Dror E. WARSCHAWSKI
dror@ibpc.fr

La RMN :
Comment ça marche ? A quoi ça sert ?

Séminaire du laboratoire de Régulation de l'Expression Génétique Microbienne (UPR 9073)
IBPC, Paris
5 avril 2013 – 11h30

Biologie Physico-Chimique des Protéines Membranaires

13 rue Pierre et Marie Curie - 75005 Paris
Tél 01 58 41 51 11 - Fax 01 58 41 50 24

http://www.ibpc.fr/UMR7099/

INSTITUT DE BIOLOGIE PHYSICO-CHIMIQUE
FRC 550

Plan

La RMN :
Comment ça marche ? A quoi ça sert ?

- **Généralités sur la RMN :**
  résolution et sensibilité, soluble *vs.* solide

- **Etude de molécules « solubles » :**
  peptides, protéines solubles, complexes, protéines membranaires

- **Etude de molécules « solides »:**
  microcristaux, fibrilles, protéines membranaires

- **Conclusions**

Pourquoi ce séminaire ?

En 2011, un consortium de laboratoires de l'IBPC, de l'Institut Pasteur et de l'Université Paris Descartes ont obtenu un financement « Equipex » pour leur projet d'étude de « systèmes complexes » intitulé CACSICE

Ce projet implique l'installation en 2014, à l'Université Paris Descartes d'un appareil de SAXS ; à l'Institut Pasteur d'un cryo-miscroscope électronique « Polara » et d'un spectromètre de masse « Synapt » ; à l'IBPC d'un mur de visualisation, d'un spectromètre de masse « Orbitrap » et... d'un spectromètre RMN « solide/liquide ».

Aujourd'hui à l'IBPC, ce spectromètre concerne principalement l'équipe de Bruno Miroux, avec Laurent Catoire (RMN « liquide ») et Dror Warschawski (RMN « solide ») au 3ème étage.

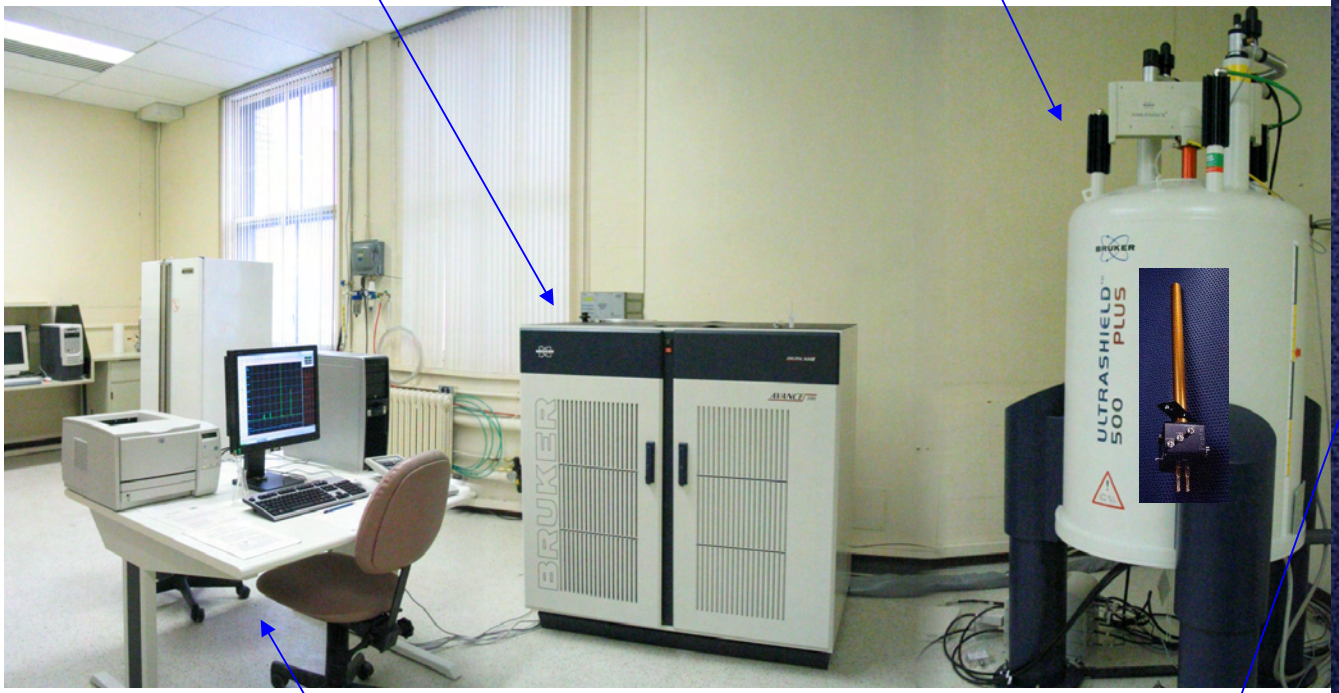
Toutefois, nous espérons que la présence d'une telle machine entraînera des collaborations et, qui sait, des vocations ?

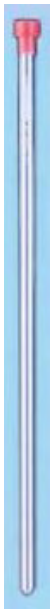
Qu'est-ce qu'un spectromètre RMN?

Console

Aimant (≈1 M€)

Tube





Station de travail

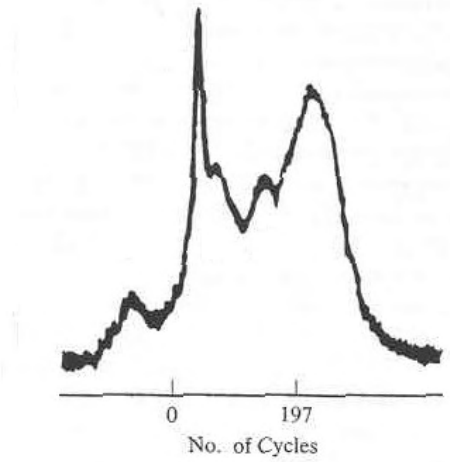
Sonde
(≈100 k€)

Pourquoi un « haut champ » ?

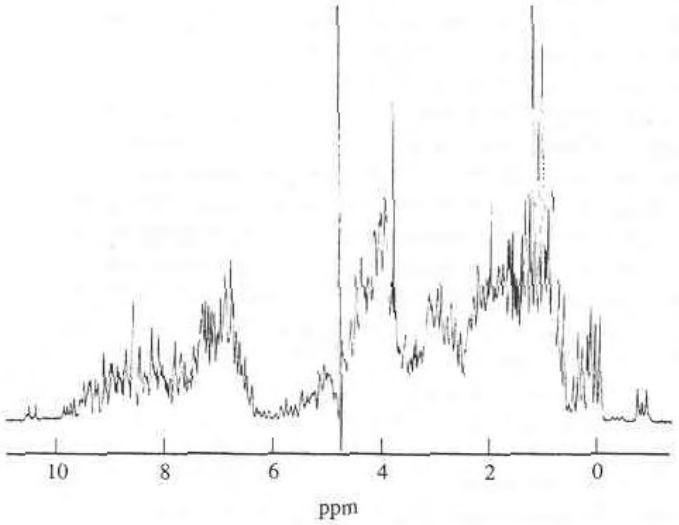
Plus le champ est élevé, mieux les expériences RMN seront résolues:

Spectre RMN de la ribonucléase (124 acides aminés):

1957, 40 MHz (1 T):



1997, 750 MHz (18 T):



Pourquoi un « haut champ » ?

Plus le champ est élevé, plus les expériences RMN seront sensibles:

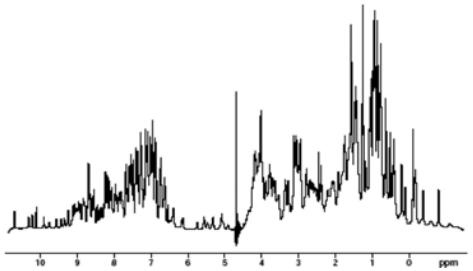
Si pour un champ de 14 T (600 MHz), la sensibilité est de 1,

alors pour un champ de 19 T (800 MHz), la sensibilité sera de 1.8

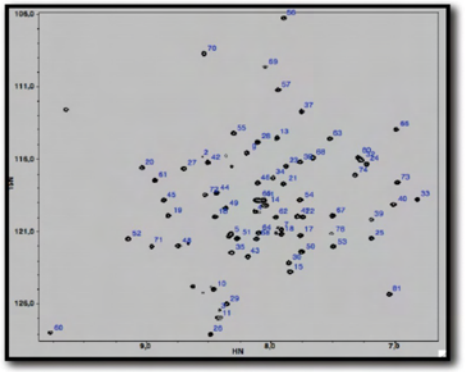
Un progrès récent consiste à refroidir la sonde (cryosonde) et permet encore de multiplier la sensibilité par 4, ce qui correspond à un gain de temps de 4^2, soit 16...

A quoi peut servir la RMN en biologie?

(i) La Résonance Magnétique Nucléaire est d'abord une technique d'analyse. Elle permet donc de détecter, d'identifier et de quantifier des molécules chimiques ou biologiques.



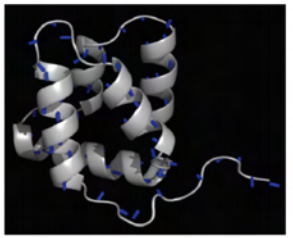
(ii) Par une utilisation plus sophistiquée, la RMN est aussi une technique de détermination structurale de molécules, et en particulier de macromolécules biologiques: protéines, acides nucléiques etc.



(iii) En biologie structurale, la RMN s'est fait une place à côté de l'autre grande technique qu'est la Cristallographie aux Rayons X:

(i) Parce que certaines protéines ne cristallisent pas

(ii) Parce que l'échantillon de RMN est plus proche des conditions natives (solution aqueuse, température ambiante, interaction avec des ligands ou des récepteurs, dynamique...)



Résolution et sensibilité en RMN

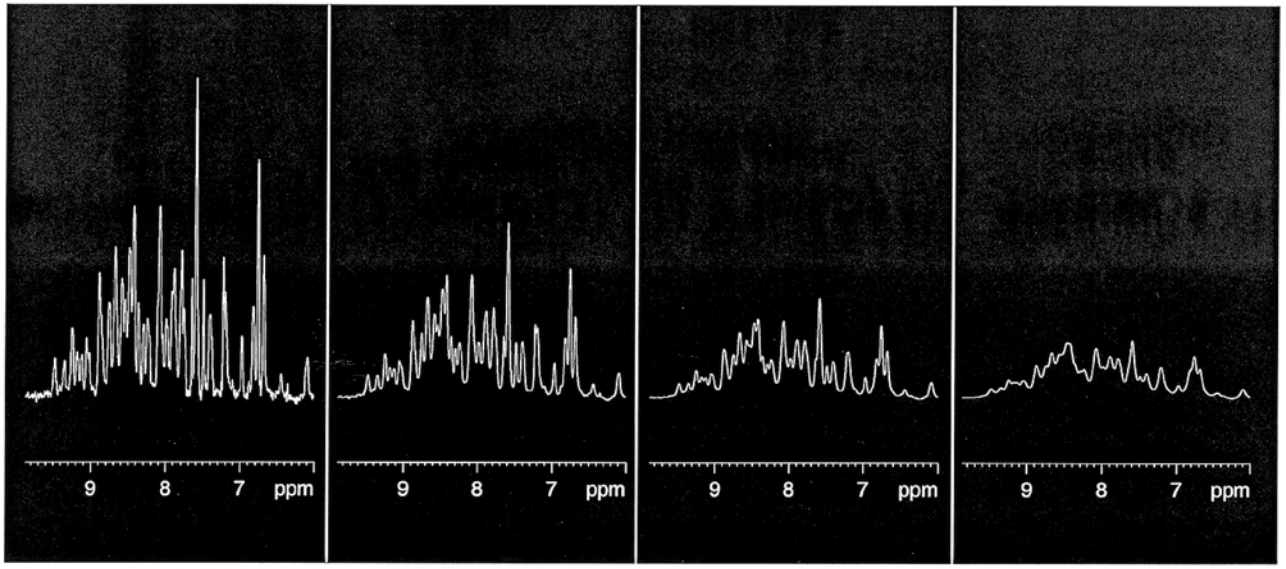
Quelle taille de protéine étudier?

Plus la protéine est grosse:

- -Moins il y aura de protéines par unité de volume: problème de sensibilité
- -Plus il y aura des résonances sur le spectre: problèmes de résolution
- -Moins elle bouge vite et les raies s'élargissent:
  problèmes de résolution **et** de sensibilité

τ_c	5ns	10ns	15ns	25ns	> µs
MW	10kDa	20kDa	30 kDa	50kDa	

à 500 MHz :



RMN des solides (...)

Résolution et sensibilité en RMN

Quel échantillon pour la RMN?

La RMN n'est pas une technique très sensible. Il faut à peu près 0.5 ml d'une solution concentrée à 1 mM

Hormis les ^{1}H, la plupart des spins obervés en RMN biologiques ne sont pas abondants naturellement. Il faut donc que l'échantillon soit **marqué isotopiquement**:

On introduit des ^{13}C et des ^{15}N pour les voir mais on introduit aussi des ^{2}H pour masquer les ^{1}H, en particulier dans les molécules du tampon.

Les ^{13}C et ^{15}N peuvent remplacer **tous** les ^{12}C et ^{14}N (marquage uniforme) ou seulement ceux d'une zone qui nous **intéresse** (marquage spécifique)

Etude structurale d'une protéine soluble par RMN

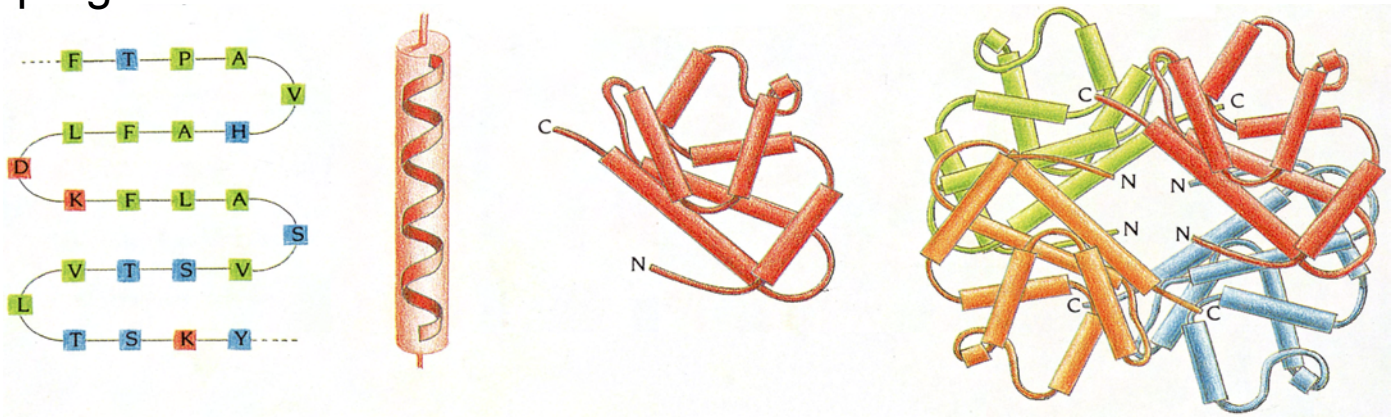
(1) On attribue les signaux RMN d'un maximum d'acides aminés de la protéine

(2) On détermine un maximum de contraintes entre noyaux, de préférence au moins 3 par acide aminé:

- -D'abord à travers les liaisons chimiques pour retrouver la séquence primaire
- -Ensuite, on déduit des éléments de structure secondaire (hélice α, feuillet β...)
- -Enfin par des contraintes à travers l'espace, on en déduit des éléments de structure tertiaire

En l'absence d'information structurale, on peut mesurer des dynamiques locales

(3) On en déduit la structure tridimensionnelle, en général à l'aide de programmes de modélisation

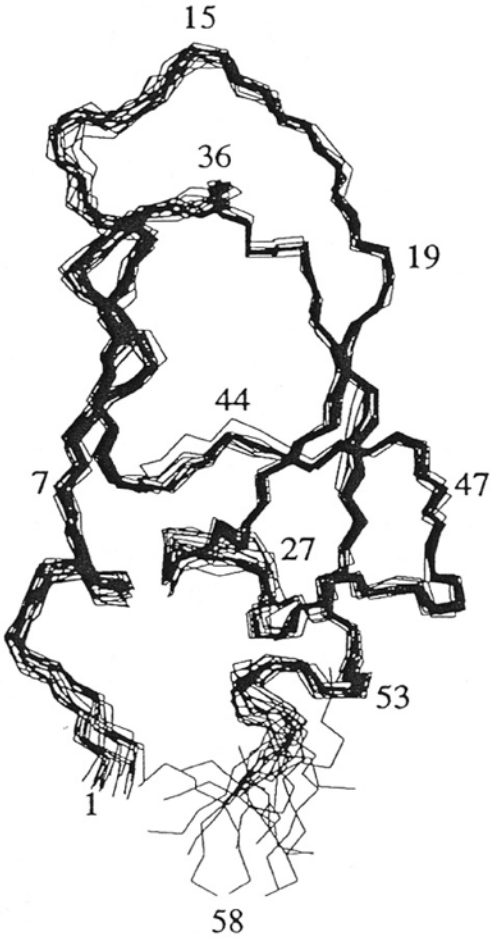


Etude structurale d'une protéine soluble par RMN

Un exemple: la BPTI (Bovine Trypsin Inhibitor), une protéine de 58 acides aminés
Wüthrich, 1992

642 contraintes de distances NOE et
115 contraintes d'angles J,
soit 13 contraintes par résidu!

20 conformères après modélisation
RMSD final: 0.43 Å



10 000 structures de protéines ont été déterminées en utilisant cette technique...

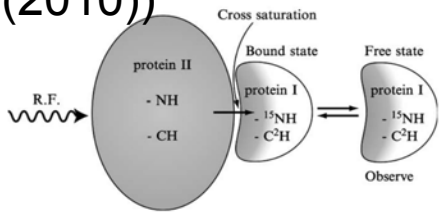
Le cas des complexes de protéines

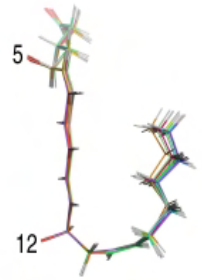
La RMN est particulièrement bien adaptée à l'étude d'interfaces entre protéines ou des interactions entre protéines et ligands:

- En comparant le spectre de la protéine avec ou sans son partenaire
- En cherchant les résidus accessibles à l'eau
- En jouant sur les transferts d'aimantation d'une molécule à l'autre, marquées différemment:

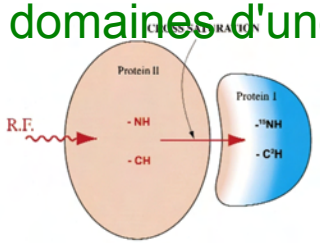
Exemples:

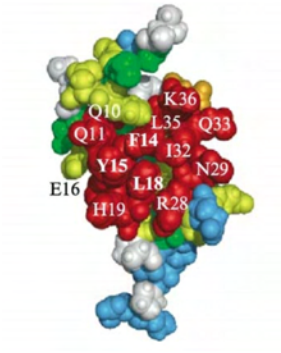
- La protéine est marquée au ^{2}H, le ligand est ^{1}H: on peut obtenir la structure du ligand dans son site actif (RCPG, Catoire *et al.* *J. Am. Chem. Soc.* **132**:9049-9057 (2010))





- Un domaine est marqué au ^{2}H et au ^{15}N, l'autre est ^{1}H: on peut obtenir l'interface entre deux domaines d'une protéine (Immunoglobuline, Takahashi 2000)





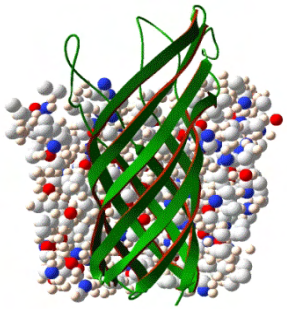
Le cas des protéines membranaires

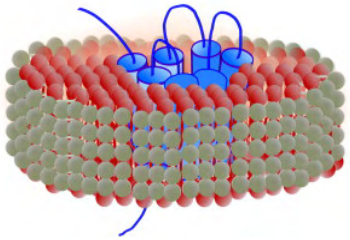
Il faut d'abord pouvoir maintenir ces protéines en solutions:

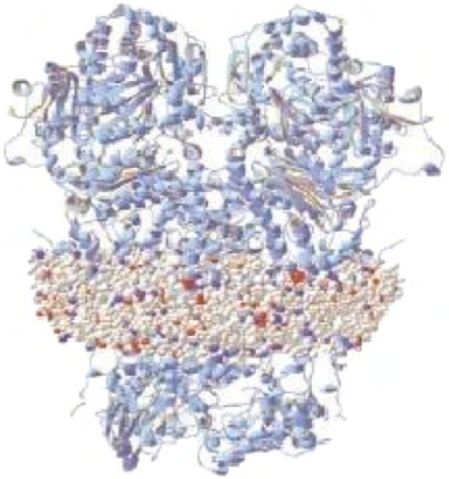
Plusieurs solutions existent : il faut que le complexe soit le plus petit possible, de préférence avec un temps de corrélation inférieur à 50 ns (soit à peu près 100 kDa):

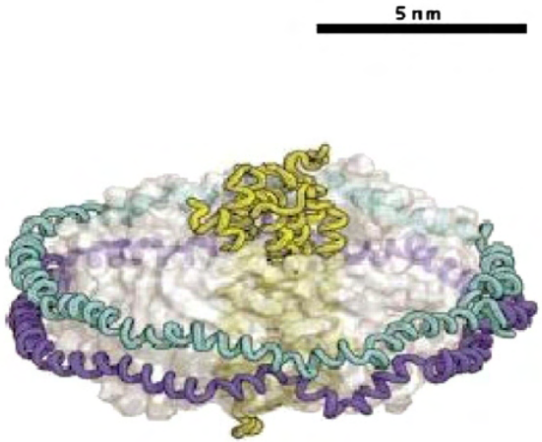
détergents, bicelles, amphipols, nanodisques...

Warschawski *et al. Biochim. Biophys. Acta* **1808**:1957-1974 (2011)









Le cas d'une protéine membranaire: VDAC

La protéine membranaire est sortie de sa membrane et maintenue en solution grâce à des micelles de détergents

VDAC (Voltage Dependent Anion Channel), une protéine membranaire de 283 acides aminés en LDAO:
Hiller *et al. Science* **321**:1206 (2008)

612 contraintes de distances NOE et

316 contraintes d'angles J,

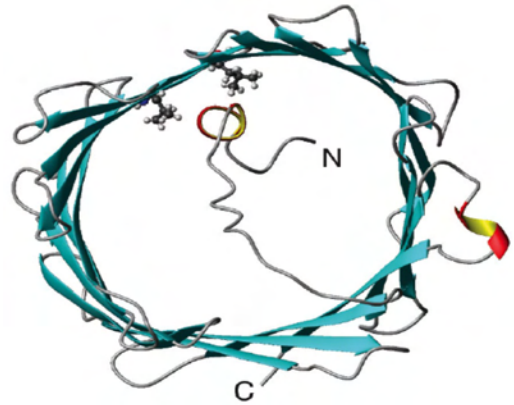
soit 3.3 contraintes par résidu

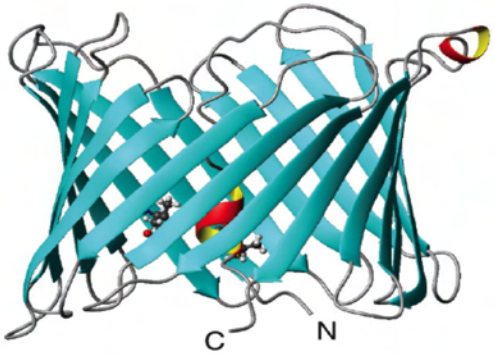
Table S1. NMR structure statistics

NOE distance restraints	All	H^N–H^N	H^N–Meth.	Meth.–Meth.
Intraresidual	69	0	69	0
Sequential	199	129	69	1
Medium range ($2 \leq \lvert i-j \rvert \leq 4$)	72	28	34	10
Long range ($\lvert i-j \rvert > 4$)	272	131	85	56
Total	612	288	257	67
Hydrogen bond restraints	139			
Dihedral angle restraints (ϕ and ψ)	2 x 158			
Ramachandran plot				
Most favored region	77.1%			
Additionally allowed region	21.6%			
generously allowed region	0.8%			
Disallowed region	0.4%			
Deviations from idealized geometry				
Bond lengths (Å)	0.005			
Bond angles (°)	0.7			
Average pairwise r.m.s.d. (backbone, Å)				
All residues except 1–25	3.6			
All residues except 1–25 and loops[a]	2.9			

20 conformères après modélisation

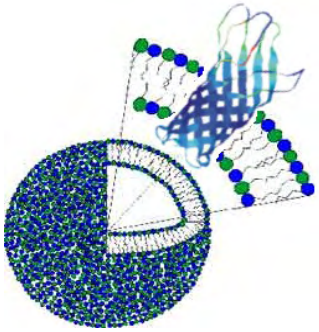
RMSD final: 2.9 Å





RMN solide

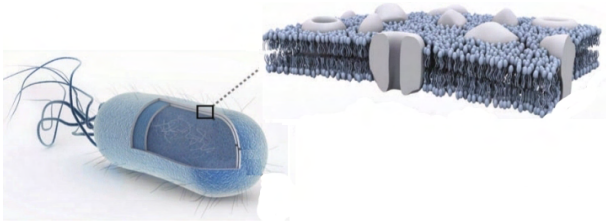
Certaines molécules biologiques sont trop grosses ou dans des environnement tels que leurs mouvements sont très lents.



Exemples:

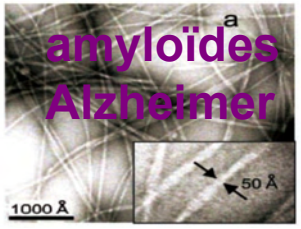
1) Les protéines membranaires dans leurs environnement natif

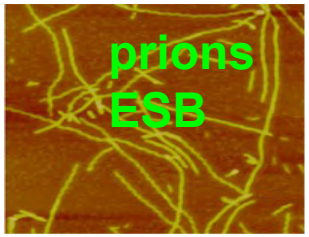
2) Les extraits cellulaires



3) Les poudres, cristaux ou micro-cristaux de protéines

4) Les protéines qui forment naturellement des fibrilles ou des plaques

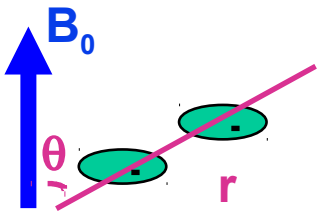




etc...

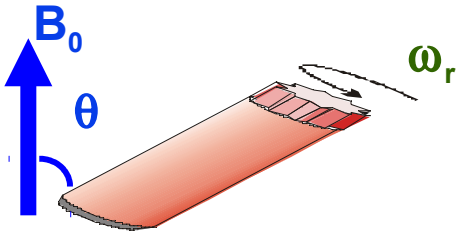
RMN solide

Sans mouvements rapides et isotropes, l'interaction dipolaire, responsable de la largeur des raies, n'est pas moyennée à zéro:



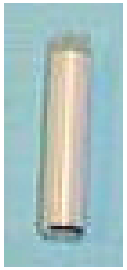
$$\Delta\nu_D = \frac{\mu_0 \gamma_i \gamma_j \hbar}{16 \pi^3 r^3}\left(3\cos^2\theta - 1\right)$$

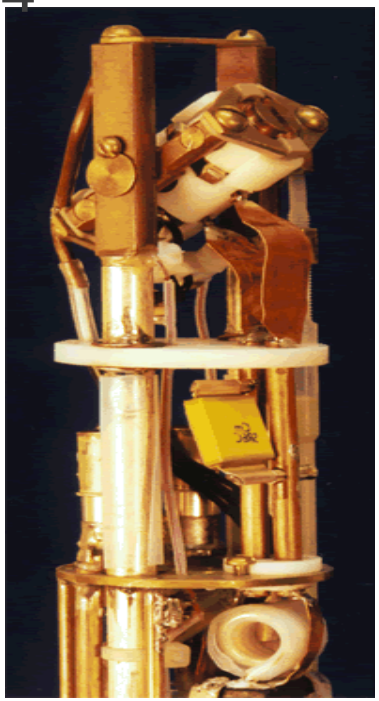
Une solution consiste à tourner rapidement l'échantillon autour d'un angle (« magique ») tel que : $\left(3\cos^2\theta - 1\right) = 0$ **soit:** $\theta = 54°\,44'$



- angle θ : 54° 44' ("magique")
- volume: 100 µl
- diamètre: 4 mm
- ω_r : 5 kHz - 15 kHz

2 cm

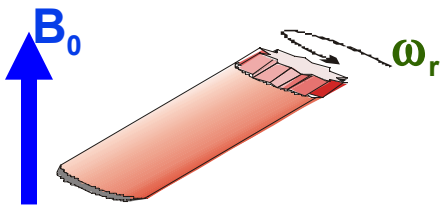


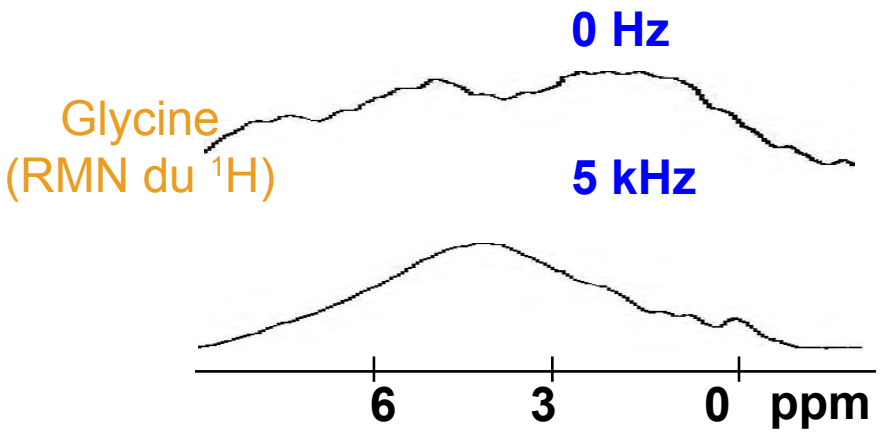


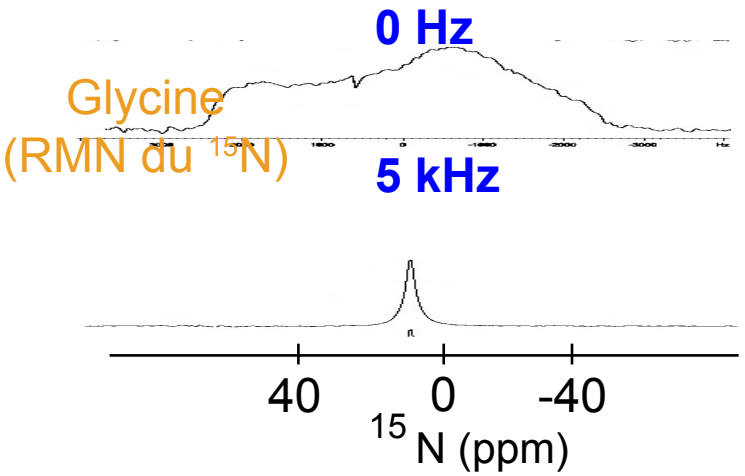
Sonde "MAS",
unité pneumatique,
compresseur,
réserve d'air comprimé,
sécheur d'air:

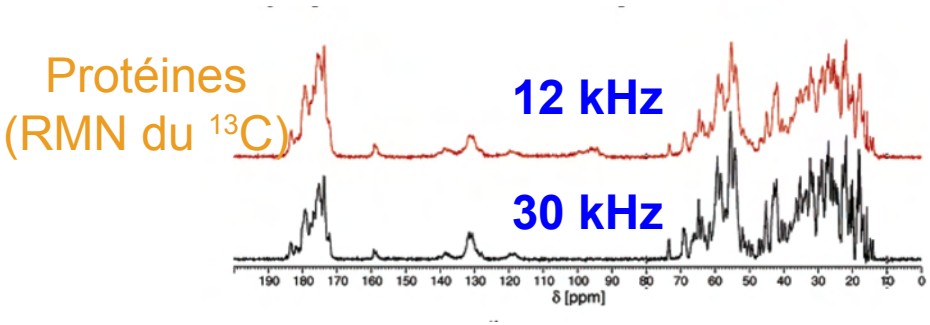
RMN solide

Exemples à une dimension:



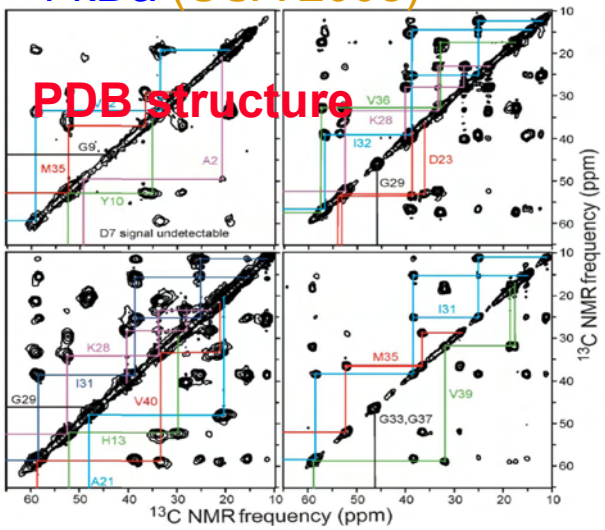




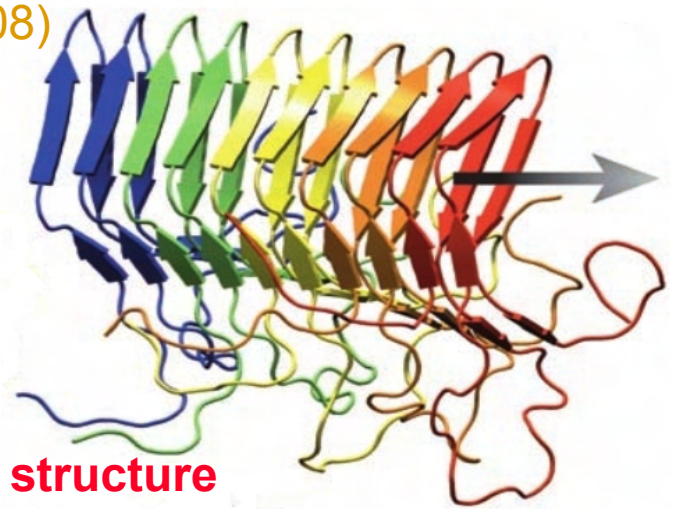


Le cas des protéines fibrillaires

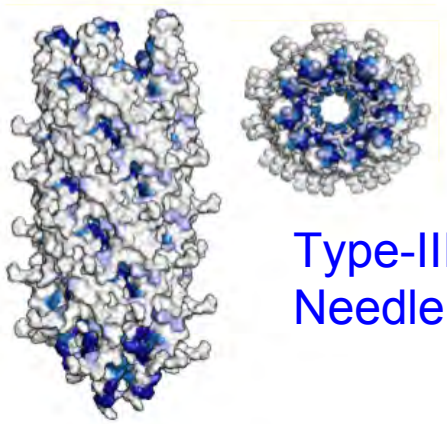
Fragment de β-amyloid (Alzheimer) de 4 kDa (USA 2008)



Fragment de HET-s (prion) de 5 x 9 kDa (Suisse 2008)



PDB structure



Type-III Secretion System
Needle de 9 x 29 kDa (Allemagne 2012)

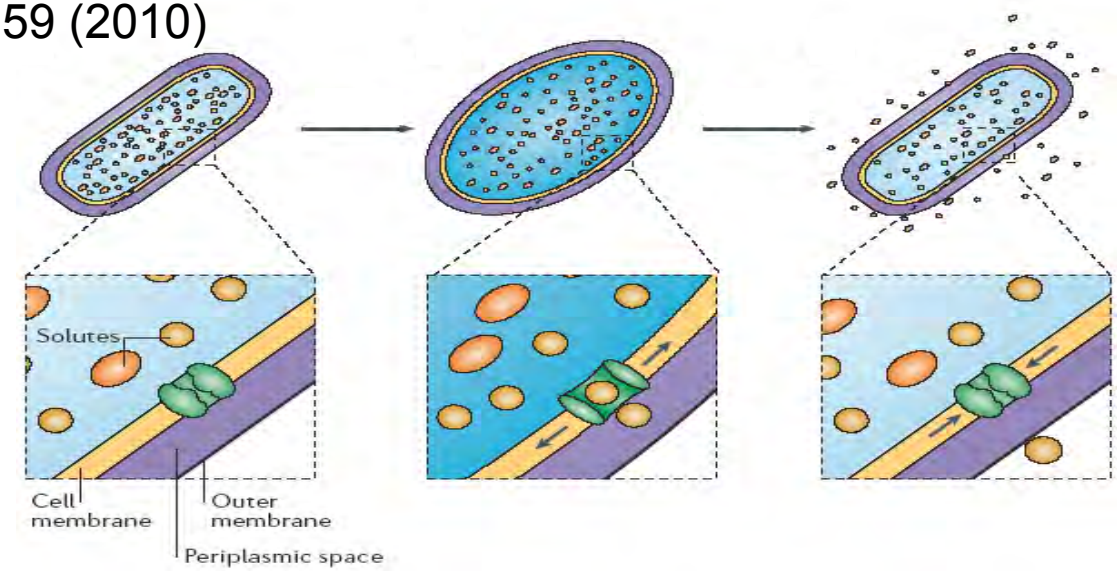
PDB structure

Le cas des protéines membranaires

La RMN des solides permet d'étudier les protéines membranaires dans un environnement proche des **conditions naturelles** (en terme de taille, forme, courbure, épaisseur, fluidité, pression latérale, constante diélectrique, hydratation...) et donc espérer que leurs structures et leur dynamiques soient celles des protéines natives et fonctionnelles...

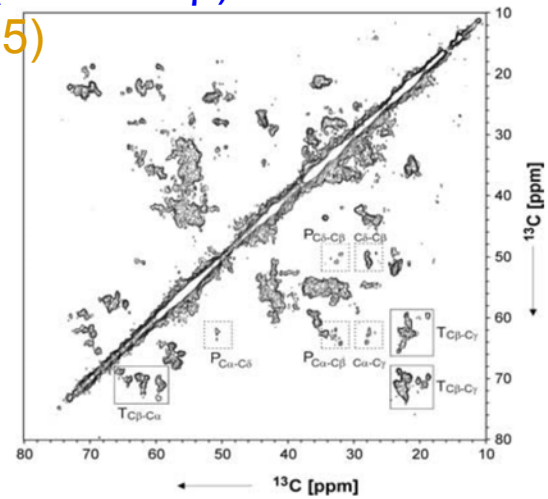
Exemple: **MscL** (canal ionique mécano-sensible bactérien)
- S'ouvre sous l'effet de tension membranaire
- 75 kDa mais 5 * 15 kDa
- Structure fermée cristallisée, mais en dehors de la membrane

Abdine *et al. J. Magn. Reson.* **204**:155-159 (2010)

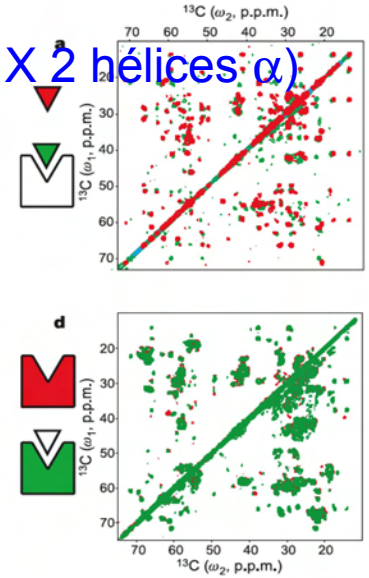


Le cas des protéines membranaires

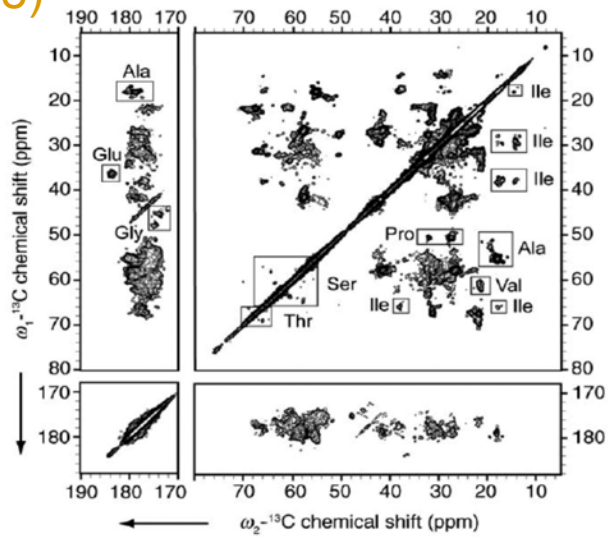
OmpG, 28kDa (tonneau β)
(Allemagne 2005)



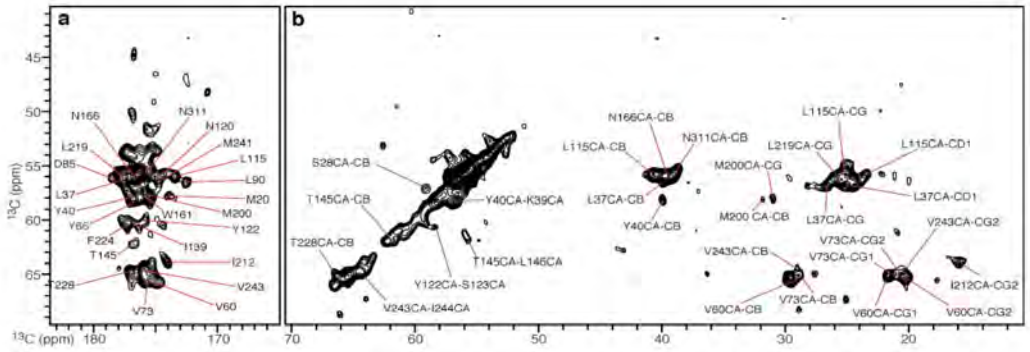
KcsA, 70 kDa (4 X 2 hélices α)
(Hollande 2006)



DsbB, 20kDa (4 hélices α)
(USA 2008)
attribuée



CXCR1, 34kDa (7 hélices α)
(USA 2012)
PDB structure



The 3rd Montagne Ste Geneviève Workshop on NMR of biological Solids

Friday, September 13, 2013
at IBPC, 13 rue et Marie Curie, 75005 Paris, France
http://www.ibpc.fr/biosolids/

Our detailed program will soon be available. The program will include a session on GPCR membrane proteins with ***Stan Opella*** *(San Diego, USA), another one on capsid and cell wall proteins with* ***Bo Chen*** *(Orlando, USA) and* ***Sabine Hediger*** *(Grenoble, France), and other applications of solid-state NMR in biological solids. Each session will contain a talk on solid-state NMR and a talk from a scientist in a different field. Since we expect a diverse attendance, we also plan a short opening course on the application of solid-state NMR to biology, and flash presentations from selected participants, to stimulate the discussions.*

Structures par RMN

-15 structures de protéines fibrillaires ou microcristallines et 19 structures de protéines membraires ont été déterminées par RMN des solides

-79 structures de protéines membranaires ont été déterminées par RMN des solutions:

PROTEINS OF KNOWN STRUCTURE

DETERMINED BY SOLID-STATE NMR

The table below provides information about proteins whose structures have been determined by solid-state NMR, to a resolution sufficient to have resulted in a file deposited at the RCSB Protein Data Bank.

Please help me to keep this table up to date by sending me missing information at: dror.warschawski at bgn.fr

Last database update : 1 March 2011

Unique proteins in database : 23

Number of coordinate files in database : 37

For more membrane protein structures: determined by NMR

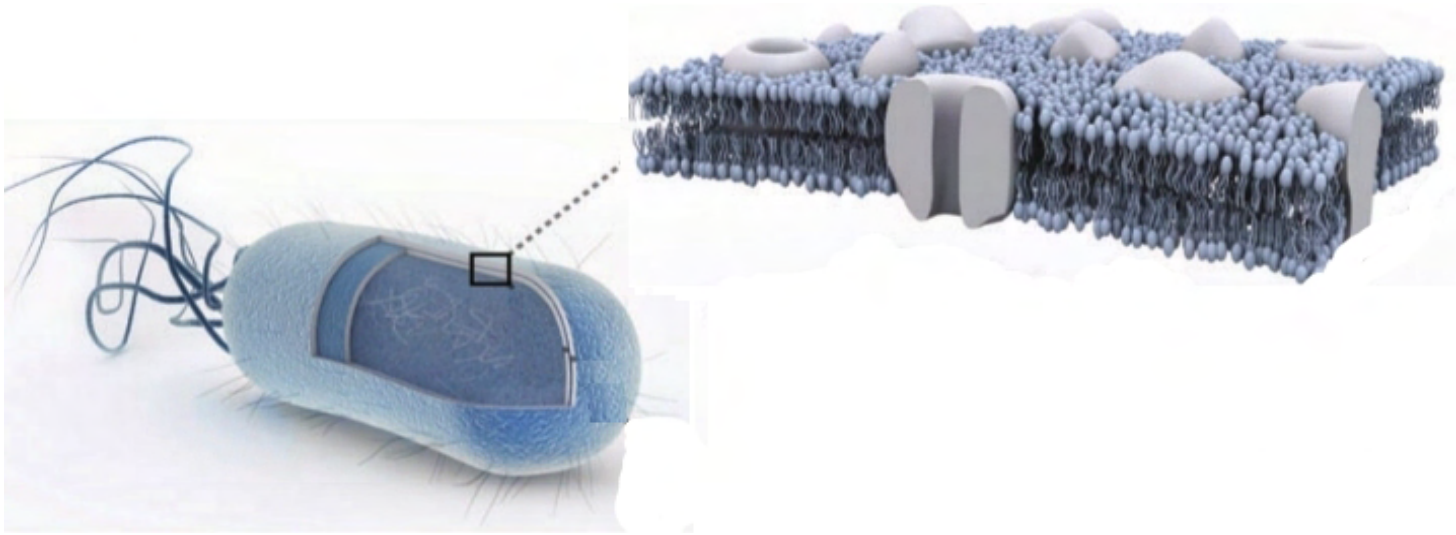
Protein	PDB Entry	BMRB Entry	Oligomer Size/State	Production Source	Reconstitution Medium	NMR Method	PubMed Reference
ALIGNABLE SOLUBLE PROTEINS							
fd coat protein in filamentous bacteriophage	1NH4	5813	5 kDa/1	*Escherichia coli*	-	**oriented**	Zeri et al. (2003)
Pf1 coat protein in filamentous bacteriophage (at various temperatures)	1FJP 1FN9 2KXM	5877	5 kDa/1	*Pseudomonas aeruginosa*	-	**oriented**	Thiriot et al. (2004) Thiriot et al. (2004) Strauss et al. (2011)
MICRO-CRYSTALLINE PROTEINS							
Spectrin SH3 domain	1M8M		7 kDa/1	*Escherichia coli*	-	**MAS**	Castellani et al. (2002)
Kalilotoxin	1X5W	6351	4 kDa/1	*Escherichia coli*	-	**MAS**	Lange et al. (2005)
beta1 immunoglobulin binding domain of protein G (GB1)	2G4 2JV 2K0P 2KOR	15136	6 kDa/1	*Escherichia coli*	-	**MAS**	Zhou et al. (2007) Franks et al. (2008) Reisenfeld et al. (2008) Nieuwkoop et al. (2009)
Crh domain, swapped dimer	2RLZ	5797	19 kDa/2	*Escherichia coli*	-	**MAS**	Loquet et al. (2008)
Ubiquitin	2LTZ		9 kDa/1	*Escherichia coli*	-	**MAS**	Marulloo et al. (2008)
MMP-12 (Macrophage metalloelastase)	2K9C 2K9J	7415	17 kDa/1	*Escherichia coli*	-	**MAS**	Bezyazyar et al. (2008) Bertini et al. (2010)
human alpha-defensin HNP-1	2KHT	16254	3 kDa/1	*Escherichia coli*	microcrystallized outside the lipid membrane	**MAS**	Zhang et al. (2009)
alpha-beta crystallin	2KLR	16391	40 kDa/2	*Escherichia coli*	-	**MAS**	Ishii et al. (2010)
PROTEINS FORMING FIBRILS							
Transthyretin in amyloid fibrils	1BVS		1 kDa/1	synthetic	-	**MAS**	Jenssen et al. (2004)
Amyloid protofilaments of beta2-microglobulin fragment	2RQ 2RQ	16202	10 kDa/4	*Escherichia coli*	-	**MAS**	Iwata et al. (2006)
WW domain of CA 150	2NNT		18 kDa/4	*Escherichia coli*	-	**MAS**	Ferguson et al. (2006)
Amyloid fibrils of the HET-s (218-289) prion	2RNM 2K3	11028 11064	43 kDa/5	*Escherichia coli*	-	**MAS**	Wasmer et al. (2008) Van Melckebeke et al. (2010)
Fragment of the human islet amyloid polypeptide (hIAPP)	2K3B	20074	6 kDa/8	synthetic	-	**MAS**	Nielsen et al. (2009)
MEMBRANE PROTEINS IN MEMBRANES							
Gramicidin	1MAG		4 kDa/2	synthetic	lipids	**oriented**	Ketchem et al. (1993)
AchR M2 channel	1C8K		3 kDa/1	*Escherichia coli*	lipids	**oriented**	Opella et al. (1999)
alpha-factor receptor M6	1F1D		2 kDa/1	synthetic	lipids	**oriented**	Valentine et al. (2001)
Influenza M2 channel residues 22-46	1NY1		11 kDa/4	synthetic	lipids	**oriented**	Wang et al. (2001)
Influenza M2 channel residues 22-46 with amantadine	2H95		8 kDa/4	synthetic	lipids	**oriented**	Hu et al. (2007)
Influenza M2 channel residues 22-46 with amantadine	2KAD 2KOT	16020 16612	11 kDa/4	synthetic	lipids	**MAS and oriented**	Cady et al. (2009) Cady et al. (2010)
Influenza M2 channel residues 22-62	2LJQ		20 kDa/4	*Escherichia coli*	lipids	**oriented**	Sharma et al. (2010)
fd coat protein	1M2T	5813	5 kDa/1	*Escherichia coli*	lipids	**oriented**	Marassi and Opella (2003)
HIV Vpu channel	1PJE 1PJF 1PJG		4 kDa/1	*Escherichia coli*	lipids	**oriented**	Park et al. (2003)
MERFT	2H30		6 kDa/1	*Escherichia coli*	bicelles	**oriented bicelles**	De Angelis et al. (2006)
Protegrin-1	1ZY8		4 kDa/2	synthetic	POPC	**MAS**	Mani et al. (2006)
human Phospholamban	2KB7		30 kDa/5	*Escherichia coli*	DPC and lipids	**solution and oriented**	Traaseth et al. (2009)
Pf1 coat protein	2K9J	5877	5 kDa/1	*Pseudomonas aeruginosa*	DMPC, bicelles and lipids	**solution and oriented**	Park et al. (2010)

http://www.drorlist.com/nmr/SPNMR.html

http://www.drorlist.com/nmr/MPNMR.html

Le cas des cellules entières

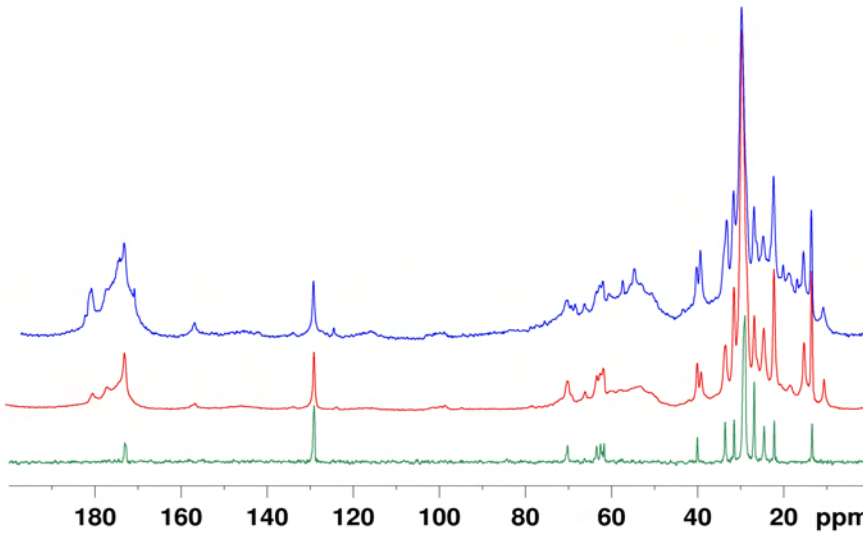
Puisque la RMN des solides n'est pas limitée par la taille de l'objet qu'elle étudie, elle peut même étudier des **cellules entières** ou des fragments de cellules dans leur environnement naturel, avec leurs partenaires membranaires et solubles, sans avoir jamais été en contact avec des détergents.



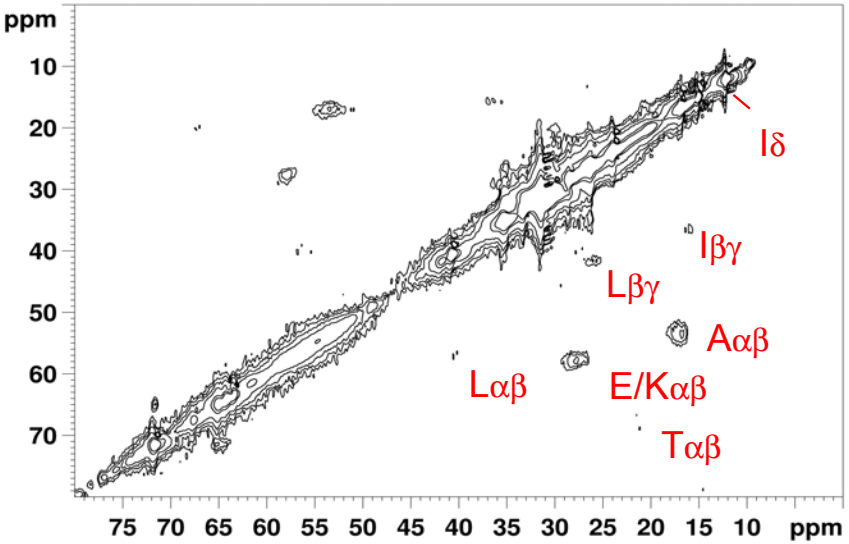
Remarque : depuis peu, la RMN des solutions permet aussi d'observer des petites molécules voire de **petites protéines solubles dans des cellules entières**...

Le cas des cellules entières

On peut aussi chercher à **voir une protéine d'intérêt** au milieu de la forêt d'autres molécules, en particulier les lipides... *(avec Xavier Warnet et Oana Ilioaia)*



^{13}C à 30°C et ω_r = 11 kHz



^{13}C-^{13}C DARR (21 heures d'acquisition)

b subunit	
AA	#
A	31
E	17
K	15
L	14
I	13
Q	10
R	10
V	10
D	9
S	5
N	4
F	3
G	3
M	3
T	3
P	2
C	1
H	1
W	1
Y	1
Total	156

Cellules entières surexprimant « b »

Membranes internes surexprimant « b »

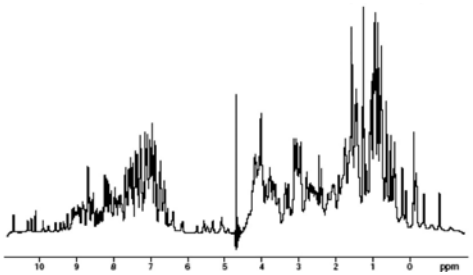
Liposomes modèles

Membranes internes surexprimant « b »

Conclusion 1/2

La RMN est:

Une technique jeune encore en plein développement



Utilisée pour l'analyse en chimie ou pour l'imagerie en médecine, elle est également utilisée pour déterminer des structures de macromolécules d'intérêt biologique

C'est une technique peu sensible, qui permet néanmoins d'observer les ^{1}H, ^{13}C et ^{15}N de molécules placées dans un champ magnétique.

Elle produit des spectres qu'il faut ensuite interpréter selon de nombreux paramètres: fréquences, intensités, formes de raies etc...

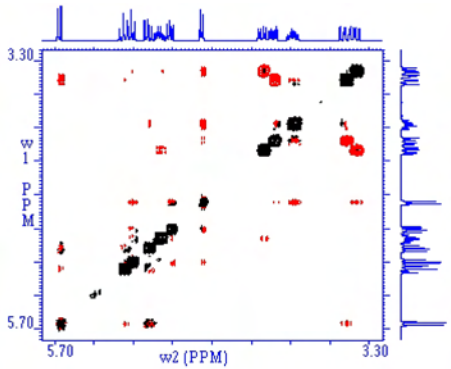
Il existe plusieurs variantes de techniques RMN, à une ou plusieurs dimensions.

La France possède quelques spectromètres à haut champ dédiés à la biologie, en particulier à Lyon et à Gif-sur-Yvette.

Conclusion 2/2

La RMN en biologie permet:

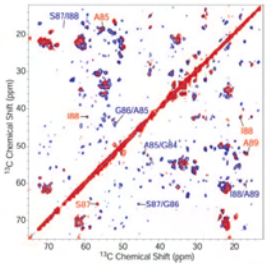
D'étudier des macromolécules biologiques: protéines, acides nucléiques, lipides, complexes protéines-ligands, protéines membranaires etc...



-D'attribuer les signaux provenant de chaque noyau de ces molécules
-De détecter des contraintes entre ces noyaux (distances, orientations...)
-D'en déduire des structures tridimensionnelles à très haute résolution

La RMN des solides en biologie est en pleine expansion et elle permet:

D'étudier des macromolécules biologiques dont l'état natif correspond à un état solide et très difficile à cristalliser, en particulier les fibriles, les protéines membranaires, et même les cellules entières :



http://www.drorlist.com/nmr/SPNMR.html