Réseau National de Formation en Biologie Structurale Intégrative

Contributions de la RMN à la biologie structurale : Approches multi-échelles spatiales et temporelles

Nathalie SIBILLE
CNRS UMR 5048, CBS
nathalie.sibille@cbs.cnrs.fr

NMR - Spatial and Time scales are highly intermingled

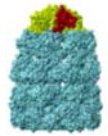
Size vs dynamics ?

Size ⟺ molecular tumbling + local dynamics

⇕ relaxation processes

⇕ linewidths & S/N

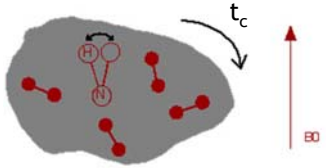
solubilized membrane proteins

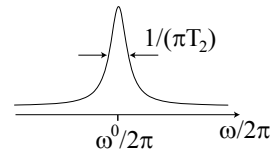


"Small" globular Protein / domain

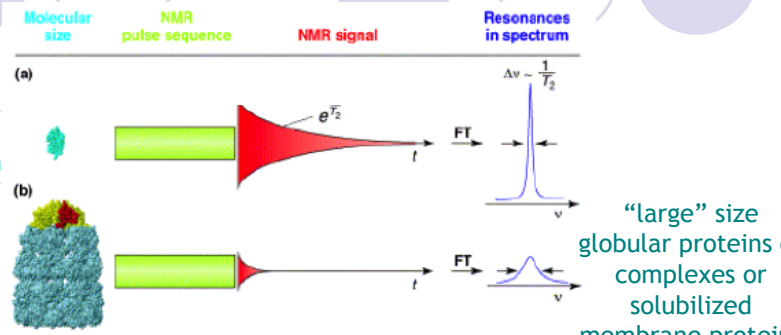
"large" size globular proteins or complexes

Global correlation time t_c (size) → Linewidth





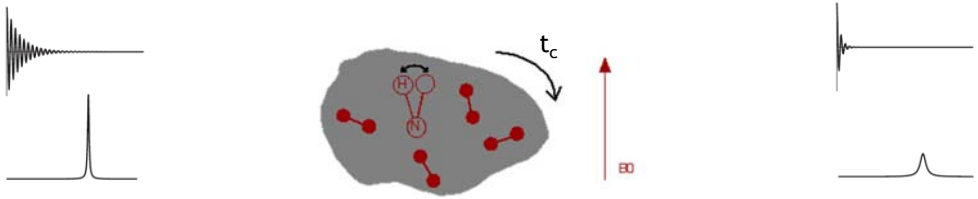
NMR - Spatial and Time scales are highly intermingled



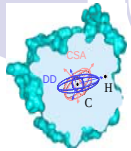
"Small" globular Protein / domain

"large" size globular proteins or complexes or solubilized membrane proteins

Global correlation time t_c (size) → Linewidth

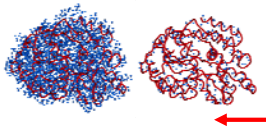


NMR - Spatial and Time scales are highly intermingled



solubilized membrane proteins

TROSY, high fields

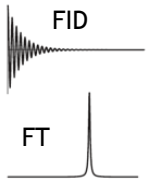


"large" size globular proteins or complexes

Labelling, Deuteration

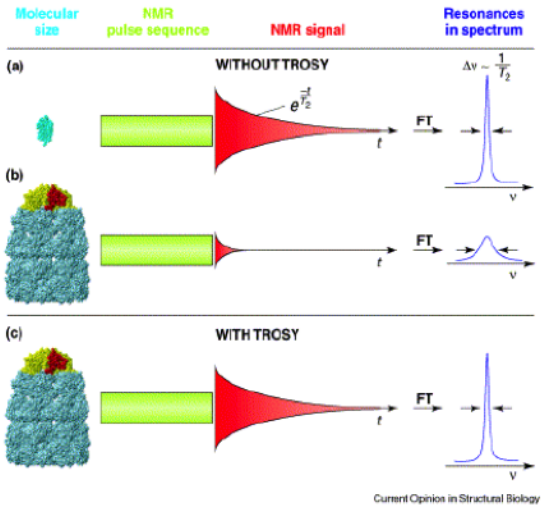
"Small" globular Protein / domain

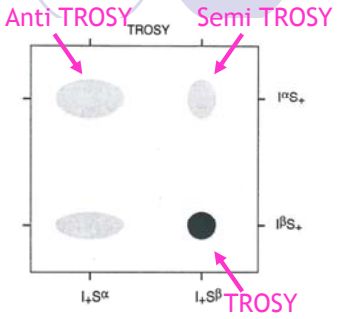
Global correlation time t_c (size) → Linewidth

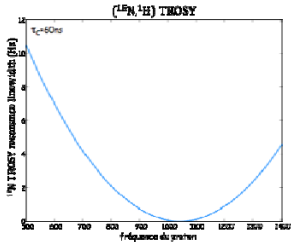


+ Local Rotational fluctuation timescale
+ density of relaxation sources

« Size » and lifetime of the signal T2





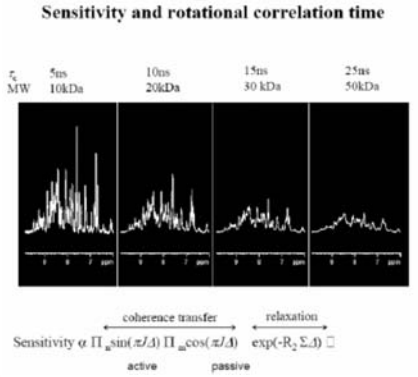


Playing with T2 : TROSY experiments
Transverse Relaxation Optimized SpectroscopY

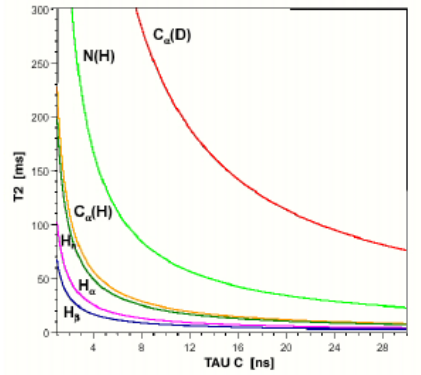
Pervushin, K. Riek, R., Wider, G. and Wütrich, K(**1997**) Proc. Natl. Acad. Sci. U. S. A. 94, 12366–12371

Nathalie Sibille - RéNaFoBis 2016

« Size » and lifetime of the signal T2



- **MW < 5 kDa**
  Homonuclear assignment (TOCSY,NOESY)
- **5 kDa < MW < 10 kDa**
  Heteronuclear assignment: ^{15}N
- **10 kDa < MW < 20 kDa**
  Heteronuclear triple resonance assignment : ^{15}N, ^{13}C
- **20 kDa < MW**
  Heteronuclear triple resonance assignment : ^{15}N, ^{13}C, ^{2}H (*E.coli* BL21+++)



The signal loss is exponential while the size of the molecule increases

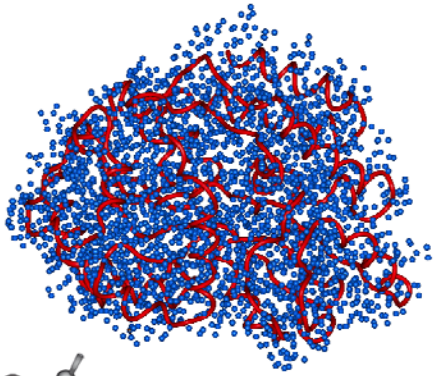
γ_H/γ_D = 6.5…. Relaxation gain !!!!

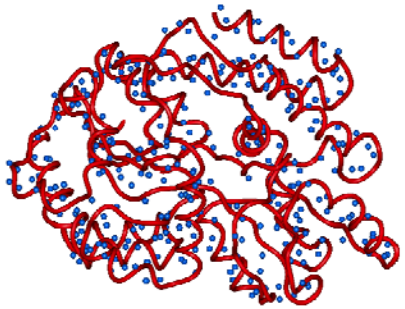
Nathalie Sibille - RéNaFoBis 2016

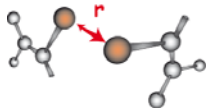
New restraints in NMR for « large » proteins

✧ **Fully protonated protein**

✧ **Deuterated protein (only $^{1}H_N$)**





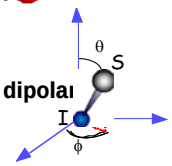


✧ **Classical method : NOE ^{1}H-^{1}H**

✧ **Ambigous NOEs restraints**

✧ **New approach : residual dipolar couplings (RDC)**

✧ **RDC non ambigous**



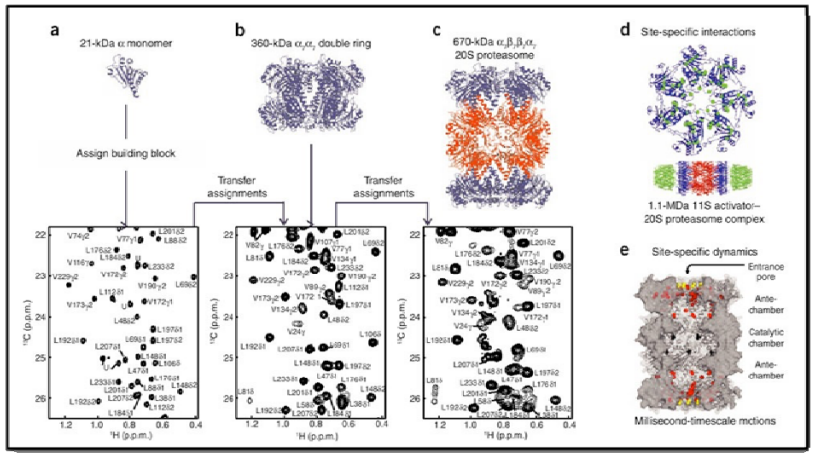
Nathalie Sibille - RéNaFoBis 2016

Towards high molecular weight complexes

Special interest of methyls :

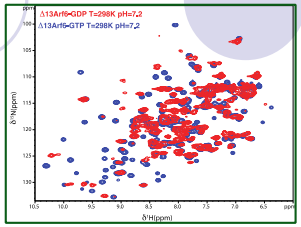
- well dispersed throughout the primary structures and localized in hydrophobic core
- favorable transversal relaxation property of CH3 (sensitivity gains)
- high sensitivity due to the presence of 3 equivalent protons

➲ Attribution des méthyls du protéasome archaebactérien (20S CP, 670kDa)
➲ Mise en évidence de mouvements concertés

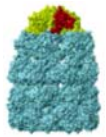


Remco Sprangers, Algirdas Velyvis & Lewis E Kay, Nat. Methods 2007 (Sept), 4, 697

NMR - Spatial and Time scales are highly intermingled



solubilized membrane proteins



"medium" size complexes

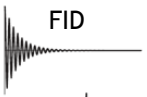
"medium" size globular proteins

"Small" globular Protein / domain

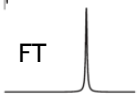
"large" size globular proteins or complexes

Global correlation time t_c (size)

Linewidth



FID



FT





+ Local Rotational fluctuation timescale
+ density of relaxation sources

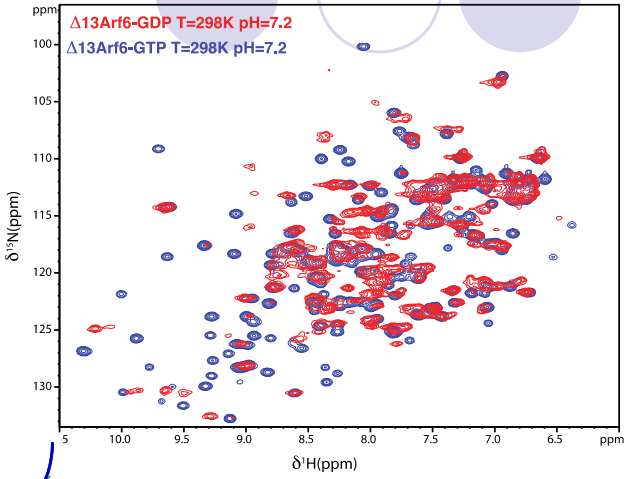
+ conformational/chemical exchange

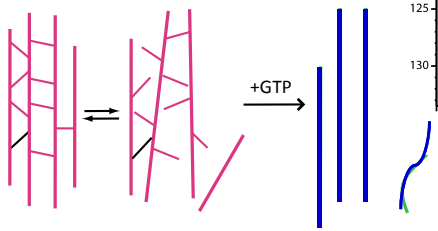
Nathalie Sibille - RéNaFoBis 2016

NMR - Spatial and Time scales are highly intermingled

Insight into the role of dynamics in the conformational switch of the small GTP-binding protein Arf1
Vanessa Buosi, Jean-Pierre Placial, Jean-Louis Leroy, Jacqueline Cherfils, Éric Guittet and Carine van Heijenoort*

JBC 2010





Δ17Arf1-GDP

DArf1-GTP

+ conformational /chemical exchange

Nathalie Sibille - RéNaFoBis 2016

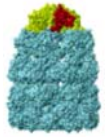
NMR - Spatial and Time scales are highly intermingled



medium size and large IDPs

TROSY, high fields

solubilized membrane proteins



Labelling, Deuteration

"medium" size complexes

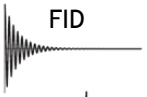
"medium" size globular proteins

"Small" globular Protein / domain

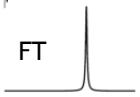
"large" size globular proteins or complexes

Global correlation time t_c (size)

Linewidth



FID



FT





+ Local Rotational fluctuation timescale
+ density of relaxation sources

+ conformational/chemical exchange

Nathalie Sibille - RéNaFoBis 2016

« Size » and disorder

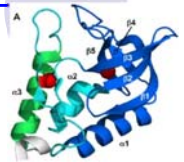


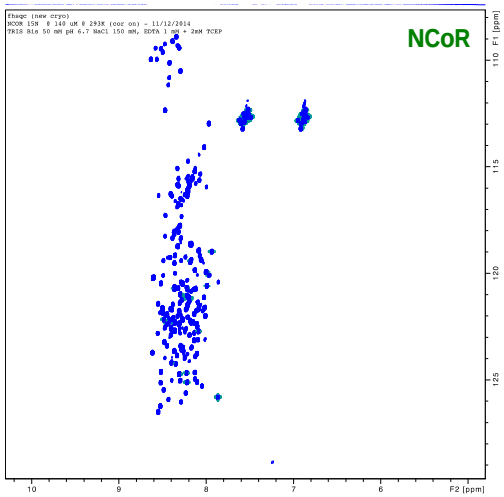
Unfolded protein

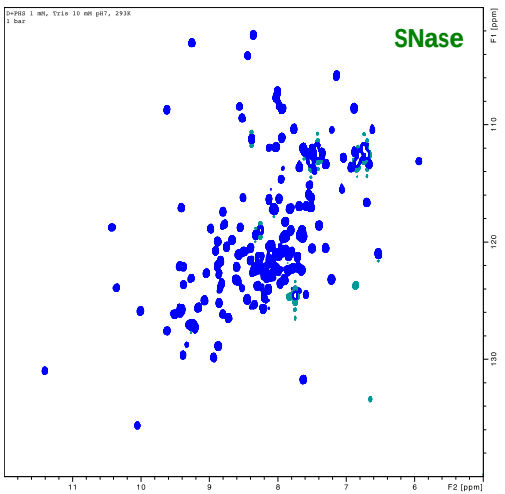
29 kDa

Folded protein

15 kDa





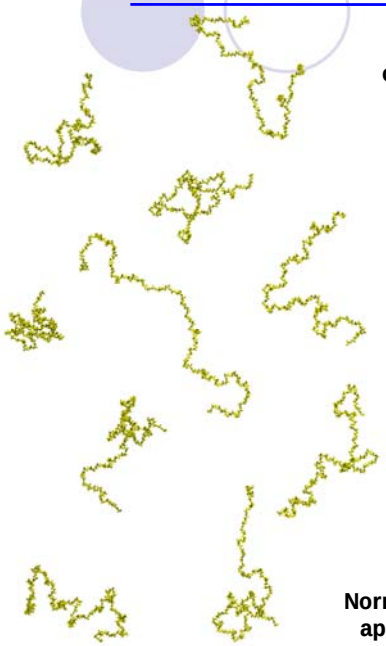


What is an IDP ?

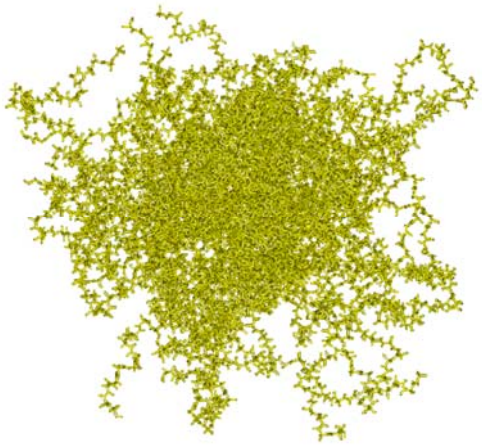
- IDPs lack stable tertiary and/or secondary structure
- Very specific amino-acid sequences… Rich in non-structuring residues and depleted in hydrophobic residues
- IDPs are more common in eukaryotes than in bacteria and archaea: Probably linked with their major biological complexity.
- Up to ~30-50% of genome in eukaryotic cells is predicted to code for natively disordered fragments (more that 30 residue fragments)
- Ordered and disordered proteins are associated to distinct biological functions
- Possibility to interact with several partners. Hubs in the interactome
- Involved in cancer, cardiovascular and neurodegenerative diseases

Chem. Rev. 2014 vol 114

Flexible Proteins, a Challenge for Structural Biology



IDPs explore an astronomical number of conformations that we assume are in fast equilibrium

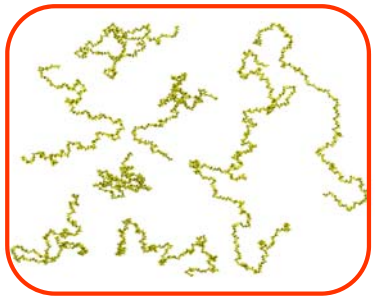


Normally averaged properties are measured. Ensemble approaches have to be used for their interpretation

Flexible Proteins, a Challenge for Structural Biology

AVERAGED

CS, RDCs, J-couplings, NOEs, PREs
SAXS, Hydrodynamic data, EPR



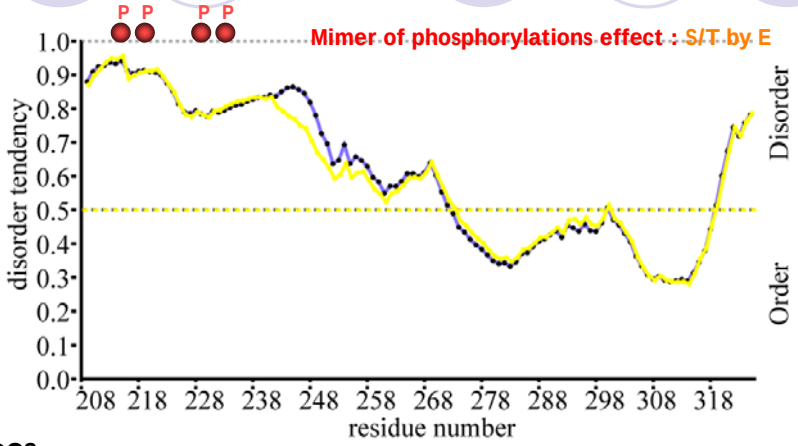
- The ensemble is underdetermined
  - Cross-validation or simplification of the structural model are required
- Structural content of the ensemble depends on the information (experimental data) introduced…
  - Residue-specific data ▸ local conformation
  - Overall data ▸ size and shape

Is our protein an IDP

- Bioinformatics
  - Identification of disordered/ordered regions
- Biophysical characterization: CD, FTIR, FRET, hydrodynamics
  - Partial Information
- Small-Angle X-ray Scattering (SAXS)/Small-Angle neutron Scattering (SANS)
  - Averaged Intensity profiles...
  - Qualitative Interpretation of averaged R_g and Kratky Plots
- Nuclear Magnetic Resonance (NMR)
  - Ensemble averaged observables: CS, J-Couplings and RDCs
  - Dynamic dependent Parameters: Relaxation Rates and PREs
- X-ray Crystallography
  - Structure determination in the bound form

BioInformatics

Predictors of disordered/ordered regions



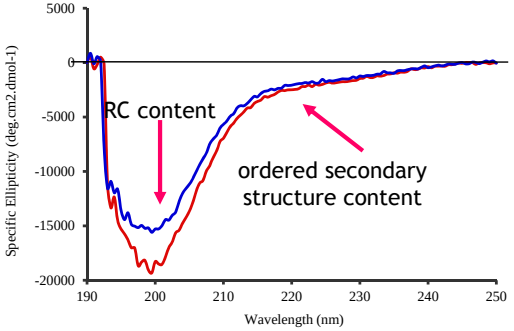
metaPrDOS integrates the results of different prediction methods
seven predictors: PrDOS (Ishida and Kinoshita, 2007), DISOPRED2 (Ward *et al.*, 2004), DisEMBL (Linding *et al.*, 2003), DISPROT (VSL2P) (Peng *et al.*, 2006), DISpro (Cheng *et al.*, 2005), IUpred (Dosztanyi *et al.*, 2005b) and POODLE-S (Shimizu *et al.*, 2007)

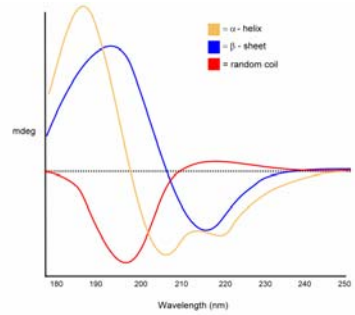
"DisProt: the Database of Disordered Proteins" Nucleic Acids Res. 2007 Jan;35 (Database issue):D786-93. Epub 2006 Dec 1.

Nathalie Sibille - RéNaFoBis 2016

Circular Dichroism

Partial/Global Information





- random coil polymer : broad minimum at 195-200 nm
- less pronounced minimum at 200 nm suggesting a reduced random coil contribution

Low level of secondary structure
$\theta_{220}/\theta_{200}$

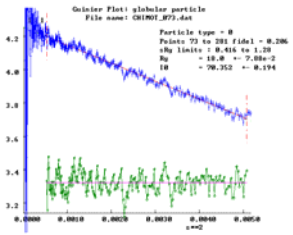
- positive (p -> p*) at 192 nm
- negative (p -> p*) at 209 nm
- negative (n -> p*) at 222 nm
- negative at 218 nm (p -> p*)
- positive at 196 nm (n -> p*)
- positive at 212 nm (p -> p*)
- negative at 195 nm (n -> p*)

Nathalie Sibille - RéNaFoBis 2016

SAXS/SANS

Qualitative interpretation of averaged Intensity Profiles

Guinier

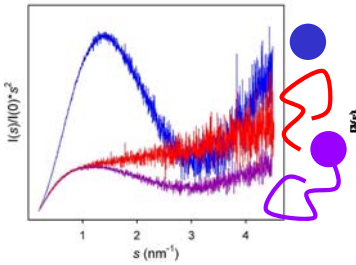


R_g , Molecular Weight

Flory's equation was parametrized for IDPs

$R_g = (2.54 \pm 0.01) \cdot N^{(0.522 \pm 0.01)}$

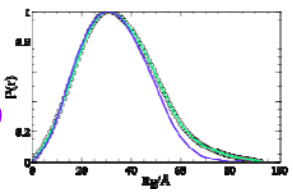
Kratky



Mass Density

- **folded**
- **unfolded**
- **folded and unfolded**

P(r)

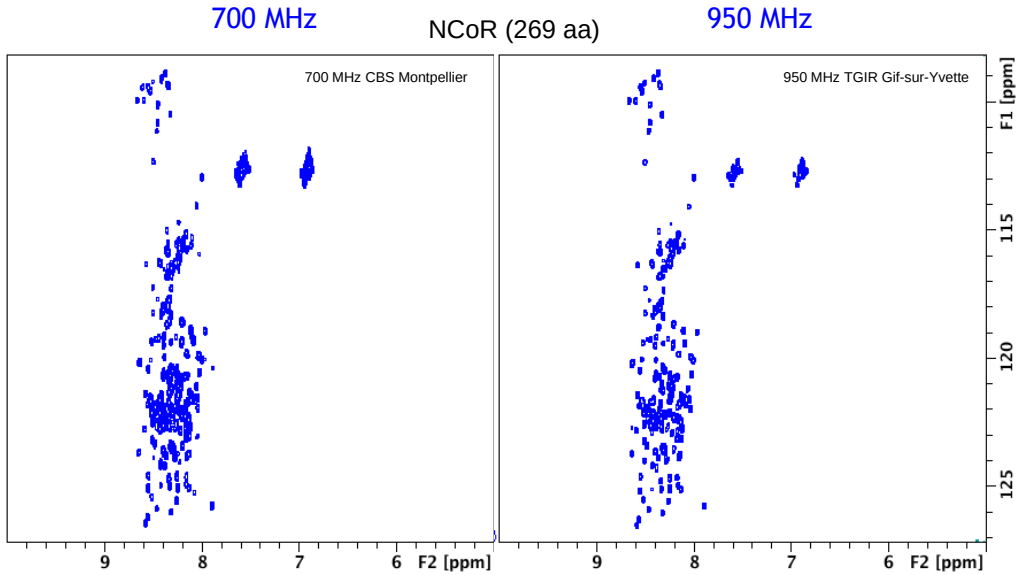


1D-Structure
Maximum Distance

Nathalie Sibille - RéNaFoBis 2016

NMR

Sensitivity and Resolution



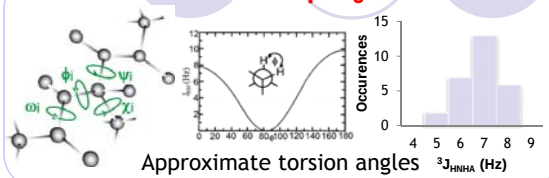
NMR is rich in structural information

Chemical Shifts

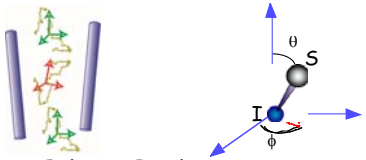


Secondary structure elements

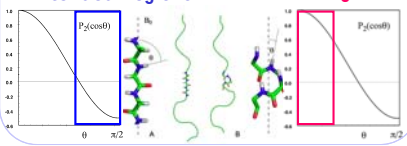
J-Couplings



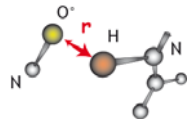
Residual Dipolar Couplings



Relative Bond orientations

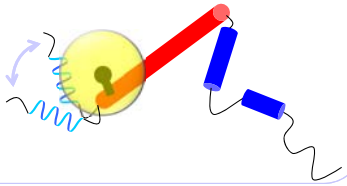


Paramagnetic Relaxation Enhancement



$$PRE \propto \frac{1}{r_{OH}^6}$$

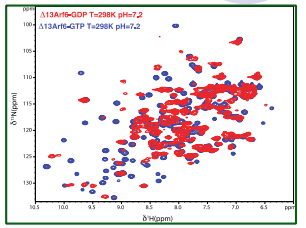
Transient long-range contacts



NMR - Spatial and Time scales are highly intermingled

medium size and large IDPs





"medium" size complexes



solubilized membrane proteins



"large" size globular proteins or complexes

"Small" globular Protein / domain

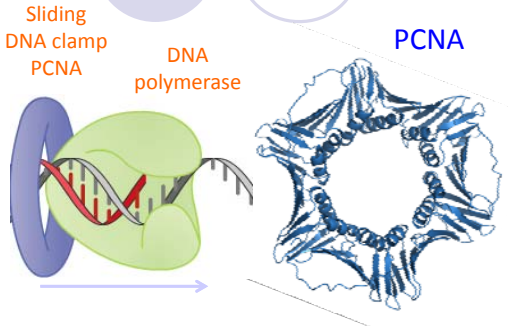
"medium" size globular proteins

Global correlation time t_c (size) → Linewidth

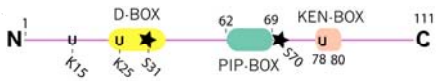
+ Local Rotational fluctuation timescale
+ density of relaxation sources

+ conformational/chemical exchange

p15^{PAF} a PCNA-Binding Protein



P15



- ✓ PCNA is a trimeric ring that can accommodate a DNA duplex through its central channel
- ✓ PCNA acts as a processivity factor for DNA polymerase ε

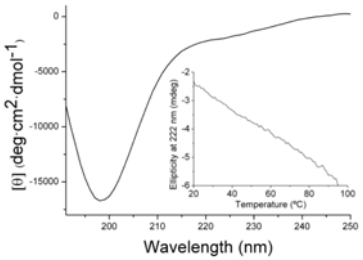
- ✓ P15 acts as a regulator of DNA repair
- ✓ is over-expressed in several types of human cancer
- ✓ It binds PCNA through its PIP-BOX
- ✓ Targeted for degradation by the ubiquitin ligase APC through its conserved KEN-BOX.

ding Clamp

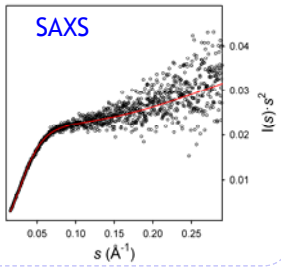
Col. Francisco Blanco (CIC-BioGune-Bilbao)

p15^{PAF} is an IDP ...

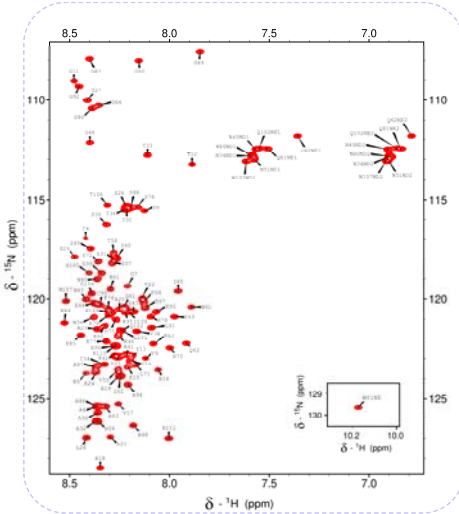
CD



SAXS

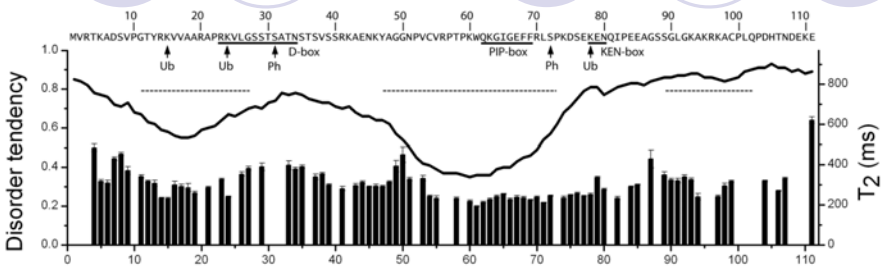


NMR



De Biasio *et al. Biophys J.* 2014

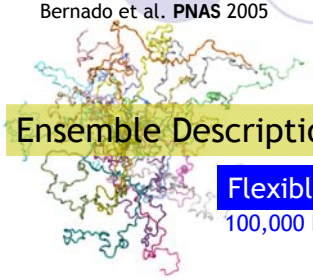
p15^{PAF} is an IDP … but not everywhere



- ✓ Disorder prediction is consistent with p15 being largely unstructured, with the exception of residues of the highly conserved PIP-box motif.
- ✓ p15 shows several relatively short T_2 values, that correspond to sequences with reduced predicted disorder.

Transient structuration monitored by RDCs

Bernado & Blackledge **Nature** 2010
Bernado et al. **PNAS** 2005

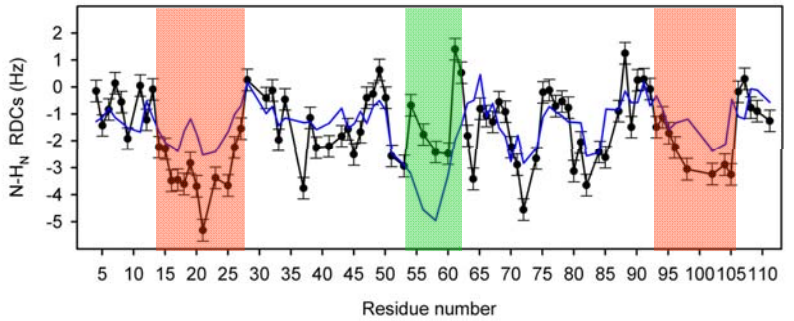


Ensemble Description

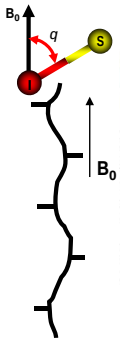
Flexible-Meccano
100,000 RC structures

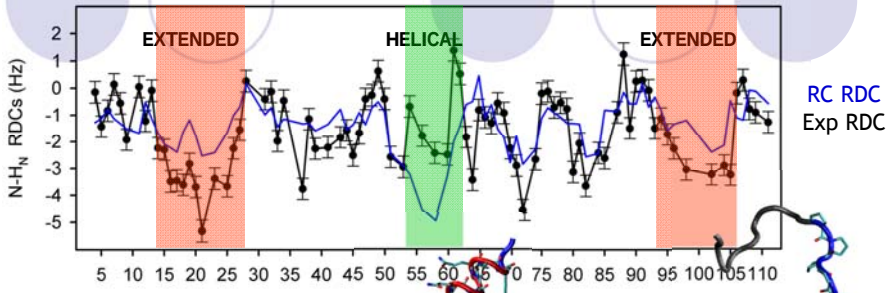
Disordered but, according to the RDCs, it presents several sites with partial structuration
FM Random Coil model helps in identifying the nature of the structural elements observed

Conformational Preferences
Secondary Structural Element : Position and Population

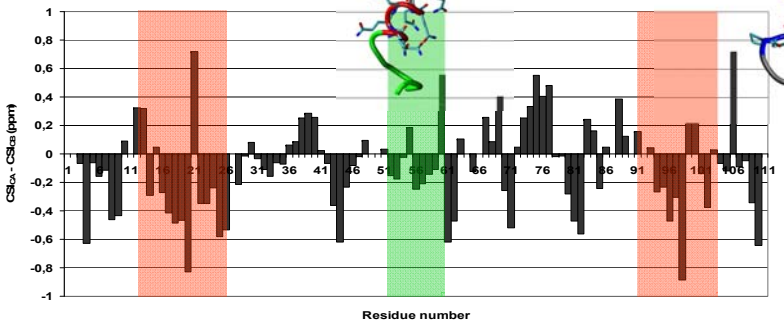


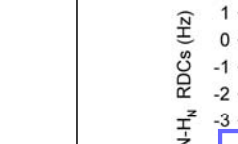
Transient structuration monitored by RDCs



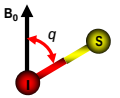


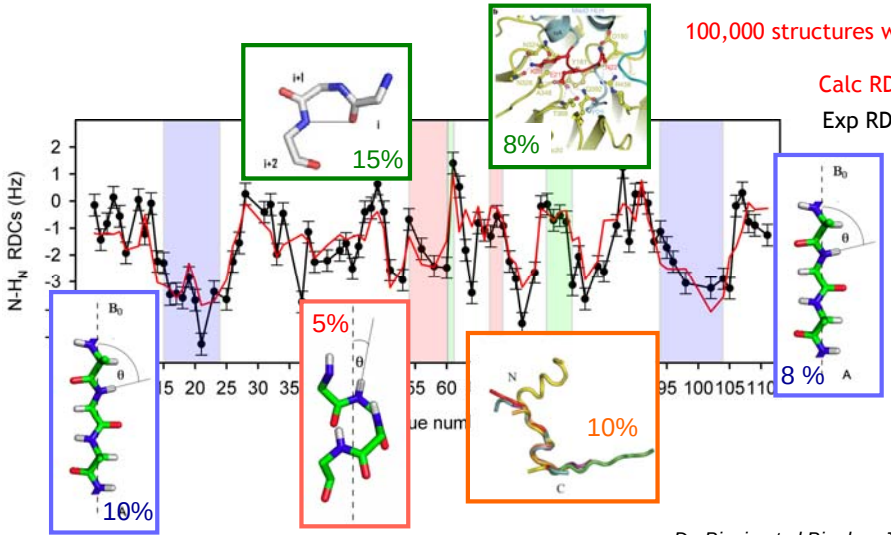
CSI with Mulder's database





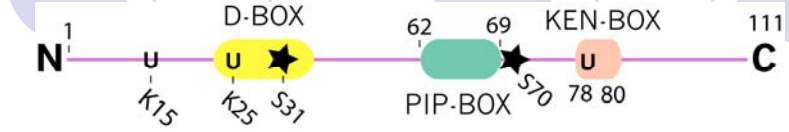
Structural Optimization

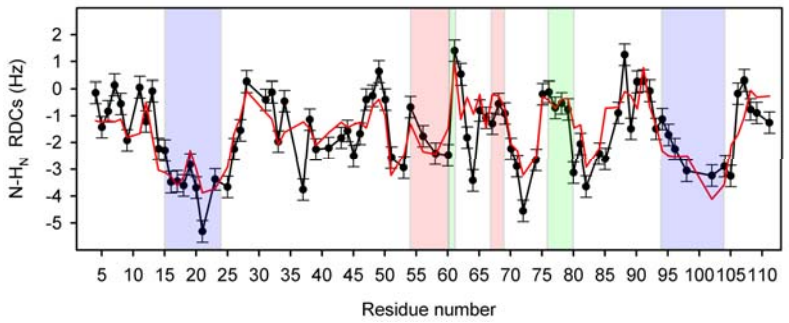




De Biasio et al Biophys J. 2014
De Biasio et al Nat. Comm 2015

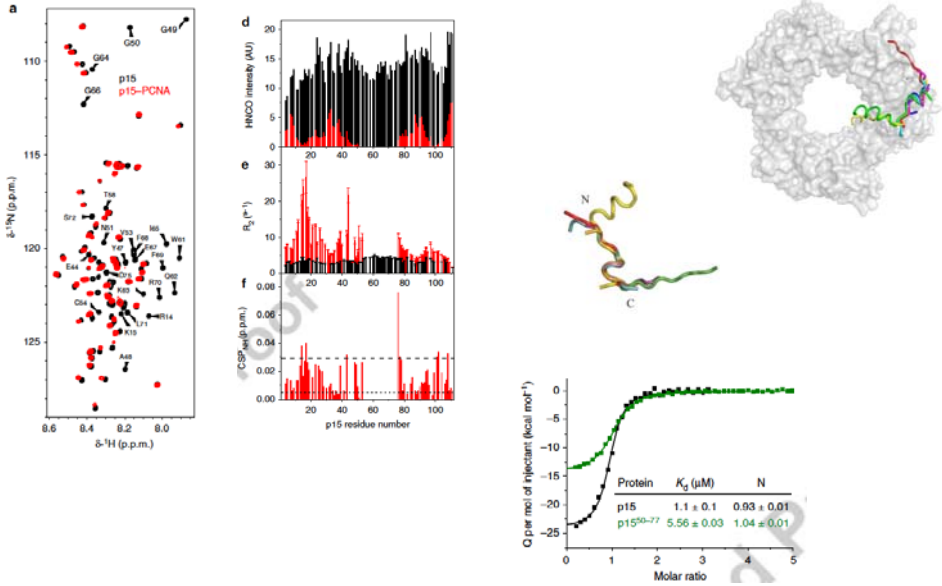
p15^{PAF} is an IDP with partial Structuration at Interaction Sites





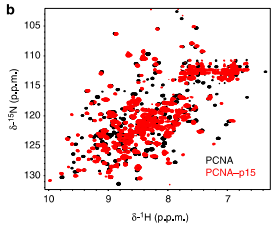
De Biasio A, Ibáñez de Opakua A, Cordeiro TN, Villate M, Merino N, Sibille N, Lelli M, Diercks T, Bernadó P, Blanco FJ. p15PAF is an intrinsically disordered protein with nonrandom structural preferences at sites of interaction with other proteins. Biophys J. 2014 Feb 18;106(4):865-74

Interaction of p15^{PAF} with PCNA

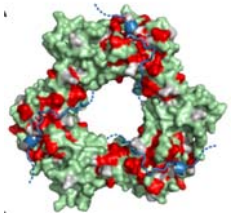


Interaction of p15^{PAF} with PCNA

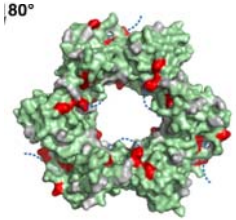
- ✓ Mapping of p15 binding on PCNA by NMR



PCNA front-side

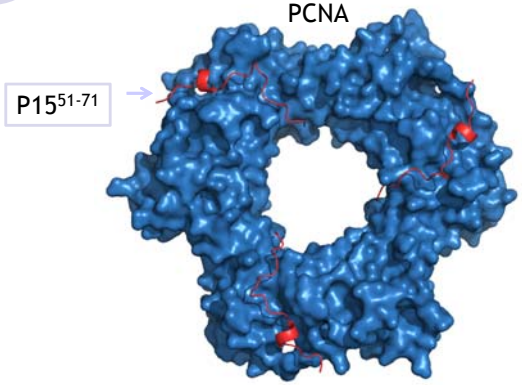


PCNA back-side



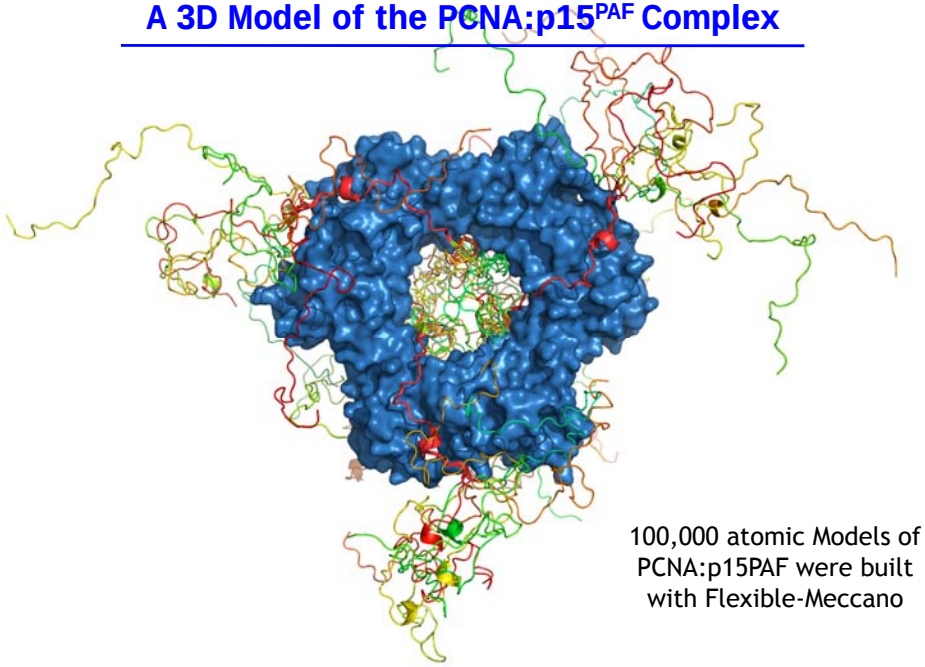
De Biasio *et al. Nature Comm.* 2015

Crystallographic Structure of p15^{PAF} with PCNA



PDB:4D2G

A 3D Model of the PCNA:p15^{PAF} Complex



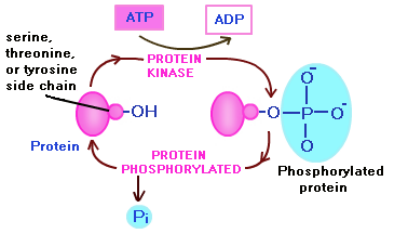
100,000 atomic Models of PCNA:p15PAF were built with Flexible-Meccano

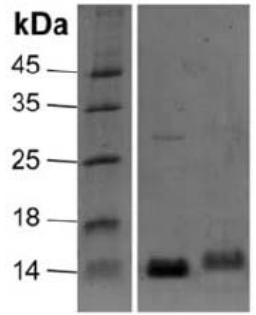
What else ?

Effect of Post-translational Modifications

Even with recombinant kinases, mass spectrometry and immunodetection are not evident for determining the full phosphorylation pattern in a qualitative and quantitative manner.

- ✧ MS: highly charged phospho-peptide
- ✧ Immunodetection by antibodies is equally limited by the requirement of a comprehensive antibody library against all possible epitopes and by the absence of a fully phosphorylated standard for every combination of sites in order to quantify the level of phosphorylation

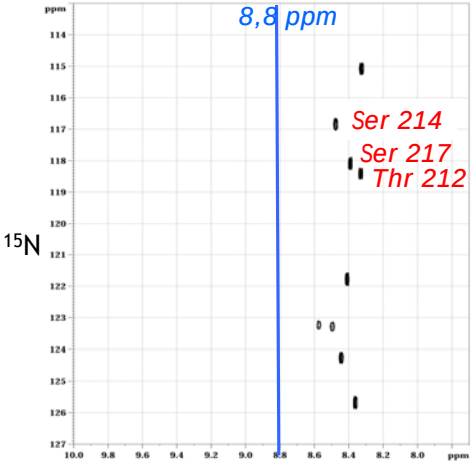




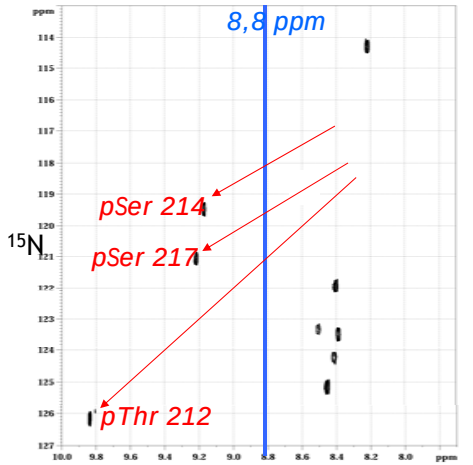
NMR is Qualitative and Quantitative

What sites? and to what extent?

➔ *In vitro* phosphorylation by recombinant kinases: Identification and quantification



SRSRTPSLPSPPTR

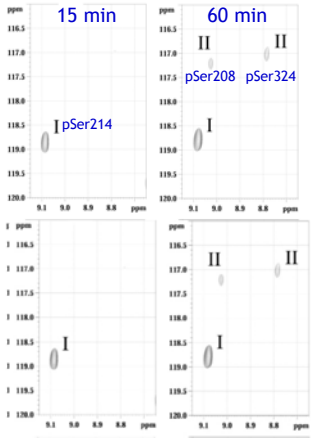


SRSRpTPpSLPpSPPTR

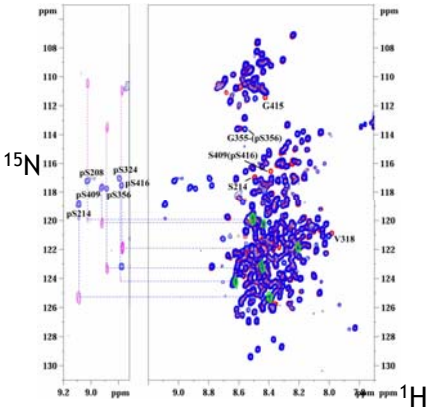
Kinetic of Phosphorylation

cAMP protein dependant kinase PKA

PKA from Prof. Langer/Schwalbe, Frankfurt, Germany



Sequential kinetic



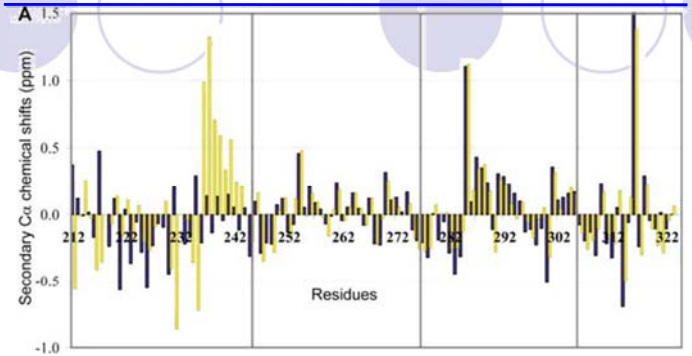
Phosphorylated residues assignment (triple resonance NMR experiments)

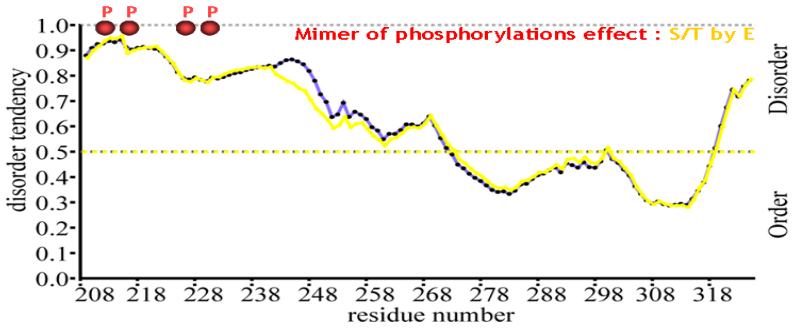
Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, N. Sibille, Schwalbe H, Saxena K, Langer T, Lippens G. *NMR analysis of a Tau phosphorylation pattern.* J Am Chem Soc 128:3575-83, 2006

Is there some degree of structuration ?

Linked to recognition events

Here linked to a loss of function





Sibille et al. **Proteins** *(2012)* 80, 454-462

Where and which kind ?

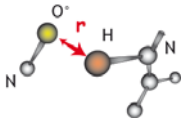
Is there some transient long range structuration ?

Dynamic dependant parameters

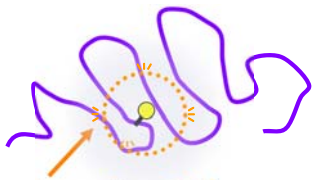
- **Paramagnetic Relaxation Enhancement (PRE)**

$$R_2 \propto \frac{\gamma_S^2 \gamma_e^2}{r^6} f(\tau_c, \tau_s \tau_i)$$

$$\frac{\gamma_e}{\gamma_H} = 658$$

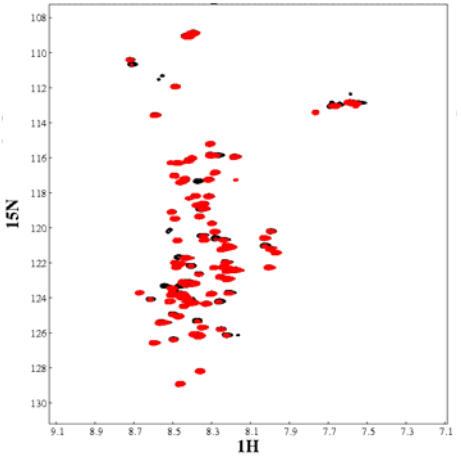


$$PRE \propto \frac{1}{r_{OH}^6}$$



Forme para I : **NO(C291) 200 uM**

Forme dia I0: **+ 5mM de vitamine C**



Is there some transient long range structuration ?

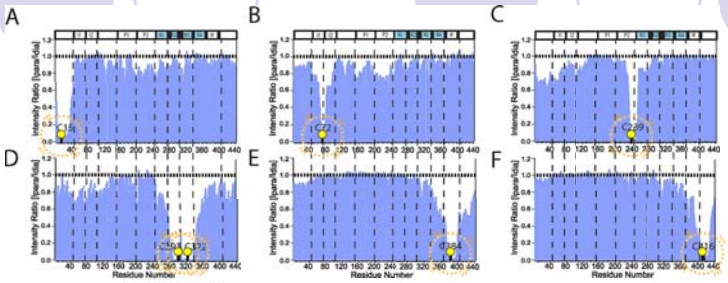
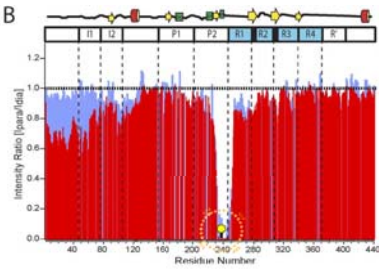
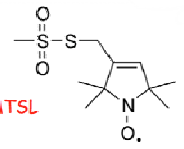
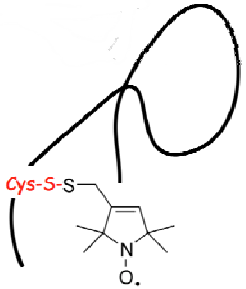


Figure 7. PRE of Amide Protons in Spin-Labelled Tau

Buffer 600 mM NaCl

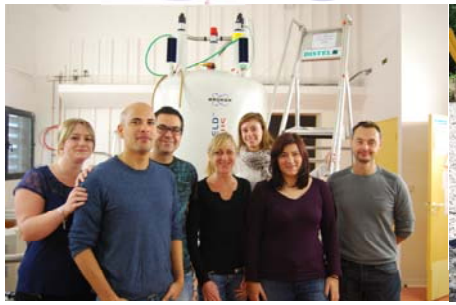






Marco D. Mukrasch, Stefan Bibow, Jegannath Korukottu, Sadasivam Jeganathan, Jacek Biernat, Christian Griesinger, Eckhard Mandelkow, Markus Zweckstetter* (2009) Structural Polymorphism of 441-Residue Tau at Single Residue Resolution. PLoS Biology, 7(2), 399

Highly Flexible Proteins Team





Pau Bernadò
Nathalie Sibille
Tiago Cordeiro
Fatima Herranz-Trillo
Frederic Allemand
Aurelie Fournet
Anna Morató
Annika Urbanek