oxoglutarate as an amine acceptor. Thus, this enzyme causes the presence of undesired ^{13}C signals from alanine in the case of incorporation of ^{12}C-Alanine (0.5 g/L) into the culture medium (Kaur *et al.* 2015).

It is worth noting that the addition of 1-[^{13}C]-alanine (labeled carbonyl) allows to obtain clean carbonyl labelling as shown by Takeuchi *et al.* (Takeuchi *et al.* 2007). Indeed, alanine is converted into pyruvate by an alanine transaminase (see above) and pyruvate is converted into acetyl-CoA by the pyruvate dehydrogenase. The ^{13}C carbonyl group of alanine is transformed into $^{13}CO_2$. Consequently, 1-[^{13}C]-alanine is directly incorporated into the protein, and no other amino acids are labelled.

The specific labelling of 3-[^{13}C]-alanine, 2-3-[$^{13}C_2$]-alanine, and U-[^{13}C]-alanine is made possible by adding specific precursors as shown by Ayala *et al*. (Ayala *et al.* 2009). For this purpose, the addition of succinate (2.5 g/L), isoleucine (60 mg/L) and α-ketoisovalerate (200 mg/L) allows for a reduction of scrambling to a level lower than 5%. This protocol was developed in order to obtain a complete incorporation of methyl-protonated alanine into a perdeuterated protein using a supplement of 2-[^{2}H]-3-[^{13}C]-Alanine in combination with the precursors succinate-d_4, isoleucine-d_{10} and α-ketoisovalerate-d_7.

On the nitrogen side, alanine ^{15}N specific labelling or reverse labelling is rather straightforward without scrambling; still, excessive amounts of alanine in the medium must be avoided ((1 g/L) (Bellstedt *et al.* 2013)), since they could induce scrambling with valines. Indeed, the presence of the valine-pyruvate transaminase is the main source of scrambling in the case of nitrogen labelling or unlabelling of alanine, as the enzyme catalyses the reversible transamination between alanine and valine. Consequently, the intensity of the valine correlation peaks can only be decreased by 50% when the medium is supplemented with ^{14}N-alanine (1g/L) (Bellstedt *et al.* 2013).

3. Arginine, lysine and histidine

These three amino acids are end products of their respective metabolic pathways represented respectively on the lines E, B and H (**Figure 22**). Moreover, as they are end products without possible reverse reactions, the specific or reverse labelling of these residues is simple, and scrambling is of no concern when they are added together or separately to the culture medium. Hiroaki and co-workers have shown that a

supplementation of 100 mg/L of arginine or lysine is enough for a correct incorporation of these residues (95 % and 82 % respectively) in the case of nitrogen (un)labelling (Hiroaki *et al.* 2011). Bellstedt and co-workers have shown that there is no scrambling both in carbon and nitrogen atoms for these three residues even when they are added in a large excess (1 g/L) to the medium (Bellstedt *et al.* 2013). However, it was shown that arginine supplementation of 175 mg/L leads to about 50% of ^{13}C unlabelling (Shi *et al.* 2009a). Finally, a ^{12}C-arginine or ^{12}C-lysine supplementation of 400 mg/L is sufficient to remove the ^{13}C resonances (Wiegand *et al.* 2017). Arginine or lysine supplementation of respectively 380 mg/L or 160 mg/L is sufficient for $^{15}N/^{14}N$ (un)labelling (Muchmore *et al.* 1989; Nishida *et al.* 2006).

Histidine has the advantage that bacteria do not degrade it. Consequently, adding histidine to the medium causes no scrambling of ^{15}N and ^{13}C. A quantity of 400 mg/L (Wiegand *et al.* 2017) or 100 mg/L (Nishida *et al.* 2006) is sufficient for reverse labelling and selective labelling (1-^{13}C-His) respectively.

4. Aspartate / asparagine and glutamate / glutamine

These four amino acids are located together on line E (**Figure 22**). Aspartate and glutamate are central players in the bacterial metabolism. Aspartate is the precursor of asparagine, isoleucine, methionine, lysine and threonine, while glutamate is the precursor of arginine, glutamine and proline. Glutamate is also the major nitrogen donor in protein bio-synthesis (Goux *et al.* 1995). As a consequence, the extensive interconnections and proximal locations of aspartate, glutamate, asparagine and glutamine on line E in **Figure 22** result in the fact that these residues remain a great challenge for a clean labelling strategy.

Still, Tate *et al.* have reported a strategy for selective labelling of the side-chain NH groups for asparagine and glutamine simultaneously (Tate *et al.* 1992) and Cao *et al.* have recently shown a strategy that is able to label selectively the side-chain NH group of asparagine residues, and also the side-chain NH groups of glutamine, asparagine and tryptophan simultaneously (Cao *et al.* 2014). We briefly describe these strategies in the following. The selective labelling of side-chain NH group of asparagine has to avoid at least three metabolic effects: (i) the transfer reaction of the side chain from glutamine to asparagine, and the reverse-reaction catalysed by asparagine synthetase B; (ii) the degradation of glutamine in glutamate and ammonium catalysed by glutaminase; and (iii) the

incorporation of ammonium into asparagine catalysed by aspartate-ammonia ligase when ammonium is abundant. Cao and co-workers obtained, with ubiquitin, a ^{15}N-labelling efficiency of the asparagine side-chain of 98% without significant scrambling by supplementing the minimal medium with 0.55 g/L $^{15}NH_4Cl$, 2 g/L ^{14}N-glutamate, 4 g/L ^{14}N-glutamine, 4g/L ^{14}N-aspartate and all other ^{14}N-amino acids (except asparagine) 0.5g/L. Moreover, they equally obtained, with ubiquitin, a ^{15}N-labelling efficiency of the asparagine and glutamine side-chain of respectively 84% and 85% without significant scrambling. The approach consists in supplementing the minimal medium with 0.27 g/L $^{15}NH_4Cl$, 2 g/L ^{14}N-glutamate, 4 g/L ^{14}N-asparagine and all other ^{14}N-amino acids (except aspartate and glutamine) 0.5 g/L. These authors have noticed that, surprisingly, despite the addition of ^{14}N-asparagine in large excess, its ^{15}N-labelling remains quite efficient. They have suggested that the ^{15}N-labelling yield of the side-chain depends mainly on the intracellular concentration of aspartate and $^{15}NH_4Cl$.

5. Methionine

Methionine is located at the end of its metabolic pathway, it is the last station on line E (**Figure 22**). Its catabolism by *E. coli* seems to be possible for nitrogen but this catabolic pathway remains unclear (Reitzer 2005). However, no scrambling was observed in ^{13}C and in ^{15}N when 1 g/L of labelled methionine was added, and when natural abundance methionine was added to the medium, the signal of methionine was strongly attenuated but not supressed (70% attenuation) (Bellstedt *et al.* 2013). The methyl group of methionine can be used as an NMR probe in methyl labelling (Plevin & Boisbouvier 2012; Stoffregen *et al.* 2012) by adding 250 mg/L of [ε-^{13}C]-labelled methionine in the culture medium (Gelis *et al.* 2007; Stoffregen *et al.* 2012).

6. Proline

Proline is represented at the end of the line E (**Figure 22**). Proline can be used as sole source of nitrogen and carbon for *E. coli* (Reitzer 2005). It can be catabolized to yield glutamate by the combination of two enzymes: proline dehydrogenase and 1-pyrroline-5-carboxylate dehydrogenase (Zhou *et al.* 2008). Consequently, the addition of natural-abundance proline in the culture medium induces general carbon scrambling, with amino acids located along line E (Gln, Glu and Arg) the most severely affected, especially when proline concentration

is 250 mg/L (Rasia *et al.* 2012)or 1 g/L (Bellstedt *et al.* 2013). This general scrambling impedes a labelled-proline addition for specific labelling due to the apparitions of resonances from Gln, Glu and Arg. On the other hand, the addition of natural abundance proline can be used for reverse labelling. Indeed, all resonances from proline will disappear, and the scrambling will induce a small signal loss of the other resonances, but will still allow 2D spectra recording (Kaur *et al.* 2015; Wiegand *et al.* 2017).

7. Phenylalanine, tyrosine and tryptophan

These aromatic residues are quite simple to label or unlabel specifically in ^{13}C. They use the end product of the pentose-phosphate pathway, erythrose-4-phosphate, a precursor of all aromatic residues. Erythrose-4-phosphate allows the fructose-6-phosphate to enter line F (**Figure 22**). As a matter of fact, there is no degradation pathway for phenylalanine and tyrosine in *E. coli*. Regarding tryptophan, *E. coli* can utilize tryptophan as the sole source of carbon. Tryptophan is degraded into indole, pyruvate and ammonium by tryptophanase. Despite that pyruvate is found at the crossroad connecting several amino acids, the labelled or natural abundance pyruvate from tryptophan will be substantially diluted by the glucose contribution. As a consequence, these three amino acids can be labelled or unlabelled together or separately. Thus, a supplementation of 150-200 mg/L (Etzkorn *et al.* 2007; Shi *et al.* 2009a; b) up to 1 g/L (Bellstedt *et al.* 2013) in the minimal medium is sufficient to perform the labelling of one of these of residues.

On the nitrogen side, phenylalanine and tyrosine are linked together by two aminotransferases: the aromatic aminotransferase and the aspartate aminotransferase. These two enzymes catalyse the reversible reaction of transamination of phenylpyruvate and 4-hydroxy-phenylpyruvate with glutamate as amino donor or 2-oxoglutarate as amino receptor (for the reverse reaction). As a result, there is non-negligible scrambling in nitrogen between these two amino acids.

Furthermore, a third enzyme, the branched-chain amino-acid transferase, can operate the same reversible reaction on phenylpyruvate or phenylalanine. The contribution of this last enzyme explains why Bellstedt and co-workers observed a diminution (about 20%) of the peak intensities from Ile, Val and Leu on a [^{1}H, ^{15}N]-HSQC when natural-abundance phenylalanine was added in the culture medium. For 3D experiments comprising a ^{15}N dimension, precursors of tyrosine or phenylalanine, respectively 4-hydroxy-phenylpyruvate

and phenylpyruvate, can be used. In point of fact, both precursors are located in the metabolism one step before transamination. Aminotransferases use the ^{15}N from the medium, which leads to ^{15}N labelling of tyrosine and phenylalanine without scrambling (Rasia *et al.* 2012).

When degrading tryptophan, tryptophanase catalyses the production of indole, pyruvate and ammonium. The latter can directly be used for the synthesis of amino acids. The action of this enzyme generates scrambling which leads to a general loss of the ^{15}N-backbone signal, when natural-abundance tryptophan is added (general loss estimated to 30% by Bellstedt and co-workers (Bellstedt *et al.* 2013)). To overcome this problem, solutions using precursors such as indole, or a mixture of indole and (un)labelled tryptophane can be used in order to reduce scrambling (Berger *et al.* 2013; Rodriguez-Mias & Pellecchia 2003; Schörghuber *et al.* 2015).

8. Glycine, serine, cysteine and threonine

Glycine and serine, neighboring stations on line C (**Figure 22**), are linked by glycine-hydroxymethyltransferase-mediated interconversion. Threonine is represented at the end of line C (**Figure 22**). Indeed, glycine and threonine are also linked by interconversion, which is, in this case, mediated by a threonine aldolase. Serine is the substrate of a deamination reaction, catalyzed by a serine dehydratase which converts the amino acid into pyruvate. Besides, this amino acid is involved in cysteine biosynthesis, as well as in tryptophan biosynthesis, where a tryptophan synthase combines a serine and an indole to form tryptophan.

In other words, (i) the addition of glycine induces scrambling with cysteine, serine and tryptophan for ^{15}N and ^{13}C; (ii) serine dilutes the labelling of all other amino acids; and (iii) threonine scrambles with cysteine, glycine, serine, tryptophan ^{15}N and ^{13}C, but only ^{13}C for isoleucine. Indeed, threonine is a precursor of the isoleucine biosynthesis pathway. Too complex to be fully represented in Figure 1, the links between these residues are presented in detail in **Figure 23**.

The addition of glycine and 2-ketobutyric acid (an isoleucine precursor), both in natural abundance, ensures an incorporation of labelled threonine without scrambling. Indeed, as labelled threonine is diverted to isoleucine and glycine (**Figure 23**), the addition of glycine and 2-ketobutyric acid in natural abundance drastically reduces the scrambling effect from

threonine. This solution has been presented for methyl labelling of threonine (Thr-γ_2[$^{13}CH_3$]) and isoleucine (Ile-d_1[$^{13}CH_3$]) (Velyvis *et al.* 2012).

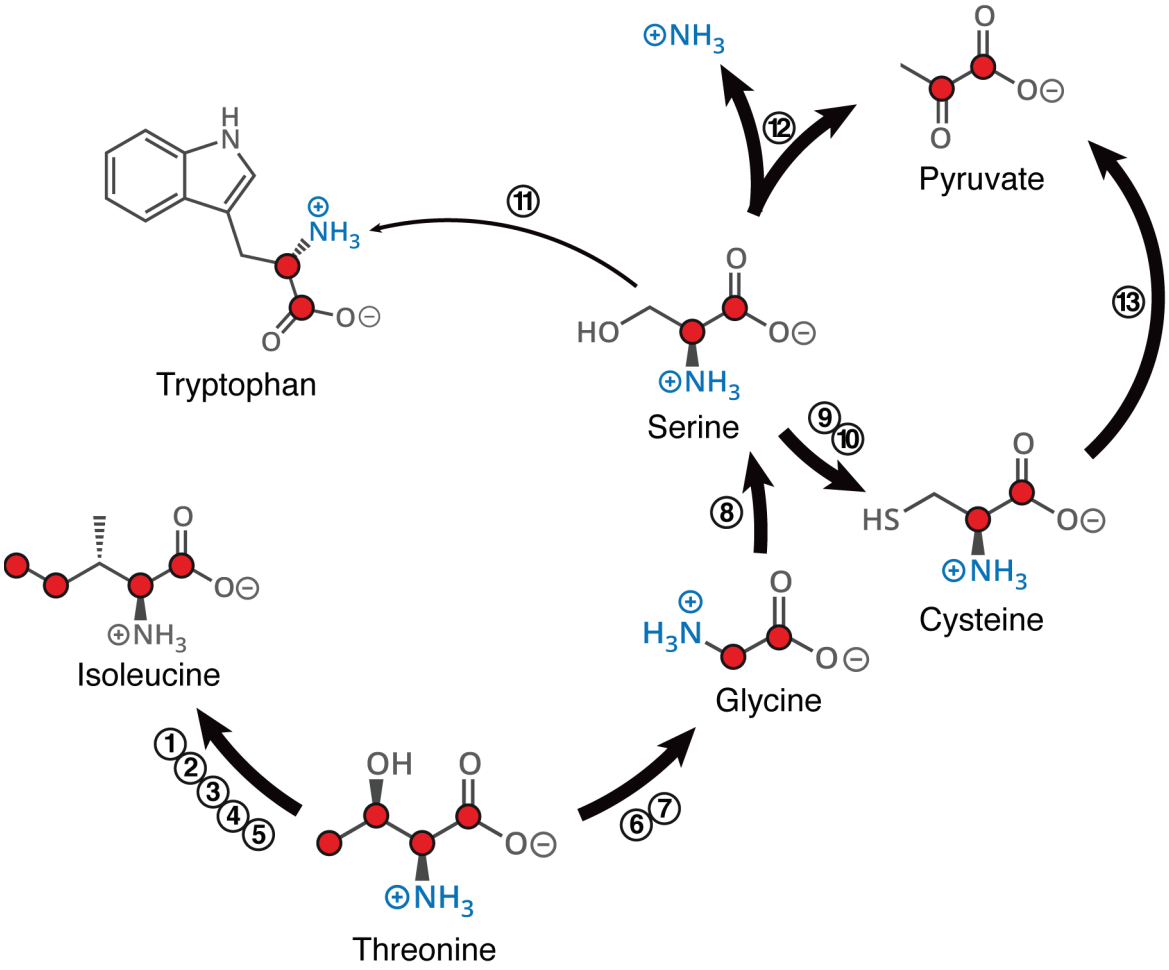


Figure 23. Metabolic connexions of threonine. Carbons from threonine are represented by red dots, whereas amines are coloured in blue. Threonine is the precursor of isoleucine in the biosynthesis pathway. Threonine deaminase (1) leads to the formation of 2-oxobutanoate and ammonia, then acetohydroxy acid synthase (2) catalizes a reaction of decarboxilyation of pyruvate and converts 2-oxobutanoate into 2-aceto-2-hydroxybutanoate. This product is reduced by an oxidoreductase (3) and dehydrated by an acid dehydratase (4) forming the next-to-last compound, 3-methyl-2-oxopentanoate. Finally, this penultimate compound forms isoleucine by a reaction of transamination catalyzed by branched-chain amino-acid aminotransferases (5). Meanwhile, the catabolism of threonine to yield glycine involves two complementary enzymes. First the threonine dehydrogenase (6) which catalyses the reaction of oxidation of threonine into 2-amino-3-oxobutanoate, then the 2-amino-3-ketobutyrate CoA ligase (7) which catalyses the reaction of condensation of 2-amino-3-ketobutyrate and coenzyme-A to form acetyl-CoA and glycine. Serine is a product of glycine catabolism, the reaction is catalysed by a serine hydroxymethyltransferase (8). Serine catabolism forms pyruvate and ammonium via the serine deaminase (12), and serine is associated with the last stage of tryptophan anabolism catalysed by a tryptophan synthase (11). The residue is also used for cysteine biosynthesis mediated by serine acetyltransferase

(9) and cysteine synthase (10). Cysteine is degraded into pyruvate, ammonia and sulfide, this reaction is catalysed by the action of cysteine desulfhydrases (13).

Figure 24 summarizes the above discussion on each amino acid, reported on the ordinate, and the scrambling induced by each other amino acid when added to the minimal medium on abscissa.

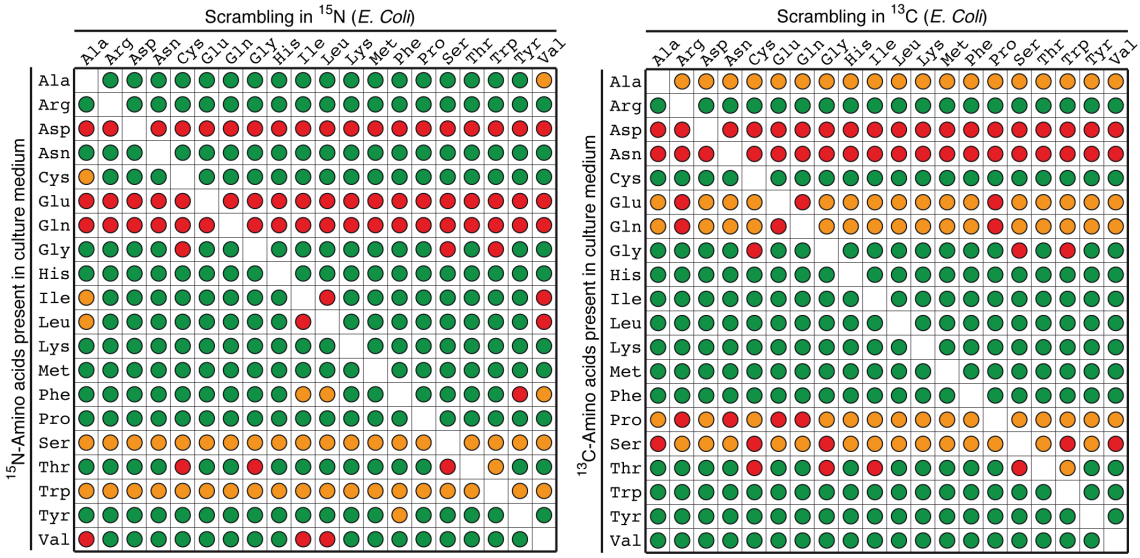


Figure 24. Influence of an added amino acid on the (un)labelling of others with respect to ^{15}N and ^{13}C labelling. Combinations without scrambling are indicated with green dots; weak scrambling, i.e. suitable for reverse labelling and specific labelling, is indicated with orange dots. Substantial scrambling is indicated with red dots.

Glycerol-2 and 1-3 labelling coupled with reverse labelling

Labelling strategies involving glycerol are well-known and have been used in the context of structure determination (Castellani *et al.* 2002a; Loquet *et al.* 2013; Tang *et al.* 2011; Tuttle *et al.* 2016; Van Melckebeke *et al.* 2010). A strategy described by Hong and Jakes in 1999 combines glycerol labelling with amino-acid reverse labelling (Hong & Jakes 1999). In the context of a membrane protein study, [2-^{13}C]glycerol was added in the culture medium containing the (not labelled) amino acids produced from the tricarboxylic acid cycle: Asp, Asn, Arg, Gln, Glu, Ile, Lys, Met, Pro and Thr. That labelling scheme was called TEASE for *ten-amino-acid selective and extensive labelling* and resulted in the labelling of amino acids produced by the glycolysis (Gly, Ala, Ser, Cys, Val and Leu) and pentose phosphate pathways (His, Phe, Trp and Tyr). Moreover, the labelling of these amino-acids produced the scheme obtained with [2-^{13}C]glycerol. However, the addition of not labelled Gln and Glu brings side-effects, as discussed previously, and induces a general dilution of the ^{15}N labelling by 50%.

The method can be extended by using [2-^{13}C]glycerol and [1,3-^{13}C]glycerol with different combinations of amino-acids. For the protein OmpG (300 residues), a combination of [1,3-^{13}C]glycerol, [2-^{13}C]glycerol, and labelled TEMPQANDSG-OmpG produced in the presence of [2-^{13}C]glycerol and [1,3-^{13}C]glycerol and labelled SHLYGWAVF-OmpG produced in the presence of [2-^{13}C]glycerol allowed the collection of distance restraints (Retel 2016). To obtain ^{13}C-^{15}N-TEMPQANDSG-OmpG and ^{13}C-^{15}N-SHLYGWAVF-OmpG, ^{15}N-RHILKMFWYV and ^{15}N-TARDNEDQIKMP (produced by tricarboxylic acid cycle) were added to the medium (50 mg/L) in order to remove the resonances of these residues from the ^{13}C-spectrum. The authors noticed the presence of scrambling in ^{13}C-^{15}N-SHLYGWAVF-OmpG, where ^{15}N-TARDNEDQIKMP was added, but peaks from QENDT were still observed in the spectrum. The amino-acid concentration in the medium may have been insufficient to fully inhibit the metabolic pathways of the remaining residues. In fact, an average concentration higher than 100-200 mg/L is necessary to achieve optimum labelling. For amino-acid concentrations under 100 mg/L in the culture medium, efficient incorporation of the desired amino acids is not guaranteed.

Application to the ABC transporter BmrA

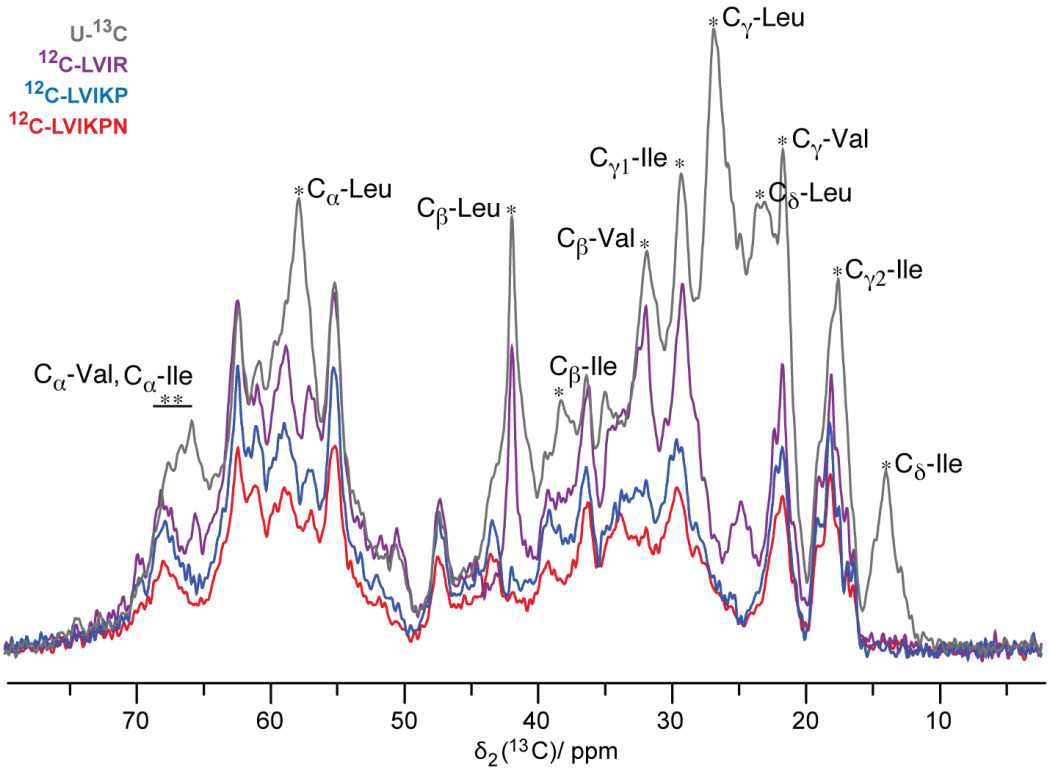


Figure 25. Reverse labelling on the 60 kDa membrane protein BmrA. Aliphatic regions of 1D spectra of uniformly labelled BmrA (grey shade), BmrA with natural abundance LVIR (purple line), BmrA with natural abundance LVIKP (blue line) and BmrA with natural abundance LVIKPN (red shade). Sample preparation and acquisition parameters were described previously in the literature (Kunert *et al.* 2014; Wiegand *et al.* 2017). Briefly, the production of reverse labelled LVIKP-, LVIR- and LVIKPN- BmrA was performed using natural abundance Leu (0.23 g/L), Val (0.23 g/L), Ile (0.23 g/L), Lys (0.40 g/L), Pro (0.10 g/L), Arg (0.40 g/L) Gln (0.40 g/L).

The ABC transporter BmrA is a dimeric membrane protein of 60 kDa per monomer. It was studied in lipid environment by solid-state NMR using a lipid-to-protein ratio of 0.5 (Kunert *et al.* 2014). Even though there is no inherent size limit in solid-state NMR, large membrane proteins are still a challenge due to low signal-to-noise ratios and signal overlap. The low signal-to-noise ratio makes the recording of 3D spectra difficult, and complicates sequential assignments of the protein. To alleviate overlap in these large systems, we also used paramagnetic relaxation enhancements in complement of selective labelling (Wiegand *et al.* 2017). Alanine, glycine and branched residues (L, V, I) are over-represented in membrane proteins due to their hydrophobic proprieties (Ulmschneider & Sansom 2001). In

the case of BmrA, they represent 45% of the protein (265 residues). In addition to this ^{13}C resonances from these residues are spread all over the aliphatic region, leading to extensive overlap with other residues. For this reason, they are targets of choice for reverse labelling and allow simplification of the spectrum. Arginine and lysine share the same ^{13}C chemical-shift regions for their side-chain resonances, and specifical labelling of only one of them allows their assignment. The branched side-chain residues, arginine and lysine do not lead to scrambling in the biosynthesis of the other amino acids, and thus allow for clean unlabelling. **Figure 25** compares, in black and purple, the 1D spectra of fully labelled and LVIR-unlabelled BmrA, and one can see that the signal in regions which should remain unaffected, for example glycine around 48 ppm, remains conserved.

The 2D ^{13}C-^{13}C spectrum of BmrA without LVIR is presented in purple shade in **Figure** 26 It is substantially less crowded than the spectrum recorded on uniformly labelled BmrA, and it displays many isolated peaks. With a view on further simplification of the spectrum, proline can be removed in order to clean up some regions in the spectrum to obtain several unique threonine, serine, glycine and alanine frequencies. This operation helps to identify certain amino acids and their neighbours in 2D spectra using longer mixing times (see below). However, as mentioned previously, adding natural abundance proline to the culture medium induces isotope dilution which affects the signal intensity of the other residues. Even though natural abundance lysine was added instead of arginine, a 25 % signal loss - due to natural abundance proline - of the other residues can be observed (**Figure 25**).

Despite the signal loss, the spectrum of BmrA with LVIKP in natural abundance (blue shade, **Figure 26**) shows sufficient S/N and, in combination with the spectrum of LVIR-BmrA, allows the assignment and the characterization of several residues (Wiegand *et al.* 2017). Natural abundance asparagine was also added to the medium together with LVIKP in order to discriminate between aspartate and asparagine. However, as asparagine is converted into aspartate, and the latter is located at a metabolic crossroad, the result is substantial isotopic dilution and as a consequence NMR signal reduction of about 45% for natural abundance LVIKPN-BmrA, as seen in the 1D spectrum in **Figure 4**.

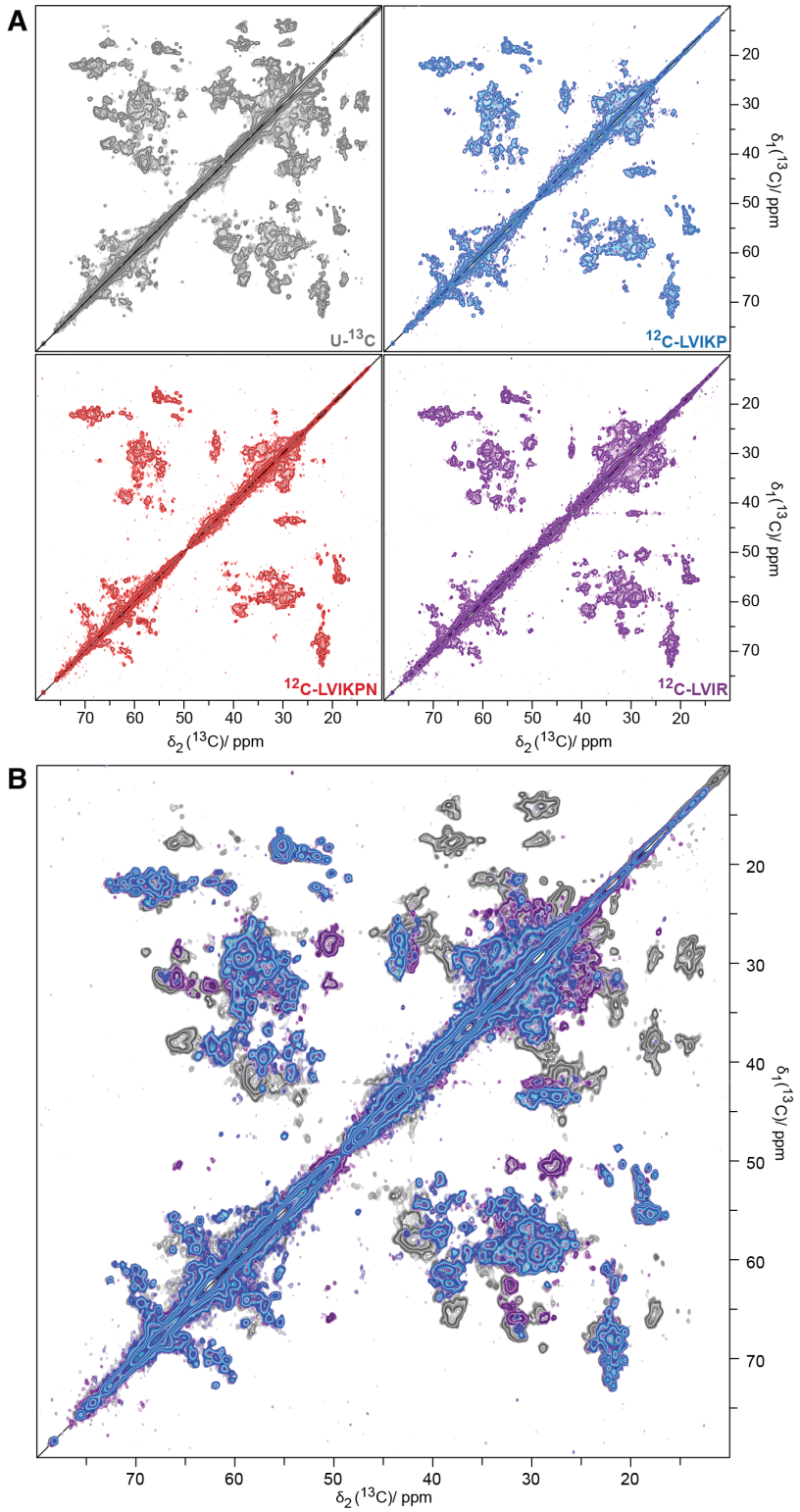


Figure 26. 2D spectra of reverse labelled BmrA. A. Aliphatic regions of 2D ^{13}C-^{13}C DARR spectra of uniformly labelled BmrA (grey shade), BmrA with natural abundance LVIR (purple shade), BmrA with natural abundance LVIKP (blue shade) and BmrA with natural abundance LVIKPN (red shade). **B.** Overlay of three spectra shown in **A** with uniformly labelled BmrA, BmrA with natural abundance LVIR (purple shade) and BmrA with natural abundance LVIKP (blue shade).

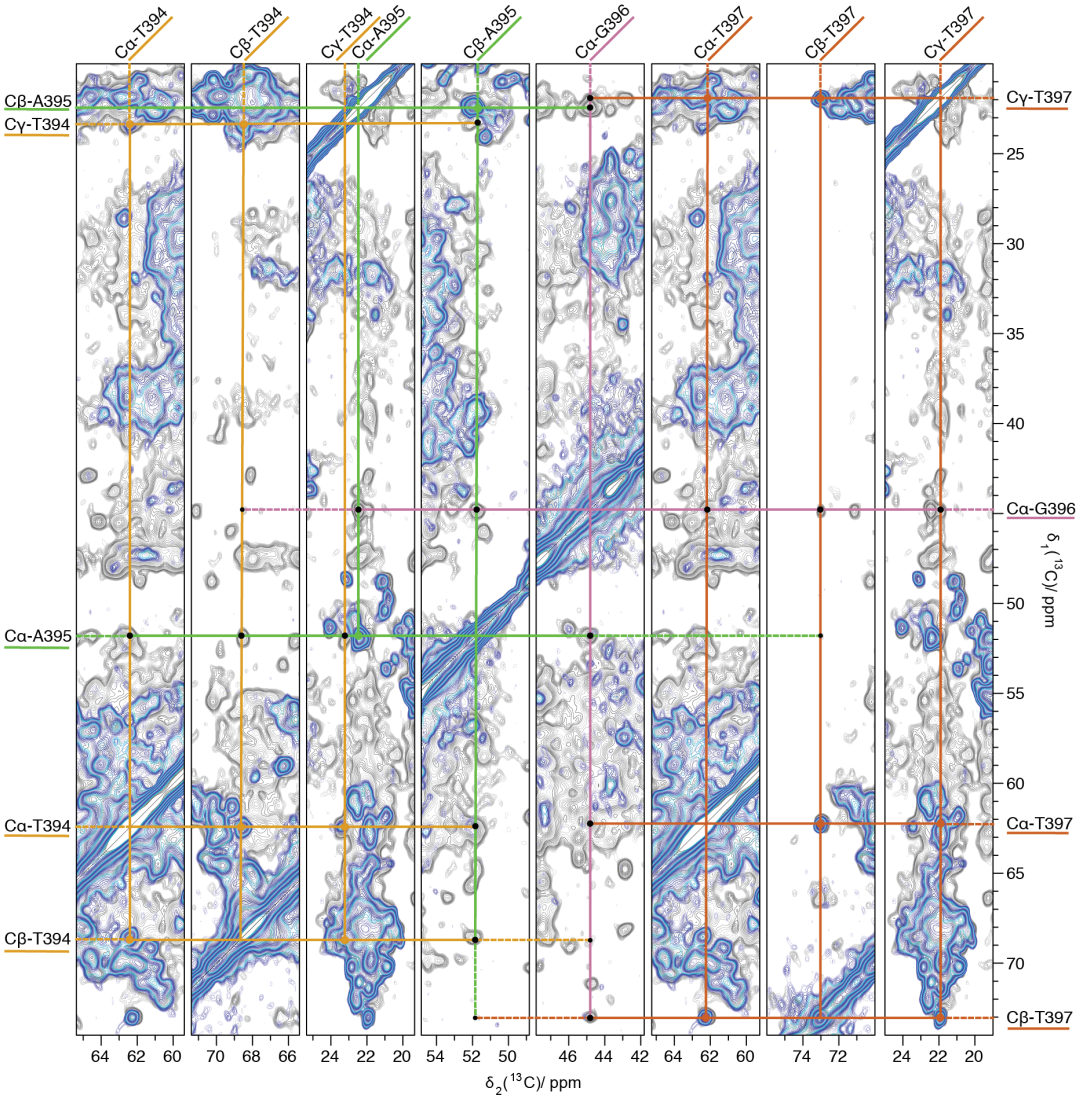


Figure 27. Example for a sequentially assigned stretch including residues T394-A395-G396-T397, using 2D ^{13}C-^{13}C DARRs recorded with 20 ms (blue shades) and a 200 ms (grey shades) mixing time on a reverse labeled ^{12}C-^{14}N-[LVIKHP]-^{13}C-^{15}N BmrA. Inter-residue correlation peaks are represented by black dots for i+1 and smaller black dots for i+2. Broken lines indicate i+2 connections. Assignments were possible for amino acids showing isolated signals, which allowed to unambiguously identify their type. Most of the inter-residual peaks i±1 and i±2 are only isolated when using selective labelling. The residue combination TAGT is present only once in the BmrA sequence, which allows to assign sequence numbers to this stretch. For residue combinations which are present in several segments of the protein, secondary structure information obtained from a homology model in combination with secondary-structure typical chemical shifts allows to reduce assignment possibilities.

The reduction of the number of resonances by unlabelling allowed for the recording of 2D ^{13}C spectra with a long mixing time displaying sequential correlations (**Figure 27**). These can be used, for isolated frequencies, to identify neighbouring residues and in some cases partial

sequential assignments. For BmrA, we used DARR spectra recorded with mixing time of 200 ms to obtain information about the i±1 neighbours and even i±2. This information enabled us to determine several sequential or amino-acid type assignments. A similar experiment recorded on fully-labelled samples yielded an overcrowded spectrum. Removal of Pro, Ile and Val cleans up the spectrum in the serine and threonine regions, whereas removing Leu, Ile and Pro gives a privileged access to the alanine and glycine regions. **Figure 6** shows an example of the sequential assignment of a 4 residue amino-acid stretch.

Conclusion

We compiled and summarized information from the literature in order to highlight the strength of amino acids selective labelling or unlabelling in the context of solid-state NMR studies. Indeed, selective isotope labelling is central for the study of large proteins, and shall contribute to the further development of NMR of biological system by pushing the size limits imposed by signal overlap. We demonstrated the use of amino-acid selective labelling for the ABC transporter BmrA, a protein composed by a large number of hydrophobic residues, which unlabelling leads to decongestion of spectra and allows to initiate sequence-specific analyses.

References

Ayala, I., Sounier, R., Usé, N., Gans, P. & Boisbouvier, J. (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *Journal of Biomolecular NMR* 43, 111–119.

Bellstedt, P., Seiboth, T., Häfner, S., Kutscha, H., Ramachandran, R. & Görlach, M. (2013) Resonance assignment for a particularly challenging protein based on systematic unlabelling of amino acids to complement incomplete NMR data sets. *Journal of Biomolecular NMR* 57, 65–72.

Berger, C., Berndt, S. & Pichert, A. (2013) Efficient isotopic tryptophan labelling of membrane proteins by an indole controlled process conduct. *Biotechnol Bioeng. 110, 1681-90.*

Cao, C., Chen, J.-L., Yang, Y., Huang, F., Otting, G. & Su, X.-C. (2014) Selective (15)N-labelling of the side-chain amide groups of asparagine and glutamine for applications in paramagnetic NMR spectroscopy. *Journal of Biomolecular NMR* 59, 251–261.

Castellani, F., van Rossum, B., Diehl, A. & Schubert, M. (2002a) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*.

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. & Oschkinat, H. (2002b) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420, 98–102.

Etzkorn, M., Martell, S., Andronesi, O.C., Seidel, K., Engelhard, M. & Baldus, M. (2007) Secondary Structure, Dynamics, and Topology of a Seven-Helice Receptor in Native Membranes, Studied by Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* 46, 459–462.

Gelis, I., Bonvin, A.M.J.J., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A. & Kalodimos, C.G. (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131, 756–769.

Goux, W.J., Strong, A.A., Schneider, B.L., Lee, W.N. & Reitzer, L.J. (1995) Utilization of aspartate as a nitrogen source in Escherichia coli. Analysis of nitrogen flow and characterization of the products of aspartate catabolism. *Journal of Biological Chemistry* 270, 638–646.

Higman, V.A., Flinders, J., Hiller, M., Jehle, S., Markovic, S., Fiedler, S., van Rossum, B.-J. & Oschkinat, H. (2009) Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively 13C-labelled proteins. *Journal of Biomolecular NMR* 44, 245–260.

Hiroaki, H., Umetsu, Y., Nabeshima, Y.-I., Hoshi, M. & Kohda, D. (2011) A simplified recipe for assigning amide NMR signals using combinatorial 14N amino acid inverse-labelling. *Journal of structural and functional genomics* 12, 167–174.

Hong, M. & Jakes, K. (1999) Selective and extensive 13C labelling of a membrane protein for solid-state NMR investigations. *Journal of Biomolecular NMR* 14, 71–74.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* 45, D353–D361.

Kaur, H., Lakatos, A., Spadaccini, R., Vogel, R., Hoffmann, C., Becker-Baldus, J., Ouari, O., Tordo, P., Mchaourab, H. & Glaubitz, C. (2015) The ABC exporter MsbA probed by solid state NMR – challenges and opportunities. *Biological chemistry* 396, 1135–1149.

Kazakov, A.E., Rodionov, D.A., Alm, E., Arkin, A.P., Dubchak, I. & Gelfand, M.S. (2009) Comparative genomics of regulation of fatty acid and branched-chain amino acid utilization in proteobacteria. *Journal of Bacteriology* 191, 52–64.

Kerfah, R., Plevin, M.J., Sounier, R., Gans, P. & Boisbouvier, J. (2015) ScienceDirect Methyl-specific isotopic labelling: a molecular tool box for solution NMR studies of large proteins. *Current Opinion in Structural Biology* 32, 113–122.

Keseler, I.M., Mackie, A. & Santos-Zavaleta, A. (2016) The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. *Nucleic Acids Res. 45, 543-550.*

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Loquet, A., Habenstein, B., Chevelkov, V., Vasa, S.K., Giller, K., Becker, S. & Lange, A. (2013) Atomic structure and handedness of the building block of a biological assembly. *Journal of the American Chemical Society* 135, 19135–19138.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. & Dahlquist, F.W. (1989) Expression and nitrogen-15 labelling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods in enzymology* 177, 44–73.

Nishida, N., Motojima, F., Idota, M., Fujikawa, H., Yoshida, M., Shimada, I. & Kato, K. (2006) Probing dynamics and conformational change of the GroEL-GroES complex by 13C NMR spectroscopy. *Journal of biochemistry* 140, 591–598.

Nuzzio, K.M., Watt, E.D., Boettcher, J.M., Gajsiewicz, J.M., Morrissey, J.H. & Rienstra, C.M. (2016) High-Resolution NMR Studies of Human Tissue Factor. *PLoS ONE* 11, e0163206.

Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 94, 12366–12371.

Plevin, M.J. & Boisbouvier, J. (2012) Isotope-labelling of methyl groups for NMR studies of large proteins. *Recent Developments in Biomolecular NMR, Royal Society of Chemistry*

Rasia, R.M., Brutscher, B. & Plevin, M.J. (2012) Selective isotopic unlabelling of proteins using metabolic precursors: application to NMR assignment of intrinsically disordered proteins. *ChemBioChem* 13, 732–739.

Reitzer, L. (2005) Catabolism of Amino Acids and Related Compounds. *EcoSal Plus* 1.

Retel, J.S. (2016) Structure Determination of Outer Membrane Protein G in Native Lipids by Solid-State NMR Spectroscopy. *Thesis Freie Universität Berlin*

Rodriguez-Mias, R.A. & Pellecchia, M. (2003) Use of selective Trp side chain labelling to characterize protein-protein and protein-ligand interactions by NMR spectroscopy. *Journal of the American Chemical Society* 125, 2892–2893.

RUDMAN, D. & MEISTER, A. (1953) Transamination in Escherichia coli. *Journal of Biological Chemistry* 200, 591–604.

Schörghuber, J., Sára, T., Bisaccia, M., Schmid, W., Konrat, R. & Lichtenecker, R.J. (2015) Novel approaches in selective tryptophan isotope labelling by using Escherichia coli overexpression media. *ChemBioChem* 16, 746–751.

Shi, L., Ahmed, M.A.M., Zhang, W., Whited, G., Brown, L.S. & Ladizhansky, V. (2009a) Three-Dimensional Solid-State NMR Study of a Seven-Helical Integral Membrane Proton Pump—Structural Insights. *Journal of molecular biology* 386, 1078–1093.

Shi, L., Lake, E.M.R., Ahmed, M.A.M., Brown, L.S. & Ladizhansky, V. (2009b) Solid-state NMR study of proteorhodopsin in the lipid environment: secondary structure and dynamics. *Biochimica et biophysica acta* 1788, 2563–2574.

Shortle, D. (1994a) Assignment of amino acid type in 1H-15N correlation spectra by labelling with 14N-amino acids. *Journal of magnetic resonance. Series B* 105, 88–90.

Shortle, D. (1994b) Assignment of amino acid type in 1H-15N correlation spectra by labelling with 14N-amino acids. *Journal of Magnetic Resonance* 105, 88–90.

Stoffregen, M.C., Schwer, M.M., Renschler, F.A. & Wiesner, S. (2012) Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. *Structure* 20, 573–581.

Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41, 207–234.

Takeuchi, K., Ng, E., Malia, T.J. & Wagner, G. (2007) 1-13C amino acid selective labelling in a 2H15N background for NMR studies of large proteins. *Journal of Biomolecular NMR* 38, 89–98.

Tang, M., Sperling, L.J., Berthold, D.A., Schwieters, C.D., Nesbitt, A.E., Nieuwkoop, A.J., Gennis, R.B. & Rienstra, C.M. (2011) High-resolution membrane protein structure by joint calculations with solid-state NMR and X-ray experimental data. *Journal of Biomolecular NMR* 51, 227–233.

Tate, S., Tate, N.U., Ravera, M.W., Jaye, M. & Inagaki, F. (1992) A novel 15N-labelling method to selectively observe 15NH2 resonances of proteins in 1H-detected heteronuclear correlation spectroscopy. *FEBS Letters* 297, 39–42.

Tuttle, M.D., Comellas, G., Nieuwkoop, A.J., Covell, D.J., Berthold, D.A., Kloepper, K.D., Courtney, J.M., Kim, J.K., Barclay, A.M., Kendall, A., Wan, W., Stubbs, G., Schwieters, C.D., Lee, V.M.Y., George, J.M. & Rienstra, C.M. (2016) Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nature Structural & Molecular Biology* 23, 409–415.

Ulmschneider, M.B. & Sansom, M.S. (2001) Amino acid distributions in integral membrane protein structures. *Biochimica et biophysica acta* 1512, 1–14.

Van Melckebeke, H., Wasmer, C. & Lange, A. (2010) Atomic-resolution three-dimensional structure of HET-s (218– 289) amyloid fibrils by solid-state NMR spectroscopy. *J Am Chem Soc. 132, 13765-75.*

Velyvis, A., Ruschak, A.M. & Kay, L.E. (2012) An economical method for production of (2)H, (13)CH3-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PLoS ONE* 7, e43725.

Waugh, D.S. (1996) Genetic tools for selective labelling of proteins with α-15N-amino acids. *Journal of Biomolecular NMR* 8, 184–192.

Wiegand, T., Lacabanne, D. & Keller, K. (2017) Solid-state NMR and EPR Spectroscopy of Mn2+-Substituted ATP-Fueled Protein Engines. *Angew Chem Int Ed Engl. 13, 3369-3373.*

Zhou, Y., Zhu, W., Bellur, P.S., Rewinkel, D. & Becker, D.F. (2008) Direct linking of metabolism and gene expression in the proline utilization A protein from Escherichia coli. *Amino acids* 35, 711–718.

Chapter IV

Production, purification and reconstitution of the ABC transporter BmrA for solid-state NMR studies

The work presented in this chapter was done in collaboration with Beat H. Meier, Jean-Michel Jault and Pierre Falson. This chapter was adapted from the following reference:

Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies

Britta Kunert[†], Carole Gardiennet[†], Denis Lacabanne, Daniel Calles-Garcia, Pierre Falson, Jean-Michel Jault, Beat H. Meier François Penin and Anja Böckmann

[†]These authors have contributed equally to this work

Front Mol Biosci. 2014 Jun 12;1:5.

Contents

Abstract

The first steps of the structural investigations into the bacterial ATP-binding cassette transporter BmrA are reported in this chapter. These steps include protein overproduction, stable isotope labelling, reconstitution into a lipid environment and the first two-dimensional solid-state NMR spectrum measurement.

This section shows that the reconstitution of BmrA in lipids from *Bacillus subtilis* at a lipid-protein ratio of 0.5 m/m allows an optimal protein insertion into lipid bilayer, as well as it complies with the two central NMR requirements: high signal-to-noise in the spectra and sample stability over a time period of months.

The obtained spectra point to a well-folded protein and a highly homogenous preparation, as witnessed by the narrow resonance lines and the signal dispersion typical of the expected secondary structure distribution of the membrane protein. It shows the way towards studies of the different conformational states of the transporter in its export cycle, as well as to interactions with substrates via chemical-shift fingerprints and sequential resonance assignments.

Introduction

Membrane proteins are essential components of biological processes and constitute a significant fraction of the proteome of all living cells. The characterization of membrane proteins remains the most challenging target in structural biology. Despite number of studies on membrane proteins characterization, they are still particularly underrepresented especially in the Protein Data Bank in which 2.5% of the data bank entries represent membrane proteins.

Indeed, there is a lack of information upon this class of proteins both in terms of structures and function. Their hydrophobic surface, flexibility or instability makes membrane proteins difficult to study. Indeed, these difficulties lead to challenges at all levels, including expression, solubilisation, purification, crystallisation, data collection and structure solution.

The bacterial ABC exporter BmrA from *Bacillus subtilis* is used as a model. The protein production and efficient reconstitution in lipids of this transporter are reported here. This

chapter can be considered as a "*materials and methods*" section of the ABC transporter BmrA preparation of this thesis. This chapter was detailed and presents the different stages of the expression, purification, reconstitution of the ABC transporter BmrA. All these stages were strictly followed in order to get an optimal preparation in term of yield, functionality and high stability over time.

Expression of ^{13}C-^{15}N-BmrA in minimal-medium

1. Culture media

Transformation, expression test and pre-culture were made in LB medium (Lysogeny Broth, Sigma-Aldrich® L3522). Labelled sample preparation was operated in minimal medium.
The minimal M9-medium is composed of 38 mM anhydrous Na_2HPO_4, 8.6 mM NaCl, 22 mM anhydrous KH_2PO_4, 2 mM $MgSO_4$, 100 µM $CaCl_2$, 50 µg.mL^{-1} ampicilline (Sigma-Aldrich® A9518), 2 g.L^{-1} D-[U-^{13}C]glucose (99%) (Cambridge Isotope Laboratories, Inc. CLM-1396-PK), 2 g.L^{-1} $^{15}NH_4Cl$ (98%) (Sigma-Aldrich® 299251) and trace element solution composed of 0.17 mM EDTA, 0.54 mM $CaCl_2$, 0.027 mM $CuSO_4$, 0.095 mM $MnCl_2$, 0.003 mM H_3BO_3, 0.024 mM $ZnSO_4$, 0.216 mM $FeSO_4$, 0.011 mM ascorbic acid with vitamin cocktail (Sigma-Aldrich® B6891).

2. Transformation and expression test

BmrA is expressed in *Escherichia coli strain C41(DE3)* - (OverExpress™ C41(DE3) Competent Cells - Lucigen®) using the expression vector pET23b(+)-bmrA (provided by Jean-Michel Jault). The resulting protein carries a hexa-histidine pseudo-affinity tag at its C-terminus and three additional amino acid residues from cloning site in N-terminus.
One µL of 0.4 µg.µL^{-1} pET23b(+)-bmrA were added to 200 µL *Escherichia coli strain C41(DE3).* The bacteria were incubated on ice for 10 min, then at 42° C for 90 sec and again 5 min on ice. 800 µL of LB medium (Lysogeny Broth, Sigma-Aldrich® L3522) were added and the bacteria were incubated at 37° C, 125 rpm for 1 h. Two agar plates containing 50 µg.mL^{-1} of ampicillin were inoculated and spread out with 50 µL and 100 µL of bacteria solution. The agar plates were incubated overnight at 37° C.

Height colonies were picked up using an inoculation loop and were mixed in height different tubes containing 3 mL of LB. The tubes were incubated at 37° C, 200 rpm until an OD_{600nm} of 0.6, then in order to save the height different colonies, two agar plates divided in height parts were inoculated with an inoculation loop, one colony for one part. After the agar plates inoculation, 0.7 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) were added in each tube (except one for control) and incubated at 37° C, 125 rpm for 4 h. After 4 hours, the OD_{600nm} was checked for each tube and diluted in order to get an OD_{600nm} of 1. One mL of each tube was collected and centrifuged at 6 000 × *g*. The supernatant was discarded, the pellet was suspended using 50 µL of Laëmmli buffer (62.5 mM Tris-HCl pH 6.5, 2 % (m/v) SDS, 10 % (v/v) glycerol, 0.01 % (m/v) bromophenol blue and 5 % (v/v) 2-mercaptoethanol) and placed in a sonicating bath during 30 min. The expression level of each clone was compared by Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) and the best one was selected (**Figure 28**).

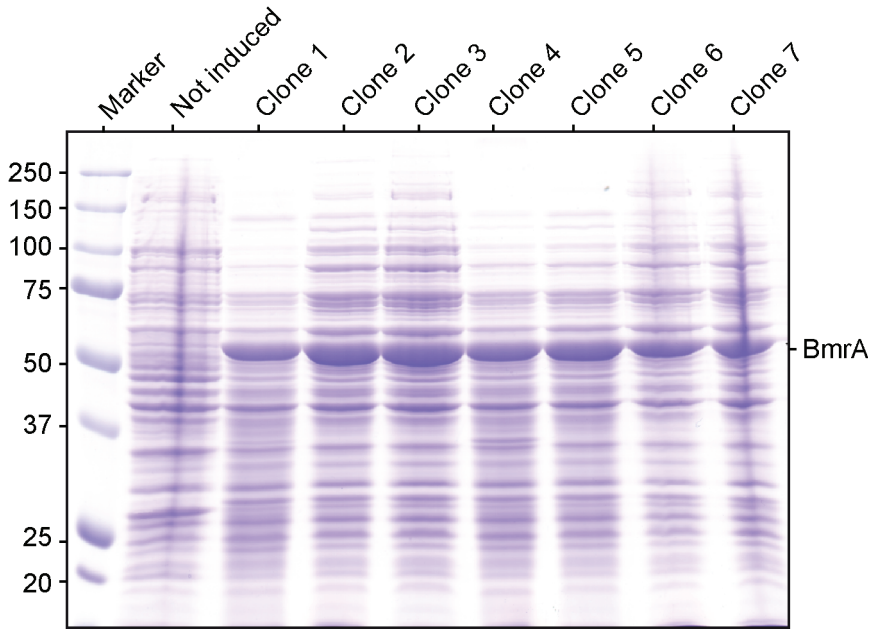


Figure 28. BmrA expression test followed by SDS-PAGE. The culture medium of each expression test was diluted to get an OD_{600nm} of 1. One mL of each tube was collected and centrifuged at 6 000 × *g*. The supernatant was discarded, the pellet was suspended using 50 µL of Laëmmli buffer (Each well was loaded with 10 µL of sample.

3. Expression of BmrA for NMR sample (2 liters)

The selected clone from the expression test was inoculated into 3 mL of LB medium and incubated 4 hours at 37° C and 200 rpm. Then, the 3 mL were used in order to inoculate 50 mL of minimal M9-medium in a baffled flask (150 mL), which is the first pre-culture. **This**

pre-culture was incubated at 37 °C and 200 rpm until an OD_{600nm} of 1.5, and then used to inoculate the second pre-culture in order to obtain an initial OD_{600nm} of 0.2**.** This second pre-culture was composed of two times 150 mL minimal M9-medium in baffled flasks (500 mL). These flasks were incubated overnight at 25° C and 130 rpm. After incubation, the final OD_{600nm} was 2.2-2.5. Finally, four baffled flasks (2 L) containing 425 mL of minimal M9-medium and not labelled amino-acids for reverse labelling were inoculated with 75 mL of the second pre-culture. Concentrations of natural abundance amino-acids are 0.25 $g.L^{-1}$ Ile, 0.25 $g.L^{-1}$ Leu, 0.25 $g.L^{-1}$ Val, 0.40 $g.L^{-1}$ Lys, 0.10 $g.L^{-1}$ Pro, 0.40 $g.L^{-1}$ His. The initial OD_{600nm} was 0.2-0.3, the four flasks were incubated at 25° C and 130 rpm. The expression of BmrA was induced with 0.7 mM IPTG when the OD_{600nm} reached 0.6-0.7, then the cultures were incubated until around 30 min after the beginning of the bacteria stationary phase. The bacteria were harvested by centrifugation at 6 000 × *g* during 20 min 4° C.

The general growth curve of *Escherichia coli strain C41(DE3)* which expresses BmrA in minimal medium is presented **Figure 29**.

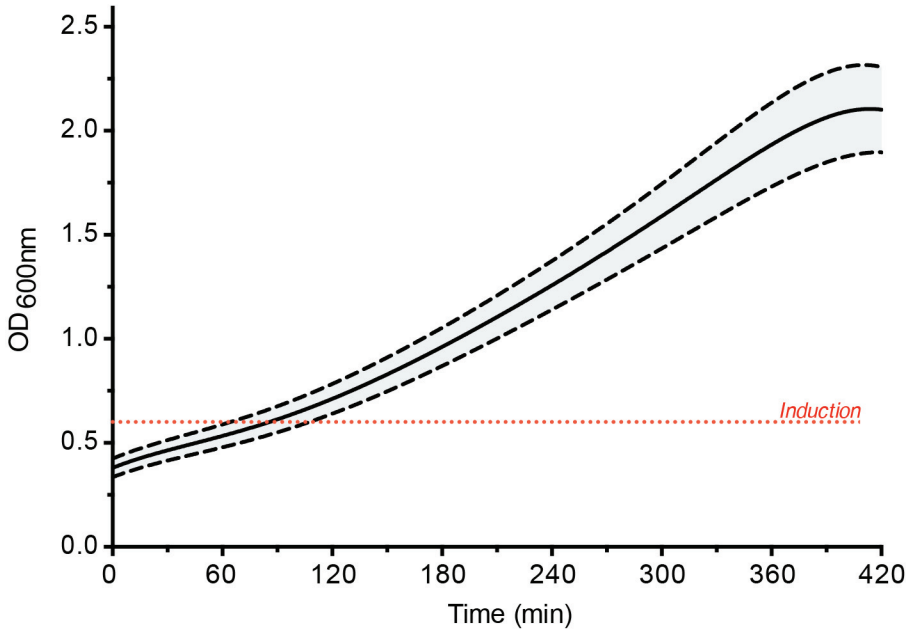


Figure 29. General growth curve of *Escherichia coli strain C41(DE3).* A black line shows the OD_{600nm} mean of 20 different BmrA expression cultures with the standard deviation represented by a grey area and delimited by dash lines. A red dot line highlights the induction (OD_{600nm} 0.6 and addition of 0.7 mM IPTG).

4. Membrane extraction

The bacteria pellet obtained previously by centrifugation is suspended with 25 mL of resuspension buffer (50 mM Tris-HCl pH 8, 5 mM $MgCl_2$, 1 mM DTT, 250 U Benzonase® Nuclease (Sigma-Aldrich® E1014) and one tablet of protease inhibitor cocktail EDTA-free (Roche 4693132001). Bacteria were disrupted using a Microfluidizer®, the solution was centrifuged at 15 000 × *g*, 4° C for 15 min in order to eliminate non-soluble particles like bacteria debris. The obtained supernatant was ultra-centrifuged at 200 000 × *g*, 4 °C for 1 h in order to harvest the bacteria membrane containing BmrA. The supernatant was discarded and the membrane pellet was suspended with storage buffer (50 mM Tris-HCl pH 8, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.3 M saccharose).

The concentration of protein inside the membrane was estimated by bicinchoninic acid assay at 150-200 $mg.mL^{-1}$.

The membrane extraction was followed by SDS-PAGE, (**Figure 30**). After the first stage of centrifugation, BmrA can be found in large quantity in the pellet. This fraction of BmrA could be due to inclusion bodies or inefficiently lysed bacteria. However, the second stage of centrifugation allows an efficient separation between soluble protein and the membrane containing BmrA.

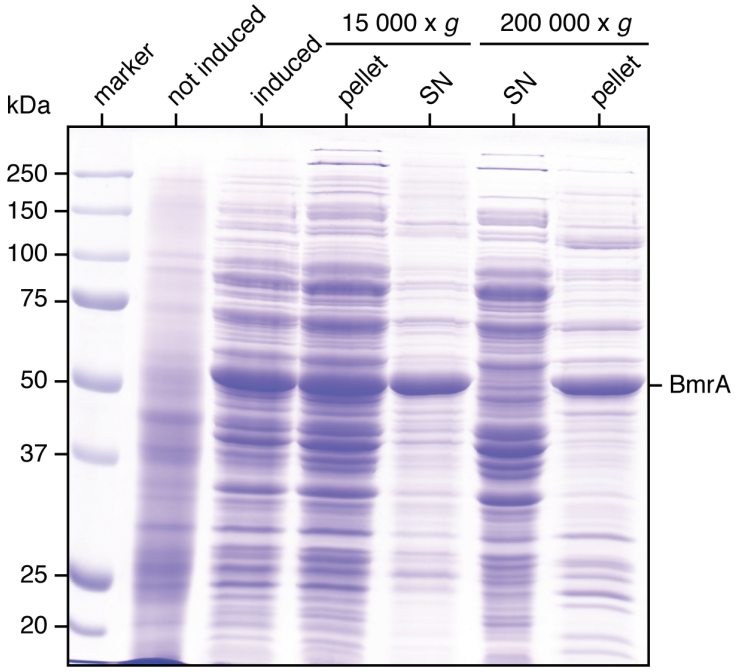


Figure 30. SDS-PAGE of the membrane extraction. Supernatant (SN) and pellet of the two centrifugation steps indicated by the associated centrifugal force: 15 000 × *g* and 200 000 × *g*. Each sample was collected and mixed with Laëmmli buffer; Each well was loaded with 10 µL of sample.

Purification of BmrA

1. Solubilisation

Membrane previously frozen were thawed and diluted at 2 mg.mL^{-1} with a solubilisation buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 1% n-Dodecyl-β-D-maltopyranoside (DDM), 1 mM DTT, 15% glycerol). The suspension was incubated and homogenised by mechanical rotation for 1 h at 4 °C. An overnight solubilisation displays the same yield than a solubilisation during 1 hour. The insoluble material was removed by ultra-centrifugation at 100 000 × *g* for 1 h at 4° C.

The solubilisation efficiency for BmrA was estimated to 80-90 % regarding the SDS-PAGE (**Figure 31**).

2. Purification by immobilised metal ion affinity chromatography

The solubilized protein was loaded onto a Ni^{2+}-nitrilotriacetic acid- (Ni-NTA) agarose resin column (Qiagen™) and incubated for 1 h at 4 °C with a mechanical rotation. The Ni-NTA resin was equilibrated with 5 column volumes (CV) of equilibrating buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 15% glycerol, 0.2% DDM and 10 mM imidazole). After the loading, the column was successively washed with 2 CV of equilibrating buffer, 2 CV of equilibrating buffer containing 0.5 M NaCl, 2 CV of equilibrating buffer containing 10 mM imidazole, 2 CV of equilibrating buffer containing 30 mM imidazole, and 2 CV of equilibrating buffer containing 40 mM imidazole. The protein was eluted with equilibrating buffer containing 300 mM imidazole. The elution was performed in batch (1 mL buffer / mL of resin) in order to maintain the protein concentration at 2-3 mg.mL^{-1} and thus avoid aggregation.

3. Desalting stage by PD-10 Desalting Columns

The imidazole was removed using desalting column, the buffer of the eluted protein was exchanged to the final buffer (50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, 0.2% DDM,

EDTA-free protease inhibitor cocktail). The desalting stage was operated using PD-10 desalting column equilibrated with 10 CV of final buffer.

4. Sample quality control

4.1. SDS-PAGE analysis

The purification was followed by SDS-PAGE (**Figure 31**). The solubilisation of BmrA by DDM from the *E. coli* membrane shows about 80% efficiency. Contaminants can be found in final fraction but represent less than 5% of sample comparing to the major band corresponding to BmrA. Regarding to the low sensitivity of NMR, the quantity of contaminants comparing to BmrA will not interfere in the NMR spectra with additional signals.

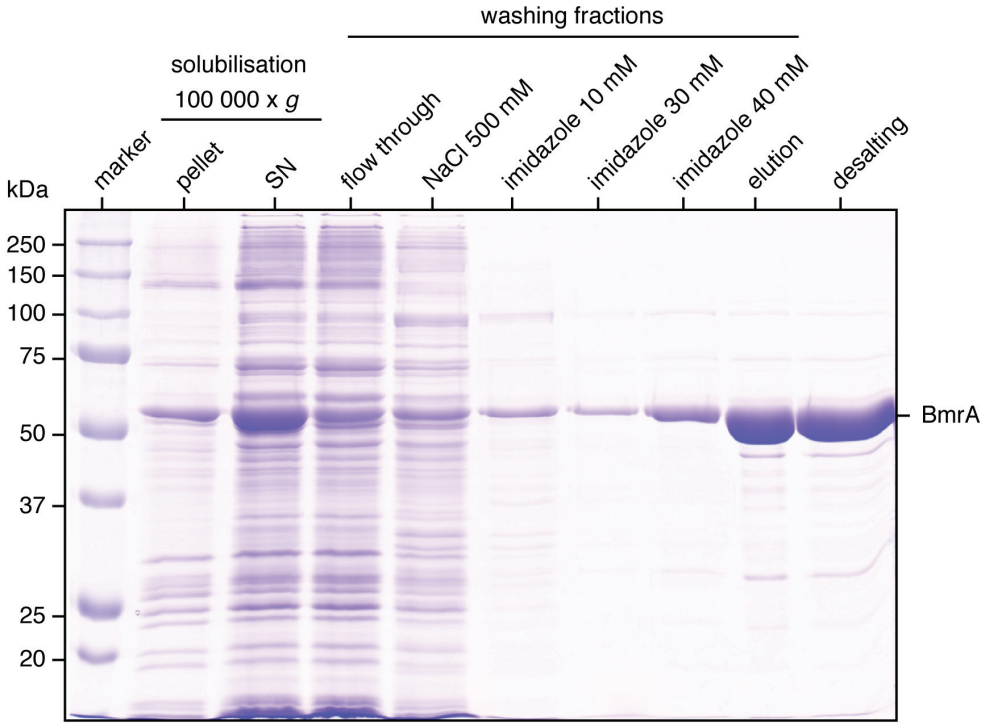


Figure 31. SDS-PAGE analysis of each purification step of BmrA. Washing buffer used is composed by 50 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM imidazole 15% glycerol, 0.2% DDM. A variation of the concentration during the washing steps is indicated with the associated compound on the corresponding fraction.. Each well was loaded with 10 µL of sample.

4.2. Ultraviolet-visible spectrophotometry

The concentration of the protein was estimated by ultraviolet-visible spectrophotometry (Thermo Scientific NanoDrop™ 2000c) using the absorbance at 280 nm and a theoretical extinction coefficient of 38 850 $M^{-1}.cm^{-1}$ (ProtParam - http://web.expasy.org/protparam).

The absorption spectrum of BmrA displayed three characterized peaks corresponding to the absorption of the carbonyl group from the protein back-bone (225 nm and 235nm) and to the absorption of the tryptophan (280 nm), **Figure 32**. The peak corresponding to the nucleotide contaminant overlaps with the peaks at 225 nm and 280 nm. For each sample of BmrA, the 280-to-260 ratio never exceeded 0.75, the canonical value for protein being 0.6. If the sample is not washed with 0.5 M NaCl during the purification stage, the 260-to-280 ratio exceeds a value of 2 due to DNA contaminations, an estimation of protein concentration with this method is not possible.

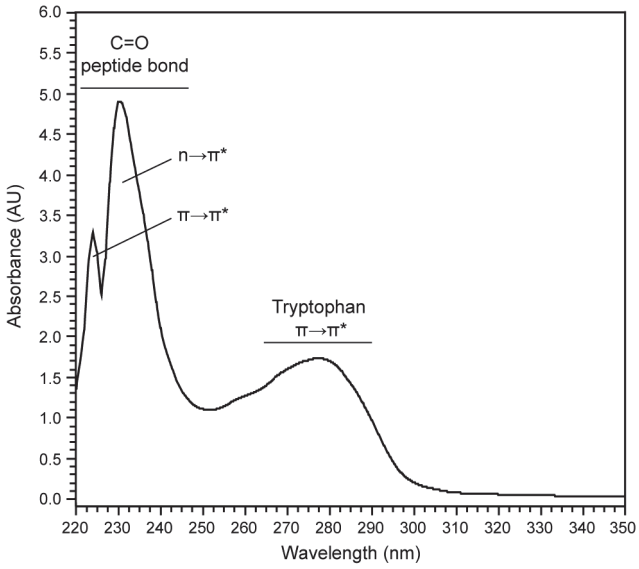


Figure 32. Absorption spectrum of BmrA. Each peak is labelled with the chromophore and with the associated transition.

4.3. Size exclusion chromatography

Size exclusion chromatography (SEC) was performed using a Superdex™ 200 Increase 10/300 GL (GE Lifesciences) on a NGC Chromatography System (Bio-Rad). The protein solution in buffer 50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, 0.2% DDM was diluted 4 fold using the same buffer without DDM. A 50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, 0.05% DDM was used as eluent. A flow rate of 300 µL/min was applied.

On the SEC-chromatogram presented **Figure 33A**, a major peak containing BmrA was observed for 11 mL of elution volume. The peak displays a shoulder at 11.7 mL. Regarding

the SDS-PAGE, both peaks can be assigned to BmrA. This shoulder seems to be the result of a sample heterogeneity which could be due to the size of the detergent-micelle containing BmrA. It is important to note that the SEC-chromatogram does not display the presence of aggregates.

Different peaks can be observed between 19.5 and 22.5 mL of elution volume. These peaks characterized species which absorb preferentially at 260 nm. As a result, these peaks could be assigned to nucleotide contaminants present in the solution.

The apparent molecular mass of BmrA was estimated using a calibration curve done with the SEC. The calibration curve was done using protein standards: cytochrome C, 12.3 kDa; myoglobin, 17.8 kDa; trypsin inhibitor, 22.1 kDa; β-lactoglobulin, 35 kDa; transferrin, 81 kDa; aldolase, 158 kDa; ferritin, 440 kDa; thyroglobulin, 669 kDa.

The partition coefficient Kav was calculated as follows: $Kav = (V_e - V_0)/(V_t - V_0)$.

The void volume (V_0) is the volume of interstitial liquid and is equal to 8.325 mL. This volume is determined using Blue dextran which is completely excluded from the SEC.

The total volume (V_t) is equal to the volume of a column cylinder, which is equal to 24 mL.

The elution volume (V_e) is the elution volume of the protein of interest.

Kav (partition coefficient) is plotted against the logarithm of molecular mass in order to determinate the apparent molecular mass of BmrA, **Figure 33B**. Kav of 0.1706 (peak 11 mL) and 0.2153 (peak 11.7 mL) correspond to 274.1 and 201.2 kDa, respectively.

The theoretical molecular mass of ^{13}C-^{15}N-BmrA-DDM complex was estimated to 349.6 kDa. Mass spectrometry showed that 410 molecules of DDM (510 Da) are bound to one molecule of BmrA (Chaptal *et al.* 2017). The molecular mass of one dimer of ^{13}C-^{15}N-BmrA is:

$140\,500 + 410 \pm 40 \times 510 = 349\,600 \pm 20\,400$ Da.

Surprisingly, the molecular mass of BmrA estimated by SEC is 25% lower than the theoretical molecular mass. This result was also found by our collaborators (Chaptal *et al.* 2017). SEC allows a separation of molecules regarding the size that is not strictly related to the molecular mass. BmrA in its DDM micelle looks like a dumbbell, and this non-globular configuration seems to interfere slightly on the apparent molecular mass estimated by SEC.

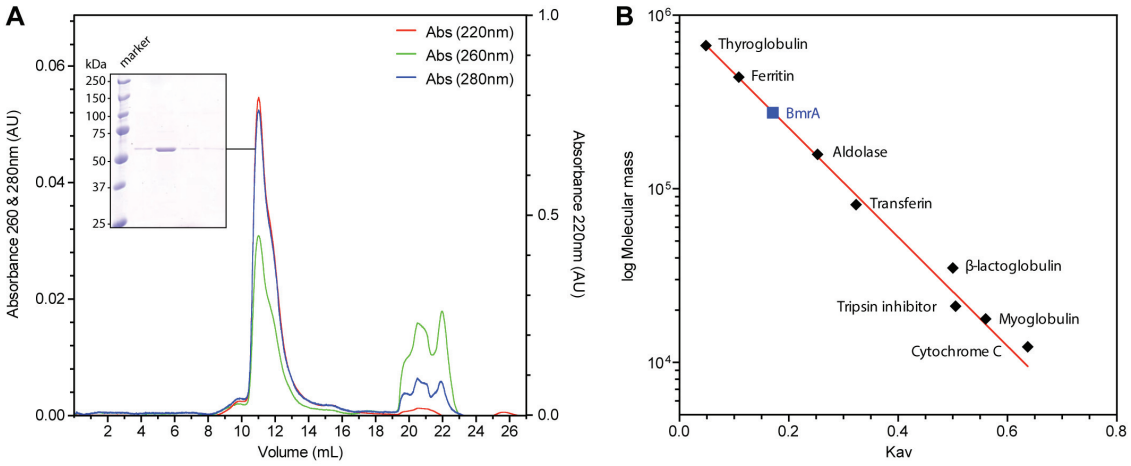


Figure 33. Size exclusion chromatography analysis. (**A**) Size exclusion chromatography of BmrA using a Superdex^{TM} 200 Increase 10/300 GL. A 50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, 0.05% DDM was used as eluent with a flow rate of 300 µL/min. The elution profiles of three wavelengths were displayed, 220 nm (red line), 260 nm (green line) and 280 nm (blue line). (**B**) Kav (partition coefficient) is plotted against the logarithm of molecular weight using protein standards. BmrA is indicated in blue on the standard curve.

4.4. Mass spectrometry

Sample containing BmrA after purification was mixed in a 10:1 ratio with matrix (10 g.L^{-1} (m/v) DHB (2,5-dihydroxybenzoic acid) in water and 1 g.L^{-1} (m/v) NaI acetone), and 1 µL of mixture was deposited on the MALDI-target and air-dried. Mass spectrometry experiments were performed with a Voyager-DE Pro MALDI-TOF (AB Sciex, Framingham, MA) Mass Spectrometer with a nitrogen UV laser (λ = 337 nm, 3 ns pulse). The instrument was operated in reflectron positive ion mode (mass accuracy: 0.008%) with an accelerating potential of 5 kV. The resulting mass spectrum is presented **Figure 34**. The molecular mass of one dimer of ^{13}C-^{15}N-BmrA is 140 500 Da. Peaks corresponding to BmrA can be identified at 70 062.3 Da (monomeric form) and 140 207.6 Da (dimeric form). Moreover, two adducts can be detected at 71 749.5 Da (+1 687.2 Da) and 73 534.5 Da (+3 472.2 Da) for the monomeric form and at 143 613.5 Da (+3 405.9 Da) and 147 178.3 Da (+6 970.4 Da) for the dimeric form of BmrA. Therefore, 1 or 2 molecules with a molecular mass of 1 720 Da are strongly bound to BmrA. The same mass spectrometry profile was observed by our collaborator with 2 or 4 molecules with a molecular mass of 1 691 Da, which are strongly bound to BmrA. The difference between both masses can be due to the isotopic labelling. However, using the *E. coli* metabolome database (ECMDB, http://ecmdb.ca) and molecular mass between 1 600 and 1 800 Da, only 10 results of potential compound can be identified,

whose 8 of them are liposaccharides. It is not surprising that lipopolysaccharides (such as lipid IVb (1 611 Da), KDO-lipid IVa (1 625 Da) or lipid A (1 800 Da)), from the *E. coli* membrane can be strongly bound to the hydrophobic segments of the protein. Peaks corresponding to ^{13}C-^{15}N-OmpF can be identified at 46 671.8 Da.

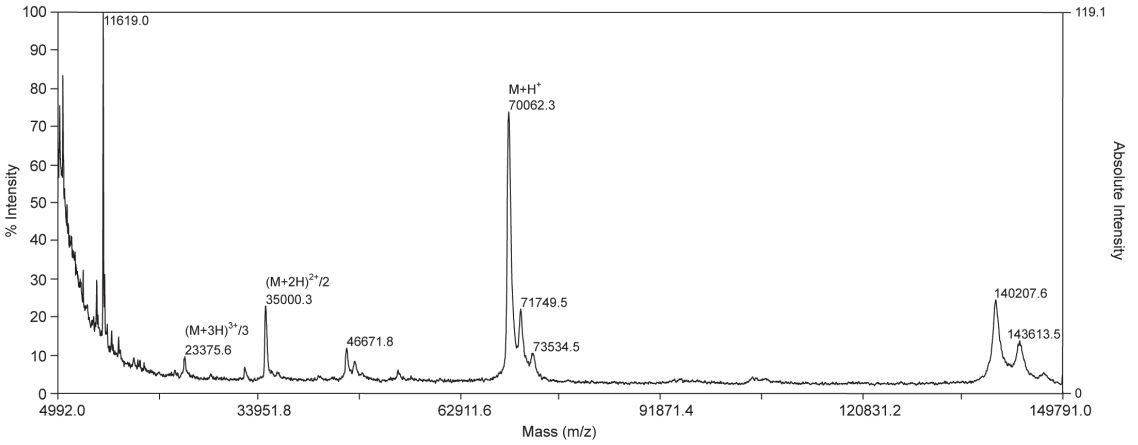


Figure 34. Mass spectrometry analysis of BmrA in DDM after purification. The major peaks are labelled with the associated masses or by the charged ions formed during the ionization. Mass spectrometry experiment was done by Frédéric Delolme SFR Biosciences (UMS3444/CNRS, US8/Inserm, ENS de Lyon, UCBL)

4.5. ATPase activity assay

BmrA was reconstituted in lipids from *Bacillus subtilis*. The reconstitution step is described in the next chapters (*Chapters V* and *VI*). Briefly, BmrA was diluted 4 fold with buffer (50 mM Tris-HCl, 100 mM NaCl, 10% glycerol) in order to get a final concentration of DDM of 0.05 % (m/v). Lipids destabilized by Triton X-100 ratio 1:10 (mol:mol) were added to the detergent solubilized BmrA to a lipid-to-protein ratio of 0.5. The detergents were removed by dialysis against dialysis buffer containing 50 mM Tris-HCl pH 8, 200 mM NaCl, 5% glycerol, 0.05% NaN_3 for 9 days at room temperature. Fourty mL of sample were dialyzed in 5 L of dialysis buffer using dialysis tubing with a cut-off of 6–8 kDa (Spectra/Por® Dialysis Membrane, spectrumlabs.com). To ensure an efficient depletion of detergents, three times the theoretical required quantity of Bio-Beads (SM2, 20–50 mesh; Bio-Rad, USA) was added to the dialysis buffer at the beginning of dialysis and again after 6 days. The calculations were based on the adsorption capacities of 105 mg DDM per 1 g Bio-beads (Rigaud *et al.* 1995) and 70 mg Triton X-100 per 1 g Bio-Beads (Bio-Rad, user manual).

After the reconstitution, an ATPase activity assay was operated in order to verify if the protein is functional. The ATPase activity assay is a colorimetric assay which allows to

determine the concentration of inorganic phosphate, according to the method described by (Doerrler & Raetz 2002). Briefly, 250 µL of reconstituted BmrA (0.2 mg.mL^{-1}) in 50 mM Tris-HCl, 200 mM NaCl, 5% glycerol was incubated 5 min at 37 °C. The ATPase reaction was started by the addition of 5 µL ATP-Mg^{2+} from a stock solution of 500 mM. Twenty-five µL of the reaction medium were pipetted and mixed with 25 µL 12% SDS. Twenty-five µL of 6% ascorbic acid in 1M HCl and 2% ammonium molybdate were added, and the colorimetric reaction was incubated 5 min at room temperature. 50 µL of 1 M HCl were then added in order to slow down drastically the colorimetric reaction. The absorbance at 850 nm was measured and the quantity of orthophosphate was determined using a Na_2HPO_4 standard (0-50 µmol).

ATPase activities of BmrA in DDM of 7 different productions and BmrA in *Bacillus subtilis* lipids of 14 different productions are presented in **Figure 35.** Activities monitored on BmrA show that each production displays a similar activity whether it is for BmrA in DDM or in lipids.

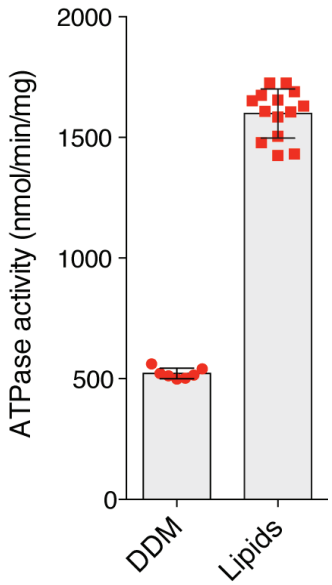


Figure 35. ATPase activity of BmrA for different sample preparations. The ATPase activity of BmrA was monitored at 521 and 1600 nmol.min^{-1}.mg^{-1} respectively in DDM and *Bacillus subtilis* lipid environment. Each sample preparation is symbolized by a red square on the histogram.

4.6. Solid-state NMR

Spectrum recording is described in the *Chapter V*. In short, NMR experiments were carried out on a Bruker Biospin AVANCE III spectrometer operating at 800 MHz ^{1}H frequency

using a 3.2 mm triple-resonance (^{1}H, ^{13}C, ^{15}N) E-free probe (Bruker Biospin). Sample temperature was determined using the chemical shift of supernatant water and was set to 278 K. All experiments were conducted at 17.5 kHz spinning frequency using 90 kHz SPINAL64 proton decoupling. The proton and carbon field amplitudes during the ^{1}H-^{13}C cross-polarization transfer step were around 66 kHz and 50 kHz respectively, using a tangential ramp with a span of ±5% on the carbon channel. The Dipolar Assisted Rotational Resonance (DARR) experiment was recorded in 48 scans per increment with a 20 ms mixing time. Acquisition times were 15 ms in t_2 and 10 ms in t_1, and the interscan delay was 2 s, corresponding to a total measurement time of 50 h. Spectra were processed with Topspin and analyzed using CcpNmr Analysis.

Spectra of ^{13}C-^{15}N-BmrA and ^{13}C-^{15}N-[^{12}C-^{14}N-LVIKHP-]-BmrA reconstituted in *Bacillus subtilis* lipids display narrow resonance lines for isolated peaks, and chemical-shift dispersion typical for a well-folded protein, even if the large number of signals from the 607 amino-acid-residue protein results in heavy resonance overlap. However, addition of the natural abundance Leu, Val, Ile, Lys, His and Pro reduces the number of signals in the spectra of this large protein and leads to obtain more isolated peaks (**Figure 36A**). Cross-peaks from many individual atom pairs can be detected (**Figure 36A**). Isolated resonances typically exhibit a line width of 90.5 Hz as shown for example for cross peaks stemming from Ala Cα-Cβ, (**Figure 36B**). The signal-to-noise ratio in this 50 h experiment is excellent if it takes into account the size of the protein and the dilution by lipids, within a factor of 2. The overlay of ^{13}C-^{15}N-BmrA and ^{13}C-^{15}N-[^{12}C-^{14}N-LVIKHP-]-BmrA spectra shows that both preparations display the same spectra. The combination of NMR spectra and ATPase activity data highlights that the procedure of the production, purification and reconstitution of the protein BmrA is highly reproducible.

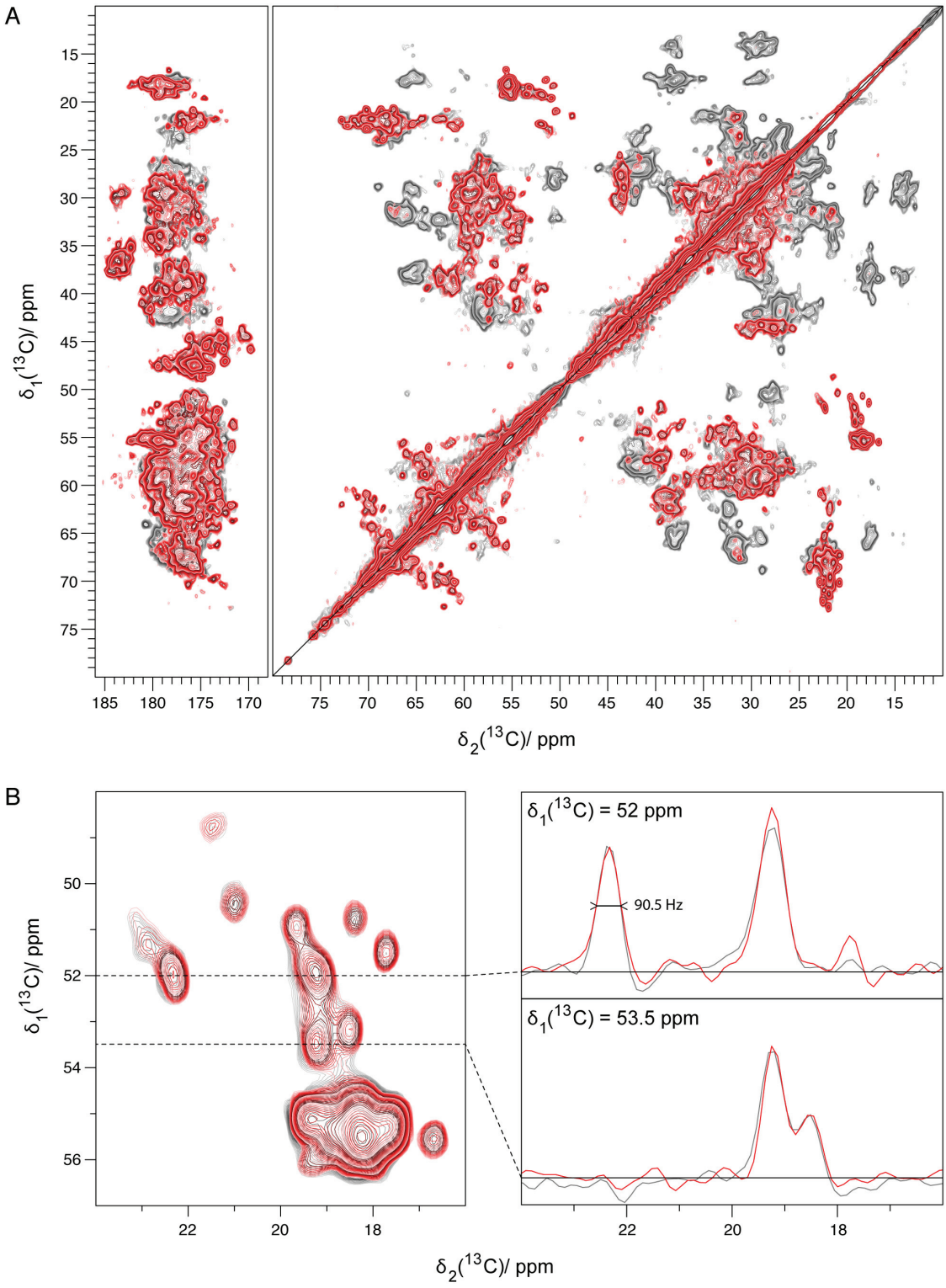


Figure 36. ^{13}C-^{13}C two-dimensional experiment (DARR, 20 ms) overlay of ^{13}C-^{15}N-BmrA and ^{13}C-^{15}N-[^{12}C-^{14}N-LVIKHP-]-BmrA. (**A**) Aliphatic and carbonyl region spectra overlay of ^{13}C-^{15}N-BmrA (grey shades) and ^{13}C-^{15}N-[^{12}C-^{14}N-LVIKHP-]-BmrA (red shades). (**B**) Alanine region with 1D traces taken at 52 ppm and 53.5 ppm, the signal dispersion and line width are similar for both preparation.

Conclusion

NMR-scale expression and lipid reconstitution of the *B. subtilis* ABC transporter BmrA leading to a lipid-inserted sample of uniformly [^{13}C, ^{15}N] enriched BmrA was reported here. The preparation yields narrow resonance lines in the NMR-spectra rivaling microcrystalline preparations. The excellent quality of the NMR spectra is a good indicator for a structurally homogeneous membrane-insertion of the protein. With a typical size encountered for most membrane proteins, BmrA is one of the largest proteins studied today by solid-state NMR. These results demonstrate the feasibility of studying BmrA, and related membrane proteins, by solid-state NMR to obtain site-resolved information about structural and dynamic features of membrane transport proteins in different states (open, closed, drug-bound), even if the *de novo* structure determination of a protein of this size is still extremely challenging.

References

Chaptal, V., Delolme, F., Kilburg, A., Magnard, S., Montigny, C., Picard, M., Prier, C., Monticelli, L., Bornert, O., Agez, M., Ravaud, S., Orelle, C., Wagner, R., Jawhari, A., Broutin, I., Pebay-Peyroula, E., Jault, J.-M., Kaback, H.R., Le Maire, M. & Falson, P. (2017) Quantification of Detergents Complexed with Membrane Proteins. *Scientific reports* 7, 41751.

Doerrler, W.T. & Raetz, C.R.H. (2002) ATPase activity of the MsbA lipid flippase of Escherichia coli. *Journal of Biological Chemistry* 277, 36697–36705.

Rigaud, J.L., Pitard, B. & Levy, D. (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochimica et biophysica acta* 1231, 223–246.

Chapter V

Membrane protein preparation for solid-state NMR studies

The work presented in this chapter was done in collaboration with Beat H. Meier. This Chapter was published with the following reference:

Sample preparation for membrane protein structural studies by solid-state NMR

Denis Lacabanne, Britta Kunert, Carole Gardiennet, Beat H. Meier and Anja Böckmann

Methods Mol Biol. 2017;1635:345-358.

Contents

Abstract

Conformational studies of membrane proteins remain a challenge in the field of structural biology, and in particular the investigation of the proteins in a native-like lipid environment. Solid-state NMR presents a valuable opportunity for this, and we present here three critical steps in the solid-state NMR sample preparation, i.e. membrane reconstitution of the protein in native lipids, rotor filling and sample quality assessment, at the example of the *Bacillus subtilis* ATP-binding cassette transporter BmrA.

Running title: Membrane proteins for solid-state NMR

Key words

Solid-state NMR; membrane protein; sample preparation; native lipids

Introduction

Solid-state NMR is today a well-established technique for protein structural investigations. It allows access to structural information of difficult-to-crystallize or insoluble proteins which escape other structural methods. Nowadays, the field of applications of solid-state NMR covers fibrillar proteins (Comellas *et al.* 2011; Gath *et al.* 2014; Habenstein *et al.* 2012; Luckgei *et al.* 2013; Van Melckebeke *et al.* 2010; Wasmer *et al.* 2008), large protein complexes (Demers *et al.* 2014; Loquet *et al.* 2012; Morag *et al.* 2015) and membrane proteins reconstituted in native or mimetic lipid environments (Emami *et al.* 2013; Franks *et al.* 2012; Kunert *et al.* 2014). As a biophysics tool, NMR brings complementary insight into other structural techniques such as X-ray crystallography and electronic microscopy, and can be used in hybrid approaches in combination with them (Demers *et al.* 2014; Sborgi *et al.* 2015).

De novo 3D structures can be obtained by using solid-state NMR for proteins up to approximately 100 amino-acid residues; it can also be used to obtain secondary structures derived from sequentially assigned chemical shifts (Wishart & Sykes 1994) for proteins of 300-400 amino acids, and spectral fingerprints can be obtained for even larger proteins (up to 700 amino acids (Luckgei *et al.* 2013)). The protein fingerprinting gives access to the comparisons between different preparations or different conformations; it can highlight interaction of the protein with partners (Fielding 2007; Williamson 2009; Zech *et al.* 2004). Spectral analysis requires well-resolved spectra with narrow resonance lines, high

sensitivity and a good dispersion of the signals. These parameters are strongly dependent on the sample preparation with respect to quality and also to quantity because NMR is a relatively insensitive method. It is well known from established techniques that sample preparation of membrane proteins for structural studies generally still presents a major challenge. As it is also the case for other structural techniques, solid-state NMR has a high demand in sample homogeneity, and in particular with respect to a very reproducible protocol allowing to prepare isotopically different, but otherwise identical samples. More specifically for solid-state NMR, in terms of sensitivity, a central requirement with respect of the analysis of membrane proteins is that a maximum amount of the protein must be reconstituted in a minimal amount of lipids without affecting the sample stability. Once the preparation of the sample is successfully achieved and high-quality spectra are obtained, multidimensional spectra can be recorded for complete resonance assignment of the protein and, *in fine*, distance restraints and torsion angles can be collected to solve the protein structure.

In this chapter, we shall focus our intention on three critical stages specifically involved in the sample preparation for solid-state NMR structural studies: the membrane reconstitution of the protein, solid-state NMR rotor loading and finally the quality assessment of the sample.

Materials

1. Lipid reconstitution

1. Stock of homemade lipids: the extraction procedure was established according to the literature (Bligh & Dyer 1959; Folch *et al.* 1957), and the extracted lipids were stored in chloroform (Sigma-Aldrich).
2. Liposomes solubilization: Triton X-100 (Sigma-Aldrich) and ultrasonic baths (Elmasonic S 30 (H), Elma Schmidbauer GmbH).
3. Buffer 2X: 100 mM Tris, 200 mM NaCl, 10% glycerol (v/v).
4. Laboratory freeze dryer (Heto PowerDry® LL3000 Freeze Dryer, Thermo Scientific).
5. Protein: BmrA (*Bascillus subtilis* Multidrug resistance ABC transporter, UniProtKB O06967) expressed in *E. Coli* C41(DE3) in minimal medium containing uniformly 99% ^{13}C-enriched glucose (2 g/L) and 98% ^{15}N-enriched ammonium chloride (2 g/L) as sole carbon and nitrogen sources. The protein was extracted from the *E. coli* membrane by using n-dodecyl-β-D-maltopyranoside (DDM) 1% (w/v) (Anatrace), purified by pseudo-affinity chromatography using metal chelate affinity chromatography resin (Ni-NTA- agarose gravity column - Qiagen™) and was desalted by using desalting column (PD-10 columns, GE Healthcare) with buffer A (50 mM Tris-HCl pH 8, 100 mM NaCl, 5% (v/v) glycerol, 0.05% (w/v) DDM). The final concentration was estimated at 1 mg.mL^{-1} as determined by measuring absorbance at 280 nm with a spectrophotometer (NanoDrop™ 2000c, Thermo Scientific).
6. Solvents: methanol (CARLO ERBA Reagents) and chloroform (Sigma-Aldrich)
7. Dilution buffer (buffer B): 50 mM Tris-HCl pH 8, 100 mM NaCl, 5% (v/v) glycerol.
8. Dialysis buffer (buffer C): 50 mM Tris-HCl pH 8, 200 mM NaCl, 5% (v/v) glycerol, 0.05% (w/v) sodium azide.
9. Dialysis tubing: cut-off of 6–8 kDa (Spectra/Por® Dialysis Membrane, SPECTRUMLABS.COM).
10. Bio-beads® (SM2, 20–50 mesh; Bio-Rad, USA).
11. ATPase activity assay: ATP-Mg^{2+} 10 mM, ammonium heptamolybdate 0.1% (w/v), ascorbic acid 0.6% (w/v) in HCl 1M, NaH_2PO_4 standard.

2. Loading a rotor using filling tools

1. Centrifuge and ultracentrifuge (Beckman Coulter) with SW 60 Ti Rotor.
2. MAS rotor 3.2 mm Zirconia with Vespel caps (Bruker HZ16916).
3. Homemade NMR filling tool as described in (Böckmann *et al.* 2009) (detailed plans can be provided by B.M.).
4. Glue epoxy adhesive Araldite® standard 2 x 15 mL.
5. Scalpel.

3. Quality assessment of the sample by solid-state NMR

1. NMR spectrometer: Bruker BioSpin AVANCE III 800 MHz (Bruker Biospin, Rheinstetten, Germany).
2. NMR probe: 3.2 mm $^{1}H/^{13}C/^{15}N$ triple channel E-free.
3. NMR analysis software: CcpNmr Analysis (Stevens *et al.* 2011).

Methods

1. Reconstitution of the protein in lipids

In solid-state NMR, it is essential that the protein preparation is homogenous and concentrated (in our case about 10-15 mg of protein reconstituted in lipids are loaded in a 3.2 mm rotor). Three parameters are crucial to protein reconstitution: the reconstitution method, the choice of the lipids and the lipid-to-protein ratio (LPR). The LPR must be as low as possible for the sake of sensitivity. There is a large variety of lipids which can form native-like liposomes (Kunert *et al.* 2014; Lange *et al.* 2010; Tang *et al.* 2013), mimetic liposomes (Kaur *et al.* 2015; Lu *et al.* 2013; Sackett *et al.* 2014), nanolipoproteins (Ding *et al.* 2015; Kijac *et al.* 2007) and bicelles (Prosser *et al.* 2006). The choice of the lipids in which the protein needs to be reconstituted is critical in order to obtain a stable preparation. Here we used a homemade preparation of lipids extracted from *B. subtilis*, since the protein shows an excellent stability in native lipids (Kunert *et al.* 2014). For the reconstitution, we removed the detergent by means of dialysis. This method is appropriate for detergents with high critical micelle concentration (CMC), but it requires the addition of Bio-Beads® for low CMCs detergents as was the case here for DDM (CMC 0.17 mM) (Rigaud *et al.* 1995; Seddon *et al.* 2004). As for our case, the addition of the Bio-beads® to the dialysis buffer yielded better results than a direct addition of the beads to the protein solution, which might induce the aggregation of the protein.

1. Solubilized lipids in chloroform from the lipid preparation or stock solution were placed into an empty glass vial whose weight was known. The aim was to add a volume corresponding roughly to 6 mg of lipids. The chloroform was removed by evaporation using a nitrogen gas flow (to avoid the oxidation of the lipids) and a freeze-drying of 30 minutes followed. The tube containing the freeze-dried lipids was weighed, after a simple subtraction with the empty tube mass, in order to estimate the final amount of lipids.
2. The lipids were solubilized in 200 µL of chloroform and mixed with Triton X-100 (MM_{av} 646 $g.mol^{-1}$) solubilized in methanol at a lipid-detergent ratio 1:10 (mol/mol). An estimated molar mass of *B. subtilis* lipids of 1000 Da was used for the calculation

(Griffiths & Setlow 2009). As done previously, the chloroform and the methanol were removed by evaporation in a nitrogen gas flow and a freeze-drying step of 30 minutes followed (see Note 1).

3. The lyophilized lipid-Triton X-100 mixture was first solubilized in half final volume of water, and then half final volume of buffer 2X is added (see Note 2). The final volume was determined so that the lipid concentration was 5 $mg.mL^{-1}$. The mixture was finally treated by sonication in an ultrasonic bath for ten minutes.
4. The lipid-to-protein ratio used for the reconstitution was 0.5 (w/w). This ratio was chosen by in relation to the ATPase activity of the protein, the yield of membrane-reconstituted protein after the dialysis step, and the homogeneity of the sample (Kunert *et al.* 2014) (see Note 3). In order to dilute the DDM, the protein solution is diluted by a factor 4 with buffer B (final protein and DDM concentration were respectively 0.25 $mg.mL^{-1}$ and 0.05%). Then, an amount of 5 mg of the lipids solubilized in Triton X-100 was added to 40 mL of the solution containing 10 mg of protein solubilized in detergent.
5. Preparation of Bio-Beads® according to Holloway 1973 (Holloway 1973): in order to deplete the detergent as much as possible, 3 g of Bio-Beads® were placed in a 25 mL beaker and were incubated with 15 mL of methanol at slow agitation. After 30 minutes of incubation, the methanol was removed and the Bio-Beads® were extensively washed in 20 mL of water for 5 minutes. The washing procedure was repeated three times.
6. The mixture containing the protein in detergent and the *B. subtilis* lipids was dialyzed against 5 L of buffer C using dialysis tubing for 9 days at room temperature and with slow agitation (see **Note 4**). The amount of added Bio-Beads® corresponded to six times the theoretical required quantity to deplete the detergent. Half of the Bio-beads® were added at the beginning of the dialysis and so was the rest after 6 days in order to be as effective as possible for the removal of the detergent. The theoretical quantity was defined considering the adsorption capacities of 1 g of Bio-beads® which is 105 mg DDM (Rigaud *et al.* 1995) and 70 mg Triton X-100 (Bio-Rad, user manual).
7. After 9 days of dialysis, the reconstituted BmrA in lipids was sedimented at 20 500 × *g* in a 1.5 mL tube by successive removal of the supernatant and addition of more

protein solution. This step is a bit lengthy but it guaranteed minimal loss of protein.

8. An ATPase activity assay was performed at 37 °C. The reaction was initiated by the addition of ATP-Mg^{2+} 10 mM to 2 µg of BmrA. Each minute or two minutes an amount of 25 µL was taken from the reaction-solution and mixed with an equal amount of 12% (w/v) sodium dodecyl sulfate in order to stop the reaction by denaturing the protein. Released Pi was determined by colorimetric assay using a method adapted from Taussiky and Shorr (TAUSSKY & SHORR 1953) and Chifflet *et* al (Chifflet *et al.* 1988). Briefly, 25 or 50 µL from a 1:1 (v/v) solution of ammonium heptamolybdate 0.1% (w/v) and ascorbic acid 0.6% (w/v) in HCl 1M were added and the mix was incubated during 5 min. The coloration reaction was drastically inhibited by adding 75 µL of HCl 1M, the absorbance was read at 820 nm with a spectrophotometer, the amount of free Pi was estimated using NaH_2PO_4 as a standard. In our case, the activity of BmrA with LPR 0.5 was about 1.3 µmol.min^{-1}.mg^{-1}.
9. The supernatant and the pellet of the sedimentation stage were analyzed by SDS-PAGE (**Figure 37A**). A stability test of the protein in native-lipids and EPC/EPA (9:1) lipids was performed and followed by SDS-PAGE, the protein was more stable in *B. subtilis* lipids (**Figure 37B**).

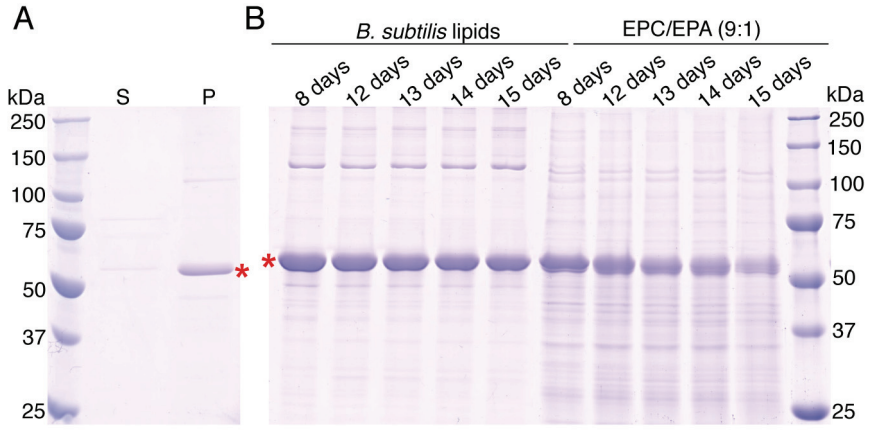


Figure 37. Reconstitution and stability of BmrA in *B. subtilis* lipids and EPC/EPA (9:1) analyzed by SDS-PAGE (10% acrylamide) from Kunert et al. 2014 (Kunert *et al.* 2014). (**A**) At the end of the dialysis, the sample was harvested and a 1 mL aliquot is centrifuged at 20 500 × *g*. The supernatant (S) and the pellet (P) were analyzed by SDS-PAGE. The gel shows that almost all the protein in lipids was pelleted. (**B**) The stability of the protein over time was also followed during 15 days at room temperature and analyzed by SDS-PAGE. Two lipids environments were used, native lipids of the protein and EPC/EPA in a ratio 9:1 (w/w). In both gels, a red star indicates the protein BmrA (66.5 kDa).

2. Loading the NMR rotor

The goal of this step is to fill the rotor with a maximum of the protein preparation with minimal loss. As the sample is highly viscous it cannot be pipetted easily. Before the design of dedicated filling tools, procedures found in the literature consisted in using a pipette tip to load the sample inside the solid-state NMR rotor (Das *et al.* 2013). Now we currently use a specific tool, which allows us to optimally fill the NMR rotors by using an ultracentrifuge (Böckmann *et al.* 2009). The tool is illustrated schematically in **Figure 38**. It is made up with a funnel which is directly in contact with the solid-state NMR rotor and can be centrifuged up to 200 000 × *g*, which leads to highly efficient packing of the protein.

1. The bottom part of the 1.5 mL tube, where the protein in lipids was previously pelleted, was carefully cut with a scalpel.
2. The NMR rotor was inserted inside the filling tools and the funnel was screwed.
3. The part of the 1.5 mL tube previously cut was turned upside down on the funnel and fixed using parafilm. The device with the sample was placed in a 15 mL tube and was gently centrifuged at 5 000 × g for 5 minutes which transferred the protein pellet into the funnel.
4. Afterward, the empty tube was removed and the filling device was centrifuged at 40 000 × *g*, 4 °C, for an hour (see Note 5). Higher speeds could be used but were not tested.
5. Finally, the rotor was removed from the tool and closed with the Vespel cap using epoxy glue spread on the cap with a pipette tip. The glue was left to dry overnight.

3. Referencing and temperature setting using 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)

Temperature is a crucial parameter in NMR experiments, with the central question being at which temperature the stability of the protein is the highest. In solid-state NMR, a large temperature range - between about -80 °C (193 K) to 60 °C (333 K) - can be covered under standard operating conditions, which makes it possible to choose the optimal temperature at which the protein is the most stable.With a protein whose behavior is unknown, it is important to start the NMR experiments carefully at temperatures around 5-10 °C (278-283 K). It is important to note that in solid-state NMR, the sample temperature - inside the rotor

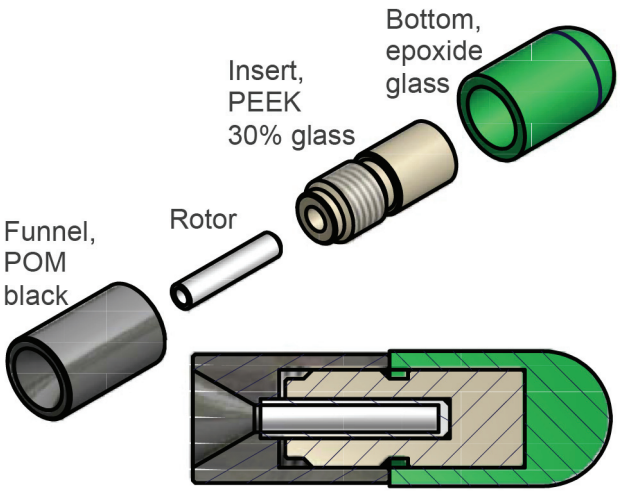


Figure 38. Rotor filling tool for ultracentrifuge from (Böckmann et al. 2009). The filling tool is composed of three parts that were designed to fit perfectly a 3.2 mm rotor. The insert part (PolyEther Ether Ketone 30% glass) containing the solid-state NMR rotor is screwed to the funnel (Polyoxymethylene) that is specially adapted for the opening side of the solid-state NMR rotor. The device can be centrifuged in the swinging buckets rotors SW 60 Ti and SW 40 Ti at 210 000 × *g*. B.M. can provide the detailed plans.

- can be drastically different from the one of the cooling gas, due to air friction when the rotor spins. For a 3.2 mm rotor, this effect is negligible up to 2 000 Hz spinning frequency, but induces a temperature increase of about 20 °C at a spinning frequency of 20 000 Hz. In order to keep the sample temperature constant, one has to increase the cooling gas flow in parallel to increasing the rotor spinning frequency.

In this part, we describe how to keep a temperature of about 5 °C using the proton chemical shift from the bulk water present in the sample. In order to internally reference the water chemical shift, a frequently used NMR standard recommended by IUPAC (Harris *et al.* 2008) in polar solvents is the 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) whose proton resonance is set to 0 ppm. Down below, we give the typical experimental values used on the 800 MHz WB magnet; they might be different for alternative setups and samples.

1. The magic angle was adjusted using KBr, observing ^{79}Br, and the shim of the spectrometer was adjusted.
2. The Variable Temperature Unit (V.T.U) was activated; the cooling gas target temperature was set to 5 °C (278 K) with a target gas flow of 1 000 L/h for a few minutes (in order to cool down the probe).

3. The rotor was inserted while temporarily reducing the gas flow to 200 L/h. The initial spinning frequency of the rotor was set to 2 000 Hz. A one-scan one-dimensional proton spectrum was recorded and the spectrum was referenced to DSS (0 ppm). The sample temperature was approximately determined using its water protons chemical shift linear dependency, following equations: $T(K) = (7.83 - \delta(H_2O)) \times 96.9$ (Cavanagh *et al.* 2010) or $T(°C) = 455 - 90 \times \delta(H_2O)$ (Gottlieb *et al.* 1997).
4. While the rotor spinning frequency was gradually increased up to 10 000 Hz, the cooling gas flow was accordingly increased to 500 L/h to reach 0 °C (273 K). The sample temperature was checked by recording a ^{1}H 1D spectrum and the gas flow or/and the target temperature was adjusted to avoid sample freezing.
5. When the temperature of 0 °C (273 K) was reached, the spinning speed was increased to 17 500 Hz (see **Note 6**) in stages of about 2 000 Hz. The temperature of the sample was checked by a ^{1}H 1D spectrum between each increase and the set temperature and gas flow were adjusted in order to reach a final sample temperature of about 5 °C.

Once the NMR experiment was finished, those stages had to be done in reverse order so as to avoid freezing or overheating of the sample.

4. Quality assessment of the sample

In order to assess the quality of the sample, three one-dimensional experiments can be used in solid-state NMR: ^{1}H-^{13}C-cross-polarization (CP) (Stejskal *et al.* 1977), ^{13}C-Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) (Morris & Freeman 1979) and ^{13}C-direct excitation (DE) (Bermel *et al.* 2008). Generally, a NMR sample can be considered of good quality if the signal-to-noise is sufficient, the line-widths at half height are of about 0.5 ppm and the sample is stable during the NMR acquisition. The CP highlights the residues of the rigid parts of the protein while the INEPT provides mainly signals for the residues located in the flexible regions of the protein. Contrary to the INEPT and the CP which select for specific properties, the direct excitation does not distinguish between flexible and rigid regions of the protein, and most residues present in the protein can be observed in this 1D spectrum (interference of dynamics with MAS and decoupling may lead to the loss of signals). Frequent repetitions of at least one of these experiments over time provide an easy

check of the stability of the sample. We repeated the most informative one, together with a proton experiment to check the temperature, before and after each longer multidimensional experiment in order to know if the sample quality was constant. A 2D DARR spectrum (Takegoshi *et al.* 2001; 2003) completed the sample assessment.

1. One-dimensional Cross-Polarization ^{13}C (see for (Gibby *et al.* 1974; Hediger *et al.* 1995; Stejskal *et al.* 1977; Yannoni 1982)) (**Figure 39**): in a first step, the 90° ^{1}H pulse was calibrated using as starting values the probe specifications and a CP-pulse sequence with high-power ^{1}H decoupling during acquisition. The cross-polarization between ^{1}H and ^{13}C was established by matching the Hartmann-Hahn condition. (Stejskal *et al.* 1977). We used SPINAL-64 ^{1}H heteronuclear decoupling (Fung *et al.* 2000). The 90° ^{1}H was determined by a nutation experiment and the proton power was adjusted to obtain a pulse length of roughly 2.5 to 3.5 μs. Finally, a reference ^{1}H spectrum was recorded. For comparison purposes, further ^{1}H spectra will be recorded with the same number of scans and dummy scans.

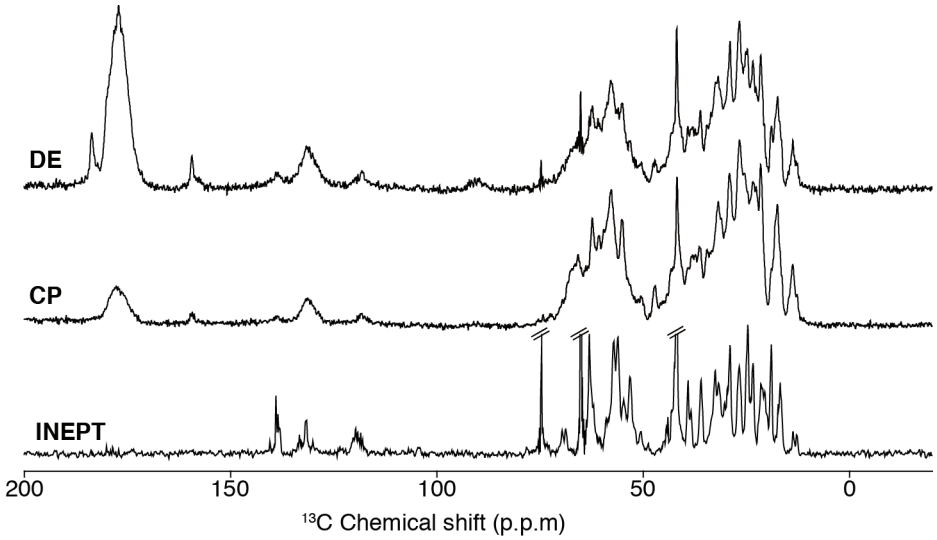


Figure 39. One-dimensional experiments at 5 °C: ^{13}C-direct excitation (DE) (Bermel *et al.* 2008), ^{13}C- Cross Polarization (CP) (Stejskal *et al.* 1977), and Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) (Morris & Freeman 1979) spectra of uniformly ^{13}C,^{15}N labelled BmrA. The number of scans was chosen in order to obtain similar signal-to-noise ratios and was respectively for DE, CP and INEPT of 8, 8 and 256.

2. One-dimensional direct-excitation acquisition ^{13}C (**Figure 33**): the ^{13}C pulse length was adjusted for the 90° pulse by using a CP pulse program followed by ^{13}C 90°-

delay-detection pulse, with no decoupling during the delay to dephase residual transverse polarization. The carrier frequency was set to the carbonyl region as to avoid frequency-offset effects. After phasing the spectrum obtained using a < 90° detection pulse, a 180° detection pulse was used to determine the zero crossing of the spectrum. The 90° pulse was calculated by dividing the 180° pulse by two. A spectrum was recorded using a 90° ^{13}C excitation pulse applied in the center of the spectrum followed by proton decoupling and an interscan delay of about 10 s.

3. One-dimensional Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) (Morris & Freeman 1979)(**Figure 39**): the previously optimized parameters (90° ^{13}C and ^{1}H) were used to setup the INEPT (^{1}H 90°- τ - $^{1}H/^{13}C$ 180° - τ - $^{1}H/^{13}C$ 90°) a SPINAL-64 ^{1}H heteronuclear decoupling (Fung *et al.* 2000) was used during the acquisition. The delay $\tau = 1/4J_{CH}$ was set to 1.25 ms, somewhat shorter than the theoretical value of 2 ms in order to take into account typical T_2' relaxation.

5. Two-dimensional ^{13}C-^{13}C Dipolar Assisted Rotational Resonance (DARR)

A 2D DARR (Takegoshi *et al.* 2001; 2003) with a short mixing time (around 20 ms) completes the quality check, as it is an experiment straightforward to set up, without a pronounced signal preference for any amino-acid-residue type, and broadband enough to display the aliphatic as well as the carbonyl and aromatic spectral regions. The necessary optimized pulse lengths were transferred from the CP experiment. The irradiation on the proton channel during the mixing time was adjusted to meet the rotary-resonance condition: $\nu_1 = n\nu_R$ (n = 1 or 2). A spectrum was recorded using a mixing time between 10-50 ms; 20 ms was a good compromise between an acceptable signal-to-noise ratio and spectral crowding. The total acquisition times can vary between one hour and one day, depending on the signal-to-noise ratio. The interscan delay was typically around 2 to 3 s.

6. Spectrum analysis

Spectra are processed using Topspin (Bruker Biospin) and analyzed using CcpNmr Analysis (Stevens *et al.* 2011). The dispersion of the resonances should reflect the expected secondary structure (**Figure 40A**) content when compared to database chemical shifts (Biological Magnetic Resonance Data Bank, (Ulrich *et al.* 2007)). Generally, this can be most easily analyzed looking at the alanine Cα-Cβ correlations (**Figure 40B**). Typical line widths of

isolated signals for samples of well-ordered protein preparations are around 0.5 ppm (**Figure 34C**).

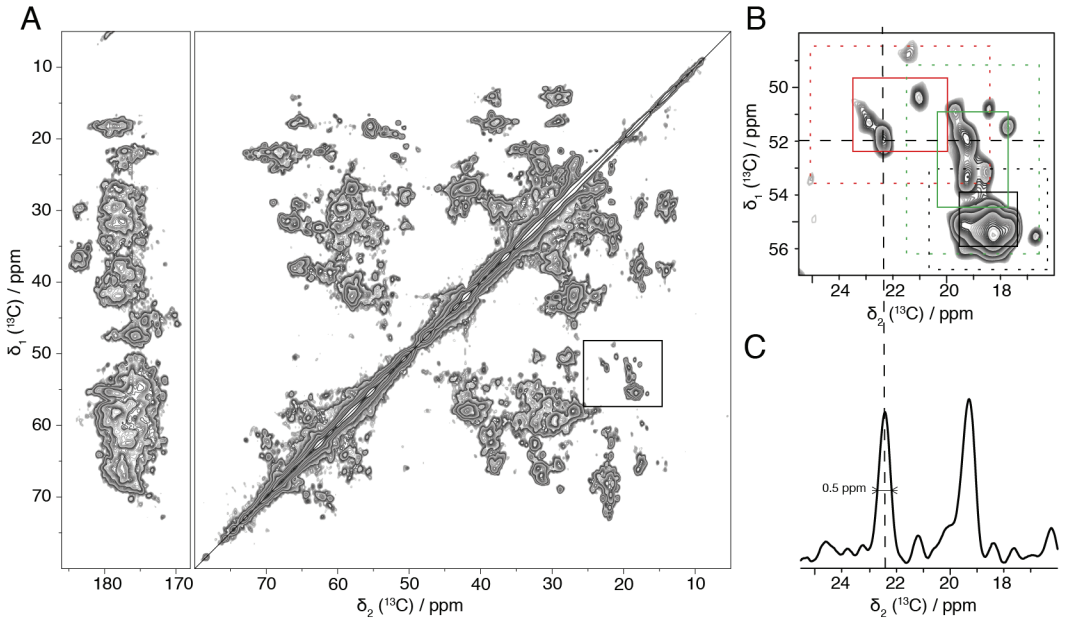


Figure 40. ^{13}C-^{13}C two-dimensional solid-state NMR correlation experiment of the protein BmrA at 5 °C (20ms DARR mixing period, 48 scans resulting in approximately 50 hours of acquisition) from Kunert et al, 2014 (Kunert *et al.* 2014). (**A**) The aliphatic and carbonyl regions are shown in the overview spectrum. (**B**) Extract of the alanine Cα-Cβ correlation, secondary structure regions for β-sheet (red), α-helical (black) and coil (green) from alanine signals are indicated. Solid lines correspond to the mean value with the standard deviation and dashed lines to the mean value with 2 times the standard deviation regarding (Wang 2002) (**C**) One-dimensional trace from the two-dimensional experiment at 52.1 ppm, the peak width at half-height is 0.5 ppm. The signal-to-noise ratio of this peak is 8.5.

Notes

1. Despite the evaporation by nitrogen gas flow, some liquid Triton X-100 remains in the tube. It is very important to increase the vacuum very slowly during the lyophilization step. Indeed, a sudden increase of the vacuum may induce a large amount of foam (from the Triton X-100) and the preparation can be lost in the freeze-dryer.
2. A direct re-suspension of the lipids with the buffer can induce the formation of salt precipitates and ultimately of lipids-salt aggregates.
3. The choice of the LPR is protein dependent but should be quite low for solid-state NMR studies. It is important to note that the protein activity can be influenced by the LPR, for example the activity of BmrA with LPR 0.5 is 1.3 $\mu mol.min^{-1}.mg^{-1}$ (Kunert *et al.* 2014) and 6.5 $\mu mol.min^{-1}.mg^{-1}$ at LPR of 20 (Steinfels *et al.* 2004).
4. Despite the long incubation times, the preparation remains stable as shown by activity tests and NMR spectra. The Bio-Beads® were placed outside the dialysis bag and the volume of the dialysis buffer was adapted according to the protein quantity and in keeping with 5L per 10 mg of BmrA.
5. The centrifugation speed depends equally on the protein.
6. The chosen spinning frequency depends on the magnetic field; here we used 17 500 Hz in order to avoid the superposition of the carbonyl spinning sidebands with the aliphatic region of the spectrum.

Acknowledgement

The authors would like to thank Thomas Wiegand for critical reading of the manuscript and Lauriane Lecoq for her contribution to the spectroscopy. This work was supported by the CNRS and by grants from the French ANRS (France Recherche, Nord & Sud, Sida-HIV et Hépatites), an autonomous agency at INSERM, France, and the ANR (ANR-14-CE09-0024B).

References

Bermel, W., Felli, I.C., Kümmerle, R. & Pierattelli, R. (2008) 13C Direct-detection biomolecular NMR. *Concepts in Magnetic Resonance Part A* 32A, 183–200.

Bligh, E.G. & Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology* 37, 911–917.

Böckmann, A., Gardiennet, C., Verel, R., Hunkeler, A., Loquet, A., Pintacuda, G., Emsley, L., Meier, B.H. & Lesage, A. (2009) Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45, 319–327.

Cavanagh, J., Fairbrother, W.J., Arthur G Palmer, I., Skelton, N.J. & Rance, M. (2010) *Protein NMR Spectroscopy*. Academic Press.

Chifflet, S., Torriglia, A., Chiesa, R. & Tolosa, S. (1988) A method for the determination of inorganic phosphate in the presence of labile organic phosphate and high concentrations of protein: application to lens ATPases. *Analytical biochemistry* 168, 1–4.

Comellas, G., Lemkau, L.R., Nieuwkoop, A.J., Kloepper, K.D., Ladror, D.T., Ebisu, R., Woods, W.S., Lipton, A.S., George, J.M. & Rienstra, C.M. (2011) Structured Regions of α-Synuclein Fibrils Include the Early-Onset Parkinson's Disease Mutation Sites. *Journal of molecular biology* 411, 881–895.

Das, N., Murray, D.T. & Cross, T.A. (2013) Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state NMR samples. *Nature Protocols* 8, 2256–2270.

Demers, J.-P., Habenstein, B., Loquet, A., Kumar Vasa, S., Giller, K., Becker, S., Baker, D., Lange, A. & Sgourakis, N.G. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nature Communications* 5, 4976–28.

Ding, Y., Fujimoto, L.M., Yao, Y. & Marassi, F.M. (2015) Solid-state NMR of the Yersinia pestis outer membrane protein Ail in lipid bilayer nanodiscs sedimented by ultracentrifugation. *Journal of Biomolecular NMR* 61, 275–286.

Emami, S., Fan, Y., Munro, R., Ladizhansky, V. & Brown, L.S. (2013) Yeast-expressed human membrane protein aquaporin-1 yields excellent resolution of solid-state MAS NMR spectra. *Journal of Biomolecular NMR* 55, 147–155.

Fielding, L. (2007) NMR methods for the determination of protein–ligand dissociation constants. *Curr Top Med Chem. 2003, 39-53.*

Folch, J., Lees, M. & Sloane Stanley, G.H. (1957) A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry* 226, 497–509.

Franks, W.T., Linden, A.H., Kunert, B., van Rossum, B.-J. & Oschkinat, H. (2012) Solid-state

magic-angle spinning NMR of membrane proteins and protein–ligand interactions. *European Journal of Cell Biology* 91, 340–348.

Fung, B.M., Khitrin, A.K. & Ermolaev, K. (2000) An improved broadband decoupling sequence for liquid crystals and solids. *Journal of Magnetic Resonance* 142, 97–101.

Gath, J., Bousset, L., Habenstein, B., Melki, R., Böckmann, A. & Meier, B.H. (2014) Unlike Twins: An NMR Comparison of Two α-Synuclein Polymorphs Featuring Different Toxicity P. van der Wel (Ed). *PLoS ONE* 9, e90659–11.

Gibby, M., Pines, A. & Waugh, J. (1974) Proton-enhanced nuclear induction spectroscopy.

Gottlieb, H.E., Kotlyar, V. & Nudelman, A. (1997) NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *The Journal of Organic Chemistry* 62, 7512–7515.

Griffiths, K.K. & Setlow, P. (2009) Effects of modification of membrane lipid composition on Bacillus subtilissporulation and spore properties. *Journal of Applied Microbiology* 106, 2064–2078.

Habenstein, B., Bousset, L., Sourigues, Y., Kabani, M., Loquet, A., Meier, B.H., Melki, R. & Böckmann, A. (2012) A Native-Like Conformation for the C-Terminal Domain of the Prion Ure2p within its Fibrillar Form. *Angewandte Chemie International Edition* 51, 7963–7966.

Harris, R.K., Becker, E.D., Cabral de Menezes, S.M., Granger, P., Hoffman, R.E. & Zilm, K.W. (2008) Further conventions for NMR shielding and chemical shifts IUPAC recommendations 2008. *Solid State Nuclear Magnetic Resonance* 33, 41–56.

Hediger, S., Meier, B.H. & Ernst, R.R. (1995) Adiabatic passage Hartmann-Hahn cross polarization in NMR under magic angle sample spinning. *Chemical Physics Letters* 240, 449–456.

Holloway, P.W. (1973) A simple procedure for removal of Triton X-100 from protein samples. *Analytical biochemistry* 53, 304–308.

Kaur, H., Lakatos, A., Spadaccini, R., Vogel, R., Hoffmann, C., Becker-Baldus, J., Ouari, O., Tordo, P., Mchaourab, H. & Glaubitz, C. (2015) The ABC exporter MsbA probed by solid state NMR – challenges and opportunities. *Biological chemistry* 396, 1135–1149.

Kijac, A.Z., Li, Y., Sligar, S.G. & Rienstra, C.M. (2007) Magic-Angle Spinning Solid-State NMR Spectroscopy of Nanodisc-Embedded Human CYP3A4 †. *Biochemistry* 46, 13696–13703.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Lange, V., Becker-Baldus, J., Kunert, B., van Rossum, B.-J., Casagrande, F., Engel, A., Roske, Y., Scheffel, F.M., Schneider, E. & Oschkinat, H. (2010) A MAS NMR Study of the Bacterial ABC Transporter ArtMP. *ChemBioChem* 11, 547–555.

Loquet, A., Sgourakis, N.G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S. & Lange, A. (2012) Atomic model of the type III secretion system needle. *Nature*, 1–13.

Lu, G.J., Tian, Y., Vora, N., Marassi, F.M. & Opella, S.J. (2013) The Structure of the Mercury Transporter MerF in Phospholipid Bilayers: A Large Conformational Rearrangement Results from N-Terminal Truncation. *Journal of the American Chemical Society* 135, 9299–9302.

Luckgei, N., Schütz, A.K., Bousset, L., Habenstein, B., Sourigues, Y., Gardiennet, C., Meier, B.H., Melki, R. & Böckmann, A. (2013) The Conformation of the Prion Domain of Sup35 p in Isolation and in the Full-Length Protein. *Angewandte Chemie International Edition* 52, 12741–12744.

Morag, O., Sgourakis, N.G., Baker, D. & Goldbourt, A. (2015) The NMR–Rosetta capsid model of M13 bacteriophage reveals a quadrupled hydrophobic packing epitope. *Proceedings of the National Academy of Sciences* 112, 971–976.

Morris, G.A. & Freeman, R. (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* 101, 760–762.

Prosser, R.S., Evanics, F., Kitevski, J.L. & Al-Abdul-Wahid, M.S. (2006) Current Applications of Bicelles in NMR Studies of Membrane-Associated Amphiphiles and Proteins †,‡. *Biochemistry* 45, 8453–8465.

Rigaud, J.L., Pitard, B. & Levy, D. (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochimica et biophysica acta* 1231, 223–246.

Sackett, K., Nethercott, M.J., Zheng, Z. & Weliky, D.P. (2014) Solid-State NMR Spectroscopy of the HIV gp41 Membrane Fusion Protein Supports Intermolecular Antiparallel β Sheet Fusion Peptide Structure in the Final Six-Helice Bundle State. *Journal of molecular biology* 426, 1077–1094.

Sborgi, L., Ravotti, F., Dandey, V.P., Dick, M.S., Mazur, A., Reckel, S., Chami, M., Scherer, S., Huber, M., Böckmann, A., Egelman, E.H., Stahlberg, H., Broz, P., Meier, B.H. & Hiller, S. (2015) Structure and assembly of the mouse ASC inflammasome by combined NMR spectroscopy and cryo-electron microscopy. *Proceedings of the National Academy of Sciences* 112, 13237–13242.

Seddon, A.M., Curnow, P. & Booth, P.J. (2004) Membrane proteins, lipids and detergents: not just a soap opera. 1666, 105–117.

Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A. & Jault, J.-M. (2004) Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in Bacillus subtilis†. *Biochemistry* 43, 7491–7502.

Stejskal, E.O., Schaefer, J. & Waugh, J.S. (1977) Magic-angle spinning and polarization transfer in proton-enhanced NMR. *Journal of Magnetic Resonance* 28, 105–112.

Stevens, T.J., Fogh, R.H., Boucher, W., Higman, V.A., Eisenmenger, F., Bardiaux, B., van Rossum, B.-J., Oschkinat, H. & Laue, E.D. (2011) A software framework for analysing solid-state MAS NMR data. *Journal of Biomolecular NMR* 51, 437–447.

Takegoshi, K., Nakamura, S. & Terao, T. (2001) 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* 344, 631–637.

Takegoshi, K., Nakamura, S. & Terao, T. (2003) [sup 13]C–[sup 1]H dipolar-driven [sup 13]C–[sup 13]C recoupling without [sup 13]C rf irradiation in nuclear magnetic resonance of rotating solids. *The Journal of Chemical Physics* 118, 2325.

Tang, M., Nesbitt, A.E., Sperling, L.J., Berthold, D.A., Schwieters, C.D., Gennis, R.B. & Rienstra, C.M. (2013) Structure of the Disulfide Bond Generating Membrane Protein DsbB in the Lipid Bilayer. *Journal of molecular biology* 425, 1670–1682.

TAUSSKY, H.H. & SHORR, E. (1953) A microcolorimetric method for the determination of inorganic phosphorus. *Journal of Biological Chemistry* 202, 675–685.

Ulrich, E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannidis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C.F., Tolmie, D.E., Kent Wenger, R., Yao, H. & Markley, J.L. (2007) BioMagResBank. *Nucleic Acids Research* 36, D402–D408.

Van Melckebeke, H., Wasmer, C. & Lange, A. (2010) Atomic-resolution three-dimensional structure of HET-s (218– 289) amyloid fibrils by solid-state NMR spectroscopy. *of the American Chemical Society* 132, 13765-75

Wang, Y. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science* 11, 852–861.

Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A.B., Riek, R. & Meier, B.H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319, 1523–1526.

Williamson, P.T.F. (2009) Solid-state NMR for the analysis of high-affinity ligand/receptor interactions. *Concepts in Magnetic Resonance Part A* 34A, 144–172.

Wishart, D.S. & Sykes, B.D. (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *Journal of Biomolecular NMR* 4, 171–180.

Yannoni, C.S. (1982) High-resolution NMR in solids: The CPMAS experiment. *Accounts of Chemical Research* 15, 201–208.

Zech, S.G., Olejniczak, E., Hajduk, P., Mack, J. & McDermott, A.E. (2004) Characterization of Protein–Ligand Interactions by High-Resolution Solid-State NMR Spectroscopy. *Journal of the American Chemical Society* 126, 13948–13953.

Chapter VI

Membrane protein reconstitution strategies for solid-state NMR studies

The work presented in this chapter was done in collaboration with Beat H. Meier, Pierre Falson, Jean-Michel Jault. This Chapter was accepted with the following reference:

Gradient reconstitution of membrane proteins for solid-state NMR studies

Denis Lacabanne, Alons Lends, Clément Danis, Britta Kunert,Marie-Laure Fogeron, Vlastimil Jirasko, Claire Chuilon,Lauriane Lecoq, Cedric Orelle, Vincent Chaptal, Pierre Falson, Jean-Michel Jault, Beat H. Meier, and Anja Böckmann

J Biomol NMR. 2017

Contents

Abstract

We here adapted the GRecon method used in electron microscopy studies for membrane protein reconstitution to the needs of solid-state NMR sample preparation. We followed in detail the reconstitution of the ABC transporter BmrA by dialysis as a reference, and established optimal reconstitution conditions using the combined sucrose/cyclodextrin/lipid gradient characterizing GRecon. We established conditions under which quantitative reconstitution of active protein at low lipid-to-protein ratios can be obtained, and also how to upscale these conditions in order to produce adequate amounts for NMR. NMR spectra recorded on a sample produced by GRecon showed a highly similar fingerprint as those recorded previously on samples reconstituted by dialysis. GRecon sample preparation presents a gain in time of nearly an order of magnitude for reconstitution, and shall represent a valuable alternative in solid-state NMR membrane protein sample preparation.

Introduction

Membrane proteins represent an important class of molecules in living organisms, and are key players in the communication between cells, the transport of different metabolites inside or outside cells and signal transduction. A genome-wide analysis indeed predicts that 20 to 30% of all open reading frame encode for membrane proteins (Wallin & Heijne 1998); they indeed represent the largest class of drug targets (Arinaminpathy *et al.* 2009) (Bakheet & Doig 2009). Paradoxically, despite their importance, they are significantly under-represented in the protein databank (www.rcsb.org) with only 2% of the structures therein (with 2150 membrane proteins structures, including 680 unique structures; http://blanco.biomol.uci.edu/mpstruc/). While most structural studies of membrane proteins were carried out by x-ray crystallography, electron microscopy and NMR are increasingly attractive near-atomic and atomic-resolution approaches respectively to study membrane proteins in a lipid environment, and possibly near physiological temperatures for NMR. Due to the anisotropy of the lipid-inserted protein phase, the NMR methods are usually characterized as solid-state NMR. It should however be kept in mind that these phases are investigated near room temperature and behave dynamically, as *in vivo*, and are far from being a real solid.

For both solid-state NMR and cryo-EM, membrane reconstitution of detergent-solubilized proteins is a central step on which further data analysis relies crucially(Wang & Ladizhansky 2014). In addition, reconstitution of membrane proteins is often important to study their properties in a more physiological environment. For instance, the ATPase activity of the maltose ABC transporter is insensitive to maltose in detergent, whereas the activity regulated by maltose when the transporter is reconstituted (Alvarez *et al.* 2015). A common approach consists in adding the protein (dissolved in detergent I) to liposomes destabilized/solubilized with the same or different detergent (detergent II) (Rigaud *et al.* 1995). These destabilized/solubilized liposomes are prone to protein insertion if detergent I is removed using either dialysis, chromatography (Seddon *et al.* 2004), addition of non-polar polystyrene beads (Rigaud *et al.* 1997) or cyclodextrin (Degrip *et al.* 1998). Identifying optimal conditions is often a lengthy process that aims at finding the best ratios of the involved reagents in order to remove the detergents. The goal is to maximize insertion of the protein into the lipid environment to form proteoliposomes and to minimize aggregation.

A novel approach, the GRecon method, developed a few years ago by Kühlbrandt and coworkers in the context of electron microscopy sample preparation (Althoff *et al.* 2012), allows for gentle and progressive encounter between the different reagents, while at the same time reducing the total experiment time by about an order of magnitude when compared to dialysis. GRecon consists in the use of ultracentrifugation of the detergent-solubilized membrane protein on a combined sucrose, lipid (destabilized with detergent II) and cyclodextrin (chosen to primarily absorb detergent I) gradient for reconstitution. The protein during ultracentrifugation migrates down the gradient encountering an increasing amount of cyclodextrin. The cyclodextrin is chosen such as to absorb the detergent I molecules stabilizing the membrane protein with higher affinity than detergent II (Degrip *et al.* 1998) needed to destabilize the lipids. Different cyclodextrins can be used for different detergents including N-Dodecyl-ß-D-maltoside (DDM), HECAMEG, $C_{12}E_{10}$ and Nonylglucose (Degrip *et al.* 1998). Remaining detergent II is mainly removed by dilution of the proteoliposomes after reconstitution. The major advantage of this method is the progressive increase or decrease in reactants, making it particularly attractive for the lipid reconstitution of large and/or fragile membrane proteins (Althoff *et al.* 2012).

Inspired by this work, we examined here the possibility to adapt this method to solid-state NMR membrane protein sample preparation where considerably larger protein amounts are needed, and close-to-quantitative recovery is a major issue. Also, typically smaller LPR (lipid-to-protein ratios) are preferred in NMR for sensitivity reasons. We here compared GRecon and reconstitution by dialysis of BmrA, a multidrug ABC transporter from *Bacillus subtilis* (Steinfels *et al.* 2004) that we used as a model here, as efficient reconstitution of this protein can be easily followed by its ATPase activity. ABC transporters can translocate a variety of molecules by coupling drug efflux with ATP hydrolysis. They are found in all forms of life and they are involved in drug resistance. We focus on the ABC transporter BmrA from *B. subtilis* as a model, which is a homologue of the human P-gp involved in drug resistance in cancer. We show how GRecon can be up scaled for solid-state NMR needs, and results in a reduction of sample preparation time of about an order of magnitude. We also evaluate the activity and superstructures of the resulting proteoliposomes, and demonstrate that NMR spectra recorded on GRecon samples yield very similar high-quality spectra to those observed after dialysis.

Material and methods

1. Production, purification and reconstitution of BmrA in B. subtilis lipids

Production, purification, and reconstitution of BmrA were done as previously described(Kunert *et al.* 2014). Briefly, uniformly [^{13}C, ^{15}N] labeled BmrA was produced in M9-medium(Studier 2005) with ^{13}C-glucose (2 g/L) and ^{15}N-ammonium chloride (2 g/L) as sole sources of carbon and nitrogen. Two precultures were performed before the main culture. The first one was done in 50 mL M9-medium, incubated at 37 °C until an OD_{600nm} of 1.5. Then the second pre-culture, 150 mL of M9-medium, was inoculated with the first pre-culture (starting OD_{600nm} of 0.2) and incubated over night at 25 °C until an OD_{600nm} of 2.5. Finally, the 150 mL from the last pre-culture were used to inoculate 850 mL M9-medium of the main culture. The latter was incubated at 25 °C and the expression was induced with 0.7mM IPTG at an OD_{600nm} between 0.6-0.7. The cells were harvested 6 hours after induction by centrifugation at 6'000 x *g*. The cells were suspended in buffer (50 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1 mM DTT with benzonase and EDTA-free protease inhibitor cocktail) and disrupted by high fluid pressure using a Microfluidizer®. The solution was centrifuged at 15'000 x *g* during one hour and the obtained supernatant was again centrifuged at 200'000 x *g* during one hour. The sedimented membranes were suspended and collected in 50 mM Tris-HCl pH 8.0, 1 mM EDTA, 300 mM sucrose. For the purification of the protein, the membranes were diluted at 2 mg/mL and incubated during one hour with 1% DDM (m/v). The insoluble material was sedimented by centrifugation at 100'000 x *g* for one hour and discarded. The soluble material was loaded on to a Ni-NTA agarose column equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 15% glycerol, 10 mM imidazole and 0.2% DDM). The column was washed using equilibration buffer with successively 0.5 M NaCl, 30 mM imidazole, 40 mM imidazole and 250 mM imidazole was used for the elution of the protein. The imidazole was removed using PD10 desalting columns equilibrated with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol and 0.2% DDM. For reconstitution by dialysis, the protein was diluted 4 times with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol, mixed and incubated one hour with a homemade preparation of *B. subtilis* lipids solubilized in Triton X-100 (Kunert *et al.* 2014). The quantity of lipids mixed follows a determined lipid-to-protein ratio (LPR) of 0.1-20 (m:m). The DDM

and Triton X-100 were removed by dialysis with Bio-beads (in the dialysis solution) during 9 days. The quantity of Bio-beads used was three times the theoretical amount required to absorb the DDM and Triton X-100 present in the solution. The theoretical amount of the adsorption capacities from Bio-beads was previously described and correspond to 105 mg DDM per 1 g Bio-beads(Rigaud *et al.* 1995) and 70 mg Triton X-100 per 1 g Bio-beads (Bio-Rad, user manual).

2. Gradient Reconstitution and upscale thereof

The 30-60% sucrose (m/v) gradients used for the small-scale reactions were performed in SW60 ultracentrifuge tubes (Seton - Open-top centrifuge tubes polyclear™; 11 x 60 mm, volume 4 mL). 1.9 mL of 30 % sucrose in 50 mM Tris-HCl pH 8, 100 mM NaCl were deposited on 1.9 mL of 60 % sucrose in 50 mM Tris-HCl pH 8, 100 mM NaCl, 0-3 % (m/v) α-cyclodextrins, and 0.08 mg/mL *B. subtilis* lipids destabilized with Triton™ X-100 (Sigma) ratio 1:1 (m/m). The continuous gradient was obtained using a Gradient Master™ (BioComp Instuments) with the following parameters: angle 86°, 18 rpm, during 65 sec. 300 µg of protein (1 mg/mL) corresponding to 4.5 nmol of BmrA monomer were loaded on the top of the gradient and the gradient was ultracentrifuged at 175'000 x *g* during 15 h. Each gradient was photographed, and 19 layers of 200 µL were collected. They were analyzed by SDS-PAGE (10 % acrylamide) and a BCA assay (Pierce™ BCA Assay Kit, Thermo Scientific), as well as an ATPase assay (see below) for each layer containing protein sample. The 30-60% sucrose gradients were operated in 12 SW32 ultracentrifuge tubes (Seton - Open-top centrifuge tubes polyclear™; 25 x 89 mm, volume 38.5 mL). 17 mL of 30 % sucrose in 50 mM Tris-HCl pH 8, 100 mM NaCl were deposited on 17 mL of 60 % sucrose in 50 mM Tris-HCl pH 8, 100 mM NaCl, 1 % (m/v) α-cyclodextrins, and 0.09 mg/mL *B. subtilis* lipids destabilized with Triton™ X-100 (Sigma) ratio 1:1 (m/m). The continuous gradient was generated using a Gradient Master™ (BioComp Instuments) with the following parameters, step 1: angle 50°, 25 rpm, during 6 min, step 2: angle 82°, 15 rpm, during 40 sec. 3 mg of protein (1 mg/mL) corresponding to 45 nmol protein monomer were loaded on the top of the gradient and the gradient was ultracentrifuged at 175'000 x *g* during 15 h. The gradient was photographed and 23 layers of 1.6 mL were collected and analyzed by SDS-PAGE (10 % acrylamide). The fractions containing protein were pooled, the proteoliposomes were harvested by sedimentation at 100'000 x *g* during 1 h and washed 8 times in order to remove

the excess of sucrose and α-cyclodextrins. The proteoliposomes were transferred into a 3.2 mm NMR rotor by centrifugation at 100'000 x *g* with a custom-made filling tool (Böckmann *et al.* 2009).

3. Quantification of DDM by Mass Spectrometry

DDM was quantified as previously described (Chaptal *et al.* 2017). Briefly, four different ratios ^{1}H-DDM:^{2}H-DDM were measured to establish a standard curve: 1:5, 1:2, 1:1 and 3:1. The concentration of ^{2}H-DDM was constant and equal to 0.025% (m/v). The different ratios were determined experimentally by MALDI-TOF using the peak intensities from ^{1}H-DDM and ^{2}H-DDM. To determine the ^{1}H-DDM:^{2}H-DDM ratios during the reconstitution, aliquots were collected directly from the dialysis bag at the following times: 0 h, 0.5 h, 1 h, 2 h, 4 h, 8 h, 1 d, 2 d, 3 d, 4 d, 5 d, 6 d, 7 d, 8 d and 9 d. A dilution by 2.5 times for the samples 0 h to 2 d was done and each aliquot were mixed with 0.025% (m/v) ^{2}H-DDM. The different ^{1}H-DDM:^{2}H-DDM ratios were determined by MALDI-TOF and reported to the standard curve in order to determine the concentration of ^{1}H-DDM. In order to avoid the pipetting error and to increase the accuracy of the method, all volumes used were weighted with a precision scale. All experiments were done in triplicate. Each sample was mixed in 10:1 ratio with matrix (10 g/L (m/v) DHB (2,5-dihydroxybenzoic acid) in water and 1 g/L (m/v) NaI acetone), 1 µL of mixture was deposited on the MALDI-target and air-dried. Mass spectrometry experiments were performed with a Voyager-DE Pro MALDI-TOF (AB Sciex, Framingham, MA) Mass Spectrometer with a nitrogen UV laser (λ = 337 nm, 3 ns pulse). The instrument was operated in reflectron-positive ion mode (mass accuracy: 0.008%) with an accelerating potential of 5 kV.

4. ATPase activity assay

To determine the ATPase activity, two assays were used: an ATP/NADH coupled ATPase activity assay, and a colorimetric ATPase activity assay. The colorimetric assay allowed determining the concentration of inorganic phosphate without dilution engendered by the ATP-regeneration buffer. This assay was useful to determine the detergent effect on the activity when the detergent concentration was unknown.

Colorimetric ATPase activity assay: We followed the method described by Doerrler and Raetz(Doerrler & Raetz 2002). Briefly, 250 µL of reconstituted BmrA (0.2 mg/mL) in 50mM Tris-HCl, 200mM NaCl, 5% glycerol was incubated 5 min at 37 °C. The ATPase reaction was started by the addition of 5 µL ATP-Mg^{2+} from a stock solution of 500 mM. 25 µL of the reaction medium were pipetted and mixed with 25 µL 12% SDS. 25 µL of 6% ascorbic acid in 1M HCl and 2% ammonium molybdate were added, and the colorimetric reaction was incubated 5 min at room temperature. 50 µL of 1 M HCl were then added in order to slow down drastically the colorimetric reaction. The absorbance at 850 nm was measured and the quantity of orthophosphate was determined using a Na_2HPO_4 standard (0-50 µmol).

NADH coupled ATPase activity assay: The ATPase activity of BmrA was quantified using a coupled enzyme assay. ATP is regenerated by a pyruvate kinase from phosphoenolpyruvate and ADP. The regeneration system was coupled with a second system using lactate dehydrogenase and NADH. For ATP one molecule of NADH was transformed into NAD^+. The absorbance of NADH was followed at 340 nm during the reaction (spectrophotometer SAFAS FLX-Xenius®). The assay was done at 37 °C in a 96 well microplate (UV-star® Microplate, 96 well – Greiner Bio-one) pre-incubated at 37 °C. In each microplate well, 0.1 µg of reconstituted BmrA (LPR 0-20) was deposited or 0.1 µg of reconstituted BmrA (LPR 0.5) was mixed with an increasing concentration of N-Dodecyl-ß-D-maltoside (DDM), 0-2 % (m/v) in ATPase buffer for both cases (50 mM HEPES-KOH pH 8, 4mM phosphoenolpyruvate, 60 µg/mL pyruvate kinase, 32 µg/mL lactate dehydrogenase, 10 mM $MgCl_2$, 0.6 mM NADH) qsq 200 µL. The reaction was started by the addition of 5 mM ATP and was followed during 20 min. The ATP activity was determined using the NADH extinction coefficient (6220 L/mol/cm), the radius of a well (3.3 mm) and the sample volume (200 µL) allow us to determine a path length of 0.585 cm.

5. Solid-state NMR experiments

NMR spectra were acquired on a Bruker Avance spectrometer operating at 850 MHz ^1H frequency using a 3.2 mm triple-resonance E-free probe. The details of NMR experiments are given in the Table S1. The sample temperature was set to 278 K using the water resonance frequency. The 2D spectra were processed using TOPSIN 3.5 with a shifted cos^2 (SSB=3) function and analyzed using CcpNMR software(Stevens *et al.* 2011).

Results and Discussion

1. Reconstitution of BmrA in lipids by dialysis

BmrA has been shown to be stable (protected against degradation) and active (ATPase competent) when reconstituted in lipids extracted from *B. subtilis* (Kunert *et al.* 2014). Also NMR spectra do not reveal differences even after three years. We first implemented this previously described approach, which includes the addition of Bio-beads to the dialysis buffer to remove the DDM (detergent I) and Triton X-100 (detergent II, 10:1 ratio to lipids) (Kunert *et al.* 2014). For BmrA, a lipid-protein ratio of 0.5 m/m (or 35 mol/mol) allows to obtain samples with high stability over a time period of months (Kunert *et al.* 2014). Lower LPRs could in principle be used, but we did not test them yet to prepare an NMR sample. When such low LPRs are used, very slow detergent removal is beneficial in order to avoid aggregation. We used dialysis during 9 days, with Bio-beads present in the dialysis buffer only. For comparison, direct addition of Bio-beads to prepare an LPR = 20 sample needs only 3 h for full detergent removal (Orelle *et al.* 2003), but leads to protein aggregation at LPR = 0.5.

In order to establish the time scale of the reconstitution, we followed the ATPase activity of BmrA and determined remaining detergent I concentrations by mass spectrometry (MS) (Chaptal *et al.* 2017) (**Figure 41A**). The calibration curve is presented in the supplementary data (**Figure S1**). After two days of incubation, half of the initial quantity of the detergent I was depleted and its concentration reached a minimum after five to six days as determined by mass spectrometry (**Figure S2**). A second addition of Bio-beads, after 6 days, and three more days of incubation, did not remove additional detergent. In this reaction, detergent I was five times more concentrated than detergent II, for which no signal was detected already after four days of dialysis. In parallel, the ATPase activity of BmrA was monitored (**Figure 41A**). It has been shown that reconstituted proteins may be substantially more active than detergent-stabilized ones (Yang *et al.* 2014),(Orelle *et al.* 2008). Experimentally, the depletion of DDM was in the first place accompanied by a decrease in ATPase activity (after one to two days of incubation), and only increases when the DDM concentration falls below its CMC (0.009% (Anatrace)). This phenomenon was also described for a progressive delipidation of sarcoplasmic reticulum Ca^{2+}-ATPase by DDM and $C_{12}E_8$ and is so far

unexplained (De Foresta *et al.* 1989). No transport activity was measured as a function of LPR, as inverted membrane vesicles are needed for this (Steinfels *et al.* 2004).

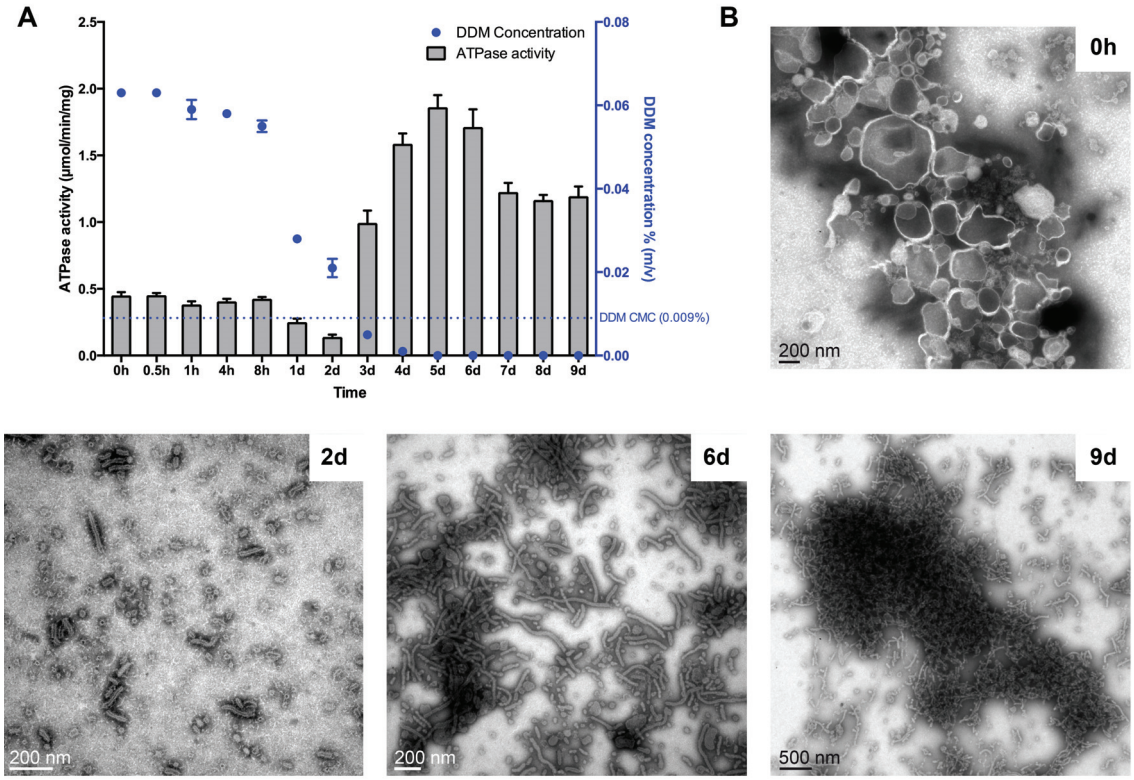


Figure 41: **Reconstitution of BmrA followed by mass spectrometry, biochemical assay, and EM** (**A**) ATPase activity of BmrA (grey histogram) and DDM concentration (blue dots) during lipid reconstitution. ATPase activities were monitored by the quantification of inorganic phosphate resulting from ATP hydrolysis by BmrA. DDM was quantified by MALDI-TOF mass spectrometry using ^{2}H-DDM as an internal standard. In parallel, each sample was observed by EM (**B**). The EM pictures at times equals to 0 h, 2 d, 6 d, 9 d are presented, and all data points are shown in **Figure S3**.

Reconstitution was also followed by EM (**Figure 41B and S3**). BmrA is known to change the structure of liposomes (*B. subtilis* lipids or EPC/EPA) to form ring-shaped protein/lipid assemblies that combine at low detergent concentrations and low LPR into annular or tubular superstructures (Fribourg *et al.* 2014; Kunert *et al.* 2014; Orelle *et al.* 2008). The liposome-protein mixture before the initiation of the detergent depletion by dialysis is shown in **Figure 41B** at 0h. After two days of incubation (2d), a mixture of rings and tubular structures appears, the latter of around 200 nm length with a diameter of 20 nm. The length of the tubular structures increases during the detergent depletion and can reach more

than 1 µm after six days of incubation. At the end of the process, the structures form a conglomerate of tubular particles with different sizes (9d).

2. GRecon

When using GRecon for NMR sample preparation, it is important that the reconstitution is as quantitative as possible, in order not to lose isotope-labeled protein, and to obtain the desired protein-to-lipid ratio in the rotor. In this context, the following parameters have to be determined:

(i) The sucrose gradient to be used as a function of the chosen lipid/protein ratio.

(ii) The amount of cyclodextrin needed per mole unit of protein for removing the detergent by a precipitation test.

(iii) The optimum amount of cyclodextrin in the gradient.

2.1. LPR and sucrose gradient

The amount of lipids n_l to be added to the initial bottom layer of the gradient depends on the desired LPR, X, and the amount of protein, n_p. The amount of lipids is calculated as $n_l = \frac{X n_p M_p}{M_l}$ where M_p and M_l denote the molecular weights of protein and lipids, respectively. After deciding on the LPR, the sucrose gradient has to be adjusted so that the protein ultimately ends up in the bottom third of the gradient once it is fully reconstituted. If the sucrose gradient is not optimal, the protein may stay too high in the gradient and not encounter enough cyclodextrin to complete its reconstitution. The lower the LPR, the further the protein will in general migrate in the gradient (Kunert *et al.* 2014). Even if the protein does not migrate all the way to the bottom of the centrifuge tube, it will encounter all lipids added as they tend to move up during centrifugation. Lipids from *B. subtilis* were mixed with Triton-X100 at a ratio of 1 (m/m).

In the original work of Kühlbrandt and coworkers (Althoff *et al.* 2012), LPRs of 3 to 4 were used in protein reconstitution for electron microscopy. For solid-state NMR sample preparation, we reduced the LPR in the BmrA sample to LPR 0.5, for signal-to-noise reasons, which has been shown to still yield high-quality NMR spectra for BmrA (Kunert *et al.* 2014). Under this condition, BmrA reconstituted by dialysis migrates to 47 % sucrose (**Figure 36A**).

In addition, the gradient should be chosen so that the protein solubilized in detergent stays

towards the top of the gradient. BmrA in DDM migrates to 32 % sucrose (**Figure 42A**) and DDM did not enter the gradient (**Figure S4)**. (The migration behavior of detergent in sucrose gradients has also been described by Hauer and co-workers who used this approach to remove excess of detergent in order to increase the quality of cryo-microscopy data (Hauer *et al.* 2015).) The gradient was thus chosen as 30-60 %, and was produced using a BioComp gradient maker device. The gradient was approximately linear as verified by coomassie blue staining (**Figure S5**) (Coombs & Watts 1985).

2.2. Precipitation test

In this second step, a precipitation test establishes how much cyclodextrin is needed to strip the detergent off the protein. An experimental procedure is then used in step (iii) to establish where this amount should be present in the gradient in order to allow for quantitative reconstitution.

While other cyclodextrins absorb detergent as well, we here used α-cyclodextrin, as it absorbs easily DDM (detergent I) and shows lower affinity for triton X-100 (detergent II) (Degrip *et al.* 1998). For the precipitation test, dissolved protein (in detergent I) was mixed with increasing amounts of α-cyclodextrin (Althoff *et al.* 2012) until complete precipitation of the protein was observed. This total amount of α-cyclodextrin (in mole) is called n_t. With known buffer detergent and protein amounts, the precipitation test allows an estimation of the quantity (in mole) of detergent bound to the protein (n_d) under the assumption that 1 cyclodextrin molecule binds one detergent molecule[11]. The amount (in mole) of protein is denoted by n_p, and we assume that one molecule represents one monomer.

Using the conditions under which BmrA was purified, i.e. n_p=0.5 nmol of protein in 27 µL of BmrA at 18.5 µM in a buffer containing n_b =105 nmol DDM (0.2 %; 3.9 mM), precipitation of the protein was complete when exceeding n_t=203 nmol cyclodextrin. (**Figure 42B).** This corresponds to a molar ratio of n_t/n_p = 406:1 α-cyclodextrin:BmrA, and n_t/n_b=203:105 α-cyclodextrin:DDM(in the buffer). A roughly twofold excess of α-cyclodextrin with respect to the DDM in the buffer seems thus necessary to remove the DDM in the sample. This is due to the fact that DDM is not only present in the buffer, but is also bound to the protein. The excess of roughly 200 nmol of α-cyclodextrin needed thus absorbs the DDM molecules which are bound to one nmol of BmrA, corresponding to n_d. This coincides well with previous estimations of 200±80 nmol (Ravaud *et al.* 2006) and 220±20 nmol (Chaptal *et al.*

2017) of DDM per nmol of BmrA. We will use in the following the approximate values of 200 detergent molecules bound to 1 nmol of BmrA (n_d/n_p = 200:1)

2.3. The cyclodextrin gradient

In this third step, we determined where the total cyclodextrin amount necessary to strip the detergent off the protein should be present in the gradient. The protein should, on its way from the top (detergent solubilized, 32 % sucrose) to the final position in the tube (membrane reconstituted, 47 % sucrose), encounter the integral amount of cyclodextrin needed to fully absorb the protein-bound detergent (**Figure 42C**). In order to determine where the protein needs to make this encounter for quantitative reconstitution, BmrA is a valuable model since one can easily follow the efficient reconstitution of the protein with its ATPase activity.

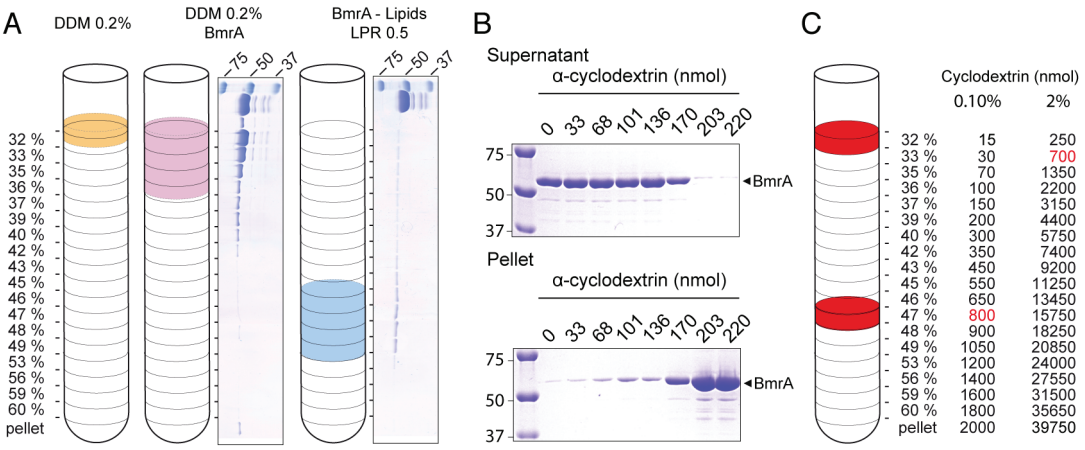


Figure 42. Migration parameters of BmrA (**A**) Migration inside the sucrose gradient of DDM, BmrA in 0.2% DDM and BmrA in lipids LPR 0.5. The top line in the gels, after the marker, corresponds to the sample load before centrifugation. (**B**) Quantity of α-cyclodextrin (nmol) needed to precipitate 0.5 nmol of BmrA by depletion of DDM. The supernatant and the pellet were obtained after centrifugation of the precipitated samples at 20'000 x *g* during 30 min. (**C**) Sketch of sucrose and α-cyclodextrin gradients corresponding to the two extreme cases of cyclodextrin concentrations (0.1 and 2%). The sucrose percentages along the tube are indicated in the left column, while the theoretical amounts of cyclodextrin along this gradient are indicated in the two right columns. The minimum quantity of α-cyclodextrin required to remove the detergent of 300 µg of BmrA and the corresponding layers in the gradient are highlighted in red. The expected needed amount of cyclodextrin should be located between these two extreme values, and they thus indicate the range of α-cyclodextrin concentrations to be tested.

The 900 nmol cyclodextrin needed to strip the detergent off the 4.5 nmol (300 µg) of BmrA deposited on the gradient is indicated for 30-60 % sucrose gradients containing 0.25%, 0.50%, 1% and 2% of α-cyclodextrin in **Figure S6**. The resulting migration profiles for these conditions, plus a control without cyclodextrin, are displayed in **Figure 43**. After ultracentrifugation, each layer of each gradient was collected and analyzed by SDS-PAGE shown next to the tubes. The ATPase activities were monitored for the layers containing protein and are shown below the tubes (**Figure 43B**).

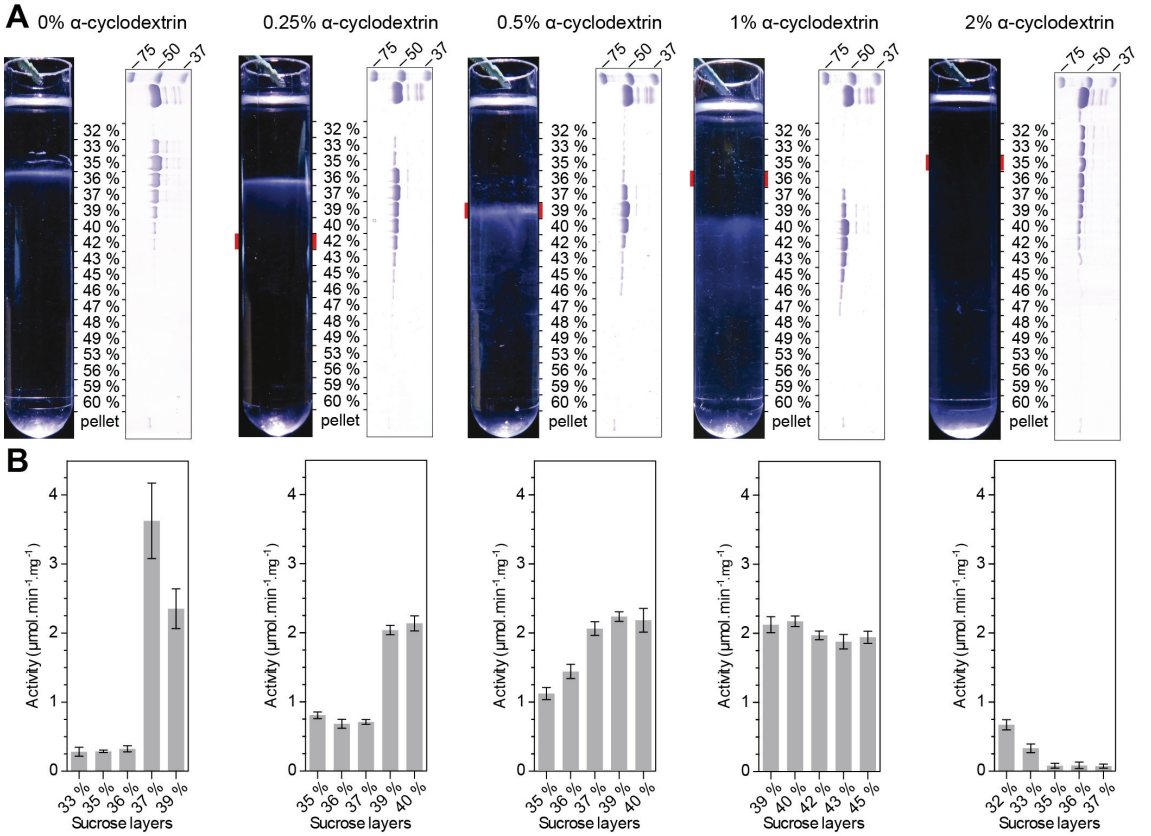


Figure 43. Determination of the best α-cyclodextrin concentration using ATPase activity readout. Gradient reconstitution of BmrA in lipids (LPR 0.5) using various α-cyclodextrin concentrations. (**A**) The image of the gradient tube, the corresponding SDS-PAGE gels for each layer of the gradient (top line in the gels, after the marker, corresponds to the sample load before centrifugation), and (**B**) ATPase activity for the layer containing a sufficient quantity of protein are shown for a α-cyclodextrin concentrations of 0 %, 0.25 %, 0.5 %, 1 % and 2 %. The approximate position in each gradient where the quantity of α-cyclodextrin necessary to fully strip the detergent is highlighted in red.

Three different migration profiles of the protein in the gradients can be distinguished. Without cyclodextrin and with the lowest α-cyclodextrin concentration (0.25%), depletion of detergent seems not efficient, and only part of the protein reconstitutes into liposomes (the part that moves the furthest along the gradient). The corresponding fractions contain little

protein but with a high ATPase activity, which, as will be shown below (**Figure S7**), is an indication that this fraction is reconstituted with a higher LPR, around 2. Interestingly, even when there is no α-cyclodextrin, BmrA can be partially reconstituted into the lipids. This might be due to partial stripping of detergent by dilution. At the highest cyclodextrin concentration (2%), no band of lipids can be observed in the centrifugation tubes, and a protein-free pellet was formed. This might indicate that α-cyclodextrin at high concentration does interact with the lipids in triton X-100 (detergent II), forming a precipitate. No reconstitution of the protein was observed, and very low ATPase activities are detected. In the intermediate range, with gradients of 0.5% and 1% α-cyclodextrin, the balance of the different concentrations seems appropriate and the ATPase activities of the collected fractions are comparable to those measured following reconstitution by dialysis, especially for the 1% gradient. The best choice of the gradient seems to be therefore 1% α-cyclodextrin, where all protein-containing fractions show active protein with a similar specific ATPase activity. One can note that this cyclodextrin concentration is the only one where the protein passes *through* the point in the gradient where the cumulated cyclodextrin quantity equals the quantity needed to strip off all detergent (red bars in **Figure 43A**, see also **Figure S6)**. Even if it might be necessary to test several cyclodextrin concentrations when applying the method to other proteins, one might consider this to be a general requirement for successful reconstitution.

The maximum rate (Vmax) of ATPase activity as well the Hill number were compared between BmrA in detergent, BmrA reconstituted in *B. subtilis* lipids by dialysis and by GRecon. Vmax values were determined to 580, 1700 and 2100 $nmol.min^{-1}.mg^{-1}$ and Hill numbers to 2.18±0.12, 1.18±0.04, 1.17±0.06 respectively (**Figure 44A)**. The Vmax of BmrA reconstituted by dialysis is comparable to previously published data for BmrA reconstituted in *B. subtilis* lipid at LPR of 0.5 [Vmax ≈ 1300 $nmol.min^{-1}.mg^{-1}$] (Kunert *et al.* 2014). However, the activity of the protein reconstituted by GRecon seems higher than after dialysis reconstitution. One can also note that the Hill number for BmrA in lipids, whether reconstituted by GRecon or by dialysis, is nearly two times smaller when compared to DDM. This strongly supports the notion that dialysis and GRecon yield the same product but detergent-solubilize protein differs significantly. This underlines the importance to study BmrA in lipid membranes. As observed also by the migration to somewhat lower sucrose concentrations in GRecon when compared to reconstitution by dialysis (**Figure**

42A), this higher activity is probably due to a slight difference in LPR. Indeed, we noticed that with increasing LPR, the activity of the protein increases (**Figure S7**) and reaches levels previously reported with a LPR of 20 (Orelle *et al.* 2003; Steinfels *et al.* 2004). Using a reference curve, we determine the LPR after GRecon to be 0.72 instead of 0.5 (**Figure 44B**). We also observed that there was no difference between reconstitution in *E. coli* lipids and *B. subtilis* lipids (**Figure S7**). Finally, the presence of residual DDM seems to be negligible in GRecon. Indeed, when plotting the ATPase activity as a function of DDM concentration, (**Figure 44C**), we observed an important decrease of the BmrA activity starting from a DDM concentration around 0.003 % DDM, no DDM was detected after GRecon (**Figure S8**).

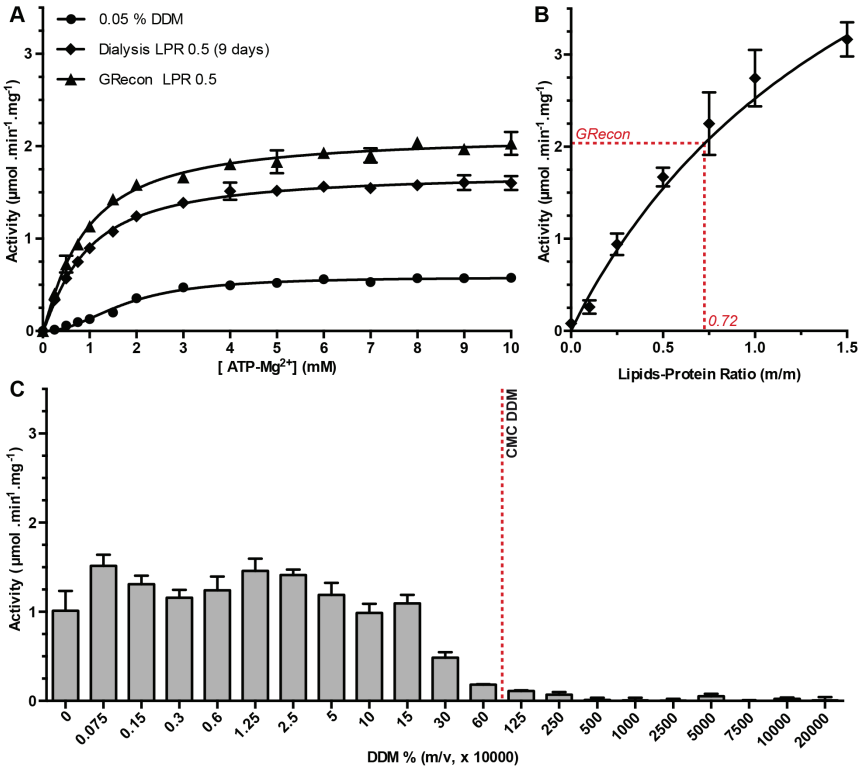


Figure 44. BmrA ATPase activities as a function of ATP-Mg^{2+} concentration and LPR. (**A**) ATPase activity of BmrA in DDM (●) and reconstituted in lipids by dialysis (◆) or GRecon (▲), as a function of ATP-Mg^{2+}. Data are shown as mean ± standard deviation of quadruplicate values. The plots were fitted using GraphPad Prism and the values obtained for the different fits for 0.05% DDM, dialysis and GRecon respectively, were as follows: Vmax= 585±9, 1712±18, 2120±40 nmol.min^{-1}.mg^{-1}; Hill number 2.18±0.12, 1.18±0.04, 1.17±0.06; $K_{0.5}$ 1.75±0.12, 0.9±0.3, 0.9±0.2 mM. (**B**) ATPase activity as a function of LPR. The full graph showing LPRs up to 20, as well as a comparison to *E. coli* lipids, are presented in the supplementary information (**Figure S7**). The LPR of protein reconstituted by GRecon was estimated to 0.72 from the ATPase activity in presence of 10 mM ATP. (**C**) Influence of increasing concentrations of DDM onto reconstituted BmrA at a LPR 0.5. Note that for convenient display, the values have been multiplied by 10000.

3. Upscaling for solid-state NMR sample preparation using GRecon method

The condition obtained previously in SW60 tubes (3.8 mL) was upscaled to SW28 tubes (37 mL). Indeed, the preparation using SW60 ultracentrifuge tubes only allows the reconstitution of 300 µg (4.5 nmol) per tube. While this would be enough for proton-detected NMR under fast magic-angle spinning(Agarwal *et al.* 2014), carbon-detected NMR requires larger amounts of sample and the method needed to be upscaled. Ten times more protein can be reconstituted using SW28 tubes where a sample of 3 mg (45 nmol) of protein was laid on the top of the gradient (**Figure S6**). All other components need to be upscaled in the same ratio, yielding then the same concentration of the gradient components as in the small-scale trial. Centrifugation was performed at 175'000 x *g* during 15 hours. After the centrifugation step, an opaque band in layers of 36% to 42% (m/v) sucrose corresponding to the protein in lipids, and confirmed by SDS-PAGE, was observed (**Figure 45A**).

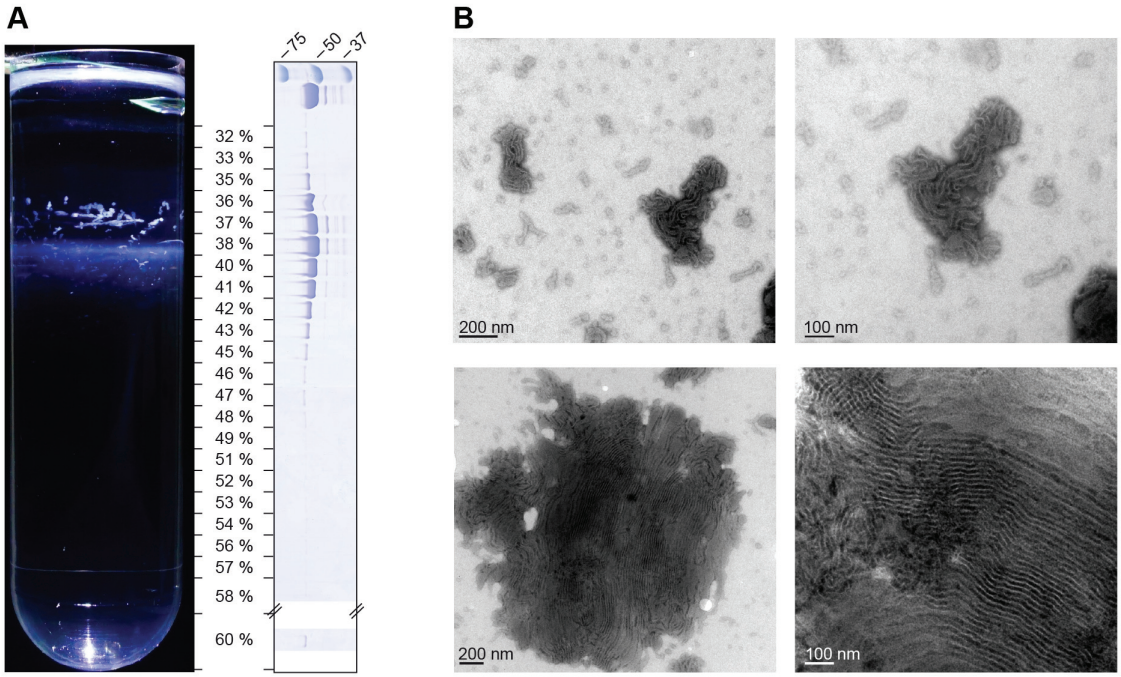


Figure 45. Reconstitution of BmrA by GRecon for a solid-state NMR sample. BmrA reconstituted in lipids by GRecon (**A**) and analyzed on a coomassie gel was characterized by TEM (**B**), which reveals the presence of the specific structural organization induced by BmrA reconstituted in the lipids. The top line in the coomassie gel, after the marker, corresponds to the sample load before centrifugation.

The layers containing the protein were collected and the material was sedimented at 100'000 x *g* during 1 h and washed 8 times in order to remove the sucrose. The protein in lipids obtained after GRecon was 1.8 mg (yield of 60 %) from one tube. The activity of the protein

in the presence of 10 mM ATP-Mg^{2+} was measured at 2.2 µmol.min^{-1}.mg^{-1}. The sample was analyzed by MALDI-TOF mass spectrometry and showed no trace of DDM or Triton X-100 (**Figure S8**). As described in the literature(Fribourg *et al.* 2014; Kunert *et al.* 2014; Orelle *et al.* 2008), tubular structures can be observed when BmrA is reconstituted in lipids with a LPR below 2. An EM analysis of two different protein-containing fractions from a 1 % α-cyclodextrin band allows the observation of tubular structures, as also observed by dialysis. Some dense structures resulting from a stacking of the tubular structures can be noticed (**Figure 45B**).

In order to fill a solid-state NMR rotor, the procedure was repeated with 12 tubes (corresponding to two centrifuge runs) for a run of 15 h resulting in a total of about 20 mg protein. When comparing this to the 6 days (144 h) needed for dialysis, this spares about an order magnitude in time. The protein pellet was loaded in a 3.2 mm solid-state NMR rotor using a rotor filling tool (Böckmann *et al.* 2009). As can be seen in **Figure 46,** the chemical-shift dispersion and line widths of the BmrA samples prepared using dialysis and GRecon look very similar. We only noticed minor differences for GRecon reconstituted sample that might be due to a slightly different LPR, as also inferred from above data.

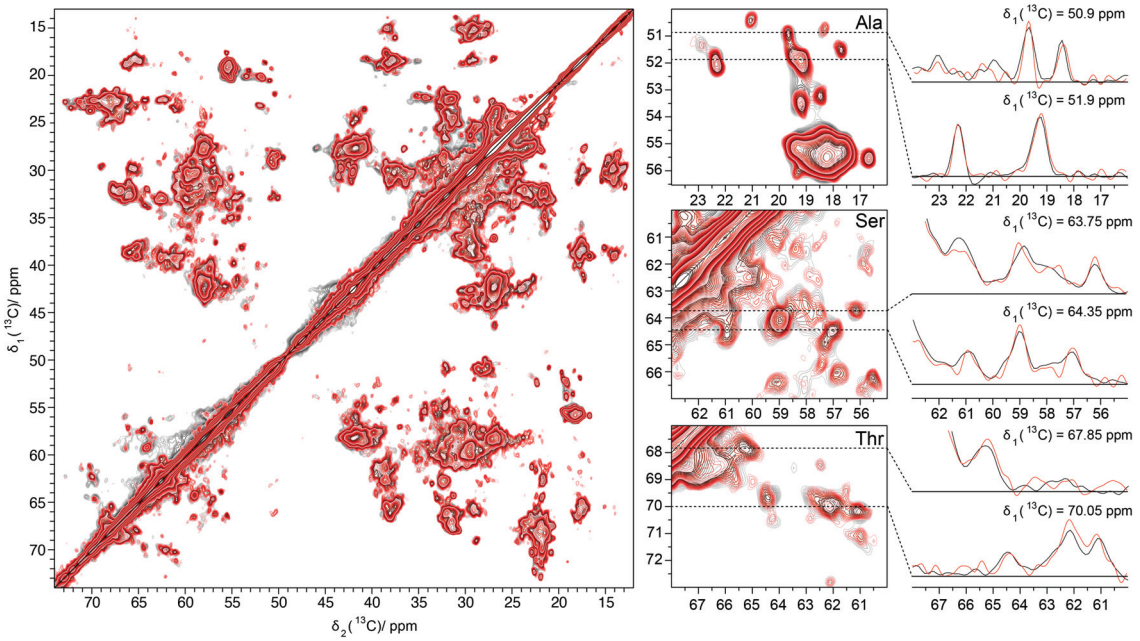


Figure 46. Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectrum of lipid-reconstituted BmrA. The aliphatic region and a zoom of the Ala, Ser and Thr regions of reconstituted BmrA are shown, with the sample resulting from GRecon in red shades overlaid on dialysis-reconstituted BmrA in grey shades. Representative 1D traces along F2 are shown for both spectra. The individual spectra are shown in **Figure S9**.

Conclusion

We here analyzed how GRecon can be applied to solid-state NMR sample preparation of the ABC transporter BmrA, and propose procedures which shall be generally applicable for the reconstitution of membrane protein NMR samples. We established optimal concentrations of lipids and cyclodextrin in the gradient, for both small- and large-scale reconstitution. The up scaling allows us to reconstitute around 20 mg of protein by two centrifuge runs. The operation reduced the reconstitution time by saving 6 days, and yielded similar NMR spectra than the dialysis method. This gain in time might be crucial in the context of proteins which prove unstable during the long reconstitution times needed when using dialysis. For both cases, the spectra displayed an optimal signal-to-noise ratio and stable samples. Compared to the dialysis reconstitution, the yield was 60 % in GRecon. The developed method shall show its highest potential when a variety of samples are produced for screening of different functional states, combined with fast MAS approaches asking for small sample quantities.

Acknowledgment

This work was supported by the French ANR (ANR-14-CE09-0024B), the LABEX ECOFECT (ANR-11-LABX-0048) within the Université de Lyon program Investissements d'Avenir (ANR-11-IDEX-0007), by the Swiss National Science Foundation (Grant 200020_159707), and the TH program of ETH Zürich (ETH-45 13-2). We acknowledge the contribution of SFR Biosciences (UMS3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) facilities and the staff of Protein Science Facility, especially Frédéric Delolme, as well as the CIQLE facility Lyon.

References

Agarwal, V., Penzel, S., Szekely, K., Cadalbert, R., Testori, E., Oss, A., Past, J., Samoson, A., Ernst, M., Böckmann, A. & Meier, B.H. (2014) De novo 3D structure determination from sub-milligram protein samples by solid-state 100 kHz MAS NMR spectroscopy. *Angewandte Chemie (International ed. in English)* 53, 12253–12256.

Althoff, T., Davies, K.M., Schulze, S., Joos, F. & Kühlbrandt, W. (2012) GRecon: A Method for the Lipid Reconstitution of Membrane Proteins. *Angewandte Chemie International Edition* 51, 8343–8347.

Alvarez, F.J.D., Orelle, C., Huang, Y., Bajaj, R., Everly, R.M., Klug, C.S. & Davidson, A.L. (2015) Full engagement of liganded maltose-binding protein stabilizes a semi-open ATP-binding cassette dimer in the maltose transporter. *Molecular microbiology* 98, 878–894.

Arinaminpathy, Y., Khurana, E., Engelman, D.M. & Gerstein, M.B. (2009) Computational analysis of membrane proteins: the largest class of drug targets. *Drug discovery today* 14, 1130–1135.

Bakheet, T.M. & Doig, A.J. (2009) Properties and identification of human protein drug targets. *Bioinformatics (Oxford, England)* 25, 451–457.

Böckmann, A., Gardiennet, C., Verel, R., Hunkeler, A., Loquet, A., Pintacuda, G., Emsley, L., Meier, B.H. & Lesage, A. (2009) Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45, 319–327.

Chaptal, V., Delolme, F., Kilburg, A., Magnard, S., Montigny, C., Picard, M., Prier, C., Monticelli, L., Bornert, O., Agez, M., Ravaud, S., Orelle, C., Wagner, R., Jawhari, A., Broutin, I., Pebay-Peyroula, E., Jault, J.-M., Kaback, H.R., Le Maire, M. & Falson, P. (2017) Quantification of Detergents Complexed with Membrane Proteins. *Scientific reports* 7, 41751.

Coombs, D.H. & Watts, N.R. (1985) Generating sucrose gradients in three minutes by tilted tube rotation. *Analytical biochemistry* 148, 254–259.

De Foresta, B., Le Maire, M., Orlowski, S. & Champeil, P. (1989) Membrane solubilization by detergent: use of brominated phospholipids to evaluate the detergent-induced changes in Ca2+-ATPase/lipid interaction. *Biochemistry* 28, 2558–2567.

Degrip, W.J., Vanoostrum, J. & Bovee-Geurts, P.H. (1998) Selective detergent-extraction from mixed detergent/lipid/protein micelles, using cyclodextrin inclusion compounds: a novel generic approach for the preparation of proteoliposomes. *Biochemical Journal* 330

(Pt 2), 667–674.

Doerrler, W.T. & Raetz, C.R.H. (2002) ATPase activity of the MsbA lipid flippase of Escherichia coli. *Journal of Biological Chemistry* 277, 36697–36705.

Fribourg, P.-F., Chami, M., Sorzano, C.O.S., Gubellini, F., Marabini, R., Marco, S., Jault, J.-M. & Lévy, D. (2014) 3D cryo-electron reconstruction of BmrA, a bacterial multidrug ABC transporter in an inward-facing conformation and in a lipidic environment. *Journal of molecular biology* 426, 2059–2069.

Hauer, F., Gerle, C., Fischer, N., Oshima, A., Shinzawa-Itoh, K., Shimada, S., Yokoyama, K., Fujiyoshi, Y. & Stark, H. (2015) GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* 23, 1769–1775.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Orelle, C., Dalmas, O., Gros, P., Di Pietro, A. & Jault, J.-M. (2003) The conserved glutamate residue adjacent to the Walker-B motif is the catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA. *Journal of Biological Chemistry* 278, 47002–47008.

Orelle, C., Gubellini, F., Durand, A., Marco, S., Lévy, D., Gros, P., Di Pietro, A. & Jault, J.-M. (2008) Conformational change induced by ATP binding in the multidrug ATP-binding cassette transporter BmrA. *Biochemistry* 47, 2404–2412.

Ostuni, M.A., Iatmanen, S., Teboul, D., Robert, J.-C. & Lacapère, J.-J. (2010) Characterization of Membrane Protein Preparations: Measurement of Detergent Content and Ligand Binding After Proteoliposomes Reconstitution. In: J.-J. Lacapère (Ed), *Membrane Protein Structure Determination*. Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 3–18.

Ravaud, S., Do Cao, M.-A., Jidenko, M., Ebel, C., Le Maire, M., Jault, J.-M., Di Pietro, A., Haser, R. & Aghajari, N. (2006) The ABC transporter BmrA from Bacillus subtilis is a functional dimer when in a detergent-solubilized state. *Biochemical Journal* 395, 345.

Rigaud, J.L., Mosser, G., Lacapere, J.J., Olofsson, A., Levy, D. & Ranck, J.L. (1997) Bio-Beads: an efficient strategy for two-dimensional crystallization of membrane proteins. *Journal of Structural Biology* 118, 226–235.

Rigaud, J.L., Pitard, B. & Levy, D. (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochimica et biophysica acta* 1231, 223–246.

Seddon, A.M., Curnow, P. & Booth, P.J. (2004) Membrane proteins, lipids and detergents:

not just a soap opera. *Biochimica et biophysica acta* 1666 (1-2):105-17

Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A. & Jault, J.-M. (2004) Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in Bacillus subtilis†. *Biochemistry* 43, 7491–7502.

Stevens, T.J., Fogh, R.H., Boucher, W., Higman, V.A., Eisenmenger, F., Bardiaux, B., van Rossum, B.-J., Oschkinat, H. & Laue, E.D. (2011) A software framework for analysing solid-state MAS NMR data. *Journal of Biomolecular NMR* 51, 437–447.

Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41, 207–234.

Wallin, E. & Heijne, von, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science* 7, 1029–1038

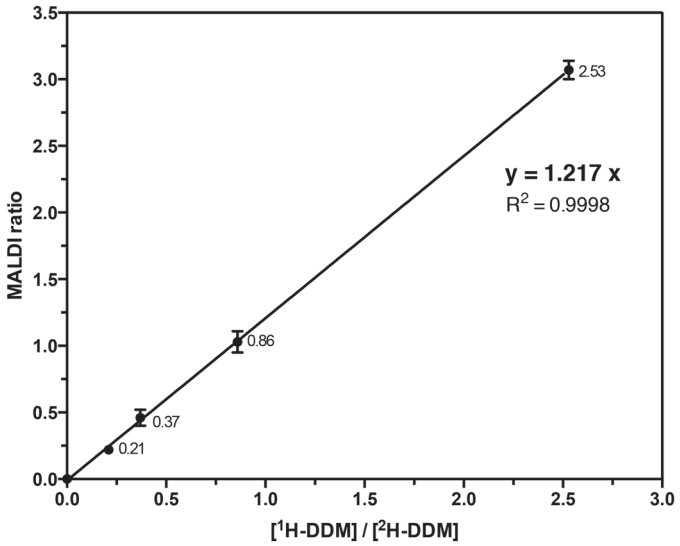
Wang, S. & Ladizhansky, V. (2014) Recent advances in magic angle spinning solid state NMR of membrane proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy* 82, 1–26.

Yang, Z., Wang, C., Zhou, Q., An, J., Hildebrandt, E., Aleksandrov, L.A., Kappes, J.C., DeLucas, L.J., Riordan, J.R., Urbatsch, I.L., Hunt, J.F. & Brouillette, C.G. (2014) Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein Science* 23, 769–789.

Supplementary material

Experiment	DARR
MAS frequency/kHz	17
Field/T	20
Transfer I	HC-CP
^{1}H field/kHz	67
^{13}C field/kHz	50
Shape	Tangent ^{1}H
Time/ms	0.65
Transfer II	DARR
^{1}H field/kHz	17
Time/ms	20
t_1 increments	2048
Windows function	QSine 3
Sweep width (t_1)/kHz	100
Acquisition time (t_1)/ms	10
t_2 increments	3072
Windows function	QSine 3
Sweep width (t_2)/kHz	100
Acquisition time (t_2)/ms	15
^{1}H SPINAL64 decoupling power/kHz	90
Interscan delay/s	2.6
Number of scans	32
Measurement time/h	48

Table S1 Overview of experimental parameters of performed MAS NMR experiments



Sample	1/5	1/2	1/1	3/1
^{1}H-DDM (%)	0.0055	0.011	0.024	0.071
^{2}H-DDM (%)	0.027	0.028	0.028	0.028
Ratio	0.21	0.37	0.86	2.53
MALDI ratio	0.22	0.46	1.03	3.07
MALDI error	0.03	0.06	0.08	0.07

Figure S1. Calibration curve for DDM quantification using ^{2}H-DDM as internal standard. The theoretical ratios are plotted against the experimental ratios (MALDI ratios) obtained by MALDI-TOF. Each ^{1}H/^{2}H ratio is indicated on the plot. The experiment was done in triplicate and the fit follows a linear regression (done using GraphPad Prism).

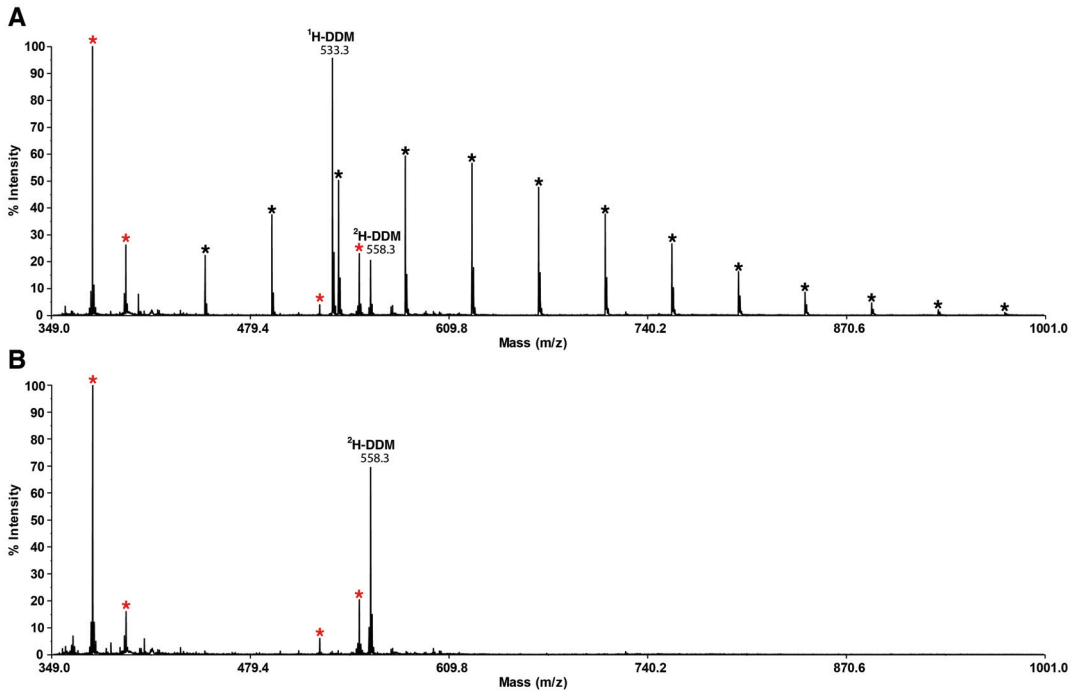


Figure S2. Mass Spectrometry analysis of BmrA before and after dialysis. (**A**) Mass Spectrometry profile of BmrA in 0.05% DDM with *B. subtilis* lipids destabilized by 0.15% Triton X-100 before reconstitution by dialysis. (**B**) Mass Spectrometry spectrum of BmrA in *B. subtilis* lipids after reconstitution by dialysis (6 days). Peaks from ^{1}H-DDM and Triton X-100 are absents after 9 days of reconstitution. ^{2}H-DDM standard was used in both samples in order to quantify the DDM presents in the sample. Black stars indicate the peaks corresponding to Triton X-100, red stars indicate the dihydroxybenzoic acid used as matrix substance for MALDI-MS.

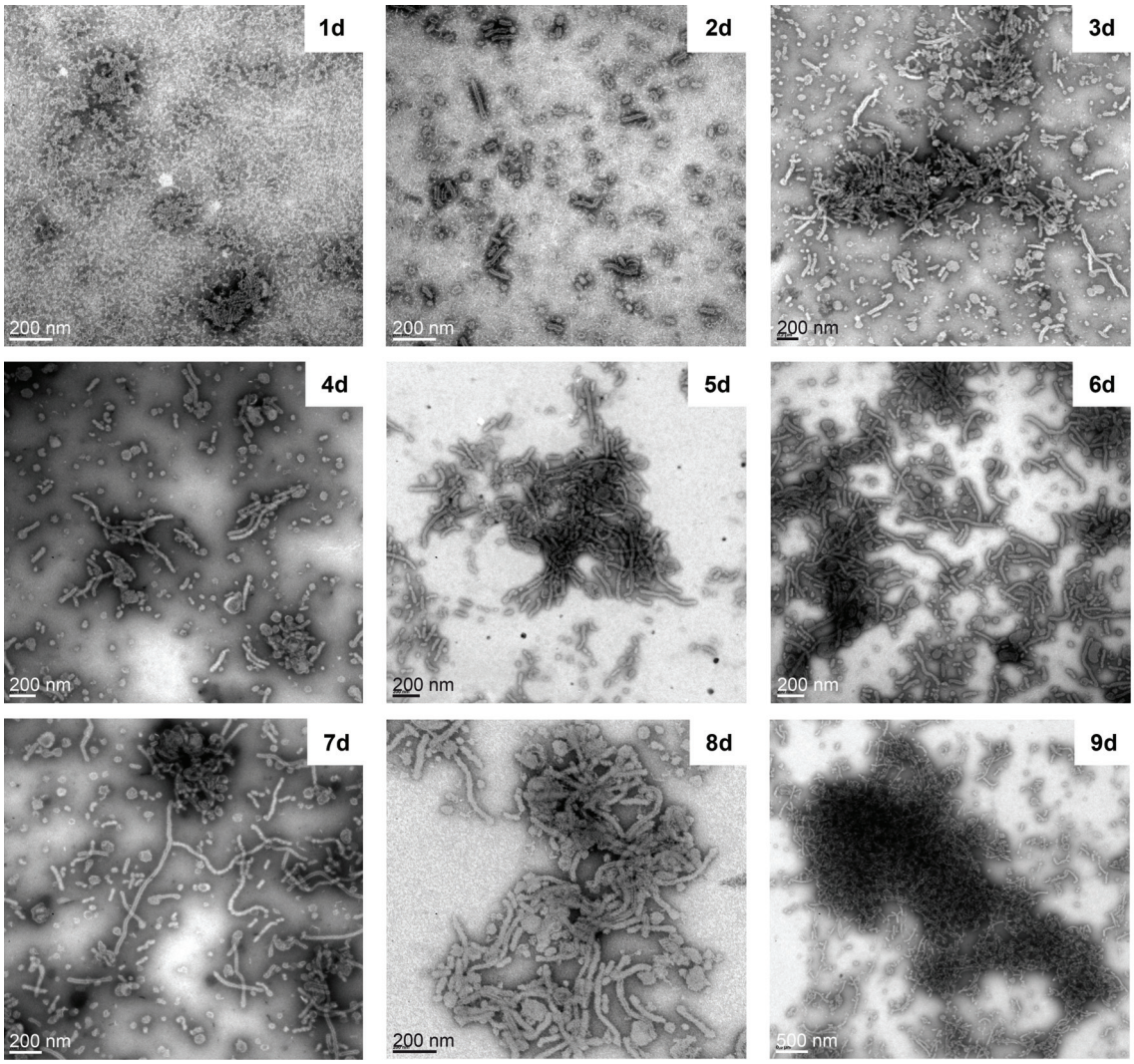


Figure S3: Electron microscopy images of BmrA recorded during reconstitution by dialysis. The samples correspond to the data points shown in Figure 1A and were collected directly from the dialysis bag each day and deposited on a 300-mesh Cu-Formvar grid, and subsequently negatively stained with 2 % phosphotungstic acid.

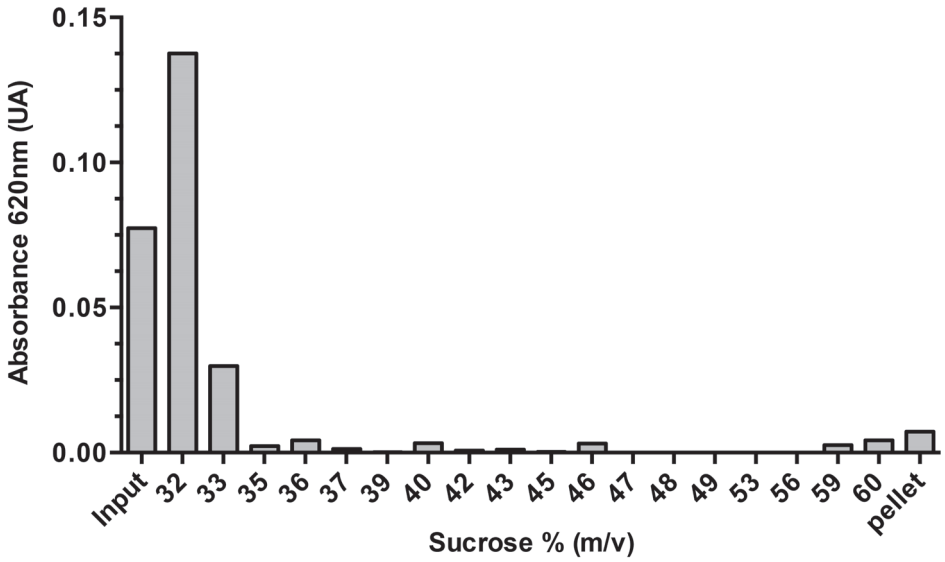


Figure S4: DDM distribution inside a 30-60% sucrose gradient. We used a coomassie (Bradford) protein assay kit (Thermo Scientific™) to probe the presence of DDM inside a 30-60% sucrose gradient. Indeed, it is well-known that colorimetric dyes in protein assay kits can be altered by the presence of detergent. The detergent interaction with the colorimetric dye can be used in order to determine the presence or the quantity of DDM(Ostuni *et al.* 2010). In a microplate, 100 µL of gradient layer were mixed with 10 µL of a coomassie (Bradford) protein assay kit (Thermo Scientific™). The absorbance was read at 620 nm (Infinite M1000 Tecan Group Ltd).

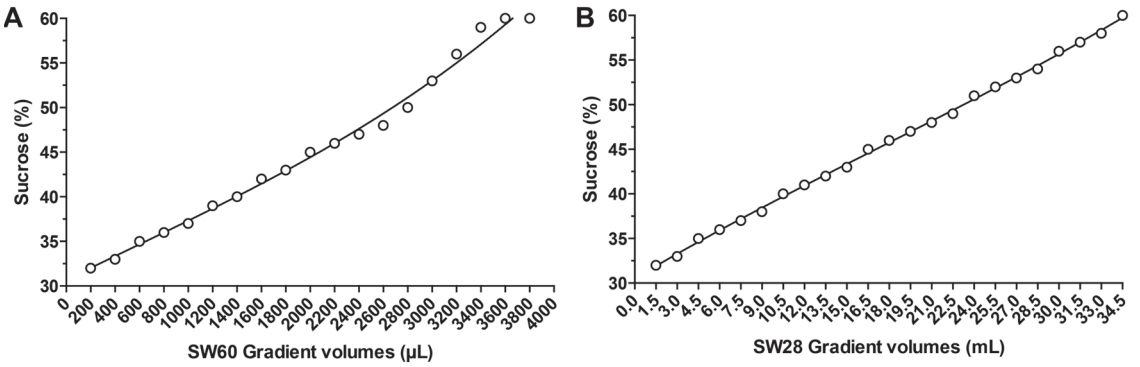


Figure S5: Sucrose concentration of each layer of gradients 30-60% SW60 and SW28. In order to determine the best gradient and also the sucrose concentration in each layer, a protocol using coomassie blue, described in the literature(Coombs & Watts 1985) and by the manufacturer (BioComp Instruments, Inc), was used. Two layers containing 30% and 60% sucrose were initially laid in the centrifuge tube. The lower solution contains equally 0.30% of Coomassie®-brilliant blue R250. Then, the upper and the lower solutions were mixed with the Gradient Master™ in order to obtain a continuous gradient. For the evaluation of gradients, the absorbance at 820 nm was determined for layers of 200 µL (SW60) and 1.5 mL (SW28). The profile obtained was converted to sucrose concentration as shown by Coombs and Watts(Coombs & Watts 1985): $C_i = (X - y)d_i + y$, with C_i the sucrose concentration at point I, X the heavy sucrose concentration (60%), y the light concentration of sucrose and d_i the absorbance of the layer.

SW60						
Layers (µl)	Sucrose (%)	Cumulative α-cyclodextrin (nmol)				
		0-0.25%	0-0.50%	0-1%	0-2%	0-3%
200	32	50	70	150	250	400
400	33	100	170	350	700	1050
600	35	150	350	700	1350	2050
800	36	250	550	1100	2200	3300
1000	37	400	800	1600	3150	4750
1200	39	550	1100	2200	4400	6600
1400	40	700	1450	2900	5750	8650
1600	42	950	1850	3700	7400	11100
1800	43	1150	2300	4600	9200	13750
2000	45	1400	2800	5600	11250	16850
2200	46	1700	3350	6700	13450	20150
2400	47	1950	3950	7900	15750	23650
2600	48	2300	4550	9100	18250	27350
2800	49	2600	5200	10400	20850	31250
3000	53	3000	6000	12000	24000	36000
3200	56	3450	6900	13750	27550	41300
3400	59	3950	7900	15750	31500	47300
3600	60	4450	8900	17800	35650	53450
3800	60	4950	9950	19850	39750	59600

SW32		
Layers (ml)	Sucrose (%)	Cumulative α-cyclodextrin (nmol)
1.5	32	1050
3	33	2550
4.5	35	5150
6	36	8200
7.5	37	11800
9	38	15950
10.5	40	21050
12	41	26750
13.5	42	32900
15	43	39600
16.5	45	47300
18	46	55500
19.5	47	64250
21	48	73500
22.5	49	83250
24	51	94050
25.5	52	105350
27	53	117200
28.5	54	129500
30	56	142900
31.5	57	156750
33	58	171150
34.5	60	186550

Figure S6: Parameter table of gradients used with SW60 and SW32. For each layer cumulative α-cyclodextrin values in nmol are estimated considering the gradient determined previously (**Figure S5**). The closest value to the theoretical 900 nmol of α-cyclodextrin needed to strip detergent from 4.5 nmol of BmrA is highlighted in red for each gradient.

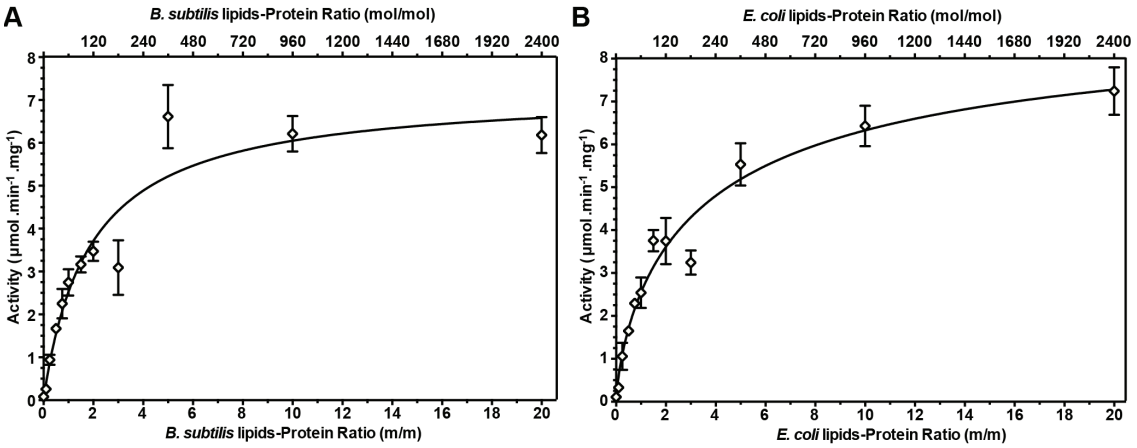


Figure S7: ATPase activity as a function of LPR. BmrA was reconstituted using different LPR between 0-20 and with lipids from **A**) *B. subtilis* (homemade preparation, see below) or **B**) *E. coli* lipids (Avanti® Polar Lipids, Inc). The ATPase activity was probed for the different LPRs and the different lipids in presence of 10 mM ATP. The activity of BmrA in *B. subtilis* lipids follows the same trend as in *E. coli* lipids. In both cases, the activity increases with the LPR following a hyperbolic curve. The catalytic activity is slightly higher in *E. coli* lipids than in *B. subtilis* lipids. The maximum velocity of BmrA in presence of 10 mM ATP and *B. subtilis* and *E. coli* lipids is estimated to 7.2 ±0.5 and 9.4 ±0.9 µmol.min^{-1}.mg^{-1} respectively. The fitting was done with GraphPad Prism 6.

The *B. subtilis* lipids homemade preparation was prepared as described in Kunert et *al*. (Kunert *et al*. 2014). Shortly, *B. subtilis* was cultivated in LB medium at 37°C during, until an OD_{600nm} of 4. Bacteria were harvested by centrifugation, 6000 × g, 30 min. For 10 g of pellet, 20 mL of chloroform and 40 mL of methanol were added directly with the pellet and the solution was incubated during 2 h at room temperature. After incubation, solid particles were removed by filtration. The filtrate was mixed with 20 mL chloroform and 20 mL water. The extraction was performed for 2 days in a separation funnel. The solvent of the lower phase containing chloroform and lipids was removed in a rotary evaporator followed by lyophilisation. The lipids were suspended in chloroform and stored at -20°C for 1 month maximum. The final lipid amount was about 20 mg per liter of *B. subtilis* culture.

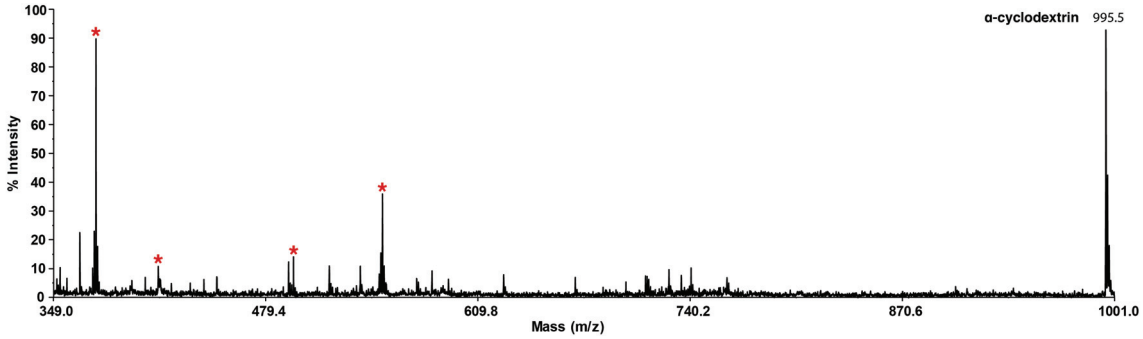


Figure S8. Mass Spectrometry analysis of BmrA after GRecon. Mass Spectrometry spectrum of BmrA in *B. subtilis* lipids after reconstitution by GRecon. The sample was diluted 100 times in the matrix in order to crystallize the deposit on the MALDI-target. The presence of DDM was not detected in MALDI-TOF. Red stars indicate the dihydroxybenzoic acid used as matrix substance for MALDI-MS.

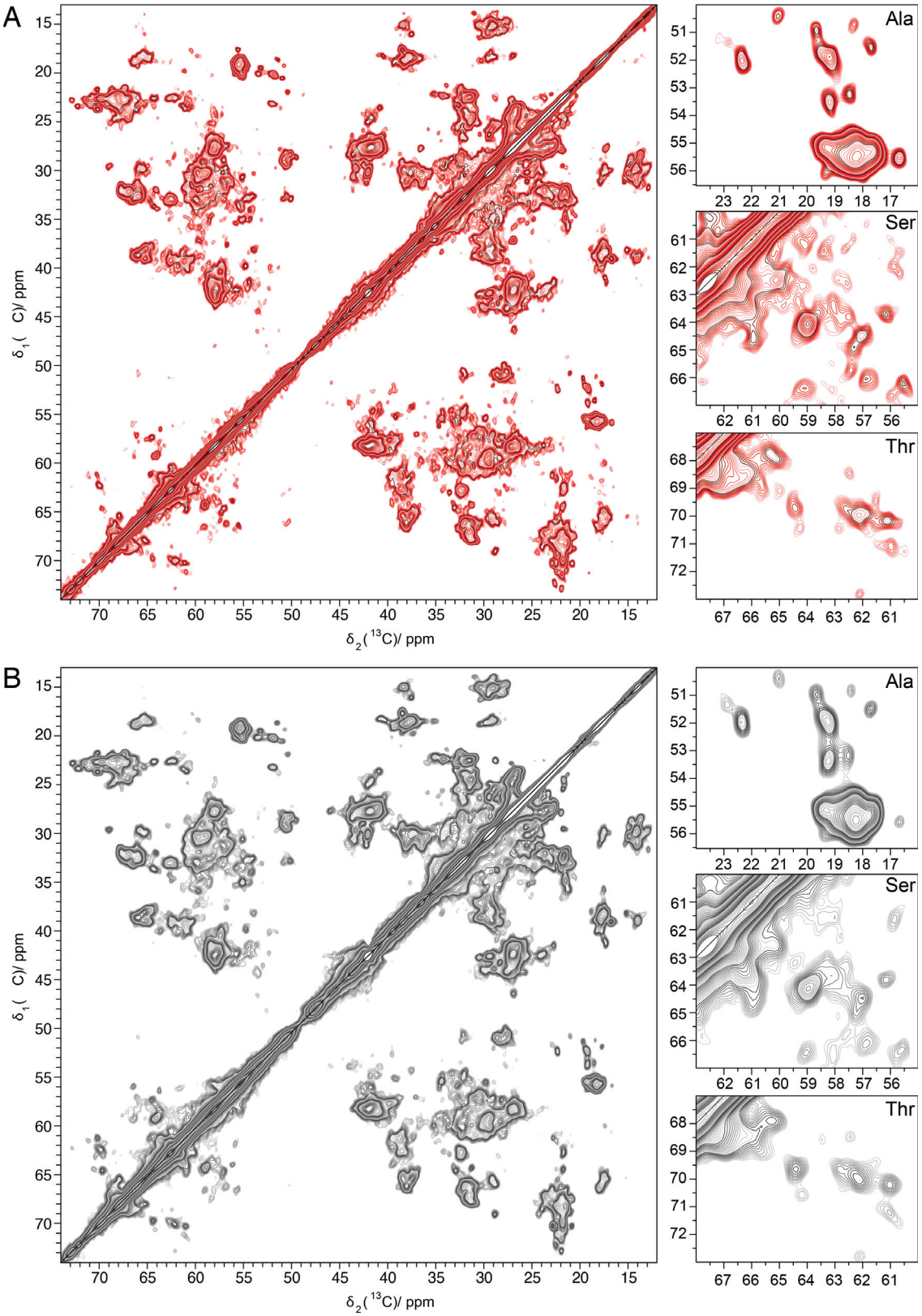


Figure S9. Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectrum of lipid-reconstituted BmrA. The aliphatic region and a zoom of the Ala, Ser and Thr regions of reconstituted BmrA are shown, with the sample resulting from GRecon in red shades (**A**) and from dialysis-reconstituted BmrA in grey shades (**B**).

Chapter VII

Solid-state NMR and EPR Mn^{2+} substituted ATP-fueled protein engines

The work presented in this chapter was done in collaboration with Beat H. Meier, Laurent Terradot and Gunnar Jeschke. This Chapter was adapted from the following reference:

Solid-state NMR and EPR of Mn^{2+} substituted ATP-fueled protein engines

Thomas Wiegand[¥], Denis Lacabanne[¥], Katharina Keller[¥], Riccardo Cadalbert, Lauriane Lecoq, Maxim Yulikov, Gunnar Jeschke, Laurent Terradot, Beat H. Meier, Anja Böckmann

[¥] contributed equally to this work.

Angew Chem Int Ed Engl. 2017 Mar 13;56(12):3369-3373.

Contents

Abstract

Paramagnetic metal ions provide structural information both in EPR and solid-state NMR experiments, offering a profitable synergetic approach to study bio-macromolecules. We demonstrate the spectral consequences of Mg^{2+}/ Mn^{2+} substitution and resulting information contents for two different ATP:Mg^{2+}-fueled protein engines, a DnaB helicase from *Helicobacter Pylori* active in the bacterial replisome, and the ABC transporter BmrA, a bacterial efflux pump. We show that, while EPR spectra report on metal binding and inform on the geometry of the metal centers in the proteins, paramagnetic relaxation enhancements identified in the NMR spectra can be used to localize residues at the binding site. Protein engines are ubiquitous and the methods described here should be applicable in a broad context.

Introduction

Nucleotide binding domains (NBDs) are multipurpose engines which convert the chemical energy released on nucleoside triphosphate hydrolysis into mechanical movement or switching. They are found in a set of proteins for which conformational change and molecular motion is needed for function (Hanson & Whiteheart 2005; Rees *et al.* 2009) with adenosine triphosphate (ATP) being a common fuel for these engines. We here inquire what information on ATP binding to NBDs can be obtained by an approach that combines data from paramagnetic solid-state NMR (Bertini *et al.* 2010; Jaroniec 2015; Pintacuda *et al.* 2007; Ullrich *et al.* 2013) and from EPR spectroscopy when Mg^{2+} as a cofactor of ATP hydrolysis is replaced by Mn^{2+}. For this, we focus on a helicase (*Hp*DnaB) (Bazin *et al.* 2015) and an ABC transporter (BmrA) (Steinfels *et al.* 2004) as two proteins whose function relies on NBDs (Geourjon *et al.* 2001), but with nucleotide-binding dissociation constants (Kd) differing by roughly two orders of magnitude (Bujalowski & Klonowska 1993; Jezewska *et al.* 1996; Siarheyeva *et al.* 2010). For ATP hydrolysis, which provides the energy source for the engine, the NBDs require a divalent metal ion, typically Mg^{2+}, as cofactor. Both proteins show multiple ATP binding sites: 12 for *Hp*DnaB, and two for BmrA. Each domain contains Walker A and B motifs, with an aspartate that coordinates Mg^{2+}. While in the helicase the NBDs maintain contact throughout the functional cycle, they are spatially separated in the

ABC transporter BmrA in the open conformation and assemble to bind and hydrolyze ATP:Mg^{2+}. In order to mimic nucleotide-bound states, poorly hydrolysable analogues can be used, such as adenylyl-imidodiphosphate (APMPNP), which binds with high affinity to *Hp*DnaB (Bazin *et al.* 2015; Wiegand *et al.* 2016a). The dissociation constants of BmrA-nucleotide complexes in the absence of substrate (Siarheyeva *et al.* 2010) is larger than the ones of DnaB (Bujalowski & Klonowska 1993; Jezewska *et al.* 1996) and large amounts of APMPNP/metal would be needed for the preparation of an NMR sample showing high occupancy of the binding sites, and we instead used ATP:VO_4^{3-}:Mg^{2+}. After ATP hydrolysis in the presence of vanadate, the NBDs are expected to remain blocked in an ADP:VO_4^{3-}:Mg^{2+} bound state, with the VO_4^{3-} ion mimicking the γ-phosphate of ATP (Orelle *et al.* 2008).

The characterization of the interaction between ATP and the NBDs is essential for understanding the functioning of such ATP-fueled proteins. By substituting the diamagnetic Mg^{2+} by paramagnetic Mn^{2+}, the binding sites become EPR observable (Bennati *et al.* 2006; Witt *et al.* 1997) and a distance-dependent paramagnetic relaxation enhancement (PRE) is induced for the NMR resonances (Bertini *et al.* 2005; Bonneau & Legault 2014; Kaur *et al.* 2015; Otting 2010; Tamaki *et al.* 2016). This provides contrast between the binding site and the part of the system remote from it. Due to similar coordination behavior, charge and ion radius of Mg^{2+} and Mn^{2+} ions, most often function is retained ((Soni 2003) and **Figure 47**).

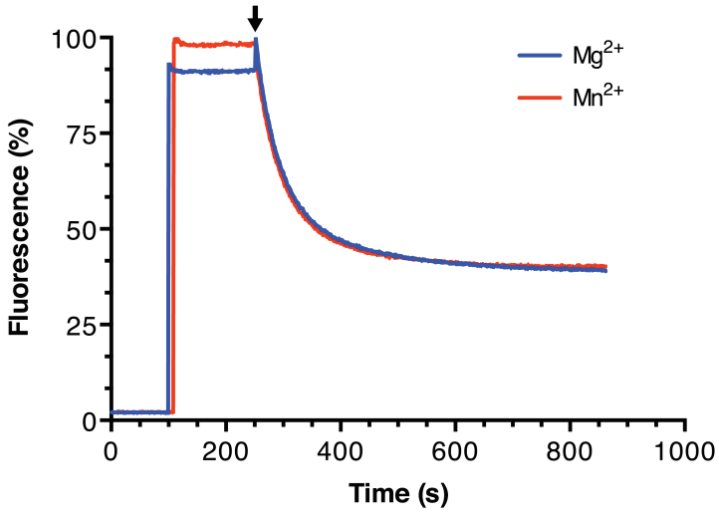


Figure 47: ATP-dependent transport of doxorubicin by inverted *E. coli* membrane vesicles. Addition of 10 µM doxorubicin with 100 µg of inverted membrane vesicles from *E. coli* C41(DE3) overexpressing BmrA incubated in presence of 2 mM ATP and ATP regeneration buffer (0-120 s). After addition of 10 µM doxorubicin (120 s), 2 mM of Mg^{2+} or Mn^{2+} were added, the addition of the metal ion is indicated by a black arrow. The transport activities of the ABC transporter show the same kinetic in presence of Mg^{2+} or Mn^{2+}.

We added Mn^{2+} and ATP analogues to the two proteins in order to demonstrate the effects of nucleotide:Mn^{2+} binding on NMR and EPR spectra. This allows to determine the occupation of nucleotide binding sites, to extract Mn^{2+}-Mn^{2+} distances (Ching *et al.* 2016), to probe the geometry of the multimeric assemblies, as well as to identify the residues in the neighbourhood of Mn^{2+}. The radius of view is on the order of 15 Å for residue identification and on the order of 35-65 Å for assembly geometry.

Material and Methods

1. HpDnaB

^{13}C-^{15}N labelled *Hp*DnaB was prepared in buffer A (2.5 mM sodium phosphate, pH 7.5, 130 mM NaCl) as described in reference (Gardiennet *et al.* 2012). Deuterated DnaB was prepared using the same protocol, but the protein was expressed in deuterated media. The buffer was exchanged by the corresponding D_2O buffer before nucleotide addition.

1.1. HpDnaB:APMPNP:Mg^{2+} complex

0.3 mM *Hp*DnaB (all concentration are given with respect to an DnaB monomer) in buffer A was mixed with 5 mM $MgCl_2 \cdot 6H_2O$ and consecutively 5 mM APMPNP (18-fold molar excess of APMPNP compared to an *Hp*DnaB monomer) and incubated for 2 h at 4°C. The protein solution was sedimented (Gardiennet *et al.* 2012; Wiegand *et al.* 2015) in the MAS-NMR rotor (16 h at 4°C with 210 000 × *g* acceleration) using home-build tools (Böckmann *et al.* 2009). Amino-acids, Na_3VO_4, APMPNP and ATP were purchased from Sigma-Aldrich for all preparations.

1.2. HpDnaB:APMPNP:Mn^{2+} complex

0.3 mM *Hp*DnaB in buffer A was mixed with 1.4 mM $MnCl_2 \cdot 6H_2O$ (5-fold molar excess compared to an *Hp*DnaB monomer) and consecutively 5 mM APMPNP (18-fold molar excess of APMPNP compared to an *Hp*DnaB monomer) and incubated for 2 h at 4°C. The protein solution was sedimented (Gardiennet *et al.* 2012; Wiegand *et al.* 2015) in the MAS-NMR rotor (16 h at 4°C with 210 000 × *g* acceleration) using home-build tools (Böckmann *et al.* 2009). Lower ratios of Mn^{2+}: *Hp*DnaB monomer were explored additionally, but only a partial occupation of APMPNP binding sites was observed under these conditions.

2. BmrA

The production, purification and reconstitution in homemade lipids of the membrane protein ^{12}C-^{14}N-[LVIKHP]-^{13}C-^{15}N-BmrA have already described by (Kunert *et al.* 2014). In order to perform the reverse unlabelling, not labelled Leu (0.23g/L), Val (0.23g/L), Ile (0.23g/L), Lys (0.40g/L), His (0.40g/L), Pro (0.10 g/L) were added in the culture medium one hour before the induction. For the purification the concentration of protein does not

exceed 1mg.mL^{-1} (17µM) to avoid aggregates. Moreover, the protein was also diluted for the reconstitution, 0.25 mg.mL^{-1} (4.25 µM).

2.1. Activity drug transport of BmrA in presence of Mn^{2+} or Mg^{2+}

The experimental procedures were already describes by (Steinfels *et al.* 2002). Briefly, in a 1 mL cuvette, 100 µg of inverted membrane vesicles from *E. coli* C41(DE3) overexpressing BmrA were incubated in presence of 2 mM ATP and ATP regeneration buffer (4mM phosphoenolpyruvate, 60µg.mL^{-1} pyruvate kinase). After 1 minute at 25 °C incubation, 10 µM doxorubicin were added and the transport-reaction is started by the addition of 2mM of Mg^{2+} or Mn^{2+}. The excitation and emission wavelengths were, respectively, 480 and 590 nm.

2.2. BmrA:ADP:Mg^{2+}:VO_4^{3-} complex

The lipid-reconstituted protein was mixed with 1 mM Na_3VO_4 and with 5 mM ATP-Mg corresponding to a 1500-fold molar excess of ATP-cation compared to a BmrA monomer. All compounds were prepared in 50 mM Tris-HCl pH 8.0, 200 mM NaCl and 5% glycerol, and the preparations were incubated for 2 h at room temperature. The protein in lipids was sedimented in the MAS-NMR rotor (1 h at 4°C with 120 000 × *g* acceleration) using home-build tools (Böckmann *et al.* 2009).

2.3. BmrA:ADP:Mn^{2+}:VO_4^{3-} complex

The lipid-reconstituted protein was mixed with 1 mM Na_3VO_4 and with 1 mM ATP-Mn corresponding to a 300-fold molar excess of ATP-cation compared to a BmrA monomer. Lower and higher ratios of Mn^{2+}:BmrA monomer were explored additionally, but a partial occupation of ATP:Mn^{2+} binding sites or global signal loss, respectively, was observed under these conditions. All compounds were prepared in 50 mM Tris-HCl pH 8.0, 200 mM NaCl and 5% glycerol, and the preparations were incubated for 2 h at room temperature. The protein in lipids was sedimented in the MAS-NMR rotor (1 h at 4°C with 120 000 × *g* acceleration) using home-build tools (Böckmann *et al.* 2009).

3. Solid-state NMR

Solid-state NMR spectra were acquired at 18.8 T and 20.0 T static magnetic field strength using 3.2 mm Bruker Biospin "E-free" probes. The MAS frequency was set to 17.5 and

17.0 kHz, respectively. The sample temperature was set to 278 K (Böckmann *et al.* 2009). The spectra were processed with the software TOPSPIN (version 3.2, Bruker Biospin) with a shifted (2.5 to 3.0) squared cosine apodization function and automated baseline correction in the direct dimension. All spectra were analyzed with the software CcpNmr (Stevens *et al.* 2011) and referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). The assignments of *Hp*DnaB and *Hp*DnaB:APMPNP:Mg^{2+} are according to references (Wiegand *et al.* 2016b). PRE effects were extracted from 2D and 3D spectra and the ratios I_{para}/I_{dia} were scaled such that the average I_{para}/I_{dia} value for resonances of the NTD is one (the NTD is not participating in nucleotide binding). This allows for a compensation of experimental errors resulting from different experimental conditions in different measurement time periods. The error bars shown in Figure 48b are based on Gaussian error propagation and estimated from the noise levels in the corresponding spectra.

4. EPR Spectroscopy

CW EPR spectra were measured for a sample with 12-fold excess of the *Hp*DnaB monomers with respect to the Mn^{2+} ions, in order to assure predominant presence of the metal ions in the form of *Hp*DnaB:APMPNP:Mn^{2+} complexes. In this case, most of protein oligomers should have only one Mn^{2+} ion bound. However, the Mn^{2+} EPR spectrum is dominated by the close vicinity of the paramagnetic metal ion, and should not differ between the studied sample and the sample with all NBDs loaded with Mn^{2+}.

Pulse EPR measurements were done on a sample with 1:1 ratio between Mn^{2+} ions and *Hp*DnaB monomers, to obtain sufficiently large fraction of multiply metal-loaded *Hp*DnaB oligomers. For significantly lower relative amounts of *Hp*DnaB the fraction of Mn^{2+} ions not bound to the NBDs would be too high, resulting in a too low dipolar modulation depth in the DEER measurements. For significantly higher *Hp*DnaB amounts the fraction of protein oligomers with only one Mn^{2+} ion bound would increase, again reducing the dipolar modulation depth in the DEER experiment. Despite the presence of, on average, six Mn^{2+} ions per protein complex, multi-spin effects are negligible due to the modulation depth of only 1% even when using ultra-wideband pump pulses.

4.1. CW-EPR

HpDnaB:APMPNP:Mn^{2+} complex : 3.6 mM *Hp*DnaB in buffer A was mixed with 0.3 mM $MnCl_2 \cdot 4H_2O$ and consecutively 5 mM APMPNP (18-fold molar excess of APMPNP compared to an *Hp*DnaB monomer) and incubated for 2 h at 4°C.

$MnCl_2$:APMPNP reference solution : A 0.3 mM $MnCl_2 \cdot 4H_2O$ and 5mM APMPNP solution in buffer A was prepared.

CW X-band EPR spectra were detected using a Bruker EMX spectrometer equipped with a Bruker Elexsys Super High Sensitive probehead. Measurements were performed at room temperature, 100 kHz field modulation, 250 to 450 mT field sweep, 4 G modulation amplitude, and 1 mW microwave power (10 dB attenuation). The time constant was 10.24 ms and the conversion time 40.96 ms. The samples were filled into 0.9 mm inner diameter glass capillaries (10-20 µl sample volume).

4.2. Pulsed EPR experiments

HpDnaB:APMPNP:Mn^{2+} complex: Buffer A was exchanged with the corresponding D_2O buffer (buffer B). 0.6 mM *Hp*DnaB in buffer B was mixed with 0.6 mM $MnCl_2 \cdot 6H_2O$ and consecutively 10 mM APMPNP (18-fold molar excess of APMPNP compared to an *Hp*DnaB monomer) and incubated for 2 h at 4°C. The solution was mixed with glycerol leading to a final protein concentration of 0.3 mM. For the DEER experiments a deuterated sample of *Hp*DnaB was used.

$MnCl_2$:APMPNP reference solution: A 0.6 mM $MnCl_2 \cdot 4H_2O$ and 10 mM APMPNP solution in buffer B was prepared. The solution was diluted with glycerol until concentrations of 0.3 mM $MnCl_2 \cdot 4H_2O$ and 5 mM APMPNP were reached.

$MnCl_2$ reference solution: A 0.6 mM $MnCl_2 \cdot 4H_2O$ solution in buffer B was prepared. The solution was diluted with glycerol (final $MnCl_2 \cdot 4H_2O$ concentration of 0.3 mM).

The protein-containing and reference samples were dissolved in buffer B:glycerol-d8 (1:1 v:v), filled into 3 mm outer diameter quartz sample tubes and subsequently shock-frozen by immersion into liquid nitrogen. Pulsed experiments were performed on a customized Bruker Elexsys E580 Q-band spectrometer (Polyhach *et al.* 2012), extended with an incoherent ultra-wide band (UWB) pulse channel based on a 12 GS/s arbitrary waveform generator (AWG, Agilent M8190A) (Doll *et al.* 2015a). The temperature was adjusted, using

a helium flow cryostat, to 10 K for distance measurements and to 20 K for echo-detected (ED) field-swept and relaxation measurements.

ED field-swept spectra were acquired with a Hahn-echo pulse sequence t_p–τ–2t_p–τ–echo using a pulse length t_p = 12 ns and τ = 400 ns. The same sequence was used to record the phase memory decay by incrementing the initial interpulse delay τ in steps of 240 ns.

Relaxation traces were detected in the valley between the 2nd and 3rd hyperfine line of the Mn^{2+} EDEPR spectra marked by an arrow in Figure S1. The reason for this choice is the larger difference in the linewidths for the central $|+1/2\rangle \leftrightarrow |-1/2\rangle$ transitions of different species as compared to the widths differences for any other transition of the high-spin Mn^{2+} ions. The six sharp lines of the Mn^{2+} EPR spectrum correspond to the central $|+1/2,M_I\rangle \leftrightarrow |-1/2,M_I\rangle$ transitions and are split by the ^{55}Mn hyperfine coupling (natural abundance 100%, nuclear spin I=5/2). Considering zero-field splitting (ZFS) term in the spin Hamiltonian as a small perturbation to the electron Zeeman interaction, in the first (linear) order of the perturbation series this transition is not affected by the ZFS term, while all other transitions of Mn^{2+} are broadened by the orientation dependent ZFS interaction. In the second order of perturbation series the central transition of Mn^{2+} gets moderately broadened by the ZFS term. Thus, the dependence of the width of this transition on ZFS is steeper than for all other transitions. As a result, the intensity of the spin echo signal per fixed amount of molecules, if measured at this transition would differ to the largest extent for different Mn^{2+} species detected in our experiments. In the valley between any two hyperfine components of the central transition the relative contributions of different species to the spin echo signal correspond more closely to their concentrations in the sample. Among all possible positions, the selected one corresponded to the maximum spin echo intensity.

4.3. HpDnaB:APMPNP:Mn^{2+} by EPR relaxation measurements.

Transverse relaxation of Mn^{2+} ions is strongly affected by the interaction with other paramagnetic centers as well as with surrounding nuclear spins, via the spin diffusion mechanism. Among all nuclear spins, matrix protons contribute by far most strongly, if they are present. For the *Hp*DnaB system this allows for a relaxation-based separation of the Mn^{2+} centers in the form of APMPNP:Mn^{2+} from the protein-a relatively low number of protons and the gyromagnetic ratio of the overwhelming number of matrix nuclei was strongly reduced for the APMPNP:Mn^{2+} species by the use of a deuterated buffer/deuterated

glycerol mixture. Binding of the complex to the protein brings the Mn^{2+} in the close vicinity of a large pool of protons. This and the presence of several Mn^{2+} ions close to each other in the *Hp*DnaB:APMPNP:Mn^{2+} species leads to a faster relaxation of the bound species as compared to unbound APMPNP:Mn^{2+}. Since in the protein-containing sample both types of species were present, the actual transverse relaxation decay trace shows a complex behavior with quick initial decay, due to the bound species, and a significantly slower decaying tail due to the unbound species.

This phenomenon allows to rather accurately estimate the fraction of Mn^{2+} ions attached at the protein NBDs. We can assume, in the same way as in an analogous longitudinal relaxation data analysis for lanthanide-radical pairs, that the additional relaxation channel for the protein-bound species is independent from all other channels, and the total relaxation rate is a sum of the rates for all relevant relaxation channels (Lueders *et al.* 2013; Razzaghi *et al.* 2013). If the transverse relaxation time trace of the protein-containing sample is divided by the corresponding trace for the sample with only APMPNP:Mn^{2+} species present, under the above assumption of independence of the relaxation channels, the ratio trace will contain a contribution from *Hp*DnaB:APMPNP:Mn^{2+}, which decays with the rate corresponding only to the additional relaxation channel, while the unbound APMPNP:Mn^{2+} species contributes constant intensity. In reality, due to the intermolecular interactions, and, perhaps, due to imperfect additivity of different relaxation rates, this latter signal is not constant, but rather very slowly decays, as it is seen in the Figure 48(c) of the main text. Still, we can rather accurately estimate the fraction of *Hp*DnaB:APMPNP:Mn^{2+} species to be somewhere between the vertical coordinate of the kink of the division trace and the point where the prolongation of the slowly relaxing part of the division trace crosses the y-axis. One can see from the Figure 48(c) that the relative intensity of the APMPNP:Mn^{2+} species is thus about 0.4, which corresponds to 60 % of the *Hp*DnaB:APMPNP:Mn^{2+} species in the stoichiometric mixture.

The four-pulse DEER data were obtained using the sequence $\pm(\pi/2)_{obs} - \tau_1 - (\pi)_{obs} - (\tau_1 + t) - (\pi)_{chirp} - (\tau_2 - t) - (\pi)_{obs} - \tau_2 -$ echo, with $\tau_1 = 400$ ns, $\tau_2 = 5000$ ns and the time increment $t = 8$ ns. The $(\pi)_{chirp}$ chirp pulse settings were analogous to the previously reported Gd(III)-Gd(III) DEER measurements (Doll *et al.* 2015b). The observation pulses had a length of 12 ns and the observation frequency was placed in the valley between the 2nd and 3rd hyperfine line of the Mn^{2+} at about 34.5 GHz. The linearly frequency swept (chirp) pump pulses

originated from the AWG. To describe the chirp pulses used in this work, we refer to the start and end frequencies of the frequency sweep as f1 and f2, respectively. Thus the sweep width Δf is the absolute difference between f1 and f2. Two chirp pump pulses were applied with lengths of 50 ns on both sides of the observer frequency. Two different sweep ranges were tested for this sample. First a sweep range from f1 = −0.6 GHz to f2 = −0.2 GHz and f1 = 0.6 GHz to f2 = 0.2 GHz with respect to the observer frequency, corresponding to a total pump bandwidth of 0.8 GHz (data shown in the top three subplots of Figure S4) and a second sweep range from f1 = −0.8 GHz to f2 = −0.2 GHz and f1 = 0.8 GHz to f2 = 0.2 GHz with respect to the observer frequency, corresponding to a total pump bandwidth of 1.2 GHz (data shown in the middle three subplots of the Figure S4). The pump pulses with a total bandwidth of 0.8 GHz resulted in a modulation depth of about 1%, while with the latter pulse settings a modulation depth of approximately 1.5 % could be achieved (see Figure S4). The data in the main text (Figure 48d,e) and in the bottom three subplots in the Figure S4 show the average of both data sets. The pulse edges were smoothened with a quarter sine wave during 10 ns. Experiments were performed on the deuterated DnaB sample to prolong the accessible distance range. A comparison of phase memory times of deuterated and protonated DnaB samples is shown in Figure S3. Due to the difference in the phase memory time between NBD-bound and free Mn^{2+} species, the dipolar modulation depth was decreasing with the increase of the transverse evolution time in the DEER experiment. Furthermore, it is possible that due to the faster transverse relaxation times of the oligomers with higher number of bound Mn^{2+} ions, the longer distance peaks are also partially reduced in intensity, similar to the earlier results of Schiemann and coauthors (Hagelueken *et al.* 2009), (Bode *et al.* 2007).

Despite the use of broad-band pump pulse and presence of multiple Mn^{2+} ions in each protein oligomer, the detected depth of dipolar modulation of only approximately 1% (Figure 48(e)) is comparable to the earlier reports on doubly Mn^{2+} labeled molecules, where modulation depths of 0.4% and 1-2% were detected for the DEER setups with narrower-band rectangular pump pulse. Here, apparently, the factors that increase the dipolar modulation depth are compensated by the rather large ZFS of the *Hp*DnaB:APMPNP:Mn^{2+} complex, and the incomplete binding of the APMPNP:Mn^{2+} species to the NBDs. Longer transverse relaxation time of the APMPNP:Mn^{2+} species further reduced the DEER modulation depth, especially for experiments with long dipolar evolution times. While low

dipolar modulation depth is more difficult to detect, it assured that in every oligomer not more than two Mn^{2+} spins were excited by the DEER pulse sequence. We can thus argue that no high-spin effects, which would be potentially possible in such oligomer systems (Valera *et al.* 2016), should be present in the reported DEER data.

Experimental data were processed using the DeerAnalysis software (Jeschke *et al.* 2006; Valera *et al.* 2016). Distance distributions were obtained by Tikhonov-regularization using a regularization parameter of 1. In the distance distribution shown in the Figure S4, the peak at about 3 nm (right panel) is responsible for the clearly visible fast oscillations of the DEER form factor signal (central panel). Multiple full oscillations were detected for this peak and, thus, its shape is determined with good accuracy. The other peak at the distance range 5-6.5 nm corresponds to the shallow oscillation with the minimum at about 2 µs (central panel). Only about one full period of oscillation could be detected for this peak, even with use of deuterated *Hp*DnaB. Furthermore, part of this oscillation could be incorporated into the background model, and variation of this model with the validation tool in DeerAnalysis resulted in changes of the amplitude and mean distance of this peak indicated by the grey error band in Figure 48e. More precisely, in the validation procedure the beginning of the background correction range was varied between 1000 and 2500 ns in 11 steps. Furthermore, 11 different artificial noise traces of the same average magnitude as the intrinsic noise in the DEER trace were added to this trace to verify the stability of the distance distribution to the level and realization of noise. Grey colored band in the Figures 48e and S4 (right three panels) indicate the variation of the Mn^{2+}-Mn^{2+} distance distribution in the described validation procedure. These variations affected the peak at 3 nm much less, then the peak at 5-6.5 nm. Still, for any background model tested, a peak at the range 5-6.5 nm was always present although it varied in width, amplitude, and mean distance. The distance range between 5 and 6.5 nm is consistent with the homology model based Mn-Mn distances for the non-nearest neighbours as shown in the main text Figure 48f.

Results

1. Mn^{2+} EPR experiments on *Hp*DnaB and BmrA

EPR spectroscopy allows the binding of Mn^{2+} to the NBD to be monitored. **Figure 48A** shows the room temperature X-band Mn^{2+} CW spectra of the *Hp*DnaB:APMPNP:Mn^{2+} complex in solution, as well as of $MnCl_2$ and APMPNP:Mn^{2+} solutions as reference. The spectrum of the $MnCl_2$ solution consists of a resolved sextet for the $|-1/2,M_I\rangle \leftrightarrow |+1/2,M_I\rangle$ transitions due to the electron-^{55}Mn hyperfine splitting of about 8 mT, as expected for the fast tumbling high-spin Mn^{2+} hexaaquo complex (S=5/2 and *I*=5/2). Addition of APMPNP leads to formation of a larger complex with longer rotational correlation time and lower symmetry and thus to an anisotropic broadening of the resonance and to additional spectral features due to the zero-field splitting. The observed spectrum indicates quantitative formation of the nucleotide-metal complex for the case of a 1:12 Mn^{2+}:*Hp*DnaB monomer ratio. The binding of the nucleotide-metal complex to the *Hp*DnaB helicase causes a further prolongation of the rotational correlation time and thus a further change of the EPR line shape as well as changes in the zero-field splitting parameters, which are expected from the changes in the ligand field (**See Figure S1** for EDEPR spectra at 10 K showing the influence of differences in the zero-field splitting on the EPR lineshape). Protein binding can even better be apprehended by the Mn^{2+} transverse relaxation curves shown in **Figure 48B** (see also **Figure S2** for results on a deuterated HpDnaB). For the protein complex we observe bimodal Mn^{2+} relaxation behavior, indicating that AMP-PMP:Mn^{2+} complexes bound to the NBDs relax significantly faster than the isolated AMP-PMP:Mn^{2+} in buffer solution. This is probably due to the close proximity of up to five other Mn^{2+} ions, as well as fluctuating hyperfine fields caused by proton spin-diffusion in the helicase that strongly exceed the fluctuating fields in the deuterated buffer matrix of the unbound complex. The binding efficiency can be determined in good approximation by dividing the normalized transverse relaxation traces of *Hp*DnaB:APMPNP:Mn^{2+} by the APMPNP:Mn^{2+} trace, see **Figure 48B**.

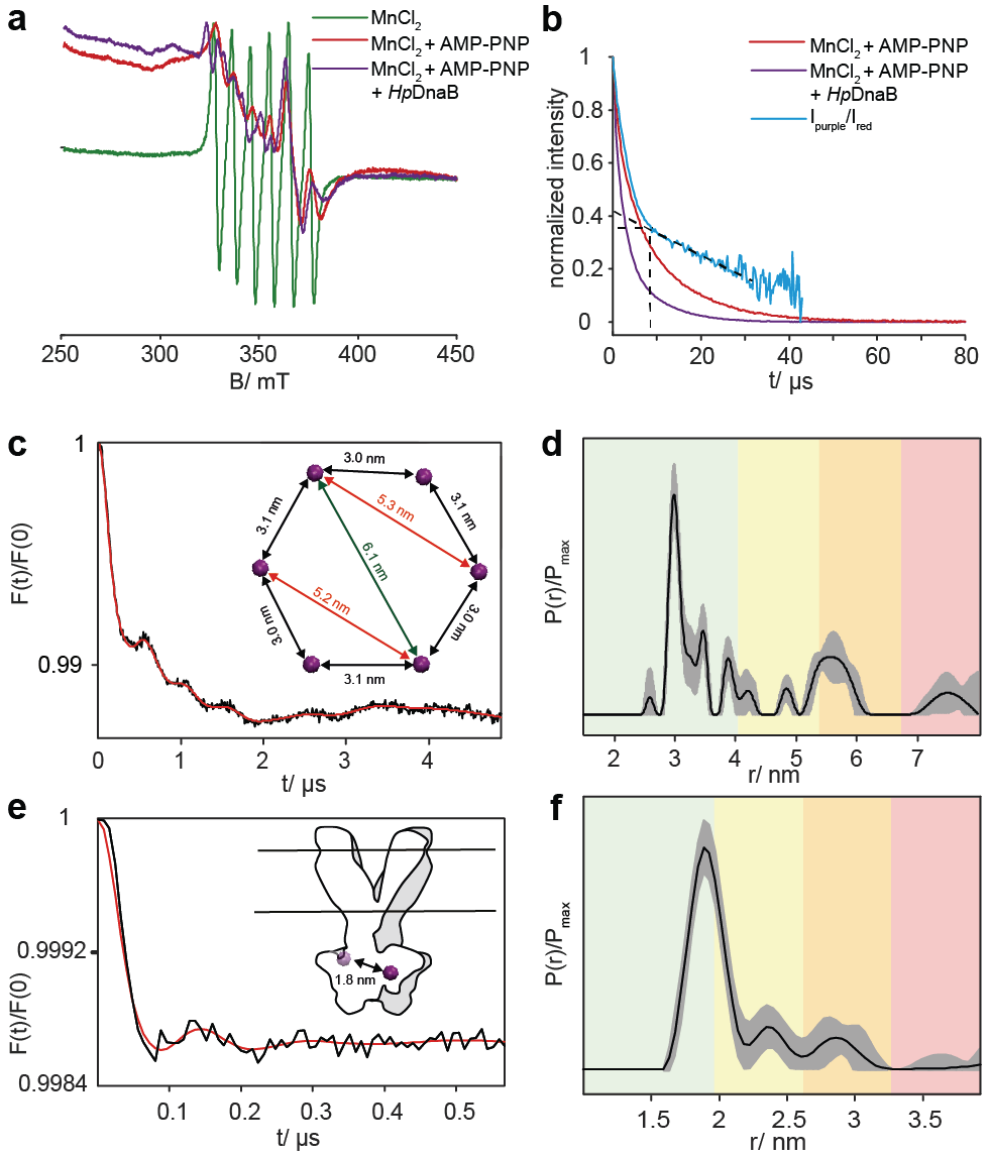


Figure 48. Mn^{2+} EPR experiments on the helicase HpDnaB and on the ABC transporter BmrA. (**A**) Mn^{2+} X-band CW spectra at room temperature of a $MnCl_2$ (green), a APMPNP:Mn^{2+} (red), and a HpDnaB:AMPPNP:Mn^{2+} solution (purple, HpDnaB:Mn^{2+} 12:1). (**B**) Mn^{2+} Hahn echo decay curves of APMPNP:Mn^{2+} and HpDnaB:APMPNP:Mn^{2+} (HpDnaB:Mn^{2+} 1:1) at 20 K and division of the normalized decaying functions of APMPNP:Mn^{2+} and HpDnaB:APMPNP:Mn^{2+} (blue). (**C**) Mn^{2+}–Mn^{2+} ultra-wideband DEER form factor (black) and simulated form factor corresponding to the distance distribution (red) and (**D**) distance distribution at 10 K for deuterated HpDnaB (1:1.5 protein:Mn^{2+} ratio; green, shape of distance distribution is reliable; yellow, mean distance and width are reliable; orange, mean distance is reliable; red, long-range distance contributions may be detectable, but cannot be quantified). Gray indicates uncertainty revealed by the validation (see the Supporting Information); 1:1.5 protein:Mn^{2+} ratio). (**E**) BmrA:ADP:Mn^{2+}:VO_4^{3-} Mn^{2+}–Mn^{2+} ultra-wideband DEER form factor (black) and simulated form factor corresponding to the distance distribution (red) and (**F**) distance distribution at a temperature of 10 K for Mn^{2+} in BmrA:ADP:Mn^{2+}:VO_4^{3-}. Colors as in (**D**).

The slowly decaying component crosses the vertical axis at a value of 0.4, indicating that about 60 % of AMP-PMP:Mn^{2+} complexes are bound to NBDs for the case of a 1:1 Mn^{2+}/*Hp*DnaB monomer ratio (this leads to a K_d value in the order of 10^{-5} M in agreement with published values for similar proteins) (Bujalowski & Klonowska 1993; Jezewska *et al.* 1996). Thus for NMR experiments we used a five-fold excess of Mn^{2+} and a 18-fold excess of APMPNP to occupy all binding sites.

The distribution of the Mn^{2+}-Mn^{2+} distances was determined by Double Electron Electron Resonance (DEER) with broadband chirp pump pulses (see Material and Methods Section in SI) in order to obtain a measurable dipolar modulation depth in spite of the large width of Mn^{2+} EPR spectrum (Banerjee *et al.* 2012; Doll *et al.* 2015b; Pannier *et al.* 2011; Vincent Ching *et al.* 2015). **Figure 48C** shows the DEER form factor trace obtained as a sum of two measurements on a *Hp*DnaB:APMPNP:Mn^{2+} sample (see **Figure S3** for a detailed description) and **Figure 48D** the resulting distance distribution obtained by Tikhonov regularization analysis (Jeschke *et al.* 2006). Distances of 3 nm and, with lower significance, of 5-6 nm are found, showing that the (double-) hexameric assembly in *Hp*DnaB (see inset **Figure 48C**) is conserved, with slight deviations from the homology model based on the *Aa*DnaB:ADP:Mg^{2+} complex (Strycharska *et al.* 2013) being expected. Peaks in the range between 3.4 and 4 nm are less significant and may be related to flexibility or to unspecifically bound Mn^{2+}. Due to the low inversion efficiency of even the ultra-wideband pump pulses, multispin contributions in the multimeric assembly (Pliotas *et al.* 2012; Valera *et al.* 2016) are not expected in our case.

Mn^{2+}-Mn^{2+} distance distributions in BmrA:ADP:Mn^{2+}:VO_4^{3-} could be determined by DEER experiments despite the low modulation depths due to the larger amount of free Mn^{2+} (see **Figures 48 E,F**). The distance distribution has a maximum at about 1.9 nm, which is in agreement with the Mn^{2+}-Mn^{2+} distances in the homology model of BmrA:ADP:Mn^{2+}:VO_4^{3-} representing the closed form (1.8 nm).

2. NMR on HpDnaB in complex with APMPNP:Mg^{2+} and APMPNP:Mn^{2+}

The 2D DARR ^{13}C-^{13}C NMR correlation spectra (Takegoshi *et al.* 2003) of the diamagnetic *Hp*DnaB:APMPNP:Mg^{2+} and paramagnetic *Hp*DnaB:APMPNP:Mn^{2+} complexes are given in **Figure S4**, and expansion of the alanine region in **Figure 49A** (for the Thr/Ile/Val regions see **Figures S5** and **S6**). Assignments are taken from ref (Wiegand *et al.* 2015; 2016a). Three

types of resonances can be distinguished in the 2D spectra: (i) residues unaffected by Mn^{2+} (e.g. 24A), (ii) residues attenuated by Mn^{2+} PRE (e.g. 228A, 310A) and (iii) resonances disappearing at the given signal-to-noise ratio because of strong PRE (e.g. 203A, 351A). **Figure 49A** further reveals that all nucleotide binding sites are occupied by APMPNP:Mn^{2+}, since resonances do not split, but shift completely. The peak intensity in the 2D spectra is influenced by the faster T_1, $T_{1\rho}$ and T_2 relaxation, resulting in a reduced signal intensity due to losses during the polarization-transfer periods and T_2 broadening, respectively (Tamaki *et al.* 2016). PRE scales with r^{-6}, r being the distance between the metal center and a protein nucleus. The radius of the sphere around Mn^{2+} with significant PRE depends on the T_1 relaxation time of unpaired electrons of Mn^{2+}, which cannot easily be measured at a temperature of 278 K. According to a homology model of *Hp*DnaB, based on *Aa*DnaB, the radius of view can be estimated to be 15 Å. Further site-specific PRE data were obtained from 3D NCACB spectra (**Figure S8**) and are summarized in **Figure 49B.** The N-terminal domain is hardly affected by Mg^{2+}/Mn^{2+} substitution, whereas parts of the NBD containing C-terminal domain (CTD) experience strong PREs (I_{para}/I_{dia}<0.4). Stretches of residues with Cα atoms closer than 15 Å to the Mn^{2+} are marked as vertical red bars and show strong PREs. The agreement is not perfect, probably due to the limitations of the homology model. **Figure 49C** shows a side view on one *Hp*DnaB hexamer model with the PRE effects plotted on the structure. Significant PRE's indeed only appear in the CTD in vicinity of the paramagnetic centers (in contrast to chemical-shift perturbations (CSPs), which can be a consequence of both proximal and allosteric changes(Williamson 2013)). An alternative view is provided in **Figure 49D** where spheres with a radius of 15 Å are plotted around the metal centers showing that resonances within this sphere are strongly affected by PRE effects upon Mn^{2+} binding. The PRE determined here correlate with the CSPs upon APMPNP:Mg^{2+} binding that were determined earlier (see **Figure S9**) (Wiegand *et al.* 2016a).

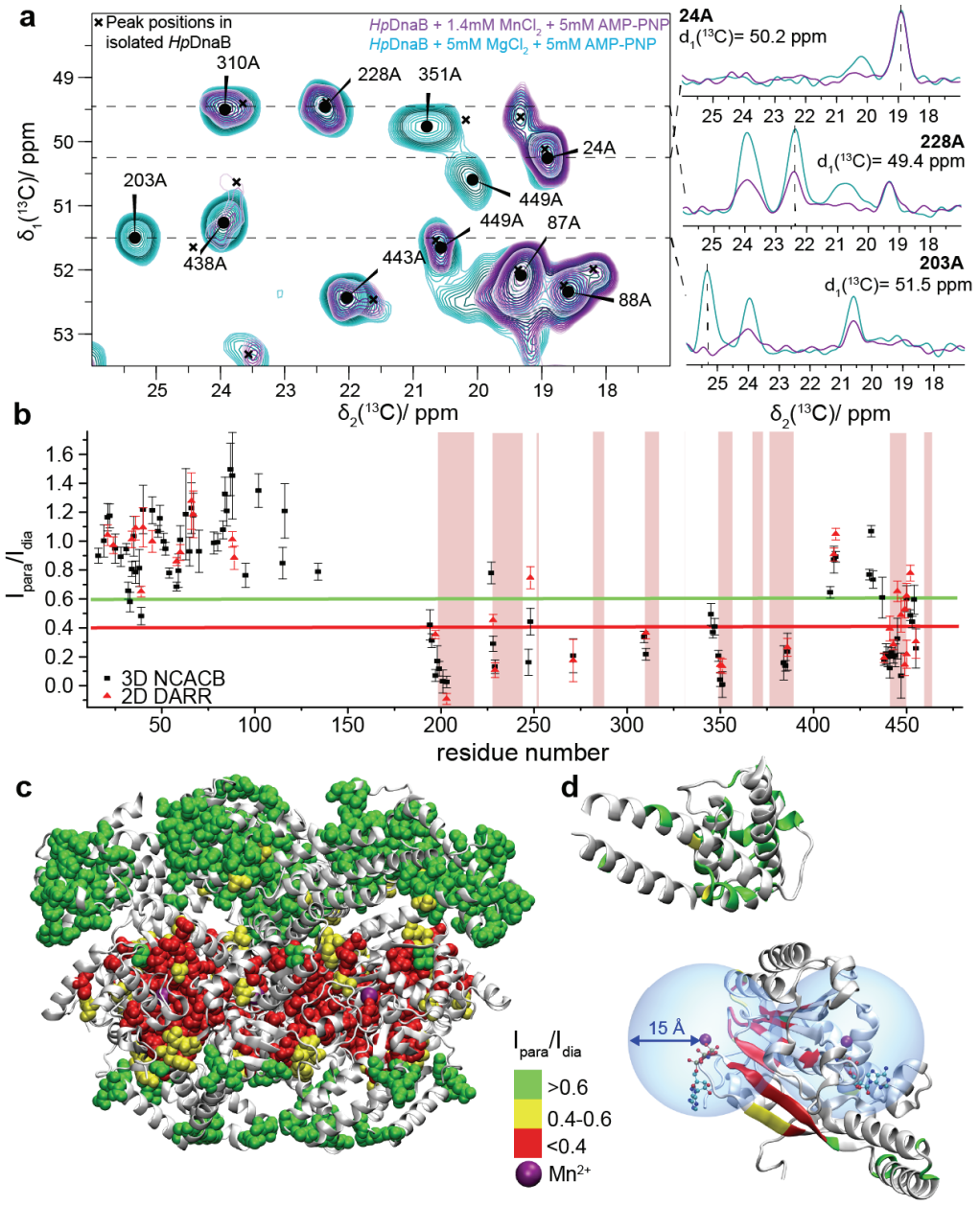


Figure 49. NMR experiments on the helicase HpDnaB in complex with APMPNP:Mg^{2+} and APMPNP:Mn^{2+} (**A**) 2D ^{13}C-^{13}C DARR spectra showing the Ala region of the *Hp*DnaB:APMPNP:Mg^{2+} (cyan) and of the *Hp*DnaB:APMPNP:Mn^{2+} (purple) complexes. 1D traces along F2 are shown for (i) unaffected resonances (24A), (ii) attenuated resonances (228A) and (iii) disappearing resonances (203A). (**B**) PRE effects (black squares: determined from 3D NCACB spectra, red triangles: determined from 2D DARR spectra). Residues with effective distances < 15 Å (distance between Cα and the two nearest metals) (see **Figure S7**) are highlighted in red. The red and green horizontal lines indicate I_{para}/I_{dia} values of 0.4 and 0.6 which we consider as bounds for residues experiencing strong (<0,4) and no PRE effects (>0.6). (**C**) *Hp*DnaB hexamer with PRE effects determined from 3D NCACB spectra plotted as vdW balls. (**D**) Zoom into one *Hp*DnaB monomer with PRE effects plotted on the structure. Spheres with R = 15 Å around the Mn^{2+} ions from two adjacent monomers are shown in blue.

3. NMR on BmrA in complex with ADP:Mg^{2+}:VO_4^{3-} and ADP:Mn^{2+}: VO_4^{3-}

To study the effects of Mn^{2+} for a protein which shows lower nucleotide binding constants, we investigated the ABC transporter BmrA, for which a larger excess of Mn^{2+} compared to a BmrA monomer is used to obtain quantitative binding (300-fold excess compared to 5-fold in *Hp*DnaB). The alanine region of the 2D DARR spectra of BmrA, recorded on a reverse labeled ^{12}C-^{14}N-[LVIKHP] sample of BmrA:ADP:Mn^{2+}:VO_4^{3-}, is shown in **Figure 50A**.

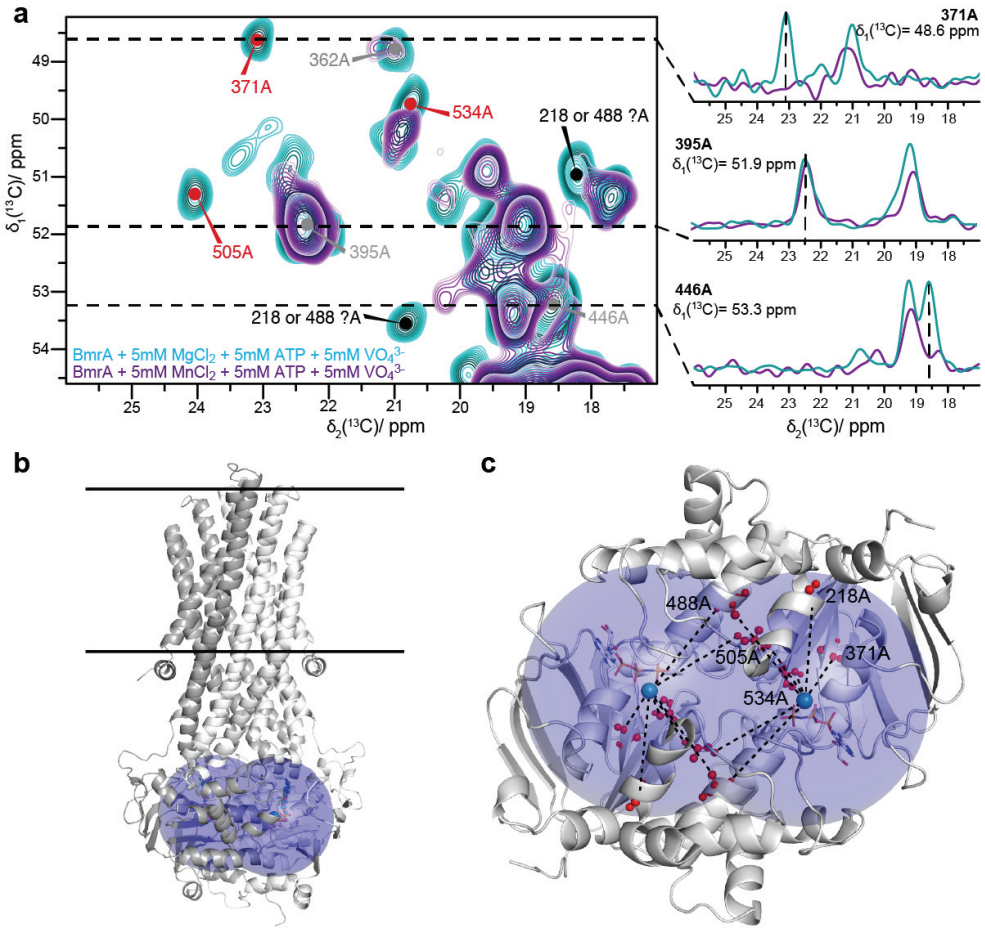


Figure 50. NMR experiments on the ABC transporter in complex with ADP:Mg^{2+}:VO43- and ADP:Mn^{2+}:VO_4^{3-} A) Alanine region from 2D ^{13}C-^{13}C DARR spectra of the BmrA:ADP:Mg^{2+}:VO_4^{3-} (cyan) and of the BmrA:ADP:Mn^{2+}:VO_4^{3-} (purple) complexes. Residues are labeled in red/black for assigned/unassigned Ala within 15 Å of the Mn^{2+}, and in grey for Ala outside this limit. Representative 1D traces along F2 are shown as insets with examples for disappearing, unaffected and attenuated resonances respectively. **B**) BmrA model obtained by homology from the ABC transporter SAV1866 x-ray structure(Dawson & Locher 2007) (2HYD) embedded in lipid bilayers represented by horizontal lines. The spheres shown in blue are centered at the Mg^{2+} and show a radius of 15 Å. **C**) Plot of the nucleotide binding domain and all Ala residues (red) located within 15 Å radius from the metal (cyan).

The full aliphatic region, as well as the Ser and Thr regions are shown in **Figures S10 and S11,** respectively. When comparing this spectrum to the one of diamagnetic BmrA:ADP:Mg^{2+}:VO_4^{3-}, one can also observe - as in the case of *Hp*DnaB discussed above - non-attenuated, attenuated and disappearing resonances in the presence of the paramagnetic ion. From a set of assigned Ala residues (**Figure 50A**) we find that the peaks of 371A, 505A and 534A show strong PRE in accordance with a homology model (based on Sav1866 (Dawson & Locher 2007)) where these residues are indeed located within a sphere of 15 Å from the metal center (**Figure 50B and 50C**). Two more alanine residues within this radius, 488A and 218A, can likely be assigned to the two additional disappearing signals labeled in black in **Figure 50A**. Two additional assigned Ala residues outside the 15 Å sphere show also attenuation by PREs. We note that 446A has four nearby negatively charged residues as neighbors, 442D, 445D, 447E and 449E. A closer look at the C' side chain region displaying mainly Glu Cδ and Asp Cγ resonances reveals that most corresponding signals are significantly attenuated. As Asp and Glu are good binders of Mn^{2+} (Khrustaleva 2014), this is likely due to the surplus Mn^{2+} interacting with the negatively charged side chains of these residues (see also **Figure 51A**). In line with this, the Glu Cγ/Cδ cross signal is attenuated by about 80 % (**Figure 51B**). The attenuation of the Cβ/Cγ cross signals is already smaller (<45 %), indicating that the effect is local and smaller than the one exerted by Mn^{2+} located in the nucleotide binding site (**Figure 51C**), probably due to fast chemical exchange of the surface-bound Mn^{2+} with Mn^{2+} in solution.

The residues located in the trans-membrane helices are not influenced by PRE, as can be seen at the signals of the α-helical Ala residues, which are the least affected peaks in the spectra, and have been used to calibrate the two spectra with respect to each other (**Figure 51C**).

Conclusion

In summary, we have demonstrated on two examples that NBD engines in proteins can be investigated by the exchange of diamagnetic Mg^{2+} with paramagnetic Mn^{2+}. No tagging is necessary, and no distribution of metal positions, like in flexible tags, exists. Pulsed EPR methods, in particular ultra-wideband DEER, allow for the evaluation of distances between paramagnetic centers in the multimers. The measurement of PREs in solid-state NMR identifies the residues close to the binding site unambiguously, in contrast to CSPs that

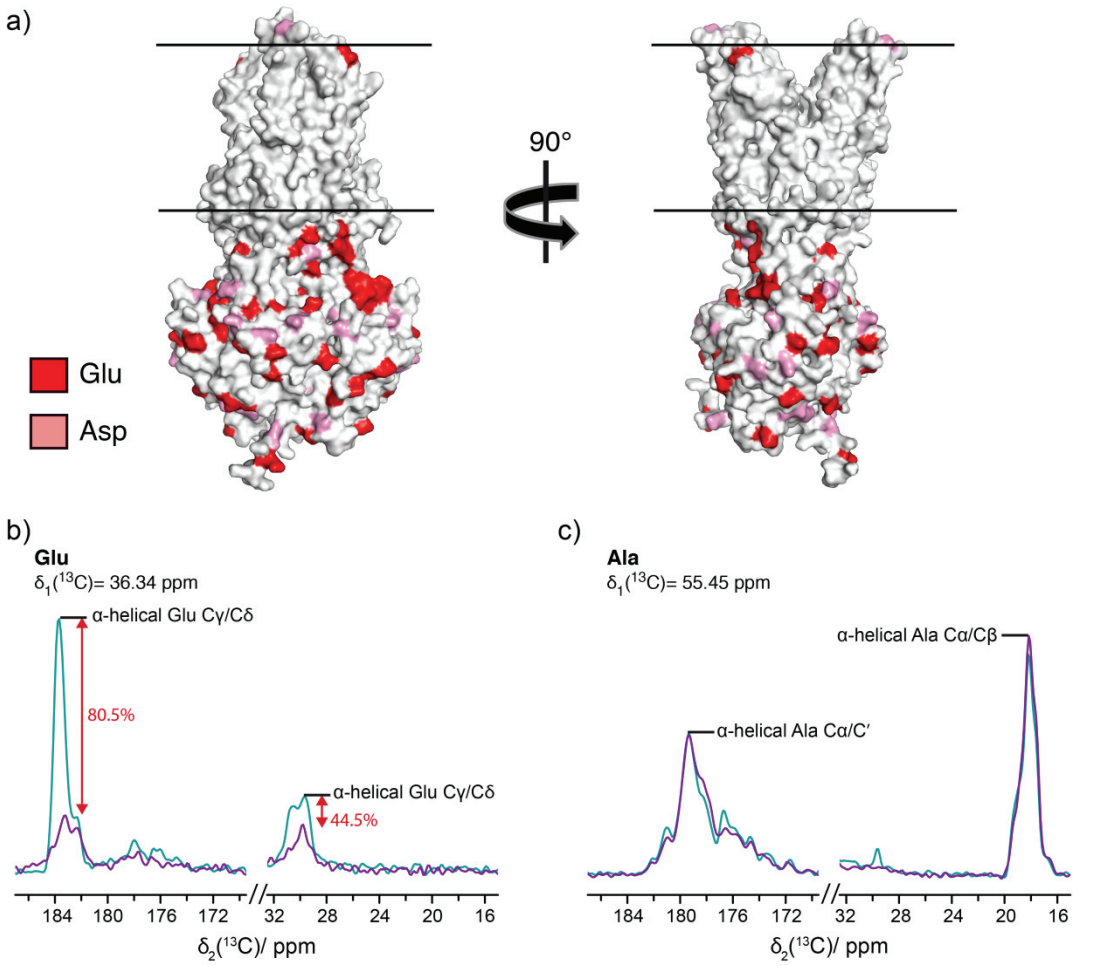


Figure 51. Glu and Asp effect on Mn^{2+} unspecific binding a) Model of BmrA with the lipid bilayer represented by two lines. Glutamate and the aspartate residues are respectively plotted in red and pink. The model used was obtained by homology from the ABC transporter SAV1866 x-ray structure (2HYD).(Dawson & Locher 2006) **b)** and **c)** 1D traces of the alpha helical alanine (**b**) and glutamate (**c**) residues of BmrA in presence of Mg^{2+} (cyan traces) and Mn^{2+} (purple traces). The traces highlight the reduction of the signal from the glutamate residues induced by the non-specific binding of the Mn^{2+} ions. In comparison, regarding the 1D trace the alpha helical alanine residues, mainly protected by the lipid bilayer, are not affected.

cannot distinguish direct and allosteric effects. Additional unspecific binding to solvent-exposed sidechains was observed in BmrA due to the higher Mn^{2+} concentrations. They act, however, in a smaller radius of a few Å and do not hinder the characterization of the binding site. The spectroscopic approaches described here should be generally applicable to the study of the many proteins possessing ATP-fueled engines, and a combination of approaches should allow deeper insights in the mechanisms powering these machines, and also the remote conformational changes driven by the engines.

Acknowledgements

We would like to thank Cédric Orelle and Jean-Michel Jault (MMSB Lyon) for guiding us for activity testing of BmrA in presence of Mn^{2+}; Simon Widler for support with the sample preparation; as well as Andrin Doll and Frauke Breitgoff for introduction to the AWG setup. This work was supported by the Swiss National Science Foundation (Grant 200020_159707, 200020_146757 and 200020_157034), the French ANR (ANR-14-CE09-0024B) and the ETH Career SEED-69 16-1.

References

Banerjee, D., Yagi, H., Huber, T., Otting, G. & Goldfarb, D. (2012) Nanometer-Range Distance Measurement in a Protein Using Mn 2+Tags. *The Journal of Physical Chemistry Letters* 3, 157–160.

Bazin, A., Cherrier, M.V., Gutsche, I., Timmins, J. & Terradot, L. (2015) Structure and primase-mediated activation of a bacterial dodecameric replicative helicase. *Nucleic Acids Research* 43, 8564–8576.

Bennati, M., Hertel, M.M., Fritscher, J., Prisner, T.F., Weiden, N., Hofweber, R., Spörner, M., Horn, G. & Kalbitzer, H.R. (2006) High-Frequency 94 GHz ENDOR Characterization of the Metal Binding Site in Wild-Type Ras·GDP and Its Oncogenic Mutant G12V in Frozen Solution †. *Biochemistry* 45, 42–50.

Bertini, I., Emsley, L., Lelli, M., Luchinat, C., Mao, J. & Pintacuda, G. (2010) Ultrafast MAS Solid-State NMR Permits Extensive 13C and 1H Detection in Paramagnetic Metalloproteins. *Journal of the American Chemical Society* 132, 5558–5559.

Bertini, I., Luchinat, C., Parigi, G. & Pierattelli, R. (2005) NMR Spectroscopy of Paramagnetic Metalloproteins. *ChemBioChem* 6, 1536–1549.

Bode, B.E., Margraf, D., Plackmeyer, J., Dürner, G., Prisner, T.F. & Schiemann, O. (2007) Counting the monomers in nanometer-sized oligomers by pulsed electron-electron double resonance. *Journal of the American Chemical Society* 129, 6736–6745.

Bonneau, E. & Legault, P. (2014) NMR Localization of Divalent Cations at the Active Site of the NeurosporaVS Ribozyme Provides Insights into RNA–Metal-Ion Interactions. *Biochemistry* 53, 579–590.

Böckmann, A., Gardiennet, C., Verel, R., Hunkeler, A., Loquet, A., Pintacuda, G., Emsley, L., Meier, B.H. & Lesage, A. (2009) Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45, 319–327.

Bujalowski, W. & Klonowska, M.M. (1993) Negative cooperativity in the binding of nucleotides to Escherichia coli replicative helicase DnaB protein. Interactions with

fluorescent nucleotide analogs. *Journal of Biological Chemistry* 32, 5888–5900.

Ching, H.Y.V., Mascali, F.C., Bertrand, H.C., Bruch, E.M., Demay-Drouhard, P., Rasia, R.M., Policar, C., Tabares, L.C. & Un, S. (2016) The Use of Mn(II) Bound to His-tags as Genetically Encodable Spin-Label for Nanometric Distance Determination in Proteins. *The Journal of Physical Chemistry Letters* 7, 1072–1076.

Dawson, R.J.P. & Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185.

Dawson, R.J.P. & Locher, K.P. (2007) Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureusin complex with APMPNP. *FEBS Letters* 581, 935–938.

Doll, A., Qi, M., Pribitzer, S., Wili, N., Yulikov, M., Godt, A. & Jeschke, G. (2015a) Sensitivity enhancement by population transfer in Gd(III) spin labels. *Physical Chemistry Chemical Physics* 17, 7334–7344.

Doll, A., Qi, M., Wili, N., Pribitzer, S., Godt, A. & Jeschke, G. (2015b) Gd(III)–Gd(III) distance measurements with chirp pump pulses. *Journal of Magnetic Resonance* 259, 153–162.

Gardiennet, C., Schütz, A.K., Hunkeler, A., Kunert, B., Terradot, L., Böckmann, A. & Meier, B.H. (2012) A Sedimented Sample of a 59 kDa Dodecameric Helicase Yields High-Resolution Solid-State NMR Spectra. *Angewandte Chemie International Edition* 51, 7855–7858.

Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A. & Jault, J.M. (2001) A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends in Biochemical Sciences* 26, 539–544.

Hagelueken, G., Ingledew, W.J., Huang, H., Petrovic-Stojanovska, B., Whitfield, C., ElMkami, H., Schiemann, O. & Naismith, J.H. (2009) PELDOR spectroscopy distance fingerprinting of the octameric outer-membrane protein Wza from Escherichia coli. *Angewandte Chemie (International ed. in English)* 48, 2904–2906.

Hanson, P.I. & Whiteheart, S.W. (2005) AAA+ proteins: have engine, will work. *Nature Reviews Molecular Cell Biology* 6, 519–529.

Jaroniec, C.P. (2015) Structural studies of proteins by paramagnetic solid-state NMR spectroscopy. *Journal of Magnetic Resonance* 253, 50–59.

Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C.R., Hilger, D. & Jung, H. (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* 30, 473–498.

Jezewska, M.J., Kim, U.S. & Bujalowski, W. (1996) Interactions of Escherichia coli primary replicative helicase DnaB protein with nucleotide cofactors. *Biophysical Journal* 71, 2075–2086.

Kaur, H., Lakatos, A., Spadaccini, R., Vogel, R., Hoffmann, C., Becker-Baldus, J., Ouari, O., Tordo, P., Mchaourab, H. & Glaubitz, C. (2015) The ABC exporter MsbA probed by

solid state NMR – challenges and opportunities. *Biological chemistry* 396, 1135–1149.

Khrustaleva, T.A. (2014) Secondary Structure Preferences of Mn 2+Binding Sites in Bacterial Proteins. *Advances in Bioinformatics* 2014, 1–14.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Lueders, P., Razzaghi, S., Jäger, H., Tschaggelar, R., Hemminga, M.A., Yulikov, M. & Jeschke, G. (2013) Distance determination from dysprosium induced relaxation enhancement: a case study on membrane-inserted WALP23 polypeptides. *Molecular Physics* 111, 2824–2833.

Orelle, C., Gubellini, F., Durand, A., Marco, S., Lévy, D., Gros, P., Di Pietro, A. & Jault, J.-M. (2008) Conformational change induced by ATP binding in the multidrug ATP-binding cassette transporter BmrA. *Biochemistry* 47, 2404–2412.

Otting, G. (2010) Protein NMR Using Paramagnetic Ions. *Annual Review of Biophysics* 39, 387–405.

Pannier, M., Veit, S., Godt, A., Jeschke, G. & Spiess, H.W. (2011) Dead-time free measurement of dipole–dipole interactions between electron spins. *Journal of Magnetic Resonance* 213, 316–325.

Pintacuda, G., Giraud, N., Pierattelli, R., Böckmann, A., Bertini, I. & Emsley, L. (2007) Solid-State NMR Spectroscopy of a Paramagnetic Protein: Assignment and Study of Human Dimeric Oxidized CuII–ZnII Superoxide Dismutase (SOD). *Angewandte Chemie International Edition* 46, 1079–1082.

Pliotas, C., Pliotas, C., Ward, R., Ward, R., Branigan, E., Branigan, E., Rasmussen, A., Rasmussen, A., Hagelueken, G., Hagelueken, G., Huang, H., Huang, H., Black, S.S., Black, S.S., Booth, I.R., Booth, I.R., Schiemann, O., Schiemann, O., Naismith, J.H. & Naismith, J.H. (2012) Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proceedings of the National Academy of Sciences* 109, E2675–E2682.

Polyhach, Y., Bordignon, E., Tschaggelar, R., Gandra, S., Godt, A. & Jeschke, G. (2012) High sensitivity and versatility of the DEER experiment on nitroxide radical pairs at Q-band frequencies. *Physical Chemistry Chemical Physics* 14, 10762–10773.

Razzaghi, S., Brooks, E.K., Bordignon, E., Hubbell, W.L., Yulikov, M. & Jeschke, G. (2013) EPR Relaxation-Enhancement-Based Distance Measurements on Orthogonally Spin-Labeled T4-Lysozyme. *ChemBioChem* 14, 1883–1890.

Rees, D.C., Johnson, E. & Lewinson, O. (2009) ABC transporters: the power to change. *Nature Reviews Molecular Cell Biology* 10, 218–227.

Siarheyeva, A., Liu, R. & Sharom, F.J. (2010) Characterization of an asymmetric occluded state of P-glycoprotein with two bound nucleotides: implications for catalysis. *The*

Journal of biological chemistry 285, 7575–7586.

Soni, R.K. (2003) Functional characterization of Helicobacter pylori DnaB helicase. *Nucleic Acids Research* 31, 6828–6840.

Steinfels, E., Orelle, C., Dalmas, O., Penin, F., Miroux, B., Di Pietro, A. & Jault, J.-M. (2002) Highly efficient over-production in E. coli of YvcC, a multidrug-like ATP-binding cassette transporter from Bacillus subtilis. *Biochimica et biophysica acta* 1565, 1–5.

Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A. & Jault, J.-M. (2004) Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in Bacillus subtilis†. *Biochemistry* 43, 7491–7502.

Stevens, T.J., Fogh, R.H., Boucher, W., Higman, V.A., Eisenmenger, F., Bardiaux, B., van Rossum, B.-J., Oschkinat, H. & Laue, E.D. (2011) A software framework for analysing solid-state MAS NMR data. *Journal of Biomolecular NMR* 51, 437–447.

Strycharska, M.S., Arias-Palomo, E., Lyubimov, A.Y., Erzberger, J.P., O'Shea, V.L., Bustamante, C.J. & Berger, J.M. (2013) Nucleotide and Partner-Protein Control of Bacterial Replicative Helicase Structure and Function. *Molecular cell* 52, 844–854.

Takegoshi, K., Nakamura, S. & Terao, T. (2003) [sup 13]C–[sup 1]H dipolar-driven [sup 13]C–[sup 13]C recoupling without [sup 13]C rf irradiation in nuclear magnetic resonance of rotating solids. *The Journal of Chemical Physics* 118, 2325.

Tamaki, H., Egawa, A., Kido, K., Kameda, T., Kamiya, M., Kikukawa, T., Aizawa, T., Fujiwara, T. & Demura, M. (2016) Structure determination of uniformly 13C, 15N labeled protein using qualitative distance restraints from MAS solid-state 13C-NMR observed paramagnetic relaxation enhancement. *Journal of Biomolecular NMR* 64, 87–101.

Ullrich, S.J., Hölper, S. & Glaubitz, C. (2013) Paramagnetic doping of a 7TM membrane protein in lipid bilayers by Gd3+-complexes for solid-state NMR spectroscopy. *Journal of Biomolecular NMR* 58, 27–35.

Valera, S., Ackermann, K., Pliotas, C., Huang, H., Naismith, J.H. & Bode, B.E. (2016) Accurate Extraction of Nanometer Distances in Multimers by Pulse EPR. *Chemistry - A European Journal* 22, 4700–4703.

Vincent Ching, H.Y., Demay-Drouhard, P., Bertrand, H.C., Policar, C., Tabares, L.C. & Un, S. (2015) Nanometric distance measurements between Mn(ii)DOTA centers. *Physical Chemistry Chemical Physics* 17, 23368–23377.

Wiegand, T., Cadalbert, R., Gardiennet, C., Timmins, J., Terradot, L., Böckmann, A. & Meier, B.H. (2016a) Monitoring ssDNA Binding to the DnaB Helicase from Helicobacter pyloriby Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* 55, 14164–14168.

Wiegand, T., Gardiennet, C., Cadalbert, R., Lacabanne, D., Kunert, B., Terradot, L., Böckmann, A. & Meier, B.H. (2016b) Variability and conservation of structural domains in divide-and-conquer approaches. *Journal of Biomolecular NMR* 65, 79–86.

Wiegand, T., Gardiennet, C., Ravotti, F., Bazin, A., Kunert, B., Lacabanne, D., Cadalbert, R., Güntert, P., Terradot, L., Böckmann, A. & Meier, B.H. (2015) Solid-state NMR sequential assignments of the N-terminal domain of HpDnaB helicase. *Biomolecular NMR Assignments* 10, 13–23.

Williamson, M.P. (2013) Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* 73, 1–16.

Witt, H., Wittershagen, A., Bill, E., Kolbesen, B.O. & Ludwig, B. (1997) Asp-193 and Glu-218 of subunit II are involved in the Mn2+-binding of Paracoccus denitrificans cytochrome c oxidase. *FEBS Letters* 409, 128–130.

Supplementary material

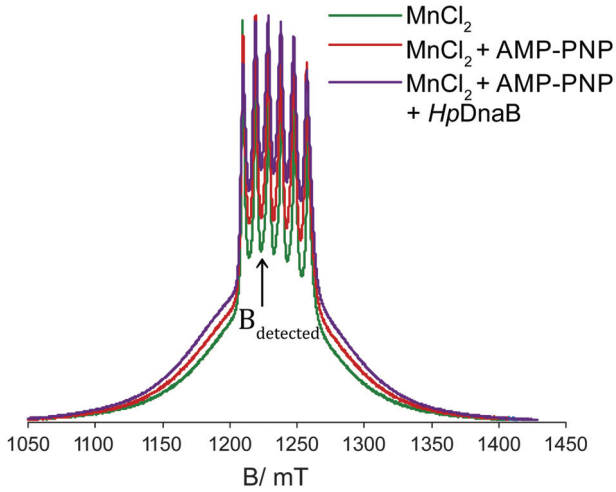


Figure S1: Anisotropic broadening of EDEPR spectra at 20 K and 34.5 GHz indicating the binding of Mn^{2+} to *Hp*DnaB: EDEPR spectra of a 0.3 mM $MnCl_2$ solution in phosphate/D_2O buffer (green), a 0.3 mM $MnCl_2$, 5 mM APMPNP solution in phosphate/D_2O buffer (red) and a 0.3 mM $MnCl_2$, 5mM APMPNP, 0.3mM *Hp*DnaB phosphate/D_2O buffer (purple). In contrast to the r.t. cw-EPR spectra in Figure 1, EDEPR experiments are performed in a frozen glassy matrix at 20 K and thus the spectral differences are dominated by changes in the zero-field splitting parameters, since molecular tumbling processes are frozen at these conditions. The observer field B_{detect} for relaxation and distance measurements is marked by an arrow.

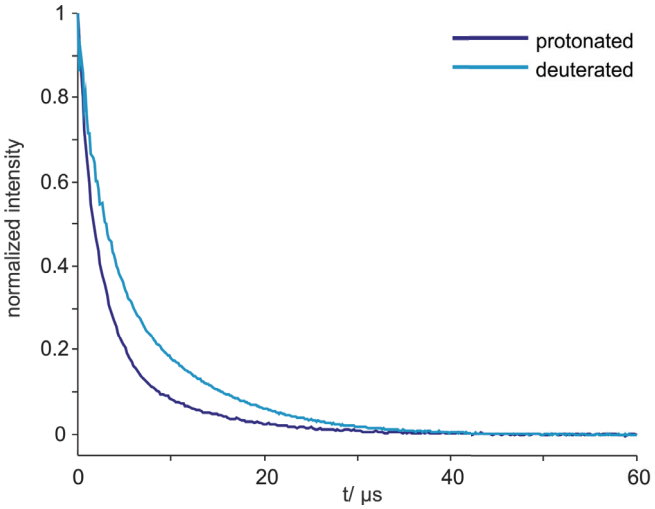


Figure S2: Hahn-echo decay of phase memory at 20 K. Comparison of 0.3 mM *Hp*DnaB phosphate/D_2O buffer expressed in protonated (purple) or deuterated (cyan) media. Expression in deuterated media allows for longer DEER traces, and, accordingly, to longer detectable Mn^{2+}-Mn^{2+}distances.

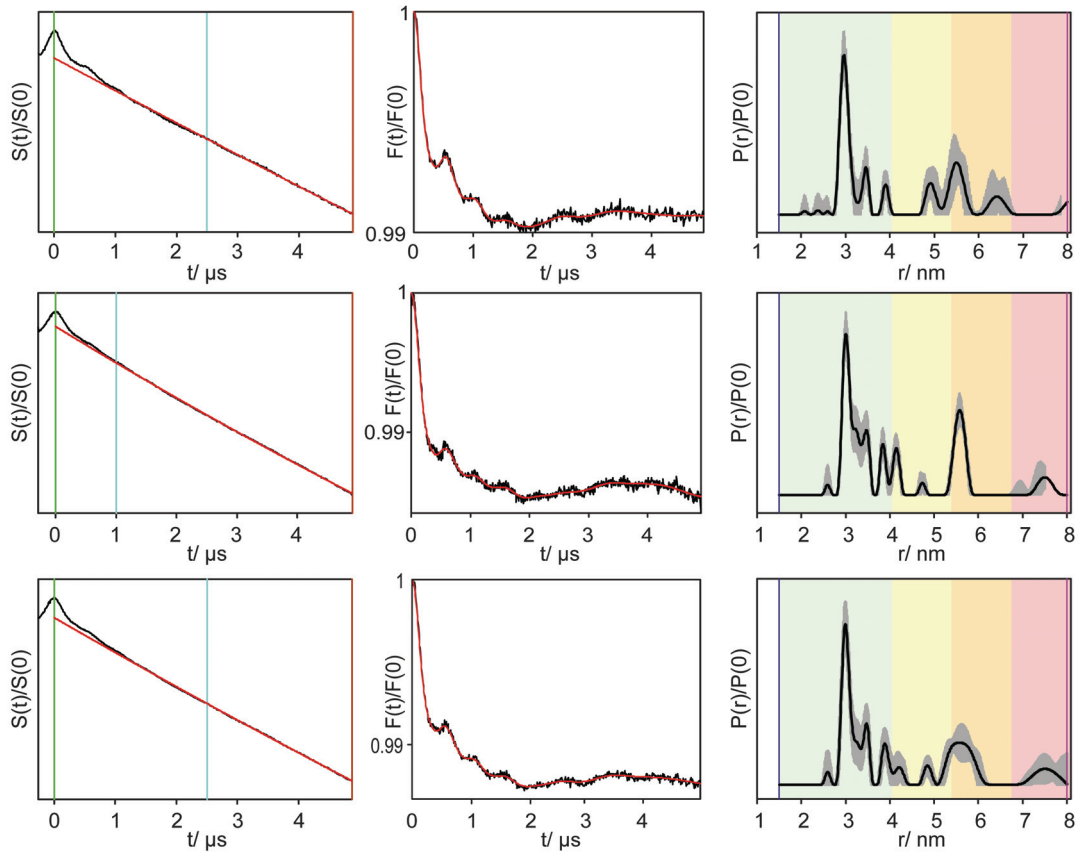


Figure S3: Overview of the Mn^{2+}-Mn^{2+} DEER data. (left three panels) Normalized primary DEER data (black line) and background fit (red line). The zero time is marked by the green vertical line, the cyan line marks the beginning of the background range fit, and the orange line labels the cutoff position. (middle three panels) Background corrected form factor in time domain (black line) and corresponding fit (red line), (right three panels) distance distribution. The shaded areas give an error estimate if the start of the background fit range is varied from 1000 to 2500 ns. The color coding indicates reliability ranges resulting from the limited length of the dipolar evolution trace. Pale green: Shape of distance distribution is reliable. Pale yellow: Mean distance and width are reliable. Pale orange: Mean distance is reliable. Pale red: Long-range distance contributions may be detectable, but cannot be quantified. [DeerAnalysis manual] Top, middle, bottom line of plots: see text above for the descriptions.

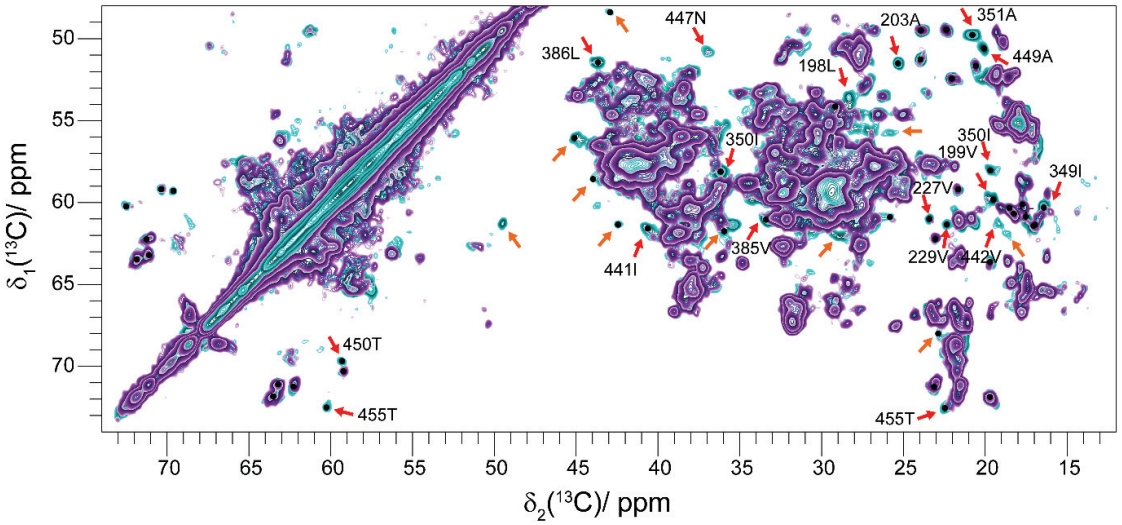


Figure S4: Aliphatic region of ^{13}C-^{13}C 20 ms DARR spectra of the HpDnaB:APMPNP:Mg^{2+} and of the HpDnaB:APMPNP:Mn^{2+} complexes. Most peaks experiencing chemical shift perturbations (CSPs) upon APMPNP:Mg^{2+} binding[10] are strongly attenuated or even disappearing in the *Hp*DnaB:APMPNP:Mn^{2+}- complex: Aliphatic region of ^{13}C-^{13}C 20 ms DARR spectra of the *Hp*DnaB:APMPNP:Mg^{2+} (18-fold excess Mg^{2+}, shown in cyan contours) and of the *Hp*DnaB:APMPNP:Mn^{2+} (5-fold excess Mn^{2+}, shown in purple contours) complexes. Strongly attenuated or even disappearing peaks are marked by red and orange arrows for assigned and unassigned peaks, respectively. Black circles indicate peaks experiencing CSPs > 0.15 ppm upon nucleotide binding compared to the isolated protein.

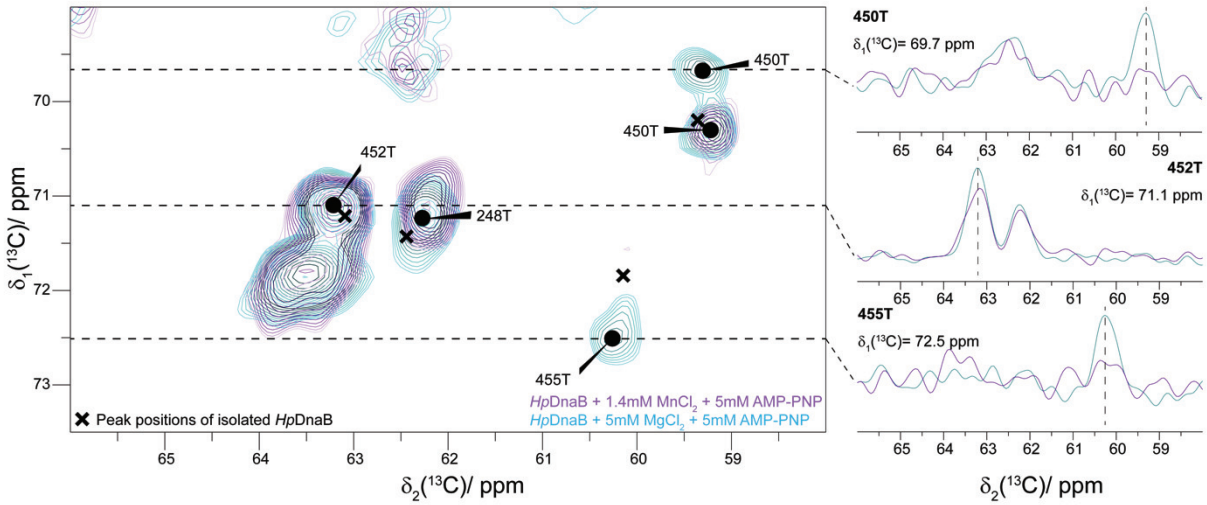


Figure S5: PRE effects influence 2D ^{13}C-^{13}C DARR spectra: Extract of threonine region from 2D ^{13}C-^{13}C DARR spectra of the *Hp*DnaB:APMPNP:Mg^{2+} (shown in cyan contours) and of the *Hp*DnaB:APMPNP:Mn^{2+} (shown in purple contours) complexes. Some representative 1D traces along F2 are shown as insets.

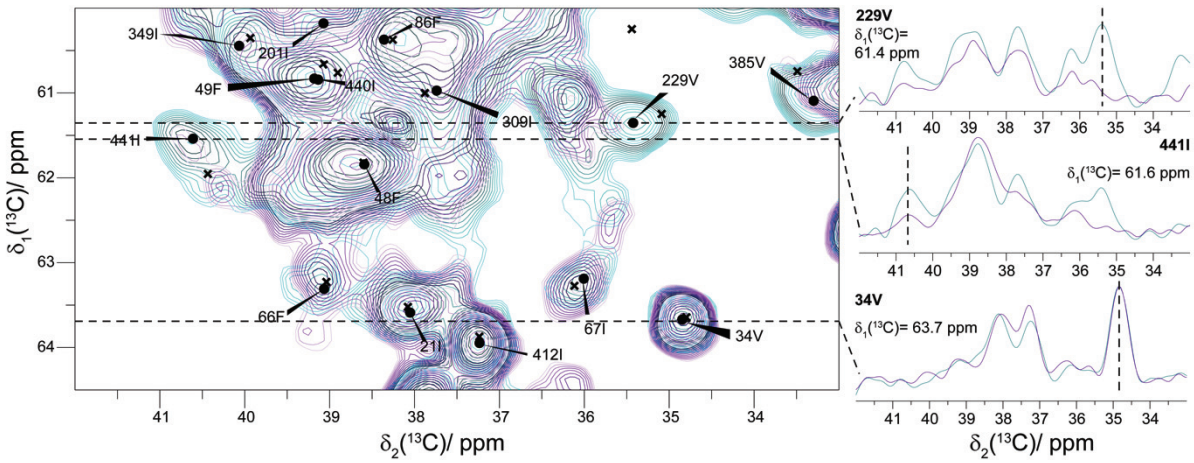


Figure S6: PRE effects influence 2D ^{13}C-^{13}C DARR spectra: Extract of isoleucine/valine region from 2D ^{13}C-^{13}C DARR spectra of the *Hp*DnaB:APMPNP:Mg^{2+} (shown in cyan contours) and of the *Hp*DnaB:APMPNP:Mn^{2+} (shown in purple contours) complexes. Some representative 1D traces along F2 are shown as insets.

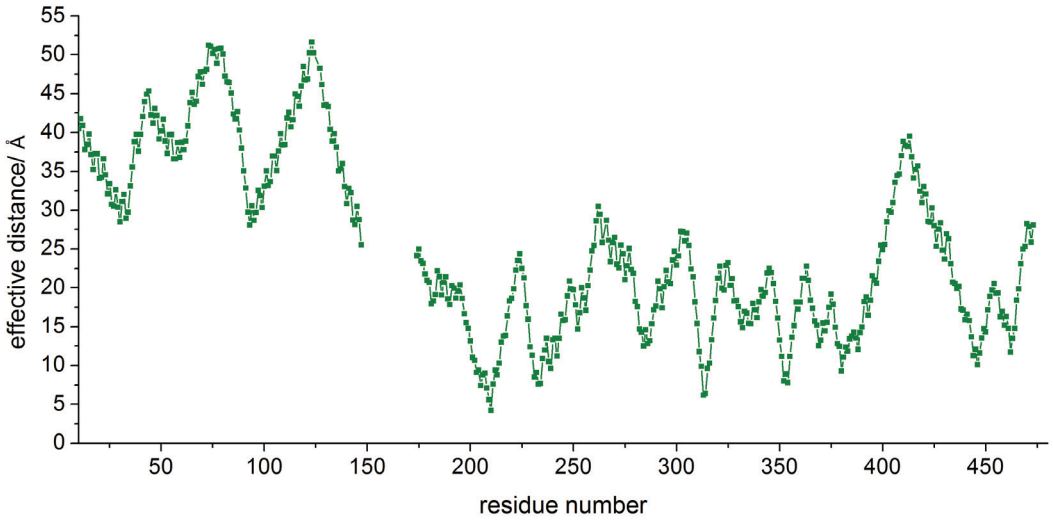


Figure S7: Effective distances between the protein and the metal center identify residues close to the NBD: Effective distances determined for *Hp*DnaB (calculated according to $\left(\sqrt[6]{\left(\frac{1}{d1}\right)^6+\left(\frac{1}{d2}\right)^6}\right)^{-1}$ with d_1 and d_2 representing the distance R between the metal ion and the Cα of the protein residue determined from the homology model). The measurements were performed on a homology model of *Hp*DnaB built from *Aa*DnaB:ADP:Mg^{2+}.

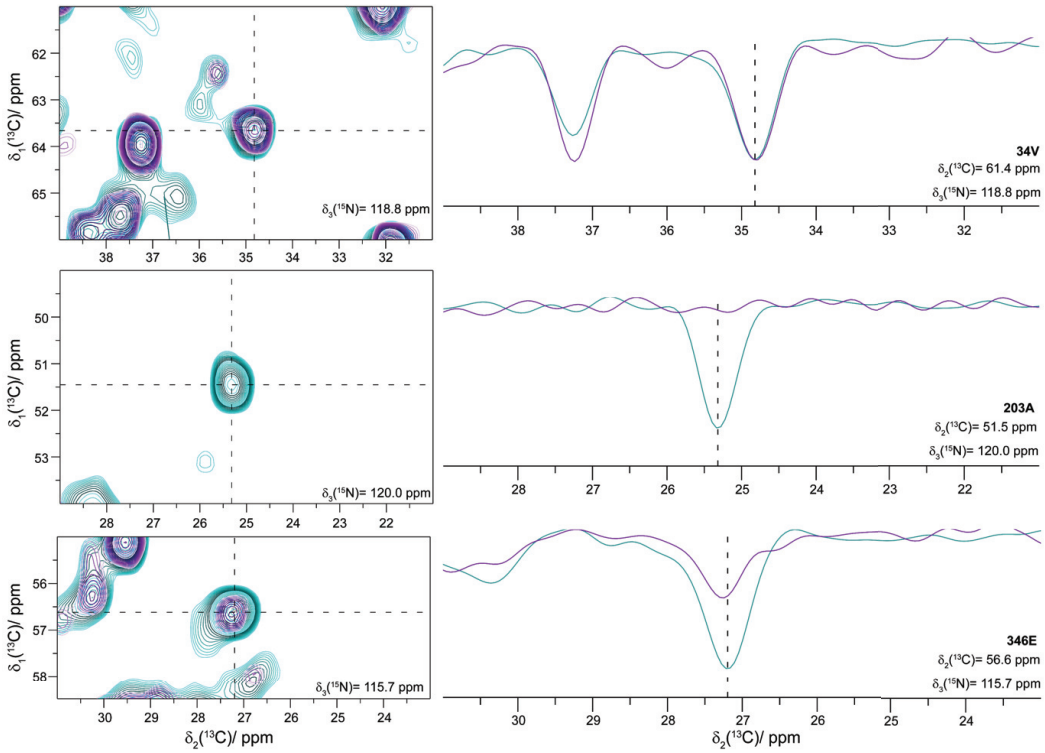


Figure S8: Site-specific determination of PRE effects from 3D spectra: Examples of 2D planes from 3D NCACB spectra of the *Hp*DnaB:APMPNP:Mg^{2+} (18-fold excess Mg^{2+}, shown in cyan contours) and of the *Hp*DnaB:APMPNP:Mn^{2+} (5-fold excess Mn^{2+}, shown in purple contours) complexes. Representative 1D traces along F2 are shown on the right.

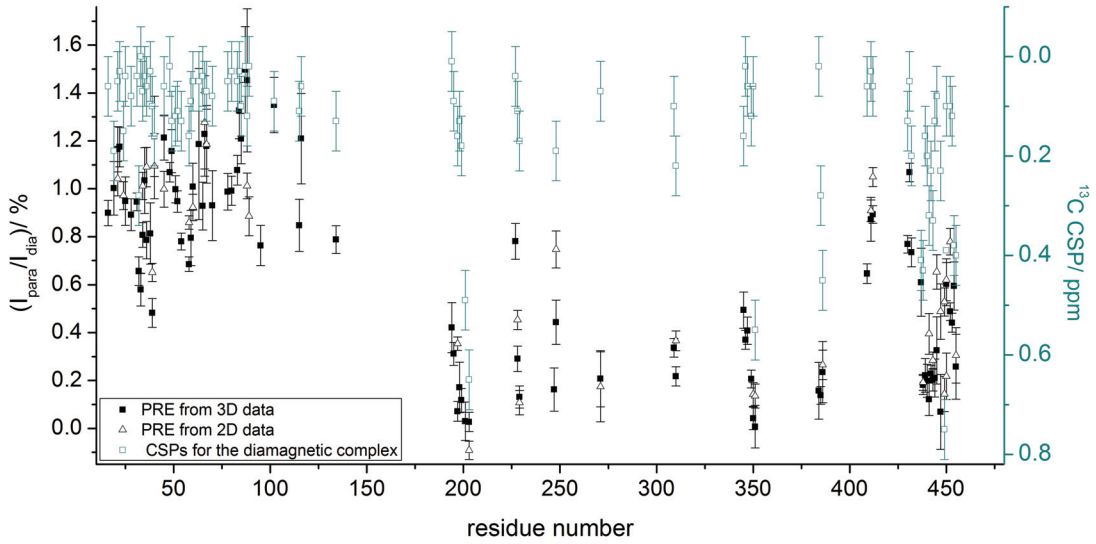


Figure S9: CSPs correlate with PRE effects: PRE effects (determined from 2D and 3D spectra shown as black squares and black triangles, respectively) are compared with ^{13}C CSPs for the diamagnetic *Hp*DnaB:APMPNP:Mg^{2+} complex (taken from reference [10]). The comparison allows the following conclusions: (i) Residues showing detectable CSPs show also sizeable PRE effects and (ii) PRE effects serve as direct probes for helicase-nucleotide interactions, whereas ^{13}C CSPs rely on adaptions in the backbone geometry.

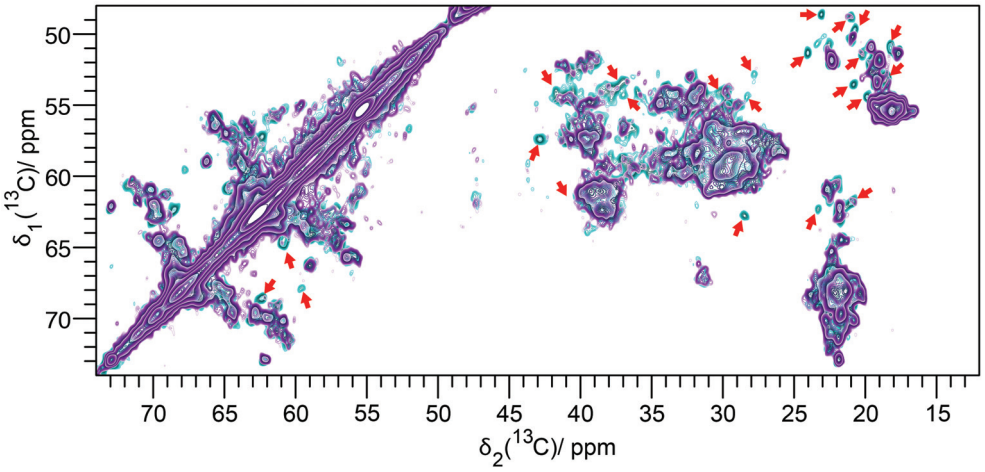


Figure S10: Aliphatic region from 2D ^{13}C-^{13}C DARR spectra of the reverse labeled ^{12}C-^{14}N-[LVIKHP]-^{13}C-^{15}N sample of BmrA:ADP:Mg^{2+}:VO_4^{3-} (shown in cyan contours) and of the BmrA:ADP:Mn^{2+}:VO_4^{3-} (shown in purple contours) complexes. Strongly attenuated or disappearing peaks are marked by red arrows.

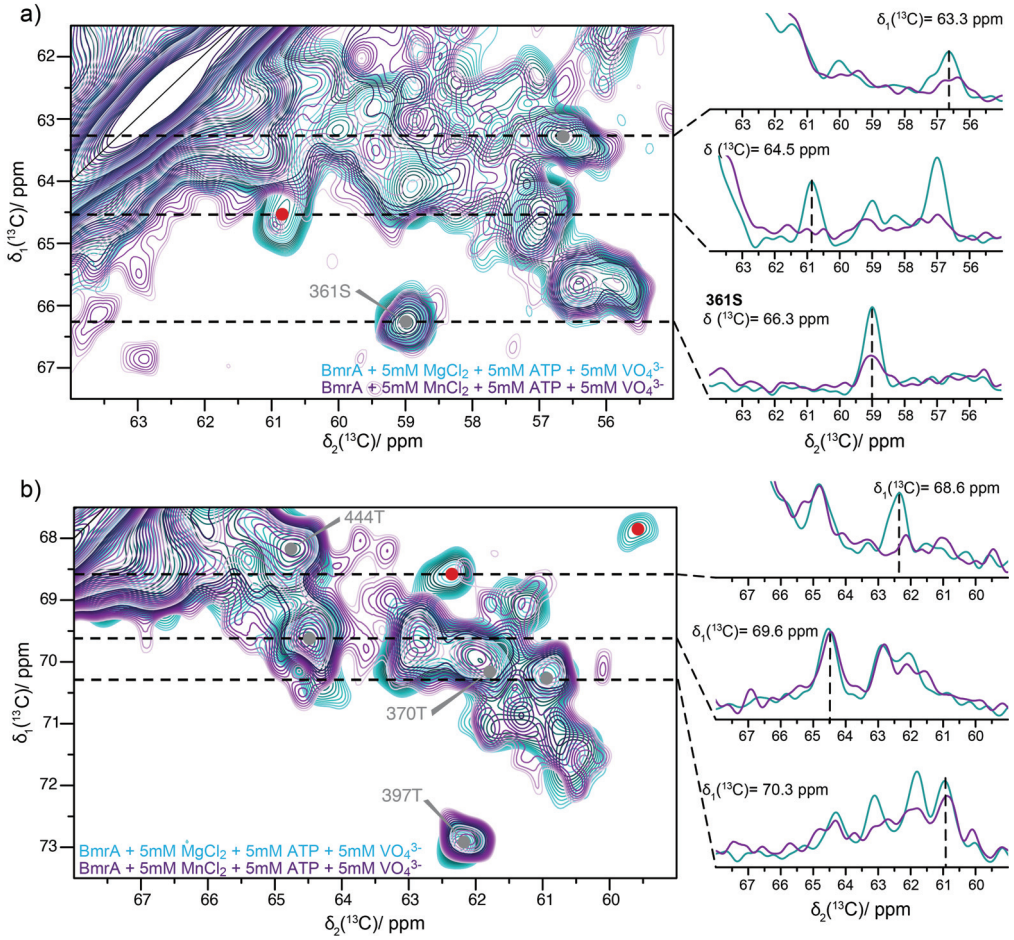


Figure S11: PRE effects influence 2D ^{13}C-^{13}C DARR spectra: Extract of threonine (**a**) and serine (**b**) region from 2D ^{13}C-^{13}C DARR spectra of the BmrA:ADP:Mg^{2+}:VO_4^{3-} (shown in cyan contours) and of the BmrA:ADP:Mn^{2+}:VO_4^{3-} (shown in purple contours) complexes. Representative 1D traces along F2 are shown as insets for disappearing (black), and unaffected or attenuated resonances (grey).

Chapter VIII

Flexible-to-rigid transition is central for transport but not for hydrolysis in the ABC transporter BmrA

The work presented in this chapter was done in collaboration with Beat H. Meier and Jean-Michel Jault. This chapter is in preparation and follows this reference:

Flexible-to-rigid transition is central for transport but not for hydrolysis in the ABC transporter BmrA

Denis Lacabanne[#], Cedric Orelle[#], Thomas Wiegand, Lauriane Lecoq, Britta Kunert, Claire Chuilon, Stéphanie Ravaud, Jean-Michel Jault, Beat H. Meier, and Anja Böckmann

[#] equal contribution

Contents

Abstract

ATP-binding-cassette (ABC) transporters are molecular pumps that work in import or export of substances across the cell membrane. They are driven by an ATP-fueled engine, the nucleotide-binding domain, that is connected to a transmembrane domain switching between inward-facing (IF) and outward-facing (OF) states for uptake and release of substances. Detailed understanding of the working mechanism of the pump remains a challenge to structural biology, as these proteins are difficult to study at molecular level in their active, membrane inserted form. We here used solid-state NMR to identify, in a first step, conformational differences between the inward-facing (IF) and outward-facing (OF) states in the ABC transporter BmrA reconstituted in its native *B. Subtilis* lipids. The very noticeable apparition of new NMR signals concomitant with large amplitude chemical shift perturbations reveal important changes in flexibility and conformation of the protein on transition to the OF state trapped by addition of the ATP:Mg^{2+}:VO_4^{3-}. Similar changes are observed for a hydrolysis-incompetent mutant protein on addition of ATP, indicating that hydrolysis is not required for the transition to the OF state. Subsequent analysis of an X-loop mutant, in which transport is abolished but which remains ATPase active, reveals an incomplete set of chemical-shift perturbations on ATP:Mg^{2+}:VO_4^{3-} addition, mainly restricted to the nucleotide binding domains, and no rigidification of the protein. This suggests that this change in protein rigidity is central for efficient transmission of the power stroke to the transmembrane domains.

Introduction

ATP binding cassette (ABC) transporters can translocate a variety of molecules across the cell membrane by driving a drug/lipid efflux-pump with an ATP fuelled engine. They are found in all forms of life and are involved in a number of drug resistance phenotypes including anti-cancer drugs and antibiotics (Davidson *et al.* 2008). Our studies focus on the drug exporter BmrA from *B. subtilis* (130 kDa) as a model system for, and homologue of the human P-glycoprotein (Steinfels *et al.* 2002) involved in multidrug resistance e.g. in the context of cancer therapy. It was shown that BmrA is able to transporte a large panel of drugs including the cytostatica daunorubicin and doxorubicin (Steinfels *et al.* 2004) and equally larger substrates of 84-100 nm (Browning *et al.* 2016). During the export cycle, ABC

transporters pass through open and closed states, referred to as inward-facing (IF) and outward-facing (OF) conformations, respectively (Seeger & van Veen 2009) (Locher 2016). In the IF state, the drug can enter the cavity located at the inner membrane leaflet. When ATP is bound, the transporter is believed to switch to the OF state, and the drug is expelled to the extracellular space (Seeger & van Veen 2009). Moreover, occluded conformations, representing possible intermediate states in the cycle, were described for the efflux pumps McjD (Choudhury *et al.* 2014) and PCAT1 (Lin *et al.* 2015). However, it is important to note that they were obtained without transported substrate. The transition from the IF to the OF state represents a key event in the transport cycle in the "pump-cycle", as it is closely related to the power stroke needed for export of the substrate. Which stage exactly in the ATP hydrolysis cycle induces this transition is still under debate, and has been described as coupled to either an ATP binding event (Timachi *et al.* 2017) or ATP hydrolysis (Verhalen *et al.* 2017) in the "motor cycle" (Doshi *et al.* 2013). The ATP hydrolysis step, where chemical energy is consumed, occurs only later in the motor cycle, and creates an energy-rich state preparing the engine for the next power stroke. The details of ATP hydrolysis itself are still under debate, with two alternative scenarios put forward: in the ATP-switch model (Higgins & Linton 2004), the transport reaction is initiated by ATP binding which brings the two nucleotide binding domains (NBDs) in contact; in the constant contact model (George & Jones 2012), the NBDs remain in contact during the cycle and ATP is hydrolyzed alternatively in each site (Jones & George 2011). Transmission of the conformational changes from the NBDs to the transmembrane domains (TMDs) has been shown to be mediated by the X-loop and the Q-loop which operate as cross-link between the NBDs and the intracellular loops 1 and 2 of the TMDs (Jones & George 2002) (Dawson & Locher 2006) (Oancea *et al.* 2009). Nucleotide analogues can be used to trap the transporter in certain conformational states assumed to represent those the transporter passes through during the export cycle. Different analogues are frequently used to mimic the prehydrolytic, posthydrolytic and transition states involved in transport, to investigate conformations and the dynamics of the protein throughout the transitions. Mutant forms can be used as alternatives to mimic for example the prehydrolytic state by using a mutant protein which is unable to hydrolyse ATP, or alternatively be induced addition of a non-hydrolysable ATP analog.

We here use solid-state NMR to analyze BmrA in presence of native lipids, to closely mimic its natural environment in the bacterial membrane. We use nucleotide analogues, as well as mutant forms of BmrA impaired in hydrolysis or transport, to highlight not only conformational differences between the IF and OF states, but also dynamical differences. Furthermore, a hydrolysis-impaired mutant confirms that transition to the OF state does not directly depend on hydrolysis and that the power stroke is indeed triggered by ATP binding. We establish the importance of a flexible-to-rigid transformation in the protein during the pump cycle which is absent in a transport-impaired but ATPase-active mutant.

Material and Methods

1. Preparation of BmrA mutants

Site-directed mutagenesis of BmrA was performed using a kit provided by Integrated DNA Technologies®. The design of the primers was done following the instruction of the provider. BmrA-K380A and BmrA-E504A were prepared as described in the literature (Orelle *et al.* 2003) (Orelle *et al.* 2008).

2. NMR sample preparation

The production, purification and reconstitution of BmrA were described previously in the literature (Kunert *et al.* 2014) (Lacabanne *et al.* 2017). Selectively labeled BmrA was produced in M9-medium supplemented with natural abundance amino acids (see below) in presence of ^{13}C-glucose (2 g/L) and ^{15}N-ammonium chloride (2 g/L) as sole sources of carbon-13 and nitrogen-15. Pre-cultures, i) 50 mL at 37 °C then ii) 150 mL at 25 °C inoculated with i) were realized before the main culture. 850 mL M9-medium containing natural abundant Leu (0.23g/L), Val (0.23g/L), Ile (0.23g/L), Lys (0.40g/L), His (0.40g/L), Pro (0.10 g/L) were inoculated with the 150 mL pre-culture at 25 °C. The expression of BmrA was induced with 0.7 mM IPTG at an OD_{600nm} between 0.6-0.7. The cells were harvested 6 hours after induction by centrifugation at 6,000 x *g*.

The cells were suspended and disrupted by high pressure with a Microfluidics Microfluidizer® in 50 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1 mM DTT, benzonase and EDTA-free protease inhibitor cocktail. The solution was centrifuged at 15,000 x *g* during 45 min, the pellet was discarded and the supernatant was centrifuged at 200,000 x g during one hour. The sedimented membranes were suspended and collected in 50 mM Tris-HCl pH 8.0, 1 mM EDTA, and 300 mM sucrose.

The protein-containing membranes were diluted at 2 mg/mL, solubilized using 1% DDM (m/v) and centrifuged at 100,000 x *g* for one hour. The supernatant was loaded onto a Ni-NTA column (Qiagen) equilibrated with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 15% glycerol, 10 mM imidazole and 0.2% DDM (m/v). The Ni-NTA column was washed using equilibration buffer with successively 0.5 M NaCl, 30 mM imidazole, 40 mM imidazole and 250 mM imidazole for the elution. The imidazole was removed using desalting columns (PD10 - GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl pH 8.0, 100 mM

NaCl, 10% glycerol and 0.2% DDM (m/v). The protein was diluted 4 times with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol. The solution was mixed with a homemade preparation of *B. subtilis* lipids (Kunert *et al.* 2014) (Lacabanne *et al.* 2017) solubilized in X% Triton X-100 and incubated for one hour. The quantity of lipid mix was adjusted at a lipid-to-protein ratio (LPR, in m:m) of 0.5 or 20. The DDM and Triton X-100 were depleted by dialysis with Bio-beads (BioRad) in the dialysis solution during 9 days (Lacabanne *et al.* 2017). The quantity of Bio-beads used was three times the theoretical amount required to absorb the DDM and Triton X-100 present in the solution.

For the preparation of the BmrA:ADP:Mg^{2+}:VO_4^{3-} or BmrA:ADP:Mn^{2+}:VO_4^{3-} complexes, the protein solutions, after 9 days of dialysis, were incubated with 1 mM Na_3VO_4, then 10 mM ATP:Mg^{2+} or 1 mM ATP:Mn^{2+} during 1 hour. For the BmrA-E504A:ATP:Mg^{2+} complex, the solution was incubated with 10 mM ATP:Mg^{2+} during 1 hour. The protein in lipids was sedimented into the MAS-NMR rotor (1 h at 4°C with 120 000 x *g*) using home-build tools (Böckmann *et al.* 2009).

3. Drug transport activity of BmrA in presence of Mn^{2+} or Mg^{2+}

Drug transport was followed using fluorescence spectroscopy (Steinfels *et al.* 2004). In a 1 mL cuvette, 100 µg of inverted membrane vesicles from *E. coli* C41(DE3) overexpressing BmrA were incubated in presence of 2 mM ATP, 4 mM phosphoenolpyruvate and 60 $\mu g.mL^{-1}$ pyruvate kinase. After 1 minute incubation at 25 °C, 10 µM doxorubicin were added and the transport reaction was initiated by adding 2 mM $MgCl_2$. The excitation and emission wavelengths used were 480 and 590 nm respectively.

4. ATPase activity assay in micro-plate

ATPase activity was monitored using an ATP/NADH coupled assay in micro-plates (Lacabanne *et al.* 2017). 0.1 µg of reconstituted BmrA (LPR 20) was mixed in 50 mM HEPES-KOH pH 8, 4 mM phosphoenolpyruvate, 60 µg/mL pyruvate kinase, 32 µg/mL lactate dehydrogenase, 10 mM $MgCl_2$, 0.6 mM NADH. The reaction was started by the addition of 5 mM ATP and the decreasing absorbance of NADH was followed at 340 nm during the reaction (spectrophotometer SAFAS FLX-Xenius®) during 20 min.

Results

1. The transporter rigidifies upon transition to the OF state

We have recorded spectra of BmrA in the inward-facing (IF) and outward-facing (OF) states, the latter induced by addition of vanadate (VO_4^{3-}) in presence of ATP and magnesium. Extracts of the alanine, serine and threonine regions of the 2D ^{13}C-^{13}C DARR correlation spectra of the two samples are shown in **Figure 52** in red for BmrA apo, and in cyan for BmrA:ADP:Mg^{2+}:VO_4^{3-} (for full aliphatic region see **Figure S1**). The overlay of both spectra reveals many peaks shifting to different resonance frequencies, which we here refer to as chemical-shift perturbations (CSPs). Several such CSPs of 0.1-2 ppm can be identified in the spectra, and pairs of peaks likely corresponding to a same residue, but with different chemical shifts in OF and IF states, are highlighted by red and cyan crosses in **Figures 52A and 52B** respectively. CSPs cannot account for all differences in the spectra, since one can notice, also in **Figure S1** where the entire aliphatic region is represented, that the spectra of the OF state show a large number of additional cross peaks, as indicated by cyan circles in the spectrum of the IF state at the frequencies where signals are observed exclusively in the OF state. Additional signals in protein spectra observed under different physico-chemical sample conditions are indicative of changes in the dynamics of the protein. Indeed, in solid-state NMR measurements, residues showing dynamics with a sizeable amplitude on a time scale of roughly 1 µs are strongly broadened and their amplitude attenuated due to interference of the protein dynamics with the heteronuclear decoupling and/or magic-angle spinning frequencies. Signals from such mobile segments thus disappear in the noise. The number of residues which show resonances only in the OF form is approximately 10 from isolated peak of the Cα-Cβ region. It can be noted that several of these resonances are located in regions indicative for β-sheets or turns. One could speculate that the new resonances originate from the doubling of existing peaks caused by an asymmetric occupation of the ATP-binding sites. We can however exclude this, since we could infer from EPR measurements on the Mn^{2+} substituted protein that both sites are occupied in the OF state.

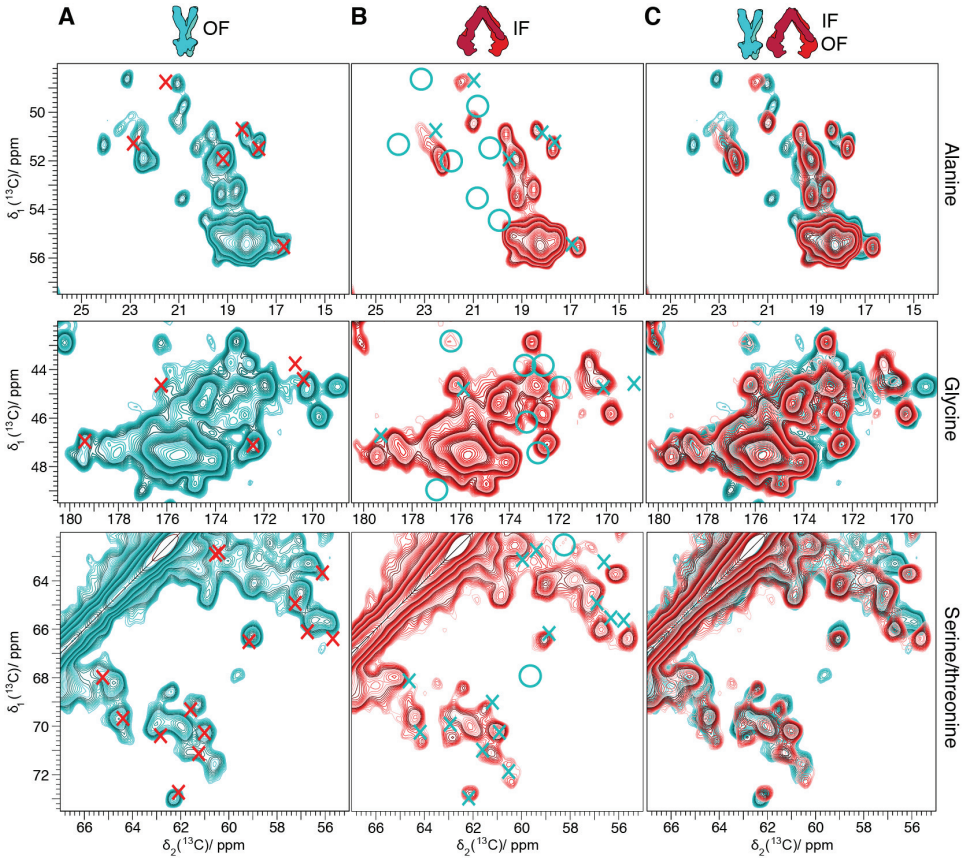


Figure 52. Many Resonances shift or disappear when going from the OF to the IF state. A: Extracts from 2D DARR spectra recorded on the OF state (cyan), with Ala resonances shown in the first row, glycine resonances in the second, and Ser/Thr in the third. Signal positions of peaks observed in the IF state and showing significant CSPs in the OF state are indicated by red crosses **B**: corresponding spectra of the IF state (red), with signal positions from the OF state indicated by cyan crosses. Cyan circles indicate peaks which are observed in the OF state only, devoid of an obvious counterpart in the IF state. **C**: Overlay of the spectra recorded on the OF and on the IF states.

2. <u>The residues only rigid in the OF state are located in the NBDs</u>

In order to identify in which region of the protein the residues rigid in the OF state are located, we substituted the Mg^{2+} ions in BmrA:ADP:Mg^{2+}:VO_4^{3-} by Mn^{2+} ions (Wiegand *et al.* 2017). Mn^{2+} induces strong paramagnetic relaxation enhancements (PRE) that wipe out all resonances from spins located within a radius of about 15 Å around the metal ion. Using this effect, residues in the vicinity of the Mn^{2+} metal ion can be identified. **Figure 53** shows extracts of spectra recorded on the different states. The signals which are not observed in the spectra of the OF state BmrA:ADP:Mn^{2+}:VO_4^{3-} but are present in the diamagnetic Mg^{2+} analogue are highlighted, in column 2, by cyan circles. In column C, the two Mn^{2+} and Mg^{2+} OF spectra are overlaid for illustration. Interestingly, when overlaying spectra recorded on

the OF BmrA:ADP:Mn^{2+}:VO_4^{3-} with those recorded on the IF state (**Figure 53** column D), one can see that almost the same residues are invisible in both spectra. This clearly indicates that the signals which are observed exclusively in the diamagnetic (Mg^{2+}) OF state are located in vicinity to the Mn^{2+} binding site, in the NBDs. One could argue that the protein did not bind the paramagnetic ion, and remained in the IF state. The observation by EPR of two Mn^{2+} ions per dimer in this preparation provides however evidence that the ions are bound, and that the protein is indeed in the OF state (Wiegand *et al.* 2017). Further evidence from a detailed comparison of spectral fingerprints is presented in **Figure S2**. We can thus conclude that the residues that are are dynamic (i.e. which are not observed) in the IF state are close (within 15 Å) to the metal binding site, and therefore must be part of the NBDs.

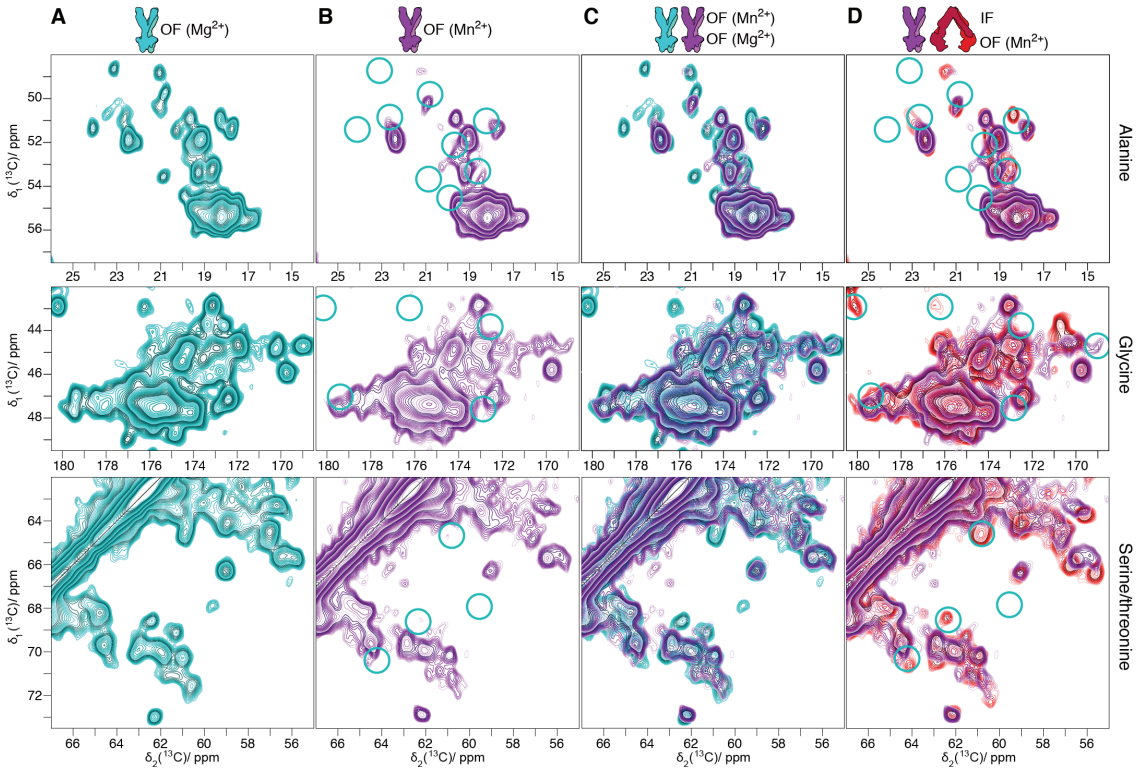


Figure 53. Parts of the NBD become dynamic in the IF form A: Extracts of 2D DARR spectra of the protein in the OF state (cyan), with Ala resonances shown in the first row, glycine resonances in the second, and Ser/Thr in the third. **B**: corresponding spectra of paramagnetic BmrA:ATP:Mn^{2+} in OF state in purple. Resonances close (within 15 Å) to the ATP binding site are strongly attenuated by PRE in the OF form, as shown by cyan circles. **C**: Overlay of the spectra recorded on the Mg^{2+}- and Mn^{2+}-bound OF states in cyan and purple respectively . **D**: overlay of the Mn^{2+}-bound OF state (purple) and the IF state (red, shown also in **Figure 52B**). The cyan circles represent peaks which are not observed in the Mn^{2+} OF state. It is clear from D that they are not present in the Mg^{2+} IF form neither.

3. ATP Hydrolysis is not required to induce the OF state in BmrA

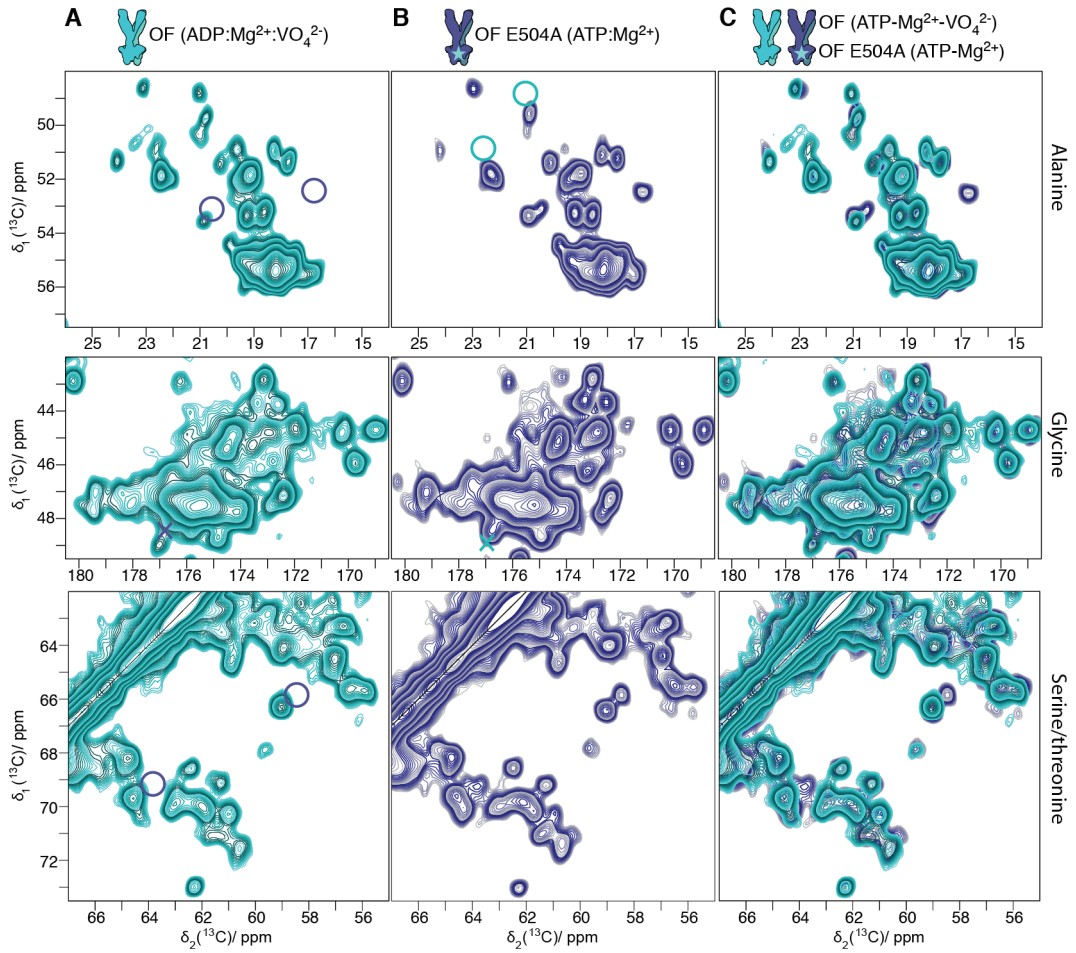


Figure 54. Mimics of the prehydrolytic and transition states display similar conformations A: Extracts of 2D DARR spectra of BmrA:ADP:Mg^{2+}:VO_4^{3-} (transition-OF state mimic) (cyan, panel A), with Ala resonances shown in the first row, glycine resonances in the second, and Ser/Thr in the third, compared to BmrA-E504A:ATP:Mg^{2+} (prehydrolytic-OF state mimic) (dark blue, panel B). One peak showing CSP is shown with crosses on both spectra, and peaks observed only in one form with circles. **C**: Overlay of the two spectra.

BmrA E504A is a mutant form which binds but does not hydrolyze ATP (Orelle *et al.* 2003), but it is still able to switch to the OF state (Orelle *et al.* 2008). While BmrA:ADP:Mg^{2+}:VO_4^{3-} is often considered as an ATP transition-state mimic, BmrA-E504A:ATP:Mg^{2+} can be considered as a snapshot of the prehydrolytic ATP state. The comparison between the spectra of the IF states of WT and E504A BmrA (**Figure S3**) reveals, as expected, an additional signal in the Ala region corresponding to the mutation, and as well a second "new" Ala residue, which likely arises from a CSP, as two Ala resonances disappear. Also,

in some other regions, several peaks slightly shift or appear, but to a much lesser extent then when comparing the IF and OF WT forms (**Figure S4**).

The comparison between the corresponding OF states of the E504A mutant and the WT protein (**Figure 54**) reveals that the spectra look highly similar. Few additional CSP (in the range of 0.2 ppm) are observed, including that due to the mutation. These peaks could be due to decreased dynamics observed for the mutant. Indeed, it is interesting to note that the Ala resonance at 23/51 ppm, which is the only residue identified showing a broad, diffuse signal in all WT preparations analyzed so far, is not observed as such in the E504A mutant spectra. The high similarity between both spectra when compared to the differences observed between OF and IF states, indicates that i) BmrA:E504A:ATP:Mg^{2+} is clearly in the OF state, confirming that no ATP hydrolysis is needed to induce the OF state in BmrA; ii) the conformation of the OF transition state mimic BmrA:ADP:Mg^{2+}:VO_4^{3-} seems to be very close to the E504A mutant bound to ATP which thus mimics an prehydrolytic state.

4. The mutant E474R is ATPase active but transport inactive

The structure of Sav1866, a homolog of BmrA revealed the presence of a so-called X-loop in the NBD that could be involved in conformational coupling with the TMD. Mutations in this loop in Tap1/Tap2 supported the essential role of this part of the NBD (Oancea *et al.* 2009). We have established the E474R mutant for assignment purposes, and on testing its activity realized that this mutant does not transport drugs. As a consequence, we assessed the doxorubicin and Hoechst 33342 transport activity for several X-loop mutants: E474D, E474Q, E474F and E474A and E474R in inverted *E. coli* membrane vesicles. We compared the transport activity of each mutant to BmrA WT, and also to a hydrolysis-deficient (and thus transport impaired) mutant (K380A). The doxorubicin transport profiles as measured by fluorescence spectroscopy are displayed in **Figure 55A**, and the initial rates are shown in **Figure 55B**. Mutants K380A and E474R feature clearly no transport activity. The mutants E474A, E474F, E474Q and E474D display 25, 34, 61 and 75 % respectively of the WT protein transport activity. The different mutants were then extracted from the *E. coli* membrane, purified and reconstituted in *E. coli* lipids. The corresponding ATPase activities are shown in **Figure 55C**, where one can see that the difference between the ATPase activity of WT and mutant proteins is rather small (**Figure 55C)**, with the E474R mutant still showing 70 % activity. The control K380A mutant shows as expected an abolished ATPase activity. The

observations that the X loop E474R mutant is able to hydrolyze ATP, but fails to transport doxorubicin, point to an impaired communication between the engine (NBDs) and the pump (TMDs) in this mutant. We thus explored the structural basis of this lack of communication.

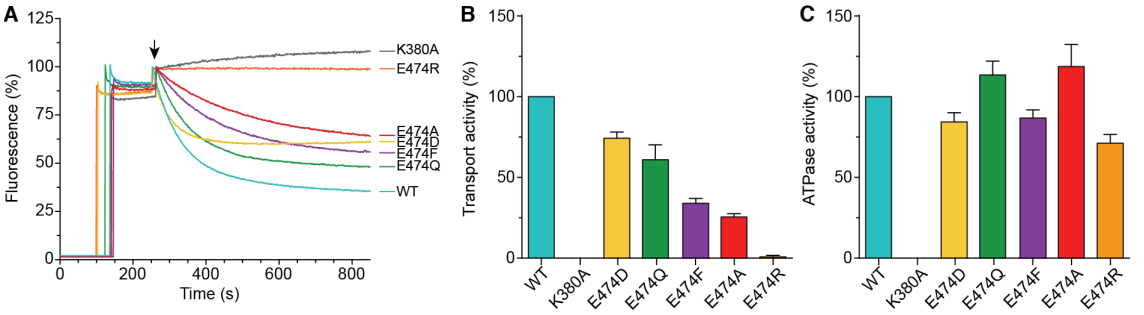


Figure 55. Drug-transport and ATPase-activity assay of BmrA WT and mutant forms (**A**) ATP-dependent transport of doxorubicin measured in inverted *E. coli* membrane vesicles. After addition of 10 µM doxorubicin (120 s), 2 mM of Mg^{2+} were added (indicated by black arrow). (**B**) Transport activities derived from initial rate fluorescence decays measured for BmrA WT and mutant forms, normalized to the rate measured in the WT form. (**C**) ATPase activity of BmrA WT and mutant forms reconstituted in *E.coli* lipids with an LPR of 20.

5. <u>The transport-disabled E474R mutant makes an incomplete switch to the OF state</u>

We recorded spectra of the IF and OF forms of E474R. A comparison of the spectrum of the BmrA-E474R IF state with that of the IF state of the WT protein (**Figure S5)** shows that both spectra are highly similar. In **Figure 50** we compare the OF state of the WT protein and the BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-} mutant state (**Panels 56 A, B**). Notably, one can see that the residues which newly appear in the spectra of the WT OF state (cf. **Figure 52**) are not observed in BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-}, indicating that no rigidification is induced by substrate binding in this mutant, in contrast to the WT protein (cyan circles in **Figure 56A**). **Figure 56C** retraces the spectra shown in **Figure 51C**, to point out the CSPs observed between the WT IF and OF states. We then compare these spectra in **Figure 56D** to the mutant BmrA-E474R: ADP:Mg^{2+}:VO_4^{3-}. It becomes immediately visible that most CSPs, highlighted by the cyan contours observable in **Figure 56C**, are covered up by the resonances from the orange mutant form in **Figure 56D**, as they coincide with the resonances from the cyan spectrum, i.e. the WT OF state. This makes clear that BmrA-E474R: ADP:Mg^{2+}:VO_4^{3-} at least partially shifts to a state comparable to the WT OF state, even if its flexible residues do not rigidify.

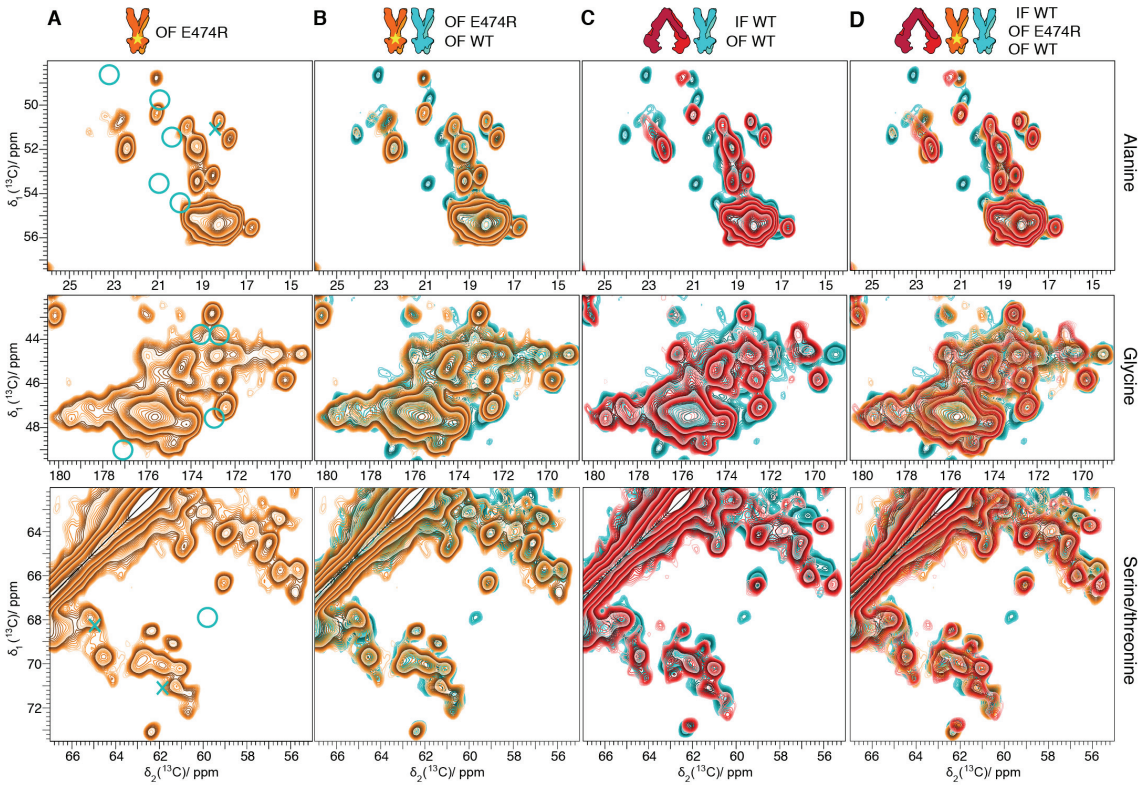


Figure 56. NMR spectra comparison between the inward-facing and outward-facing of BmrA WT with the outward-facing form of BmrA E474R. **A:** extracts of 2D DARR spectra of BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-} in the OF state (orange), with Ala resonances shown in the first row, glycine resonances in the second, and Thr/Ser in the third. Significant CSPs or with the BmrA:ADP:Mg^{2+}:VO_4^{3-} in OF state are shown with cyan crosses. Cyan circles represent peaks which are appearing in BmrA:ADP:Mg^{2+}:VO_4^{3-} OF spectrum. **B** and **C** Overlay of, respectively, the spectra recorded on the BmrA OF state with BmrA-E474R OF state and the spectra recorded on BmrA IF state with BmrA OF state. **D** Overlay of the spectra recorded on BmrA IF with BmrA-E474R OF and BmrA OF state.

In order to analyze whether the TMDs go through the conformational transition leading to an OF state, we searched for signals from residues far from the ATP binding site, as indicated by their conservation in the Mn^{2+}-substituted OF state spectra. We considered signals which fulfilled all following points: i) a CSP must be observed between WT IF and OF; ii) the corresponding signal of the OF state must not be attenuated in the paramagnetic OF state; and iii) the CSP must not be observed in BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-}, i.e. the shift must be similar to the IF state. There are only few such isolated peaks, and as expected, one finds them mainly in the rather α-helical regions where peaks overlapped. Four of them are shown in **Figure 57**, and thus correspond to residues which remain in an IF conformation. The DARR spectrum of the E474R mutant thus presents a peak distribution

that is an intermediate between the IF and the OF states. From this follows that the rigidification is not the only missing feature in the E474R mutant which would characterize an OF state in that the conformational transition also remains incomplete.

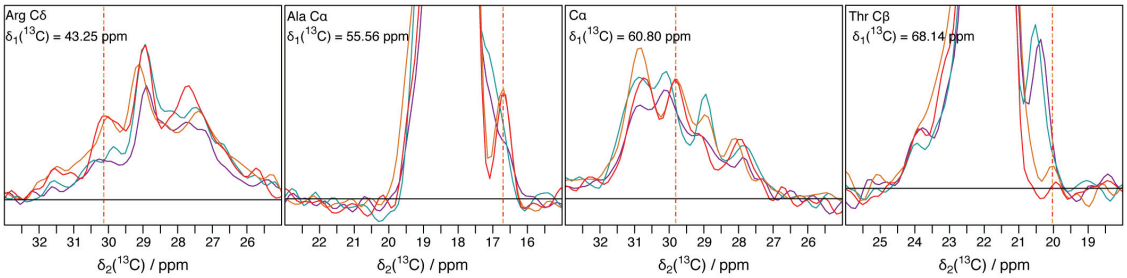


Figure 57. The E474R mutant makes an incomplete transition to the OF state. Spectra from the 474 mutant (orange) show signals coinciding with the IF state (red) and different from the OF state (cyan). The corresponding signals do not dissappear in the Mn spectra, indicating that the related residues are remote from the ATP:Mn binding site. 1D slices from 2D DARR spectra α-helical regions of BmrA:ADP:Mg^{2+}:VO_4^{3-} (cyan), BmrA:ADP:Mn^{2+}:VO_4^{3-} (purple), BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-} (orange) and BmrA Apo (red). The indirect dimension is specify for each traces, resonances similar between BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-} (orange) and BmrA Apo (red) are highlighted with red-orange dashes.

6. <u>^{31}P NMR of ADP/ATP highlights two asymmetric binding sites in the dimer</u>

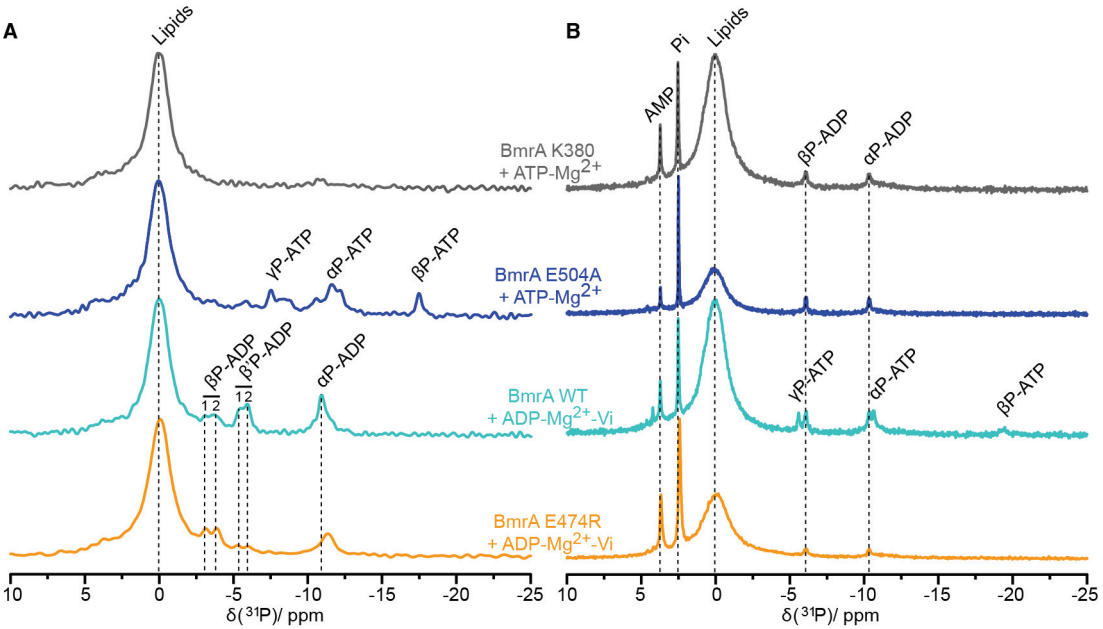


Figure 58. ^{31}P spectra of nucleotides in the different BmrA preparations. A ^{31}P CPMAS spectra and **B**: direct-pulse MAS spectrum on BmrA (cyan) and the mutants K380A (grey), E504A (blue) and E474R (orange). Each peak is labelled with the corresponding phosphorus for ATP or ADP. The peak at 0 ppm comes from the lipid membrane.

In order to characterize the ATP binding properties of BmrA, ^{31}P NMR cross-polarization MAS (**Figure 58A**) and direct-pulse spectra (**Figure 58B**) were recorded for BmrA:ADP:Mg^{2+}:VO_4^{3-} of the WT protein and the transport-disabled E474R mutant, and of BmrA:ATP:Mg^{2+} of the hydrolysis-deficient K380A and E504A mutants. ^{31}P CPMAS experiments are selective for relatively rigid ^{31}P moities of ADP or ATP, meaning that they are bound to the protein, whereas ^{31}P MAS shows dynamic phosphorus from free ADP or ATP (relaxation filter). The spectra shown in **Figure 58** reveal that actually no ATP at all was bound to the mutant K380A, as no signal is observed in the corresponding CPMAS spectrum shown in grey. This means that the K380A mutant, rather than being hydrolysis impaired, is actually unable to even bind ATP with an affinity that would allow observing it in the NMR spectra. Still, in the MAS spectrum in **Figure 58B**, one can see that ATP has been hydrolyzed to ADP and AMP despite this; since BmrA is unable to do so, this clearly points to the presence of contaminants able to hydrolyze ATP and ADP. The ^{31}P CPMAS spectrum of the mutant E504A, in contrast, display signals of ATP bound to the protein, as expected and also observed by the conformational switch revealed in the ^{13}C spectra (**Figure 54**). The E504A spectrum actually reveals more than one peak for the γ and α phosphates, which indicates that the ATP molecules do not show a well defined structure when bound, and are found in slightly different environments. It should be mentioned that the ^{31}P chemical shifts are highly sensitive to conformational heterogeneities. And indeed, no indications for such an asymmetry has been observed in the 13C spectra. Still, binding of ATP is mediated via side-chain resonances, with CSPs which might be small in 13C, as chemical shifts are mainly sensitive to dihedral angles, and less to binding. It is thus possible that the minor asymmetries revealed in 31P spectra remain undetected in 13C spectra. In the ^{31}P CPMAS of BmrA WT and the E474R mutant one can observe a clearer pattern of multiple shifts for the βP-ATP, which seems split into four peaks (in cyan in **Figure 58A**), two βP-ADP and β'P-ADP, with a chemical shift difference of 2.3 ppm, and βP- and β'P-ADP which are further splitting into a pair of resonances referred to as β_1P-, β_2P-, β'_1P- and β'_2P-ADPwith a chemical shift difference around 0.5 ppm. BmrA E474R yields essentially the same pattern, but with somewhat different intensities. It is unclear at the moment what these four different signals correspond to, and further experiments using different ATP analogues shall yield more insight.

The ^{31}P NMR MAS experiments indicate hydrolysis of ATP and ADP by the observation of AMP signals for each sample. For K380A, E504A and E474R, no trace of free ATP was detected, only free AMP, ADP and Pi. For the WT BmrA sample, traces of ATP were detected in addition. The spectra indicate that ATP is efficiently hydrolyzed in every sample, despite the non-catalytic mutations K380A/E504A or the presence of vanadate.

AMP has also been detected in NMR spectra of the ABC transporter MsbA, and was imputed to an adenylate kinase activity of the protein (Kaur *et al.* 2016). For BmrA, the observation of AMP in the spectra of the K380A and E504A mutants excludes an adenylate kinase activity (i.e. enzymatic conversion of ADP to AMP). Indeed, these mutants are not able to hydrolyse ATP or even to bind ATP/ADP in the case of mutant K380A. These results suggest that AMP does not come from BmrA and is instead a clear indicator for the presence of contaminating enzymes able to hydrolyze ATP and ADP in the preparation.

Discussion

ABC transporters go through different steps in the transport cycle, which can be split up into different conformations of the transporter as a function of different occupancy of the ATP-binding site (and possibly also catalyzed by the presence of a drug). If this cycling is difficult to observe directly, different ATP analogs can be used to mimic certain steps in the cycle which is driven by ATP binding, its hydrolysis and release. Alternatively, mutant forms can be used to arrest the protein in a certain state. If it remains unclear how close these mimics actually represent the reaction pathway followed during an actual transport cycle, they constitute a convenient means to further explore the conformational and dynamic space sampled by the transporter during the cycle. Also, the investigation of bacterial transporters as models is advantageous in that large quantities can be produced as recombinant proteins for structural studies. We here explored different conformational states of the *B. subtilis* protein BmrA using Apo, ATP:Mg^{2+} and ADP:Mg^{2+}:VO_4^{3-}bound states, as well as different mutant forms, including the non-catalytic K380A and E504A mutants. We produced a previously unknown mutant form, E474R, which we show to be hydrolysis competent, but transport disabled. Other intermediate states do exist, and are worth exploring, but certain remain challenging to trap in vitro, as for example the post-hydrolytic state.

Addition of ATP:Mg^{2+}:VO_4^{3-} is a widely used strategy to mimic a state with a nucleotide trapped in its pre-hydrolysis form in ABC transporters, and also in other NBD-driven molecular motors (Ward *et al.* 2007) (Kaur *et al.* 2016). The idea is that on addition of ATP:Mg^{2+}:VO_4^{3-}, ATP is hydrolyzed, and the place left by the leaving γ-Phosphate (Pi) is occupied by the vanadate. Another explanation would be that ATP is hydrolyzed in solution to ADP, and that ADP:Mg^{2+}:VO_4^{3-} is the binding species. Yet, addition of ADP instead of ATP does not lead to the same result in BmrA (Orelle *et al.* 2008) which indicates that ATP binds with higher affinity, and also that the protein remains after hydrolysis in the same conformation for at least the laps of time necessary to accommodate the incoming VO_4^{3-} and to trap the state. It has been shown by EPR recently that different ATP analogues can induce very different states of the protein (Timachi *et al.* 2017). In BmrA, addition of ATP:Mg^{2+}:VO_4^{3-} induces the transition to the OF state, as has already been highlighted by

the fact that superstructures, only observed under the microscope when the protein is in the IF state, dissolve on addition of ATP:Mg^{2+} (Orelle *et al.* 2008).

NMR analyses of proteins the size of BmrA remain challenging, and the work described here was enabled by the use of high-field spectrometers, combined with amino-acid selective unlabelling approaches which reduce the number of resonances in the spectra in order to obtain a spectral fingerprint displaying more isolated resonances. The NMR spectra recorded on BmrA show a clearly distinct chemical shift fingerprint for the IF and OF states, and an important difference lies in the appearance of additional peaks in the OF state spectra, pointing towards a rigidification of the protein induced by ADP:Mg^{2+}:VO_4^{3-}binding and transition to the OF state. While one could argue that the arising peaks are due to heterogeneities, no concomitant decrease of other peaks could be identified, clearly stating that the peaks are new, and not coming from the splitting of one peak. The absence of peak doubling also indicated that, within the detection limit of ^{13}C chemical shifts, the conformations of the two monomers are identical. This absence of heterogeneity is evidenced in both IF and OF states. This feature also tends to support the ATP switch model in which only symmetrical states are involved (Higgins & Linton 2004) (Linton & Higgins 2007), and excludes binding of a single ATP or ADP molecule for BmrA as postulated in the constant contact model which NBDs would be asymmetric: either closed or open (George & Jones 2012).

We could locate the residues showing this decrease in dynamics to within 15 Å of the Mn^{2+} binding sites using PREs. This locates the corresponding residues to large parts, but not exclusively, of the NBD, where a homology model of the protein highlighting the vanished regions is shown. And indeed, the ^{13}C DARR spectrum of the ADP:Mn^{2+}:VO_4^{3-} OF state is not devoid of signals showing β-strand typical chemical shifts, which indicates that not the entire NBD is flexible in the IF state. EPR on MsbA (Zou *et al.* 2009) (Mittal *et al.* 2012), P-gp (Wen *et al.* 2013), LmrA (Hellmich *et al.* 2012), BmrCD (Mishra *et al.* 2014) and TM287/288 (Timachi *et al.* 2017) highlighted the flexible behavior of these proteins in absence of substrate. Indeed, a wide distance-distribution between the two NBDs has been observed in these studies, which suggests large amplitude motions in absence of any substrate. In the case of BmrA, the protein has been shown by H/D exchange coupled with mass spectrometry to be quite dynamic in the IF state as opposed to the OF state (Mehmood *et al.* 2012). Moreover, the flexibility of the NBDs induces poorly-resolved densities in the NBDs

as observed for CFTR by cryo-electron microscopy (Zhang & Chen 2016). Dynamics probed using deuterium solid-state NMR on LmrA displays motional flexibility of the transporter in the apo state, whereas motional flexibility is restricted in the LmrA:APMPNP:Mg^{2+} state (Siarheyeva *et al.* 2007). The time scales of motion revealed in these NMR experiments are however likely faster than the movement which leads to extinction of the signals in the NMR spectra observed for BmrA. Indeed, deuterium experiments determine a motion of the NBSs with a correlation time on the order of microseconds (Siarheyeva *et al.* 2007), while disappearing signals in solid-state NMR spectra point to correlation times roughly in the millisecond range. In addition to dynamic changes, conformational differences can be observed between the two spectra, both near the ATP binding site and also remote changes induced by the IF to OF transition.

The spectra of BmrA-E504A:ATP:Mg^{2+} (pre-hydrolytic mimic) are very similar to those recorded on BmrA:ADP:Mg^{2+}:VO_4^{3-} (transition-state mimic), both in the IF and OF states of the proteins when compared to the WT spectra, including conformational and dynamic features reflected in the spectra. Some different chemical shifts are observed, which are likely mainly due to the mutation, as for example the additional Ala resonance, which shows a chemical shift corresponding to a rather peculiar coil conformation in the OF spectra. Thus, the state before and after hydrolysis (but before release of the nucleotide) seem to look, within the technical limits (i.e. spectral overlap, insensitivity of ^{13}C chemical shifts to surface effects) very similar. And indeed, it was shown for BmrA that the E504A mutant operates the same conformational switch in the presence of ATP:Mg^{2+} as the WT protein on addition of ATP:Mg^{2+}:VO_4^{3-}: the ring-shaped super structures formed during lipid reconstitution are destabilized in both cases, indicating a transition to the OF state (Orelle *et al.* 2008). Our observations confirm this behavior at the atomic level, and implies that the prehydrolytic and transition states display rather similar conformations. This would mean that the power stroke has already proceeded during the ATP binding event, and that the transition to the high-energy form did not take place yet, i.e. happens only on release of ADP:Pi or ADP:VO_4^{3-}.

The use of the hydrolysis-enabled but transport-inactive E474R mutant allows to a certain extent to separate the conformational differences induced by hydrolysis and transport, and the spectra let assume that most of the larger CSPs correspond to the ATP-binding event, as they occur as well in the mutant protein. Still, bias is introduced in that chemical shifts of

the NBDs are more dispersed due to their β-sheet components, and also by the amino-acid-selective unlabelling of the transporter excluding many residues from the TMDs. One should notice that ^{13}C chemical shifts are more sensitive to conformational changes induced by ligand binding events, and less to the interaction itself, and that the large shifts of about 2 ppm observed for example in the Gly region are certainly due to structural rearrangements of parts of the proteins. The different E474 substitutions underline the importance in drug transport of the X-loop in the communication between the NBDs and the TMDs. All BmrA E474 mutants show a decrease in drug transport, and introduction of a positively charged residue completely abolishes it. A similar mutation was introduced in the heterodimeric TAP ABC transporter, and was shown to abolish transport in this protein as well (Oancea *et al.* 2009). The role of the X-loop thus seems to be central whether the protein is homo- or heterodimeric, and in BmrA, interaction between the X-loop and the Intra-Cellular Loop 1 (ICL1) mediated by the conserved glutamate 474 seems to be important for the transmission of the conformational change from the NBDs to the TMDs. Loosening the link between the TMDs and the NBDs still allows ATPase activity to be fulfilled by the NBDs, which seem to work independently from the rest, as ATP hydrolysis is reduced at most by 25%. The NMR spectrum of BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-} shows a fingerprint which is rather close to the one observed for the WT OF state spectrum. Still, several residues which are remote from the Mg^{2+}- binding site, possibly in or near the TMDs, display an IF state fingerprint. Importantly, the residues which rigidify in the OF state of BmrA-WT:ADP:Mg^{2+}:VO_4^{3-} are not rigid in BmrA-E474R:ADP:Mg^{2+}:VO_4^{3-}. The fact that ATPase activity is only slightly affected in the E474R mutant suggests that the residues which remain dynamic are not directly involved in the ATP hydrolysis, as for ATP hydrolysis, backbone and side chain atoms take highly specific positions necessary to act efficiently in catalyzing hydrolysis.

Figure 53 summarizes our findings. We observed that in the inward-facing state several residues located in the NBDs are flexible, (**Figure 59A**) in contrast to the outward-facing state (**Figure 59B**) in which they become rigid, likely to be able to transmit action to the TMDs. In the case of the X-loop mutant E474R, in which drug transport is disabled, these residues remain dynamic even in presence of ATP:Mg^{2+}:VO_4^{3-}. Furthermore, several residues remote from the ATP:Mg^{2+} binding site remain as in the IF state, suggesting that only an incomplete conformational transition takes place in the mutant, which impedes proper transport (**Figure 59C**).

This is remnant of a car, in which the engine turns the input shaft, but no gear is engaged which allows the transmission to the output shaft driving the wheels, or here the pump. In the WT protein, the gear is engaged by rigidification of parts of the protein which show ms motion in the IF state, allowing for transport.

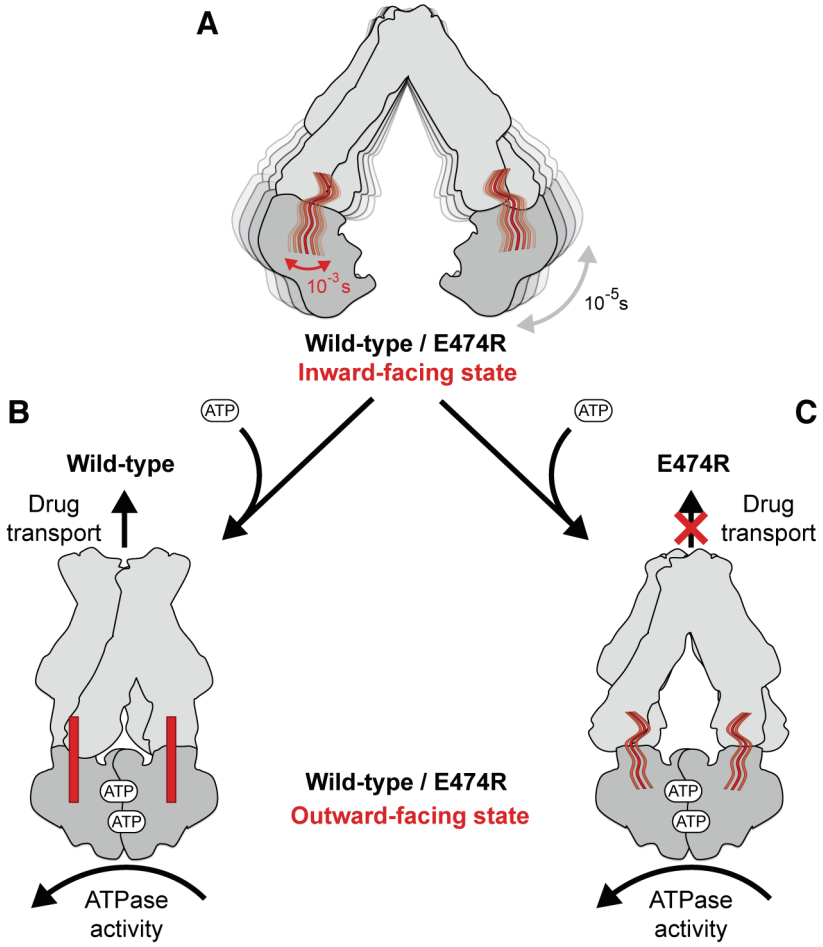


Figure 59. Summary of our finding (highlighted in red). A. BmrA in its IF state with a subset of residues showing dynamics on the ms time scale, which are mainly located within 15 Å of the ATP:Mg^{2+} binding site (signals appearing in the OF state coincide with those that show PREs). **B.** BmrA in its OF state, in which the residues rigidify, and conformational changes occur around the ATP-binding site, but also remotely in an allosteric manner (CSPs are observed as well for residues not showing PREs in Mn2+ substituted preparations). **C.** BmrA-E474R mutant with ATP:Mg^{2+}:VO_4^{3-} which makes an incomplete conformational and dynamic transition to the OF state, with flexible residues remaining flexible, and only partially observed CSPs.

Conclusion

We characterized by solid-state NMR the inward-facing and the outward-facing state of the ABC multidrug transporter BmrA. We show that in BmrA, ATP binding is sufficient to induce the OF state, i.e. no hydrolysis is needed to obtain the OF spectral fingerprint. Using selective labelling strategies and paramagnetic Mn^{2+} substitution, we show that several residues near the ATP:Mg^{2+} binding site rigidify upon NBD closure. The absence of peak doubling in the spectra highlighted that both monomers are structurally symmetrical in the OF state. We established the X-loop BmrA-E474R mutant which is ATPase active but which lacks transport activity. We showed that this mutant only makes a partial transition to the OF state, both with respect to conformation and dynamics, providing a structural rational for the failure of transport. This provides structural insight into the mechanistic necessary successful transport in this important class of proteins.

References

Böckmann, A., Gardiennet, C., Verel, R., Hunkeler, A., Loquet, A., Pintacuda, G., Emsley, L., Meier, B.H. & Lesage, A. (2009) Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45, 319–327.

Browning, L.M., Lee, K.J., Cherukuri, P.K., Nallathamby, P.D., Warren, S., Jault, J.-M. & Xu, X.-H.N. (2016) Single Nanoparticle Plasmonic Spectroscopy for Study of the Efflux Function of Multidrug ABC Membrane Transporters of Single Live Cells. *RSC advances* 6, 36794–36802.

Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H.W. & Beis, K. (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 111, 9145–9150.

Davidson, A.L., Dassa, E., Orelle, C. & Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and molecular biology reviews : MMBR* 72, 317–64– table of contents.

Dawson, R.J.P. & Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185.

Doshi, R., Ali, A., Shi, W., Freeman, E.V., Fagg, L.A. & van Veen, H.W. (2013) Molecular disruption of the power stroke in the ATP-binding cassette transport protein MsbA. *The Journal of biological chemistry* 288, 6801–6813.

George, A.M. & Jones, P.M. (2012) Perspectives on the structure-function of ABC transporters: the Switch and Constant Contact models. *Progress in biophysics and molecular biology* 109, 95–107.

Hellmich, U.A., Lyubenova, S., Kaltenborn, E., Doshi, R., van Veen, H.W., Prisner, T.F. & Glaubitz, C. (2012) Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *Journal of the American Chemical Society* 134, 5857–5862.

Higgins, C.F. & Linton, K.J. (2004) The ATP switch model for ABC transporters. *Nature Structural & Molecular Biology* 11, 918–926.

Jones, P.M. & George, A.M. (2002) Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. *Proceedings of the National Academy of Sciences* 99, 12639–12644.

Jones, P.M. & George, A.M. (2011) Molecular-dynamics simulations of the ATP/apo state of a multidrug ATP-binding cassette transporter provide a structural and mechanistic basis for the asymmetric occluded state. *Biophysical Journal* 100, 3025–3034.

Kaur, H., Lakatos-Karoly, A., Vogel, R., Nöll, A., Tampé, R. & Glaubitz, C. (2016) Coupled ATPase-adenylate kinase activity in ABC transporters. *Nature Communications* 7, 13864.

Kunert, B., Gardiennet, C., Lacabanne, D., Calles-Garcia, D., Falson, P., Jault, J.-M., Meier, B.H., Penin, F. & Böckmann, A. (2014) Efficient and stable reconstitution of the ABC transporter BmrA for solid-state NMR studies. *Frontiers in molecular biosciences* 1, 5.

Lacabanne, D., Lends, A., Meier, B.H., Böckmann, A., Orelle, C., Jault, J.-M. & Falson, P. (2017) Journal of Biomolecular NMR *Gradient reconstitution of membrane proteins for solid-state NMR studies*. 1–20pp.

Lin, D.Y.-W., Huang, S. & Chen, J. (2015) Crystal structures of a polypeptide processing and secretion transporter. *Nature* 523, 425–430.

Linton, K.J. & Higgins, C.F. (2007) Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Archiv : European journal of physiology* 453, 555–567.

Locher, K.P. (2016) Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural & Molecular Biology* 23, 487–493.

Mehmood, S., Domene, C., Forest, E. & Jault, J.-M. (2012) Dynamics of a bacterial multidrug ABC transporter in the inward- and outward-facing conformations. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 109, 10832–10836.

Mishra, S., Verhalen, B., Stein, R.A., Wen, P.-C., Tajkhorshid, E. & McHaourab, H.S. (2014) Conformational dynamics of the nucleotide binding domains and the power stroke of a heterodimeric ABC transporter. *eLife* 3, e02740.

Mittal, A., Böhm, S., Grütter, M.G., Bordignon, E. & Seeger, M.A. (2012) Asymmetry in the homodimeric ABC transporter MsbA recognized by a DARPin. *The Journal of biological chemistry* 287, 20395–20406.

Oancea, G., O'Mara, M.L., Bennett, W.F.D., Tieleman, D.P., Abele, R. & Tampé, R. (2009) Structural arrangement of the transmission interface in the antigen ABC transport complex TAP. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 106, 5551–5556.

Orelle, C., Dalmas, O., Gros, P., Di Pietro, A. & Jault, J.-M. (2003) The conserved glutamate residue adjacent to the Walker-B motif is the catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA. *Journal of Biological Chemistry* 278, 47002–47008.

Orelle, C., Gubellini, F., Durand, A., Marco, S., Lévy, D., Gros, P., Di Pietro, A. & Jault, J.-M. (2008) Conformational change induced by ATP binding in the multidrug ATP-binding cassette transporter BmrA. *Biochemistry* 47, 2404–2412.

Seeger, M.A. & van Veen, H.W. (2009) Molecular basis of multidrug transport by ABC transporters. *Biochimica et biophysica acta* 1794, 725–737.

Siarheyeva, A., Lopez, J.J., Lehner, I., Hellmich, U.A., van Veen, H.W. & Glaubitz, C. (2007) Probing the molecular dynamics of the ABC multidrug transporter LmrA by deuterium solid-state nuclear magnetic resonance. *Biochemistry* 46, 3075–3083.

Steinfels, E., Orelle, C., Dalmas, O., Penin, F., Miroux, B., Di Pietro, A. & Jault, J.-M. (2002) Highly efficient over-production in E. coli of YvcC, a multidrug-like ATP-binding cassette transporter from Bacillus subtilis. *Biochimica et biophysica acta* 1565, 1–5.

Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A. & Jault, J.-M. (2004) Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in Bacillus subtilis†. *Biochemistry* 43, 7491–7502.

Timachi, M.H., Hutter, C.A., Hohl, M., Assafa, T., Böhm, S., Mittal, A., Seeger, M.A. & Bordignon, E. (2017) Exploring conformational equilibria of a heterodimeric ABC transporter. *eLife* 6, e20236.

Verhalen, B., Dastvan, R., Thangapandian, S., Peskova, Y., Koteiche, H.A., Nakamoto, R.K., Tajkhorshid, E. & McHaourab, H.S. (2017) Energy transduction and alternating access of the mammalian ABC transporter P-glycoprotein. *Nature* 543, 738–741.

Ward, A., Reyes, C.L., Yu, J., Roth, C.B. & Chang, G. (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 104, 19005–19010.

Wen, P.-C., Verhalen, B., Wilkens, S., McHaourab, H.S. & Tajkhorshid, E. (2013) On the origin of large flexibility of P-glycoprotein in the inward-facing state. *The Journal of biological chemistry* 288, 19211–19220.

Wiegand, T., Lacabanne, D. & Keller, K. (2017) Solid-state NMR and EPR Spectroscopy of Mn2+-Substituted ATP-Fueled Protein Engines. *Angewandte Chemie …*.

Zhang, Z. & Chen, J. (2016) Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* 167, 1586–1597.e9.

Zou, P., Bortolus, M. & McHaourab, H.S. (2009) Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. *Journal of molecular biology* 393, 586–597.

Supplementary material

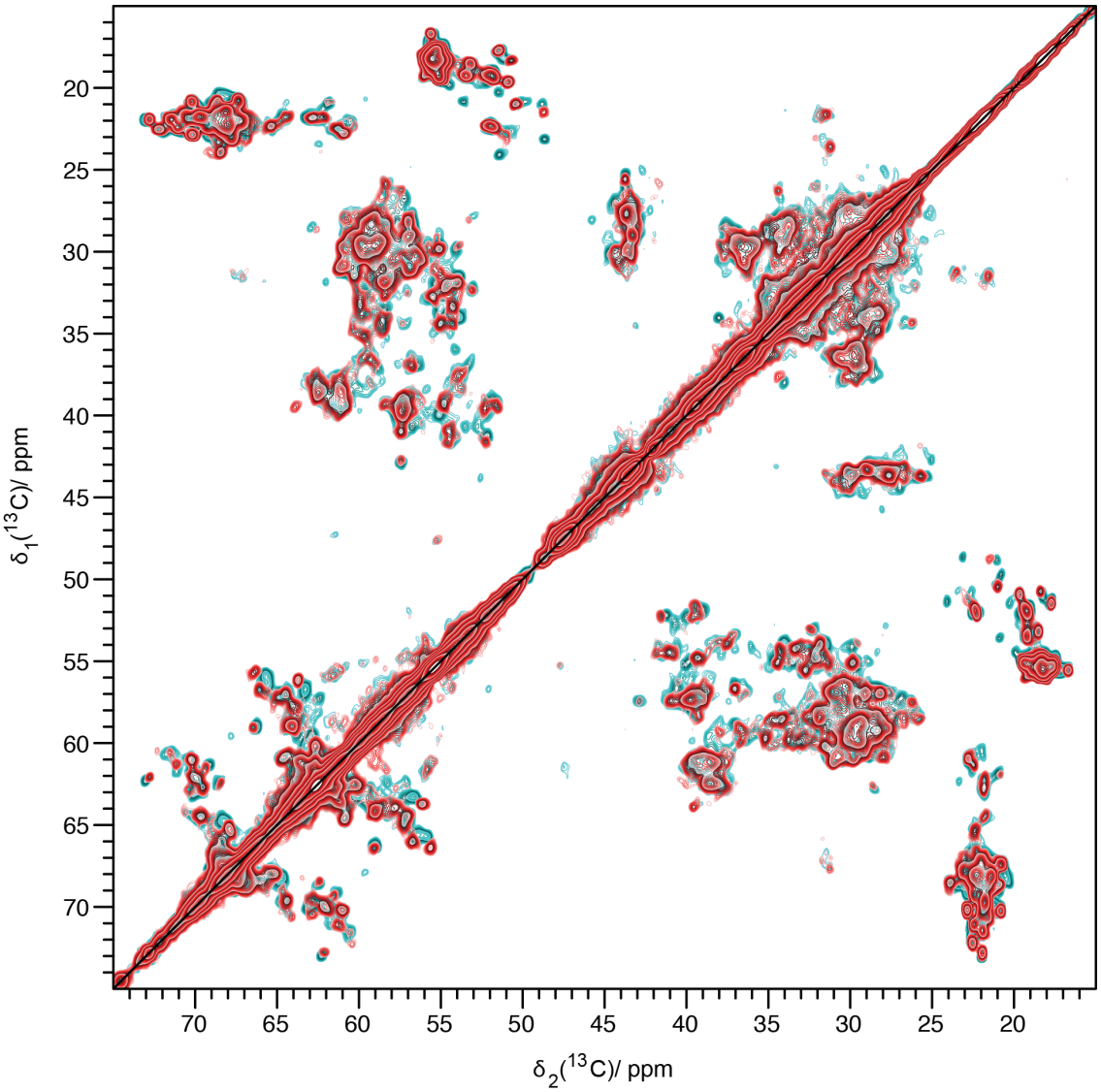


Figure S1. 2D DARR of BmrA and BmrA in presence of ADP:Mg^{2+}:VO_4^{3-} Aliphatic Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectra of lipid-reconstituted [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA (red shades) overlaid with [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-ADP:Mg^{2+}:VO_4^{3-} (cyan shades).

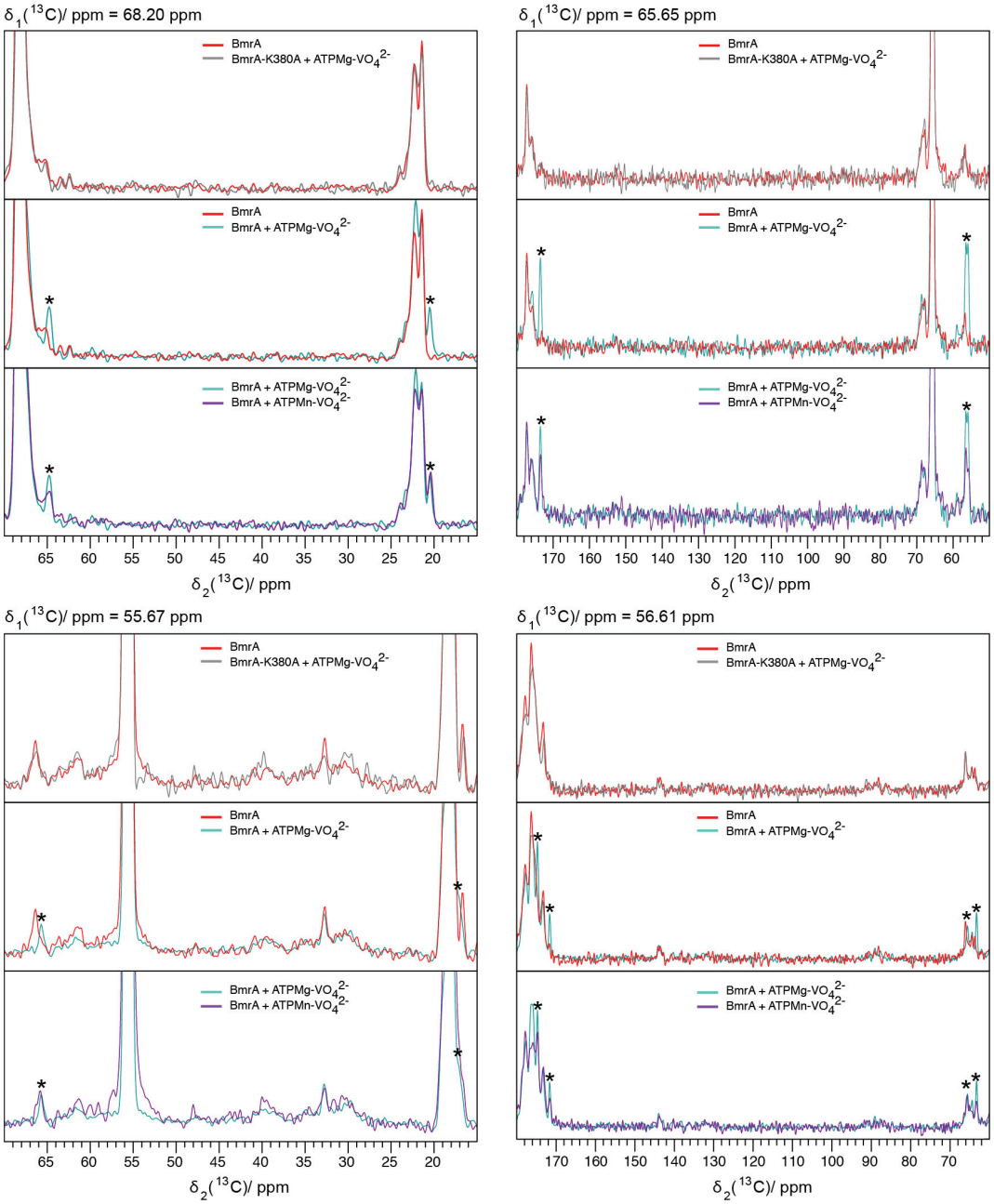


Figure S2. Traces overview of BmrA Traces from a 2D ^{13}C-^{13}C 20 ms DARR spectra of BmrA-WT (red), of the non-catalytic mutant K380A which remains in IF state in presence of ATP-Mg (Orelle *et al.* 2008). (gray), of BmrA-WT in presence of ATP-Mg-VO_4^{3-} (cyan) and ATP-Mn-VO_4^{3-} (purple).

For each traces, in the first panel BmrA-WT traces are overlaid with BmrA-K380A in presence of ATP-Mg-VO_4^{3-} to characterize the IF state. In the second panel BmrA-WT traces are overlaid with BmrA-WT in presence of ATP-Mg-VO_4^{3-} to highlight the difference due to the OF state switch, significant CSP are labelled with a black star. In the third panel BmrA-WT traces in presence of ATP-Mn-VO_4^{3-} are overlaid with BmrA-WT in presence of ATP-Mg-VO_4^{3-}, the similarity between the CSP indicates that BmrA is in the same conformation in presence of ATP-Mn-VO_4^{3-}.

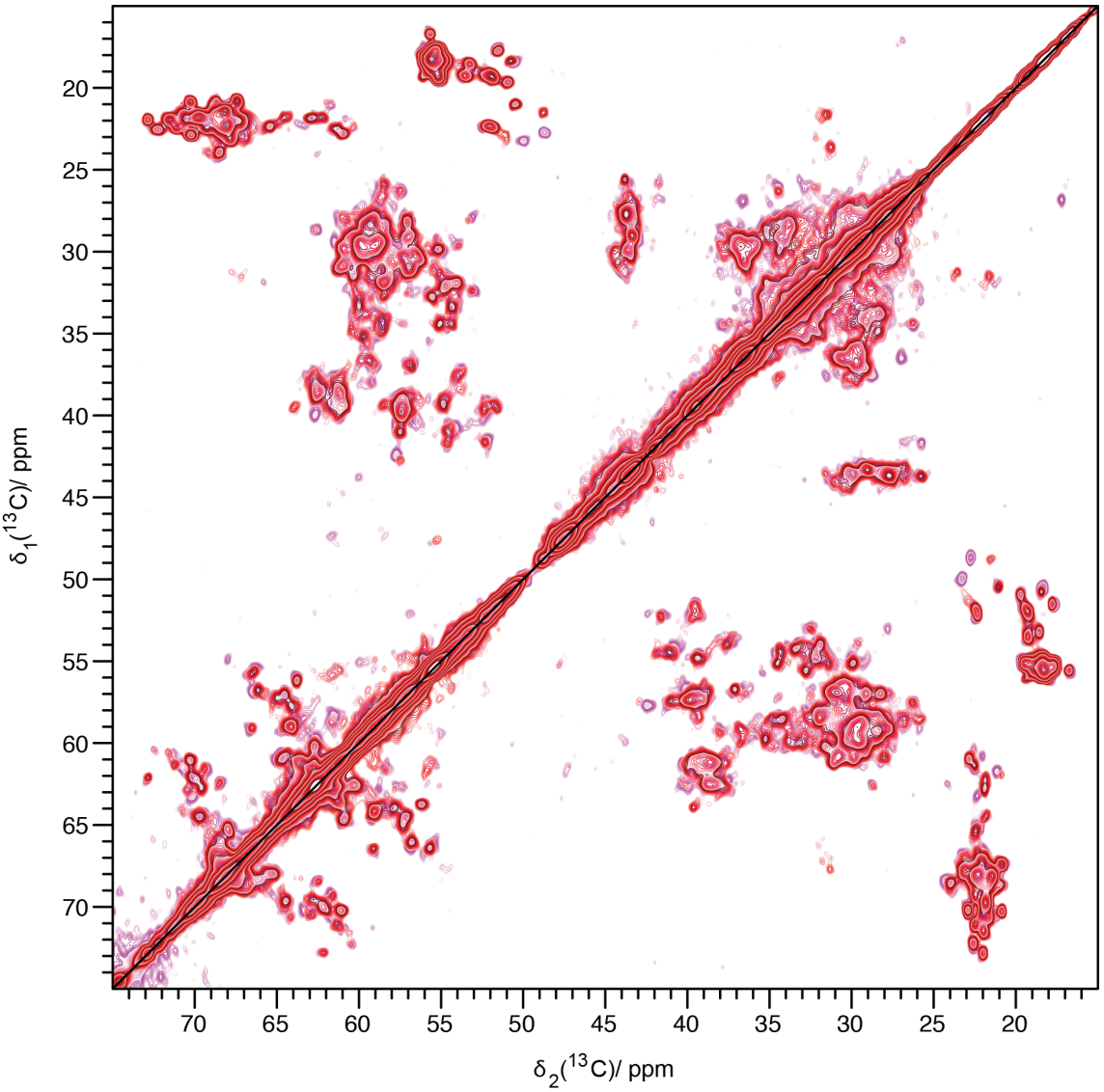


Figure S3. 2D DARR of BmrA and BmrA-E504A Aliphatic Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectra of lipid-reconstituted [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA (red shades) overlaid with [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-E504A (pink shades).

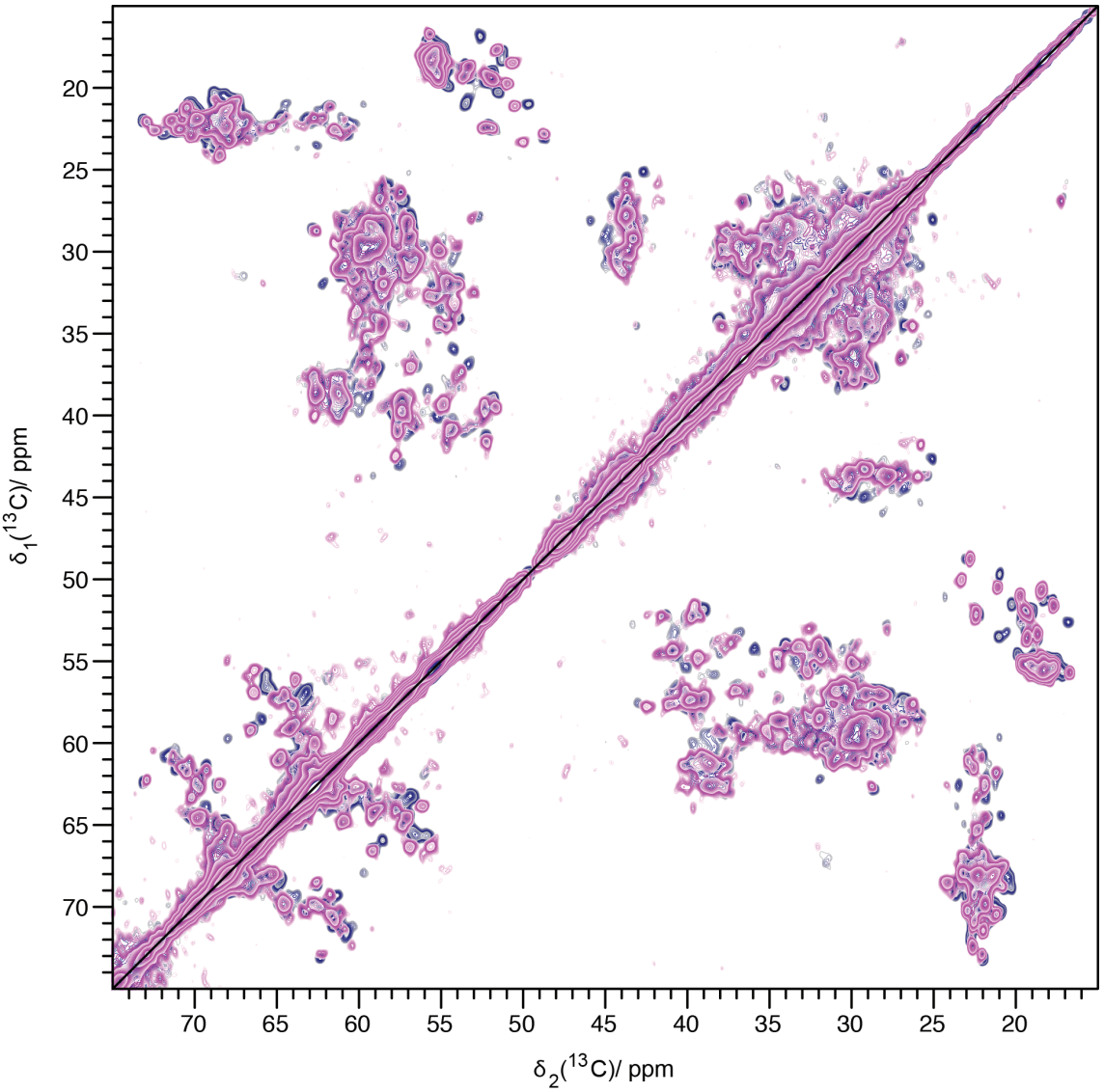


Figure S4. 2D DARR of BmrA-E504A and BmrA-E504A in presence of ADP:Mg^{2+}:VO_4^{3-} Aliphatic Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectra of lipid-reconstituted [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-E504A (pink shades) overlaid with [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-E504A-ADP:Mg^{2+}:VO_4^{3-} (blue-gray shades).

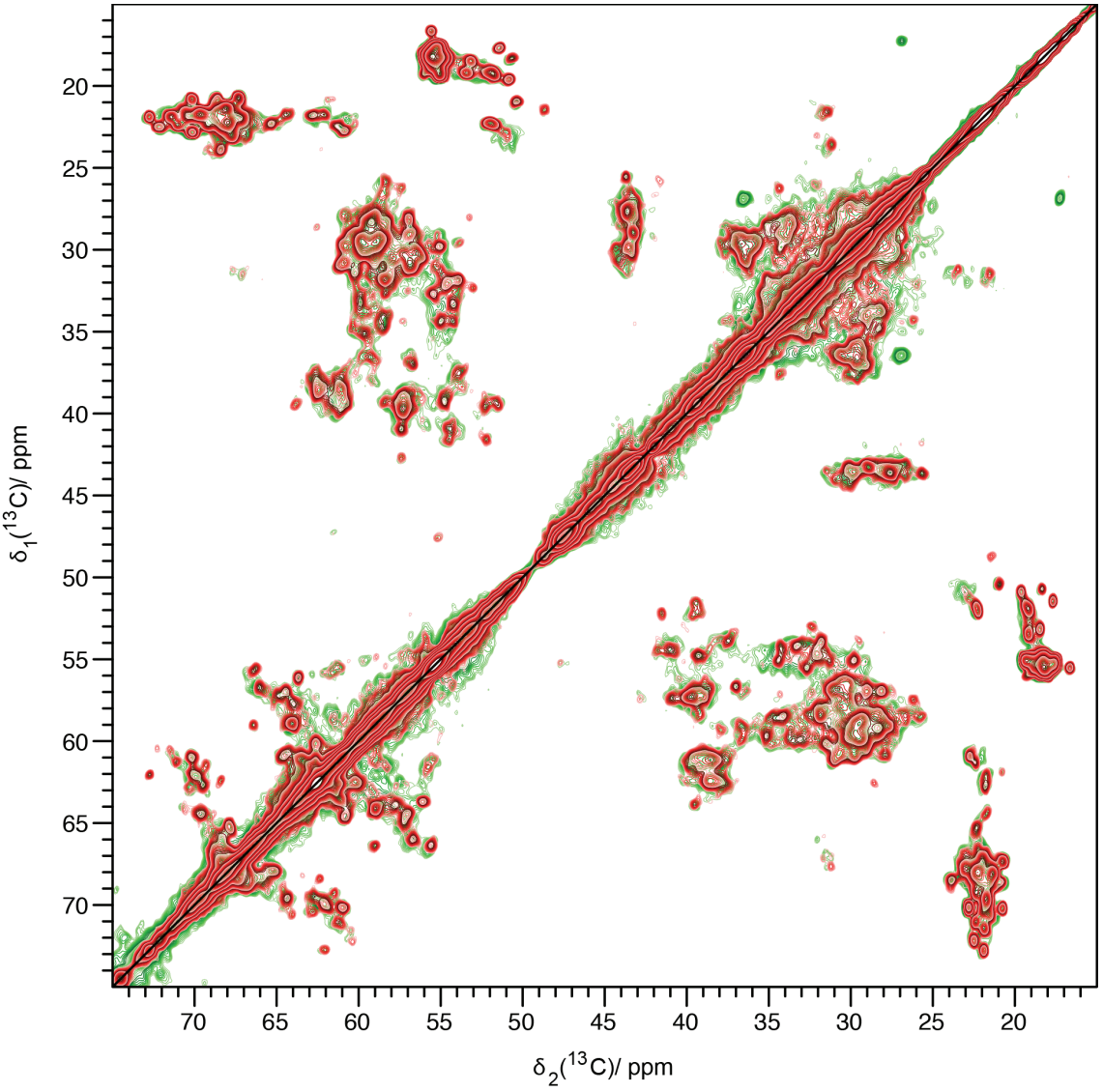


Figure S5. 2D DARR of BmrA and BmrA-E474R Aliphatic Extracts of a 2D ^{13}C-^{13}C 20 ms DARR spectra of lipid-reconstituted [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-WT (red shades) overlaid with [^{12}C-^{14}N-LVIKHP]-^{13}C-^{15}N-BmrA-E474R (green shades).

Chapter IX

Conclusion and perspectives

Conclusion

In the present work we have investigated the ABC transporter BmrA by the means of NMR. The protein production, purification and reconstitution in lipids have been optimized as far as possible. The yield of 10 mg per liter of culture is one of the highest yields for an ABC transporter considering the discussions about it.

The reconstitution of the ABC transporter BmrA in lipids by detergent depletion, using dialysis and Bio-beads, was set up. Lipid reconstitution was followed using detergent quantification, electronic microscopy and enzymatic assay.

Moreover, GRecon was set up using BmrA as a model. This method allows for a lipid reconstitution of the protein by using ultra-centrifugation for 8 hours instead of 6 days with the classical dialysis method. Both methods led to similar and well-resolved spectra of the ABC transporter BmrA in its lipid environment.

Due to its large size – 600 residues – the chemical shift assignment of uniformly labelled BmrA based on sequential NMR assignment is still challenging. However, we successfully used two different strategies in order to get information about the type of residue associated with its resonance and its localization in the protein. These two different approaches were paramagnetic relaxation enhancement using Mn^{2+} substitution and selective unlabelling. The Mn^{2+} substitution which induced paramagnetic relaxation enhancement led to a disappearance of the resonances located around the paramagnetic species. By contrast, selective unlabelling allowed for a remarkable decrease in peaks overlap. The combination of the two approaches allowed us to identify the resonances located in the binding site and some resonances assignments using 2D DARR with 200 ms mixing time. The implementation of those strategies brought about a better analysis of the spectra and provided insight in the functioning of the ABC transporter BmrA. Indeed, NMR fingerprints of the two main ABC transporter conformations were recorded: the inward-facing and the outward-facing states.

Moreover, two different outward-facing states were obtained: a pre-hydrolytic state using non-catalytic mutant and a hydrolysis transition-state using ATP:Mg^{2+} and vanadate. Both outward-facing states revealed that each monomer is structurally symmetric to each other, suggesting that BmrA seems to follow the processive clamp model. Furthermore, ^{31}P NMR pointed to an asymmetry of the ADP:VO_4^{2-}.indicating two different environments around

the two ADP bound with the protein. We postulated that one ADP lost its VO_4^{2-} which seemed to have been released from the protein. This last feature is again in favor of the processive clamp model in which the ATP is hydrolyzed sequentially with phosphate releasing after each ATP-hydrolysis.

The last point concerns the X-loop mutant which is ATPase active but is inactive in its drug-transport activity. NMR experiments highlighted a loss of rigidity of the outward-facing state in the X-loop mutant compared with the wild-type form. Furthermore, in the presence of ATP:Mg^{2+} and vanadate, an analysis of resonance from the X-loop mutant TMDs suggests that the TMDs may remain in the inward-facing state. In contrast, the NBDs display the specific fingerprint of an outward-facing state. These results reveal that the X-loop motif is crucial for the conformational transmission from the NBDs to the TMDs. A mutation of the motif leads to a dissociation of the TMDs from the NBDs.

Perspectives

All these results Show that NMR provides structural and mechanistic information. It is worthwhile remembering that the ABC transporter BmrA was studied in its *Bacillus subtilis* lipid environment, making our investigations particularly relevant. However, TMDs and the identification of the residues involved in drug-binding have not yet been studied with solid-state NMR.

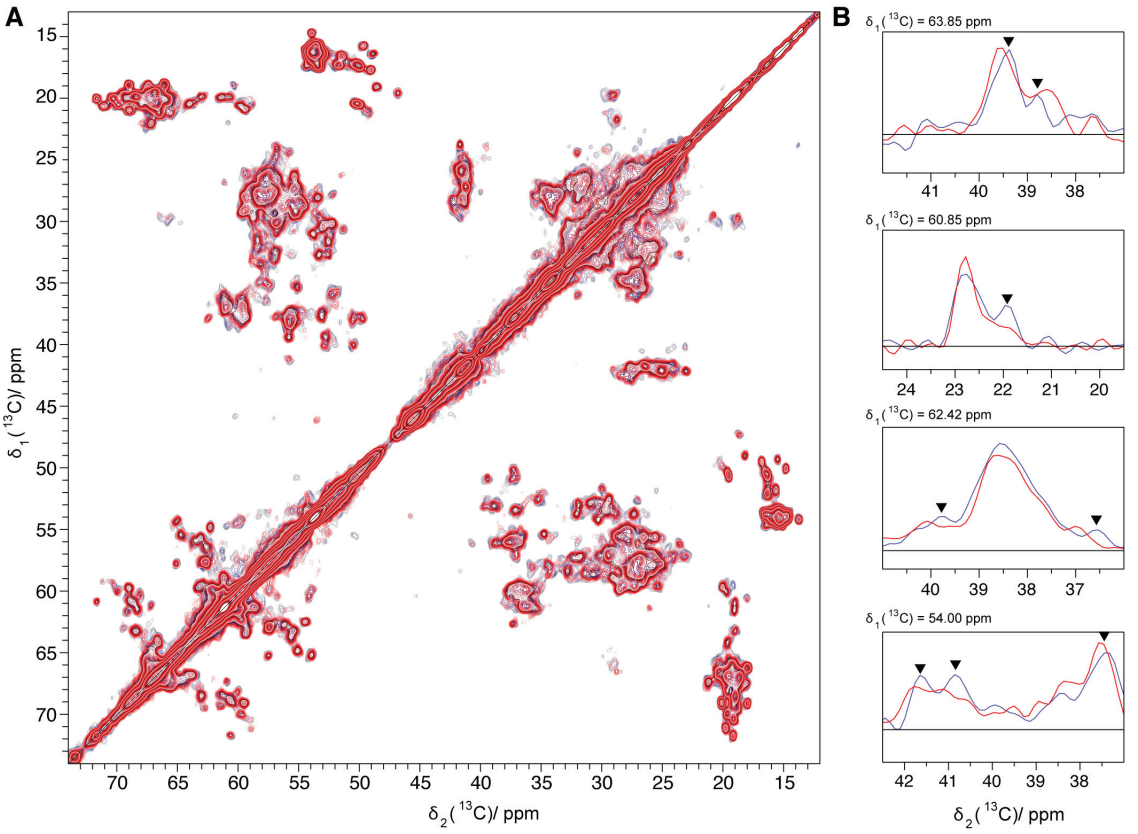


Figure 60. Extracts of a 2D ^{13}C-^{13}C 20ms DARR spectrum of BmrA and BmrA in presence of Hoechst 33342. A. Extracts of a 2D ^{13}C-^{13}C 20ms DARR spectrum of BmrA (red shades) and BmrA in presence of Hoechst 33342 (blue shade). **B.** one dimensional traces of a 2D ^{13}C-^{13}C 20ms DARR spectrum of BmrA (red lines) and BmrA in presence of Hoechst 33342 (blue lines). Chemical shift perturbations are highlighted by a black arrow.

Sample of ^{13}C-^{15}N-[^{12}C-^{14}N-LVIKHP]-BmrA (140 residues labelled) reconstituted in *Bacillus subtilis* lipids was incubated with 100 µM of Hoechst 33342 which is a drug-substrate of BmrA with a Kd of 9.5 µM. As presented in **Figure 60**, the 2D spectrum of BmrA Apo and BmrA drug-bound was almost the same. Minor CSP can be found for few resonances. These

minor CSP are due to the facts that carbons are less surface exposed than ^{1}H or ^{15}N as a result minor effect on Cα and Cβ shifts and the drug-binding site is localized in the TMD of BmrA. Minor CSP in areas in which ^{13}C-^{13}C resonances were severely overlapped make CSP identifications difficult. However, the overall signal from BmrA drug-bound appears to be slightly higher comparing to the Apo form, despite that both spectra were scaled on the α-helical resonances from alanine. This slight difference could be caused by a diminution of the protein flexibility induced by the drug-binding.

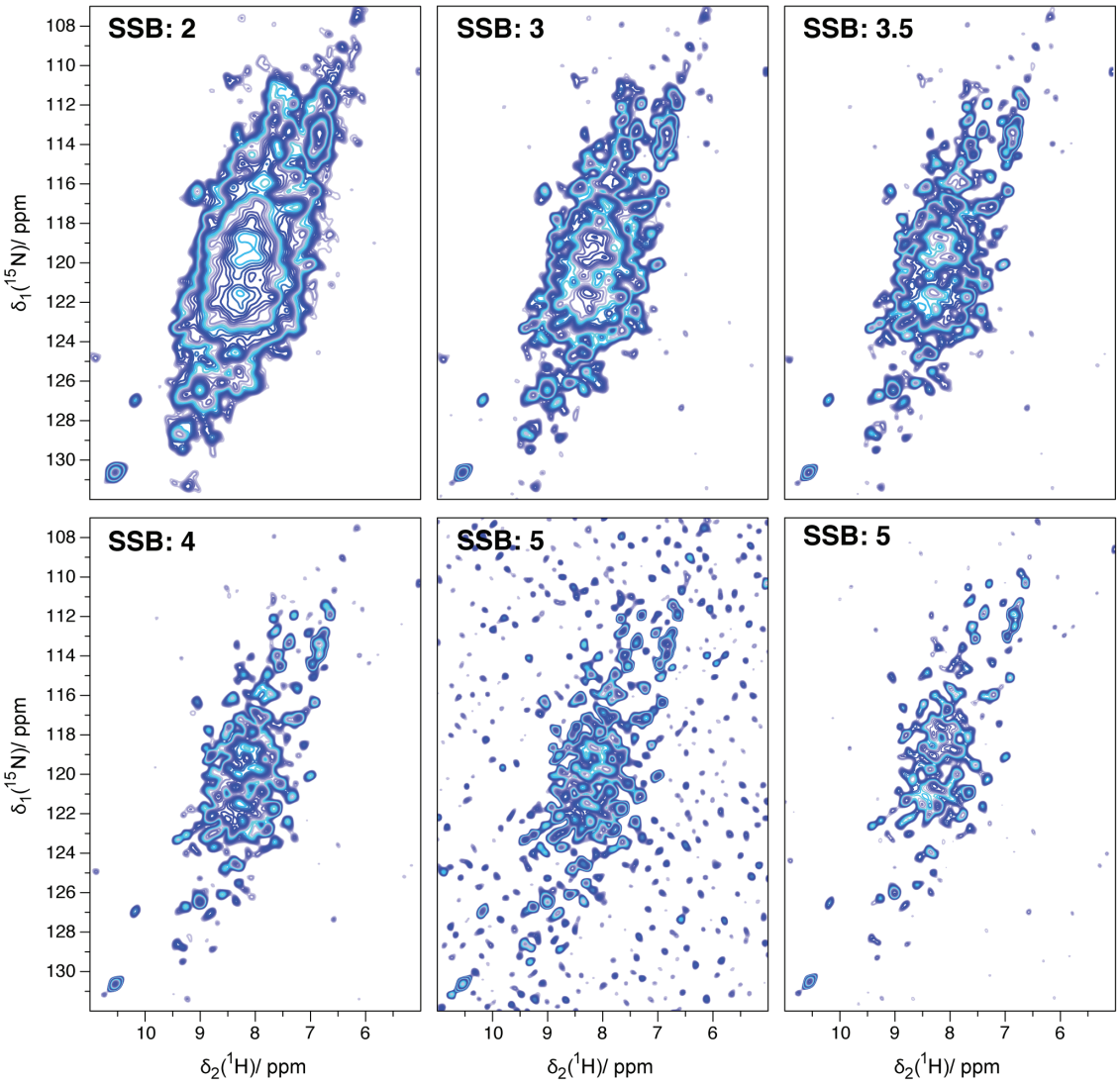


Figure 61. HNH 2D spectrum ^{15}N-^{13}C-ADEQP-$W_{side\ chain}$-BmrA 110 kHz. Spectrum was recorded during 14 h. Various sine bell shift (SSB) have been shown between 2 to 5. Two spectra with SSB 5 are presented, one with the same base level of SSB2 to 4 and one with a higher base level in order to remove the noise. The spectrum shows isolated peaks but the majority of peaks are overlapped due to the a-helical organization of the protein. The ^{13}C signal seems to be too low for 3D recording, probably a consequence of ^{13}C scrambling due to the adding of too much unlabelled residue.

Despite 40 years of ABC transporter studies, the drug-binding pocket was never strictly identified. ABC transporter structure in presence of a drug was never solved. In spite of real progress in crystallography or in cryo-microscopy, solid-state NMR at >110 kHz rotor spinning frequency remains the best suited method for an in-depth study of the ABC transporter BmrA.

Steady progress in this field makes it possible to record HNH 2D spectrum with a resolution near to liquid-state NMR. The first HNH 2D spectrum of BmrA in which only ADEQP-$W_{side\ chain}$ residues were ^{15}N labelled, was recorded and is really promising for this challenging protein in terms of size and for its hydrophobic nature, **Figure 61**.

This first HNH 2D spectrum of BmrA paves the way to further prospects for solid-state NMR of large membrane protein embedded in its natural lipids. Indeed, using the sensitivity of the proton, the combination of selective labelling/unlabelling and the deuterium/proton exchange, only resonances from residues localized in the TMDs can be selectively HNH labelled. Characterization of the drug-binding site and identification of key-residues involved in drug binding of an ABC transporter are within our reach using solid-state NMR at >110 kHz.