

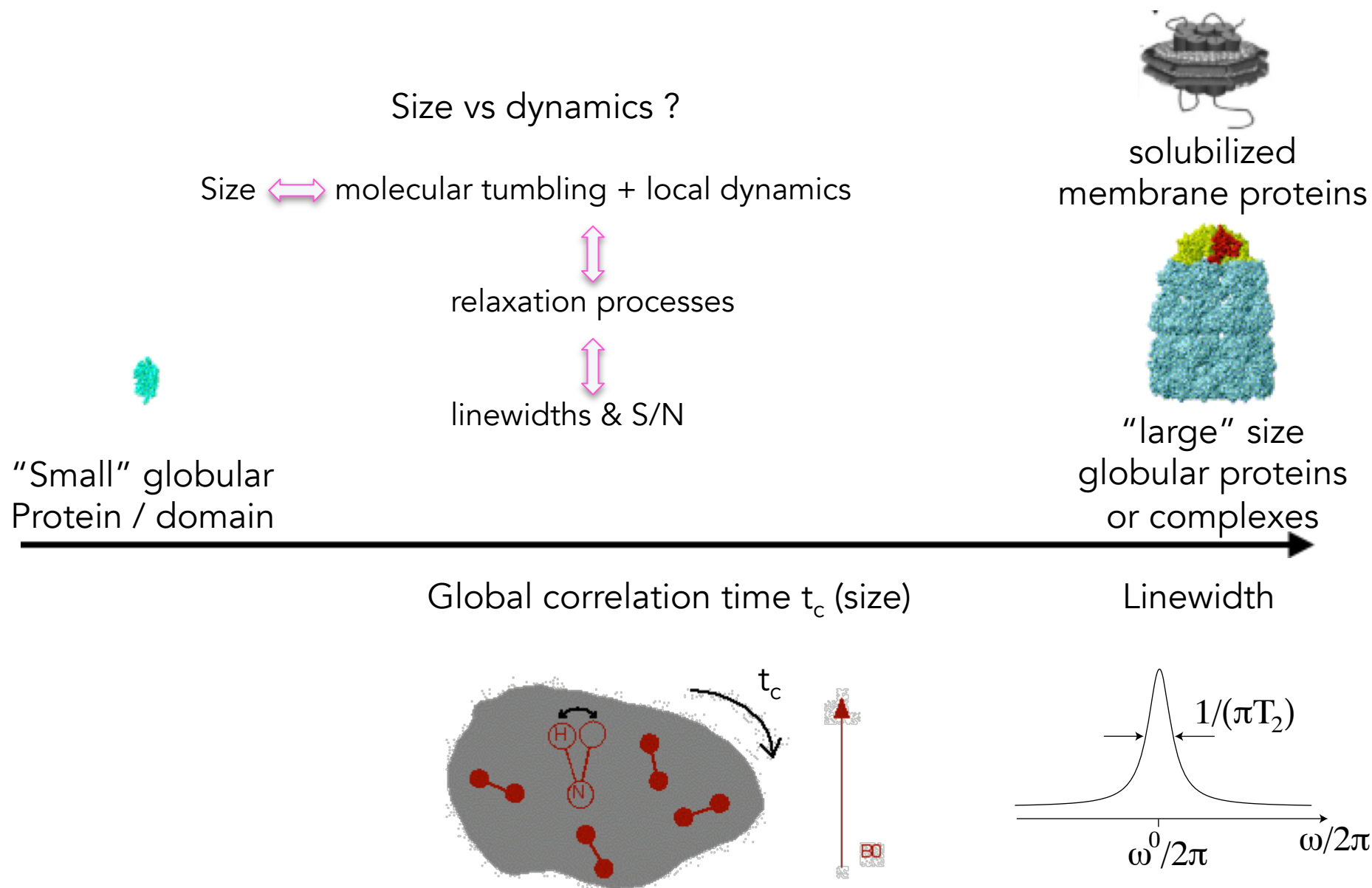
Contributions de la RMN à la biologie structurale : Approches multi-échelles spatiales et temporelles



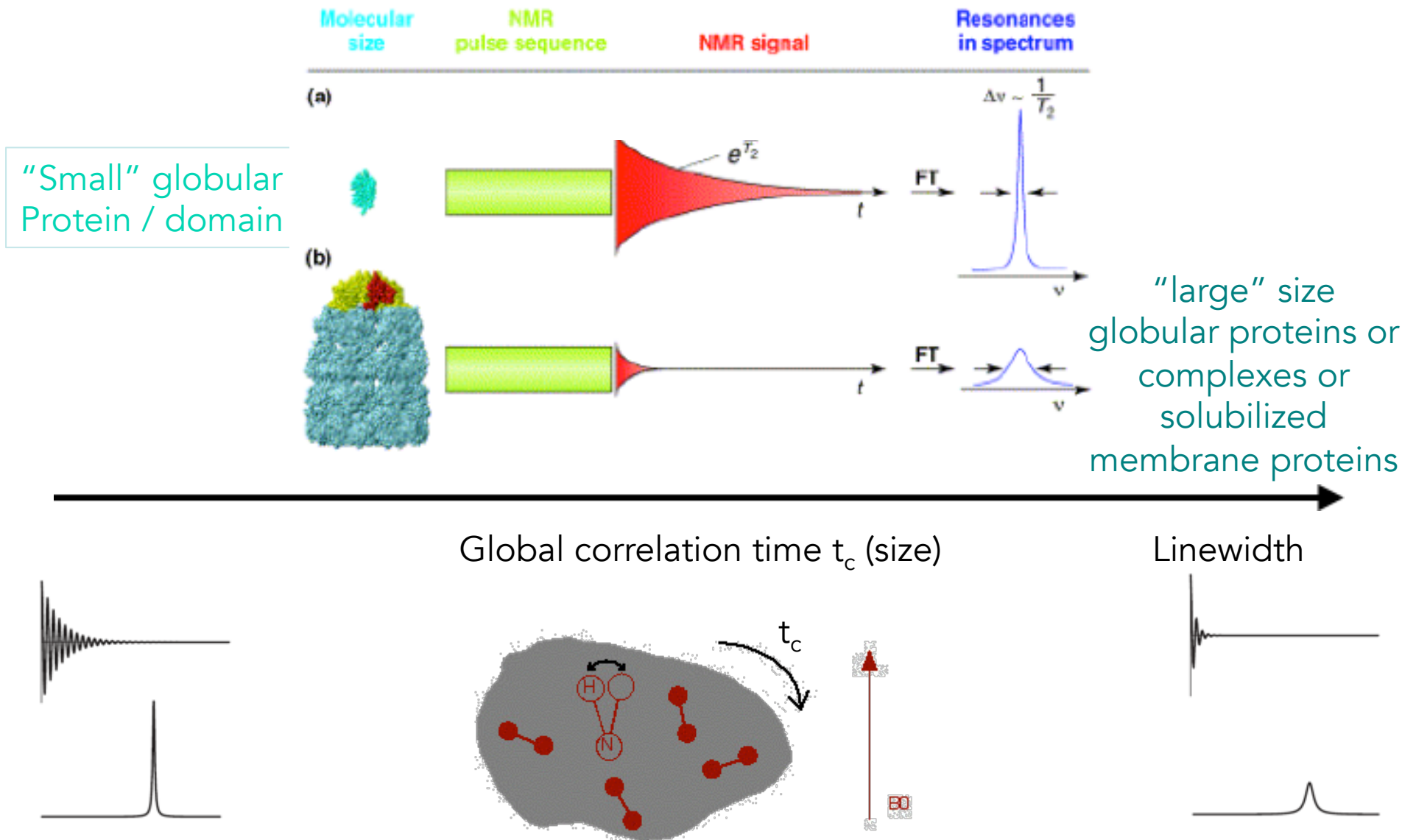
- Nathalie SIBILLE
- CNRS UMR 5048, CBS
- nathalie.sibille@cbs.cnrs.fr



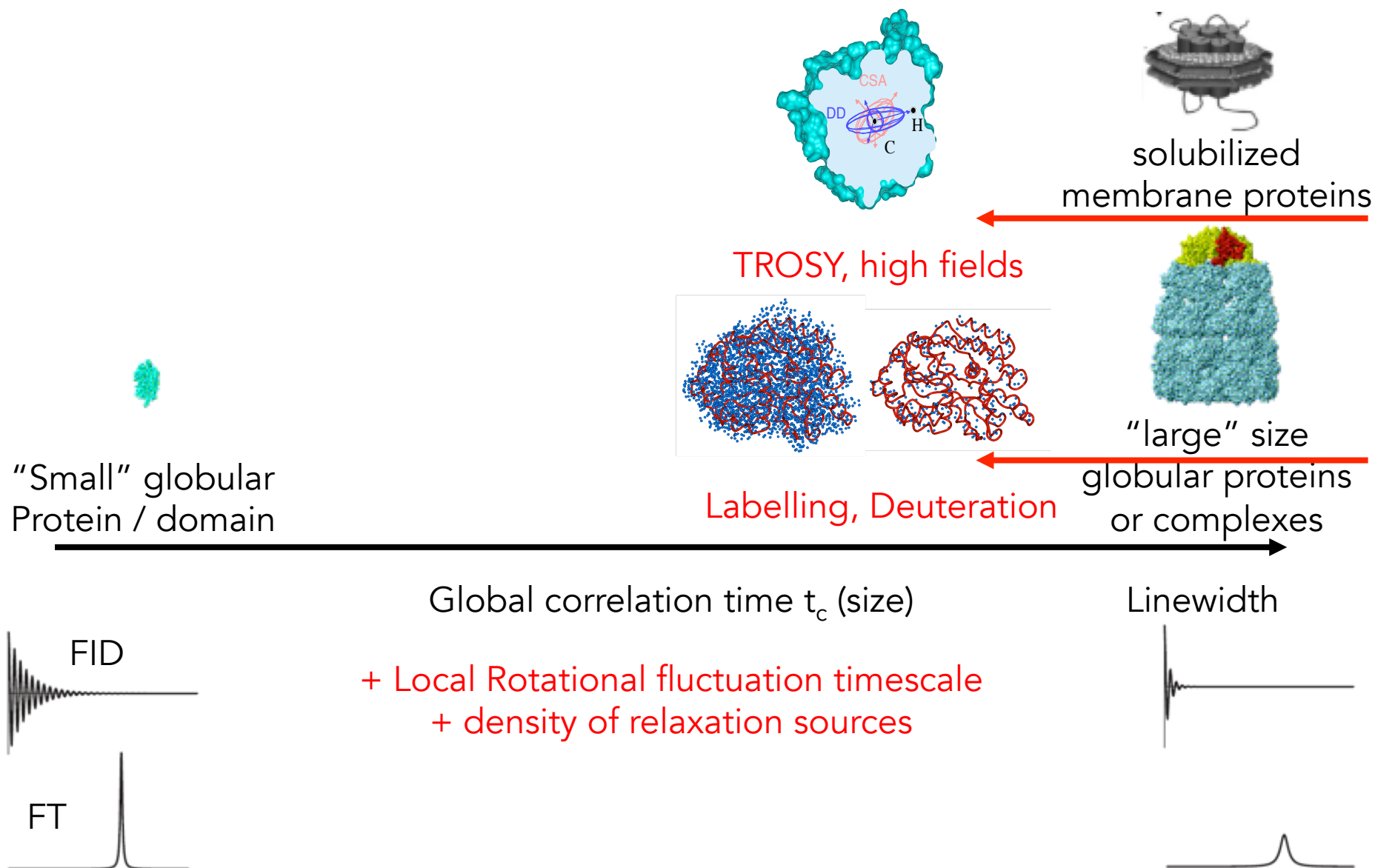
NMR - Spatial and Time scales are highly intermingled



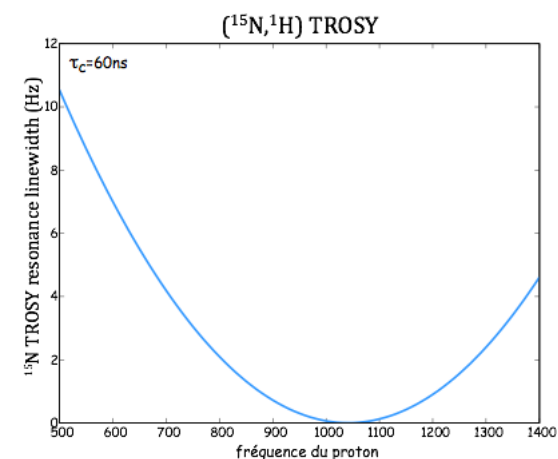
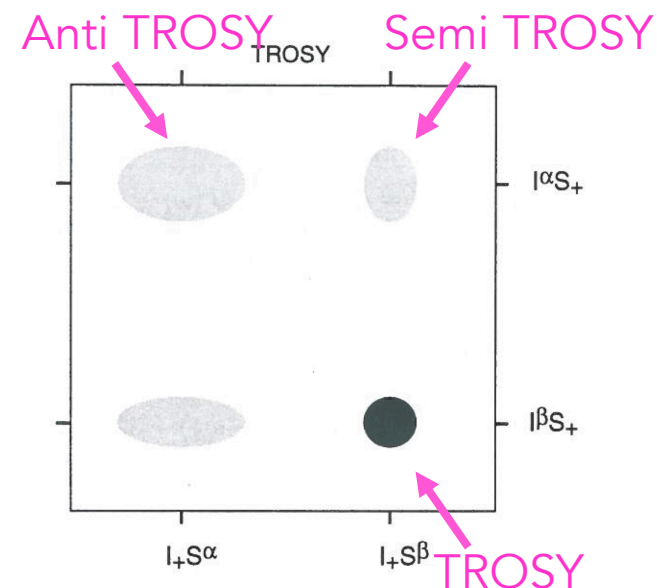
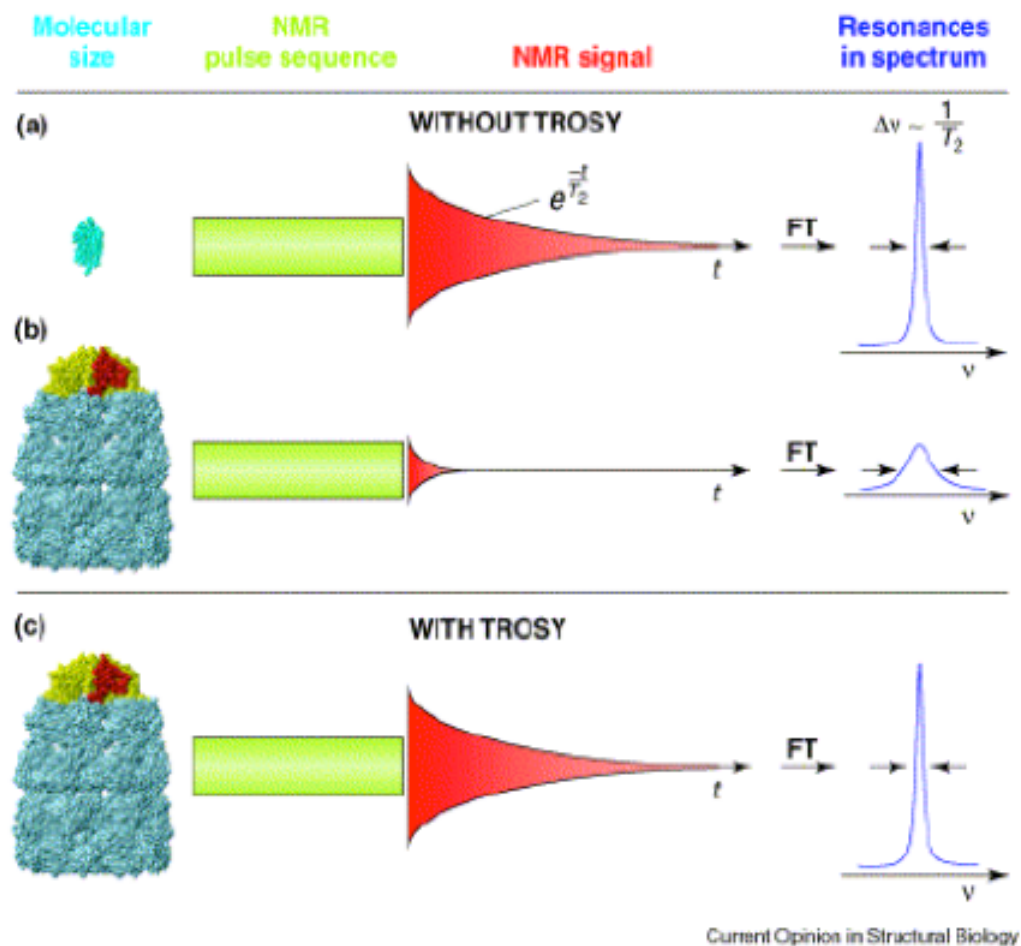
NMR - Spatial and Time scales are highly intermingled



NMR - Spatial and Time scales are highly intermingled

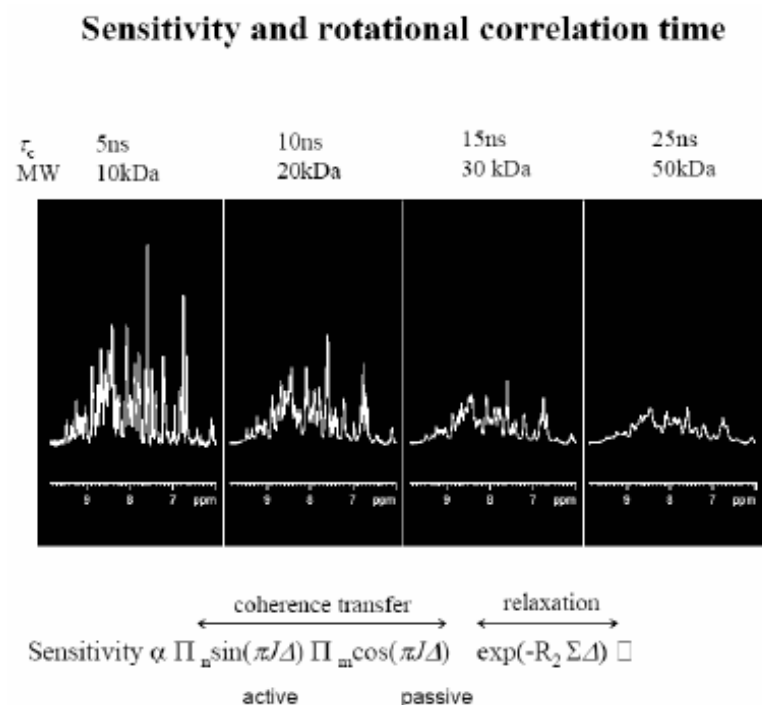


« Size » and lifetime of the signal T2



Playing with T2 : TROSY experiments
Transverse Relaxation Optimized Spectroscopy

« Size » and lifetime of the signal T2



- MW < 5 kDa

Homonuclear assignment (TOCSY, NOESY)

- 5 kDa < MW < 10 kDa

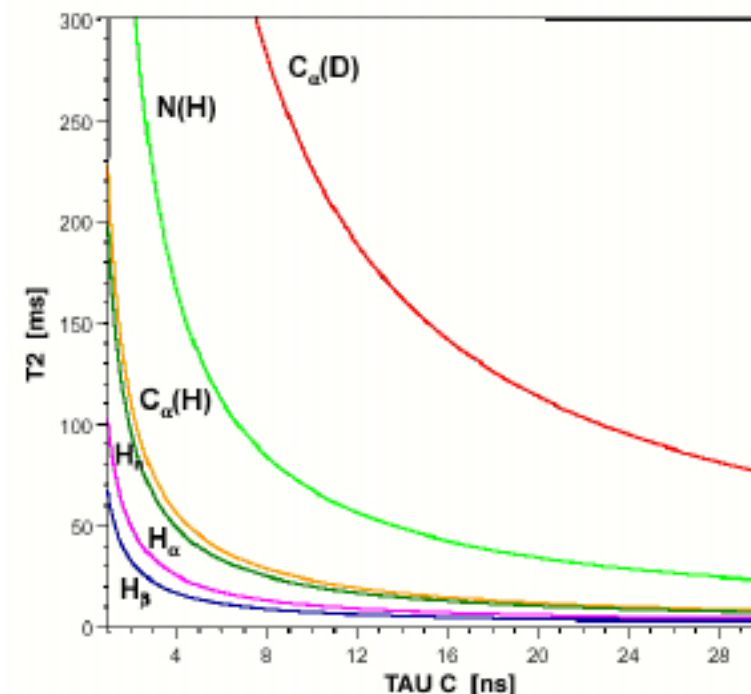
Heteronuclear assignment: ^{15}N

- 10 kDa < MW < 20 kDa

Heteronuclear triple resonance assignment : ^{15}N , ^{13}C

- 20 kDa < MW

Heteronuclear triple resonance assignment : ^{15}N , ^{13}C , ^2H (E.coli BL21+++)



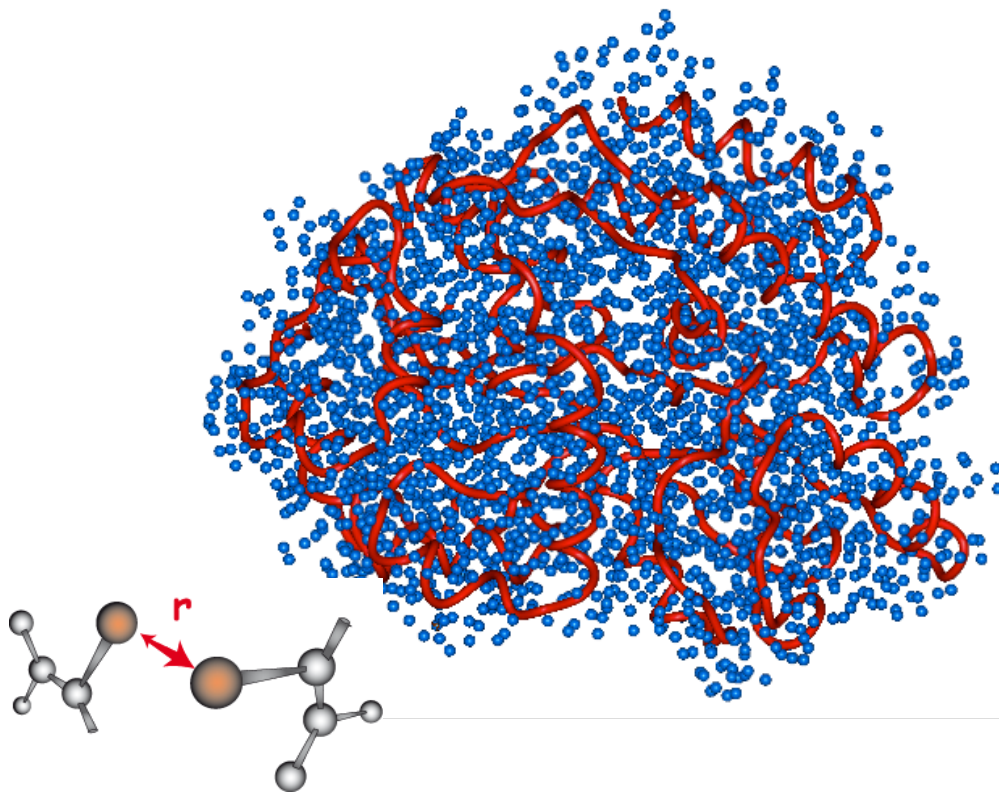
The signal loss is exponential while the size of the molecule increases

$\gamma_{\text{H}} / \gamma_{\text{D}} = 6.5 \dots$ Relaxation gain !!!!

New restraints in NMR for « large » proteins

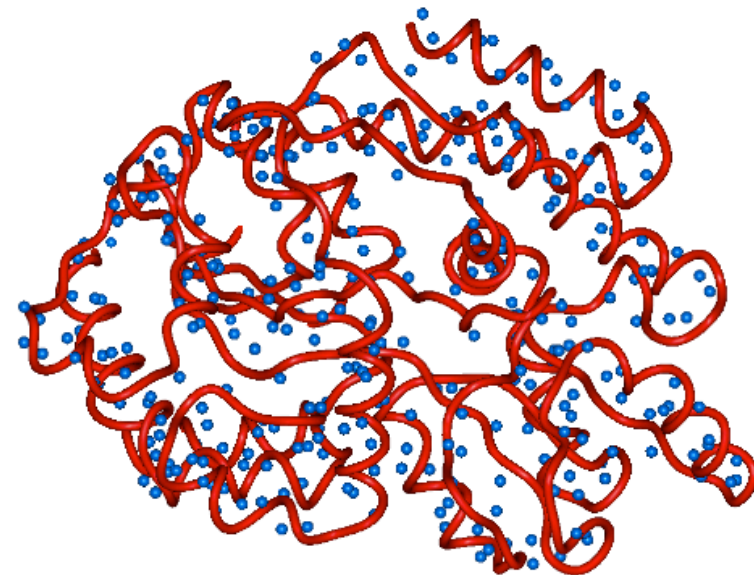
✧ Fully **protonated** protein

✧ **Deuterated** protein (only $^1\text{H}_\text{N}$)



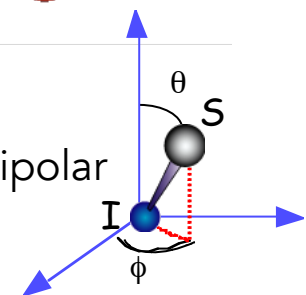
✧ Classical method : NOE ^1H - ^1H

✧ Ambiguous NOEs restraints



✧ New approach : residual dipolar couplings (RDC)

✧ RDC non ambiguous



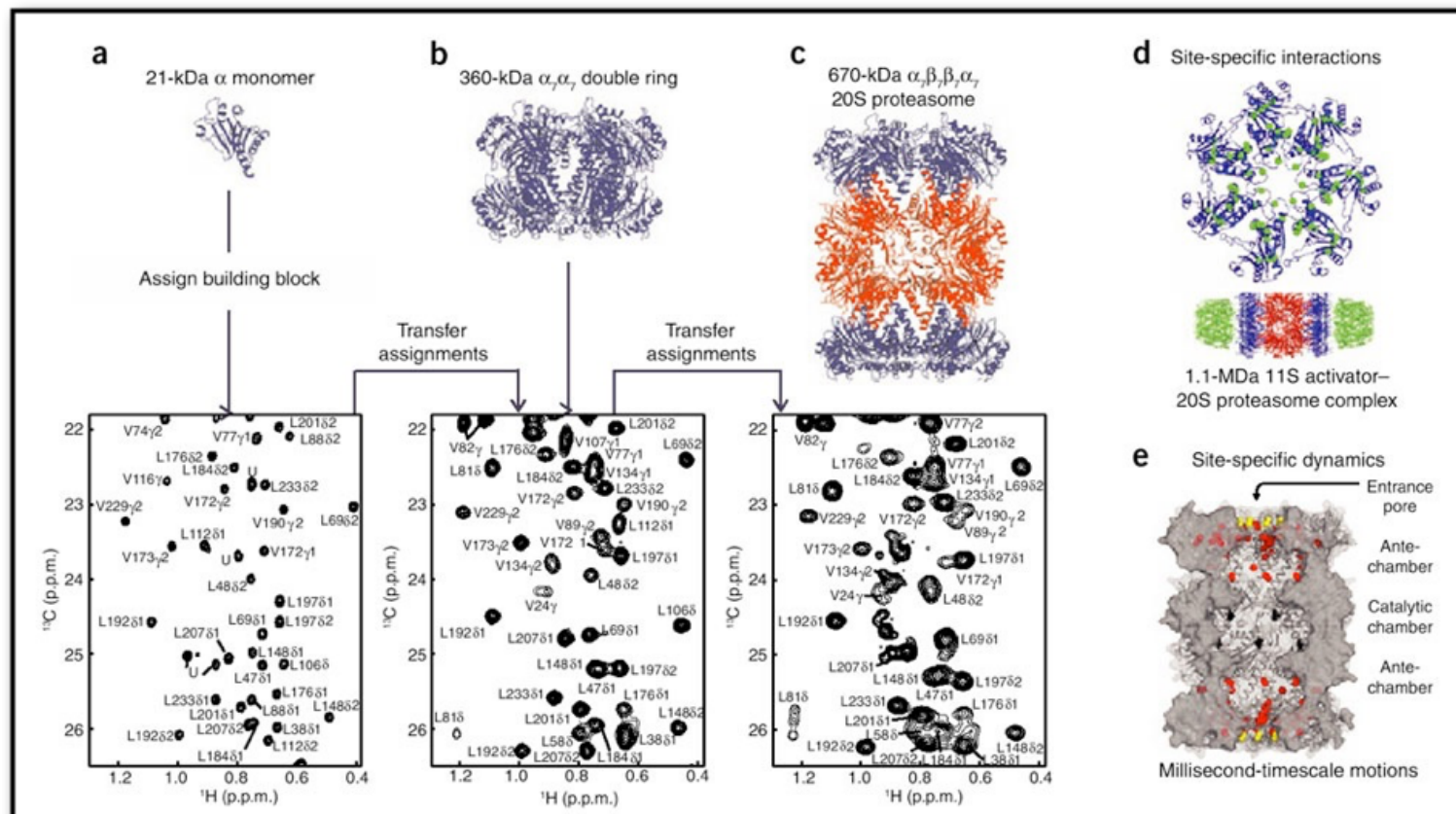
Towards high molecular weight complexes

Special interest of methyls :

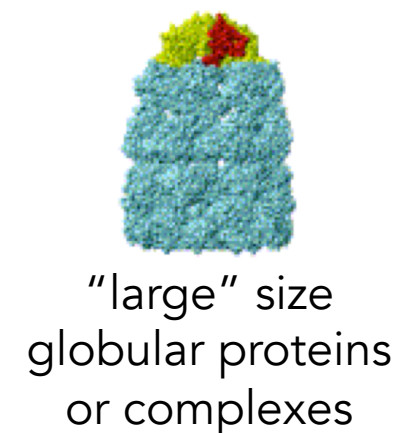
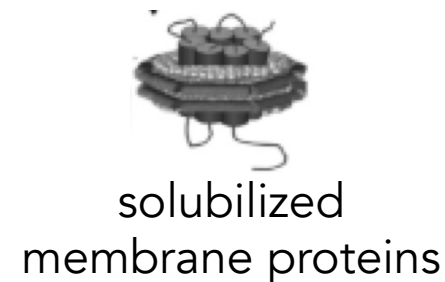
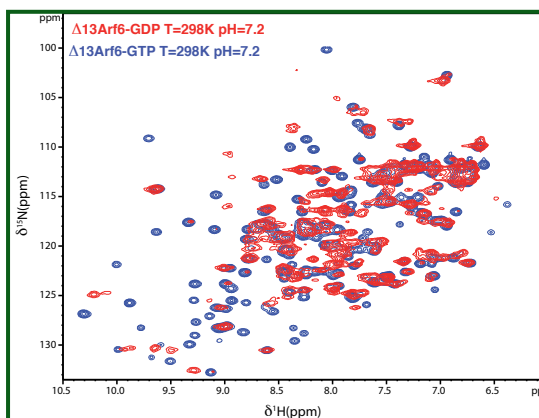
- well dispersed throughout the primary structures and localized in hydrophobic core
- favorable transversal relaxation property of CH₃ (sensitivity gains)
- high sensitivity due to the presence of 3 equivalent protons

⇒ Archaeon proteasome methyl assignment (20S CP, 670kDa)

⇒ Methyl groups with concerted millisecond-timescale motions



NMR - Spatial and Time scales are highly intermingled



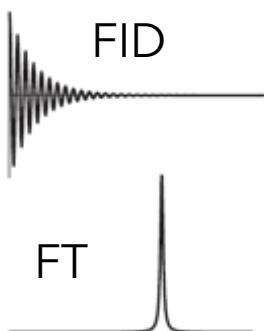
"Small" globular
Protein / domain

"medium" size
complexes

"medium" size
globular proteins

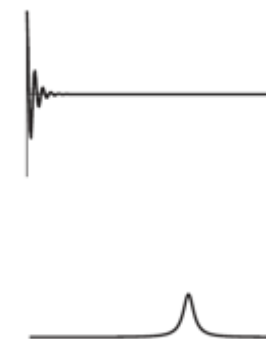
Global correlation time t_c (size)

Linewidth



+ Local Rotational fluctuation timescale
 + density of relaxation sources

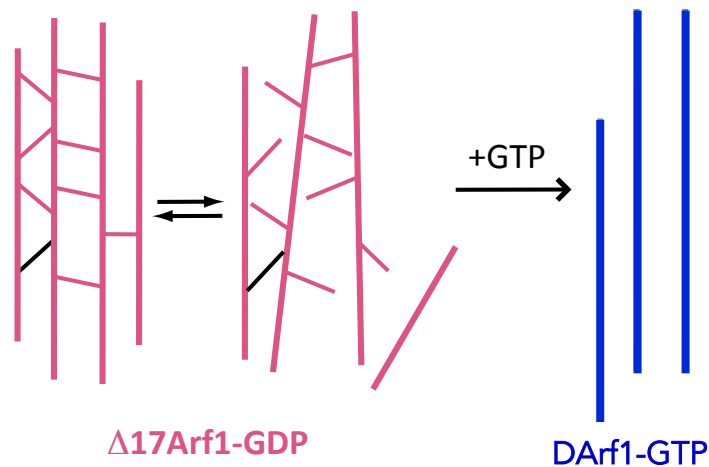
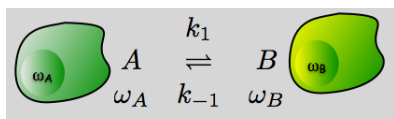
+ conformational/chemical exchange



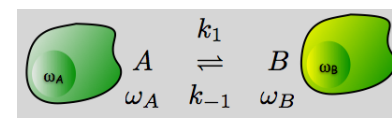
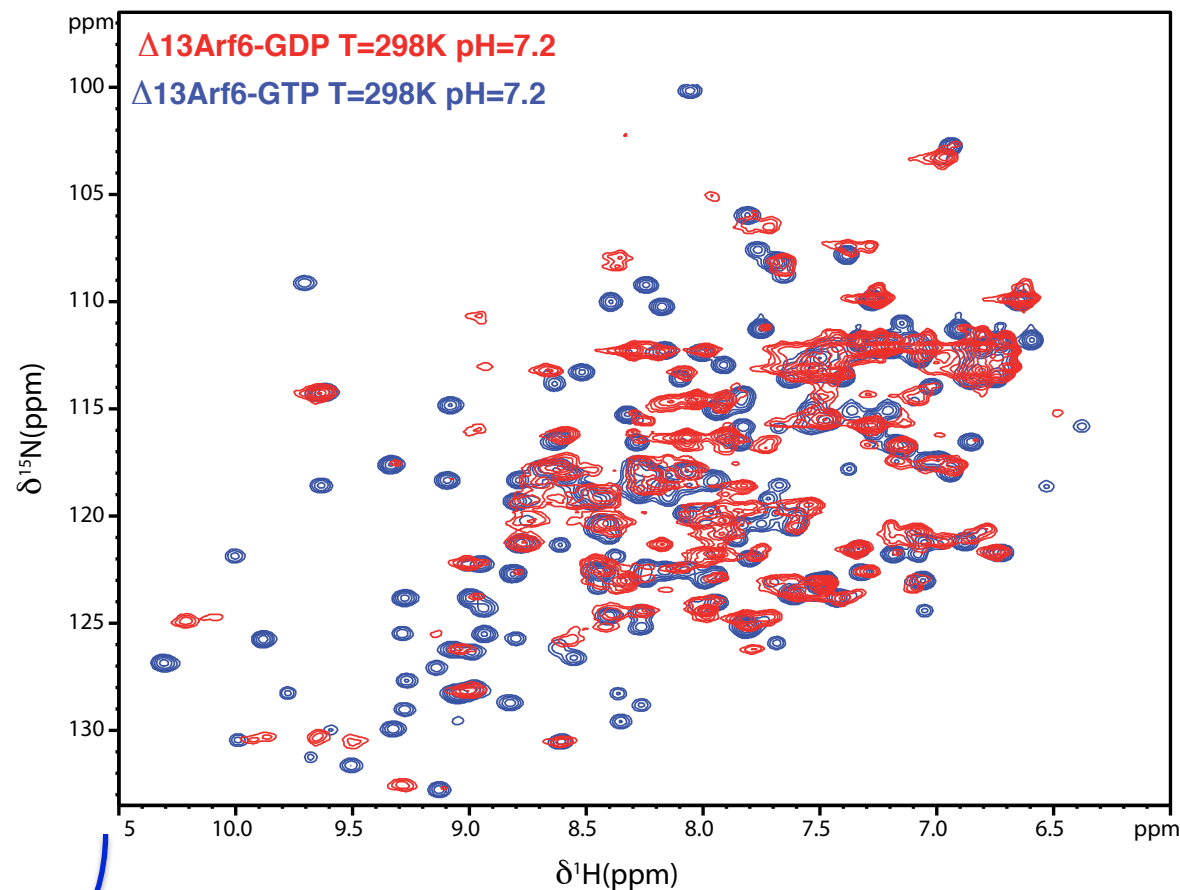
NMR - Spatial and Time scales are highly intermingled

Insight into the role of dynamics in the conformational switch of the small GTP-binding protein Arf1

Vanessa Buosi, Jean-Pierre Placiat, Jean-Louis Leroy, Jacqueline Cherfils, Éric Guittet and Carine van Heijenoort*
JBC 2010



+ conformational /
chemical exchange



“Slow”
exchange

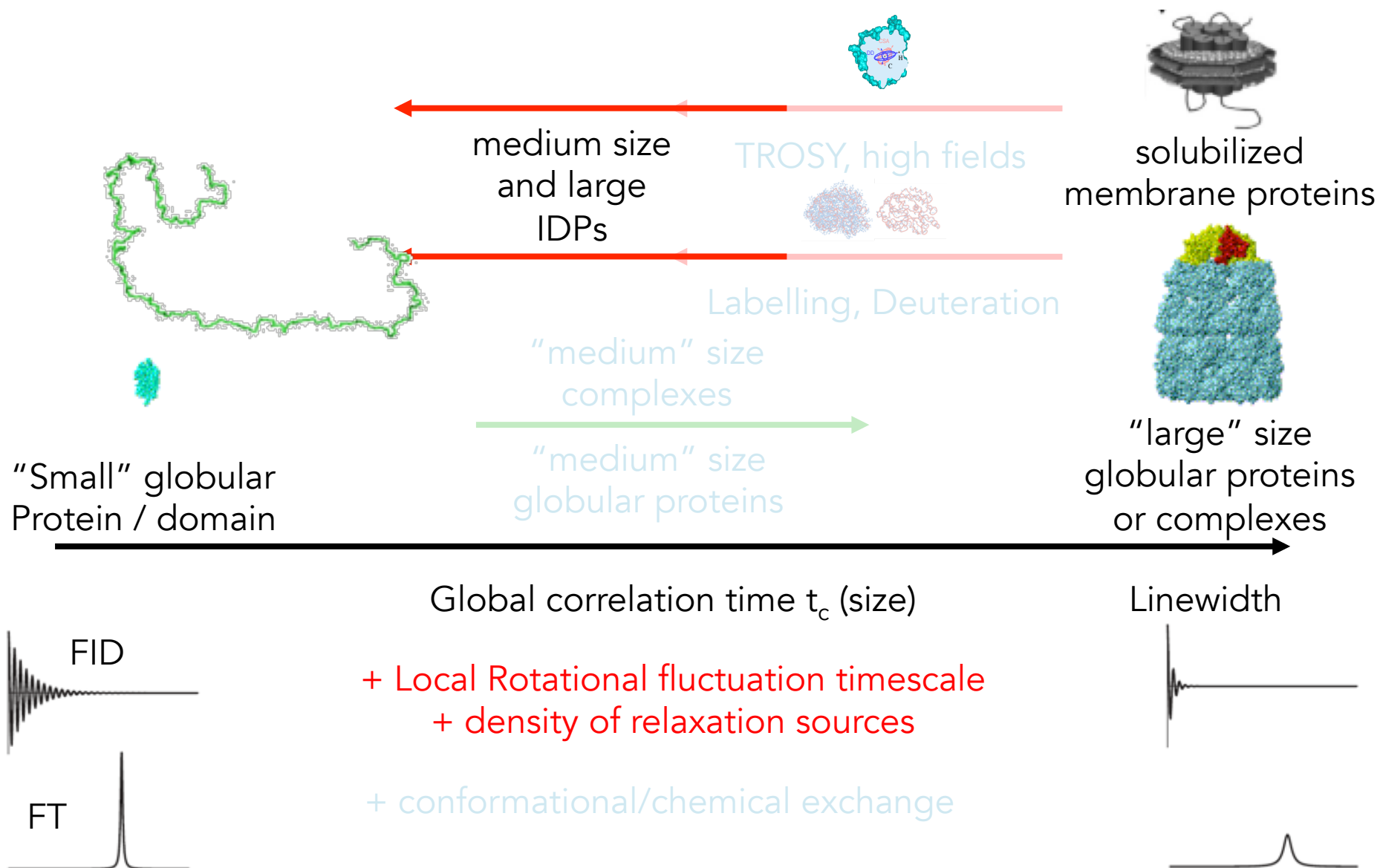


“Intermediate”
exchange

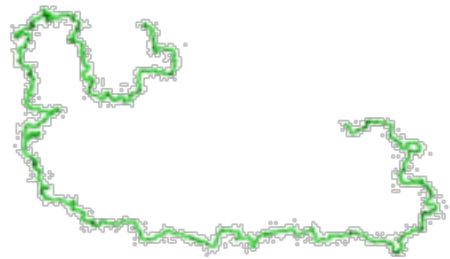


“Fast”
exchange

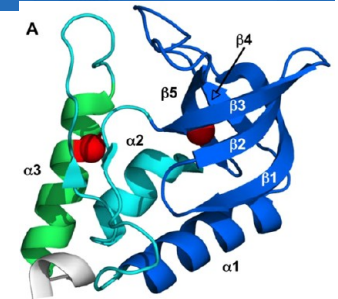
NMR - Spatial and Time scales are highly intermingled



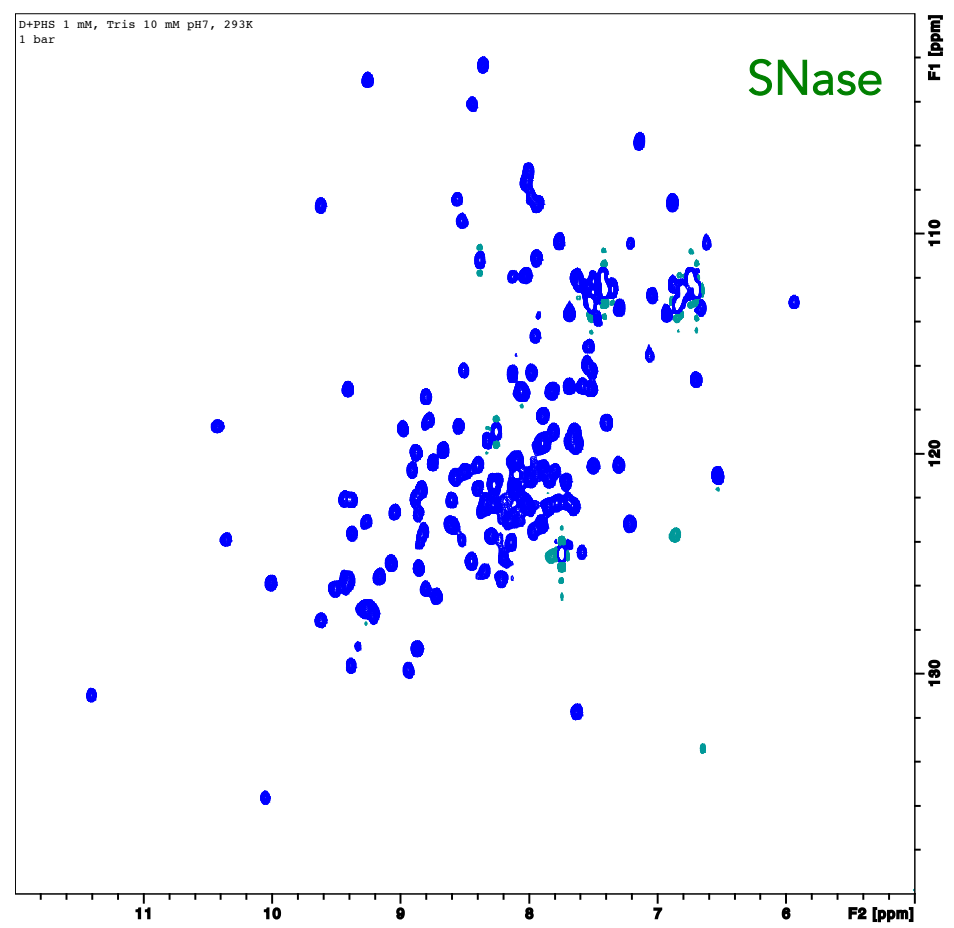
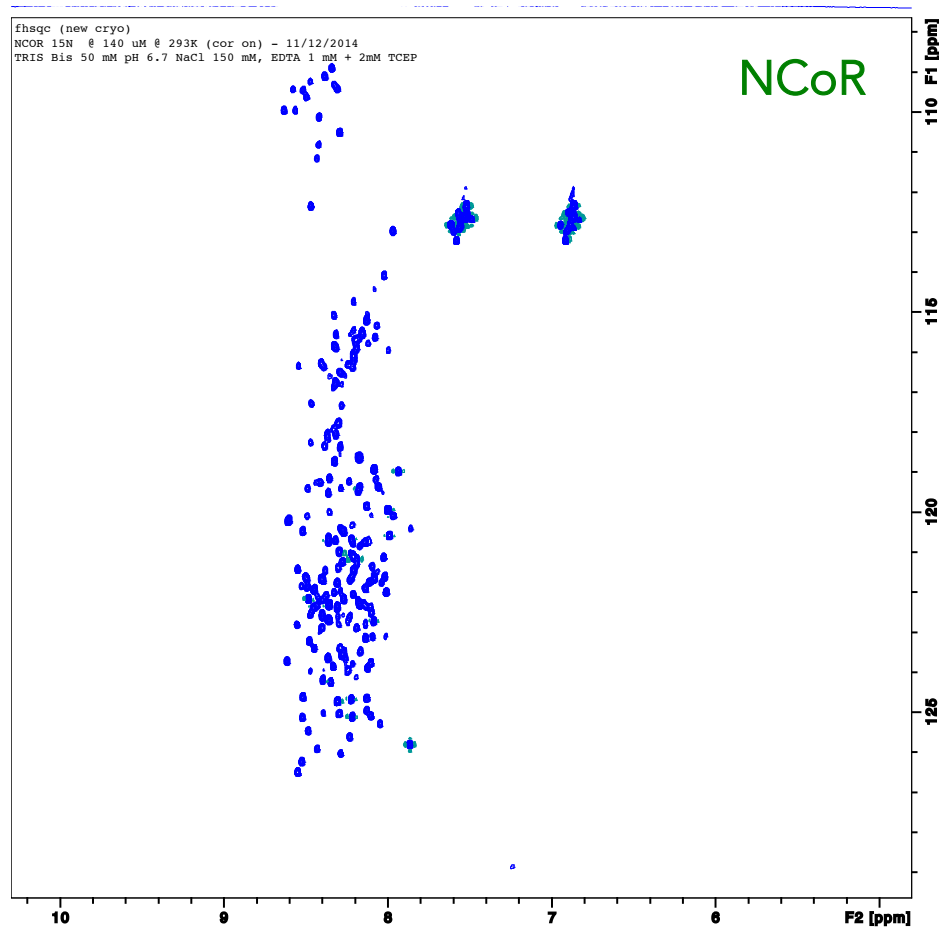
« Size » and disorder



Unfolded protein
29 kDa



Folded protein
15 kDa

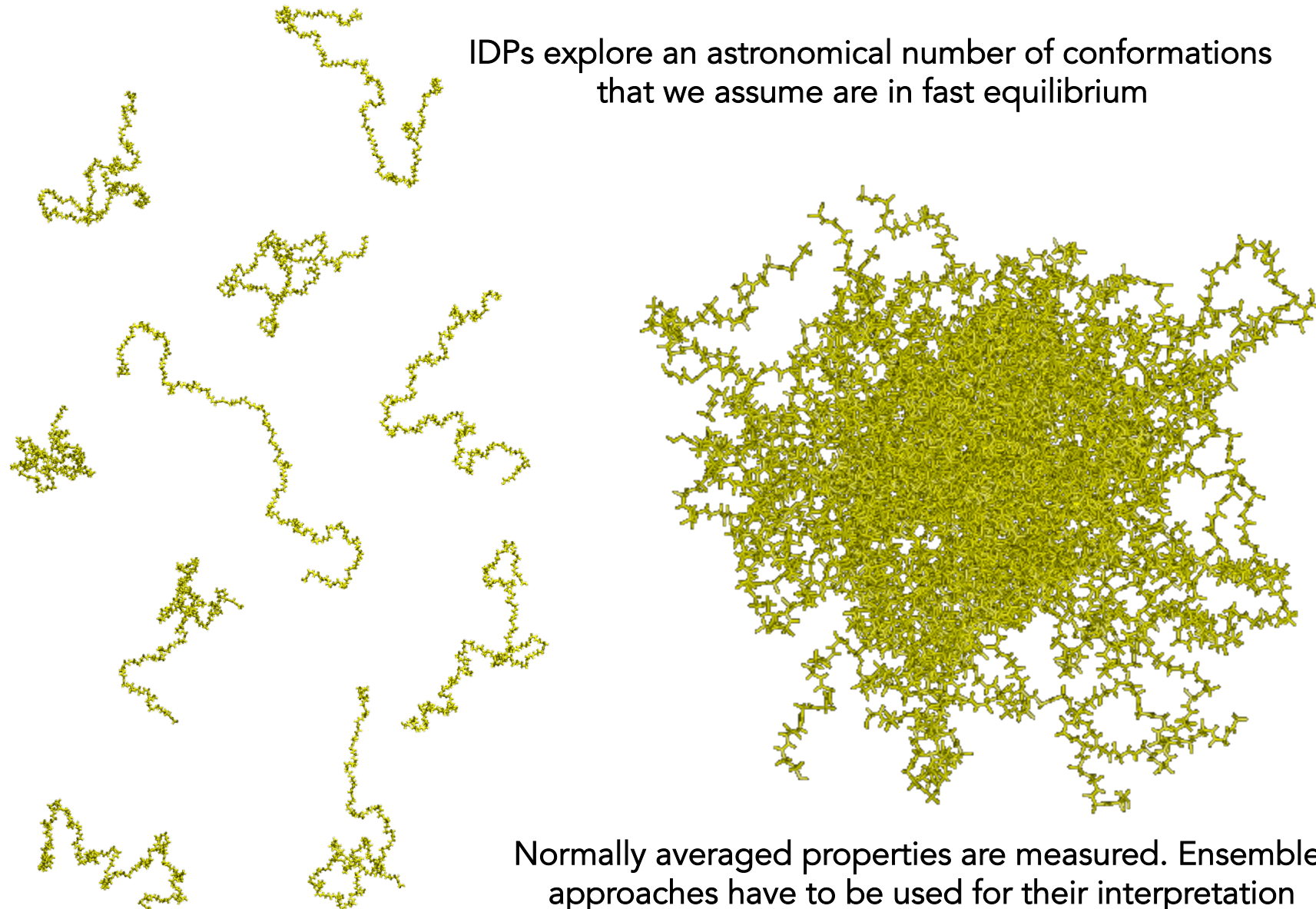


What is an IDP ?

- ✧ IDPs **lack stable** tertiary and/or **secondary structure**
- ✧ IDPs are **more common in eukaryotes** than in bacteria and archaea: Probably linked with their major biological complexity... Up to ~30-50% of genome
- ✧ Very specific amino acid sequences rich in P G R K and E, and **depleted on hydrophobic residues**
- ✧ IDPs are **the target of the vast majority of PTM**: Phosphorylation, Glycosylation...
- ✧ **Interactome hubs** are enriched in **disordered regions**
- ✧ **Partner recognition** of IDPs is normally performed through **short linear motifs** (SLiMs) that are embedded in disordered poorly conserved regions

Flexible Proteins, a Challenge for Structural Biology

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

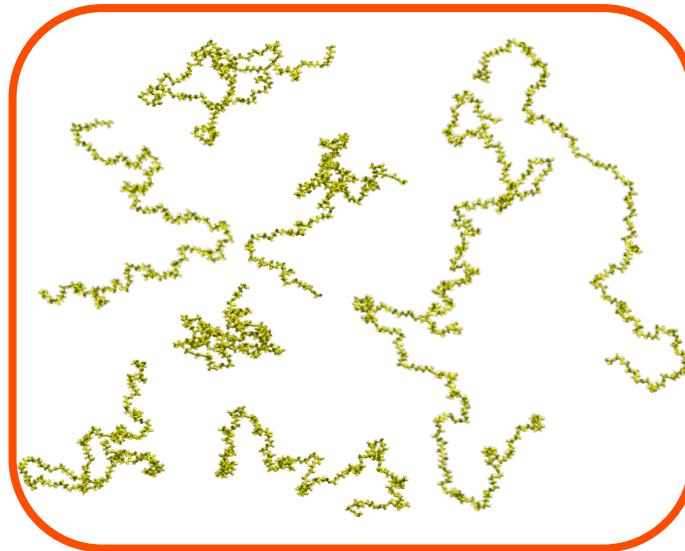


Normally averaged properties are measured. Ensemble approaches have to be used for their interpretation

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 ?

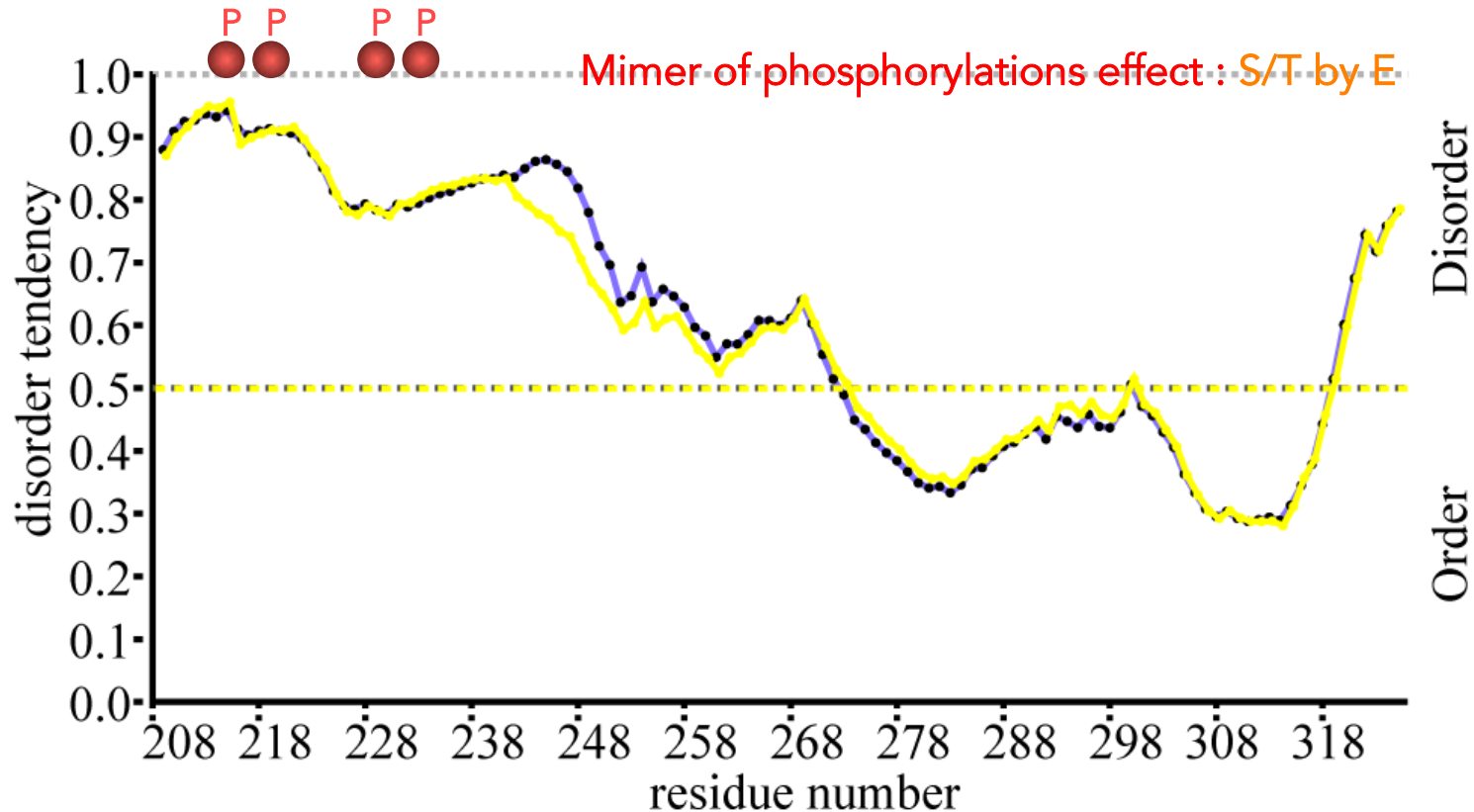
- ✧ Bioinformatics
Identification of disordered/ordered regions
- ✧ Biophysical characterization: CD, FTIR, FRET, hydrodynamics
Partial Information
- ✧ Small-Angle X-ray Scattering (SAXS)/Small-Angle neutron Scattering (SANS)
Averaged Intensity profiles...
Qualitative Interpretation of averaged R_g and 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

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.

Bioinformatics

- ✧ Other predictors of disordered/ordered regions
 - ✧ the meta-predictor PONDR-FIT (<http://disorder.compbio.iupui.edu/pondr-fit.php>) that combines the results of six different methods (PONDR-VLXT, PONDR-VSL2, PONDR-VL3, FoldIndex, IUPred, and TopIDP)
 - ✧ PrDOS <http://prdos.hgc.jp> a structure-based method
 - ✧ SPOT-disorder [http://sparks-lab.org/server/ SPOT-disorder/](http://sparks-lab.org/server/SPOT-disorder/) based on a window-based neural network (SPINE-D)
 - ✧ DISOPRED3 (<http://bioinf.cs.ucl.ac.uk/psipred/>) that combines three machine learning models: support vector machine, neural network and nearest neighbor

- ✧ Secondary structure predictions
 - ✧ PSIPRED v3.3 <http://bioinf.cs.ucl.ac.uk/psipred/>
 - ✧ Jpred <http://www.compbio.dundee.ac.uk/jpred4>
 - ✧ SOPMA https://npsa-prabi.ibcp.fr/NPSA/npsa_sopma.html
 - ✧ alpha- turns predictor ALPHAPRED crdd.osdd.net/raghava/alphapred/
 - ✧ PSSpred <https://zhanglab.ccmb.med.umich.edu/PSSpred/>

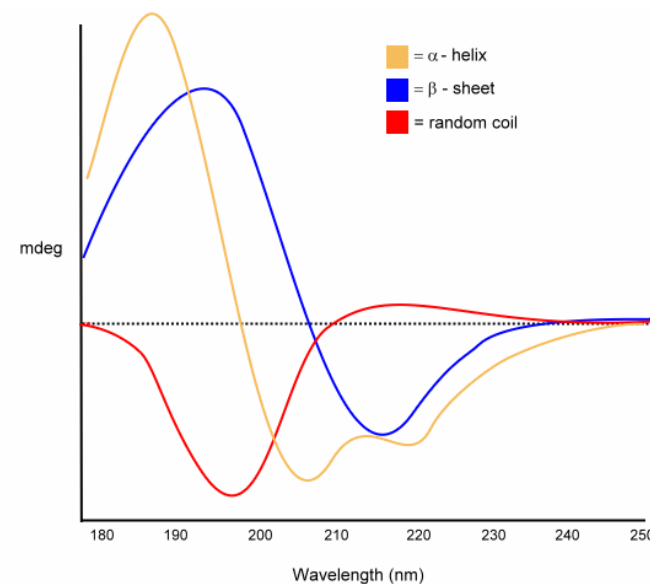
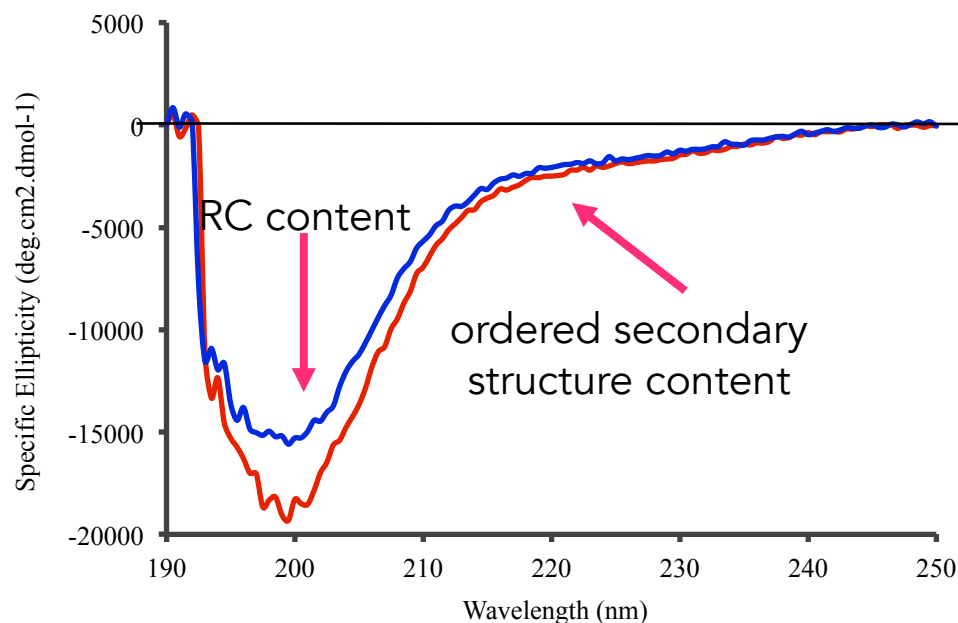
- ✧ Sequence conservation
 - ✧ GREMLIN (Generative Regularized Models of proteins) software

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

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 $\rightarrow$ p*) at 192 nm
- ✧ negative (p $\rightarrow$ p*) at 209 nm
- ✧ negative (n $\rightarrow$ p*) at 222 nm
- ✧ negative at 218 nm (p $\rightarrow$ p*)
- ✧ positive at 196 nm (n $\rightarrow$ p*)
- ✧ positive at 212 nm (p $\rightarrow$ p*)
- ✧ negative at 195 nm (n $\rightarrow$ p*)

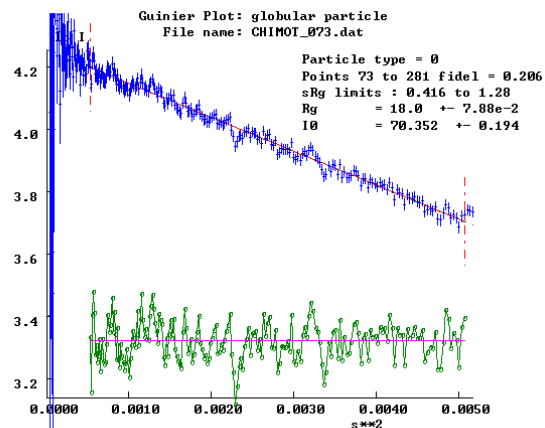
Is our protein an IDP ?

- ✧ 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

SAXS/SANS

Qualitative interpretation of averaged Intensity Profiles

Guinier

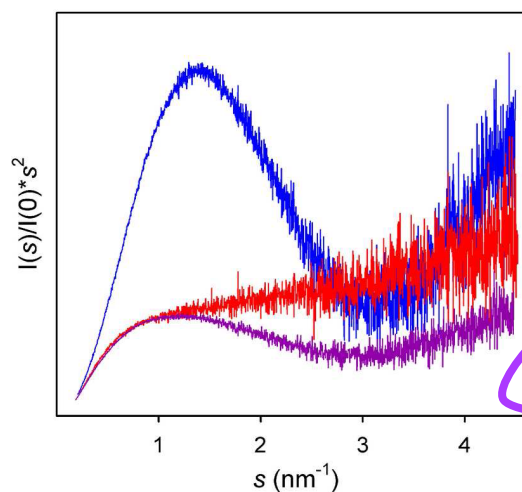


R_g , Molecular Weight

Flory's equation was parametrized for IDPs

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

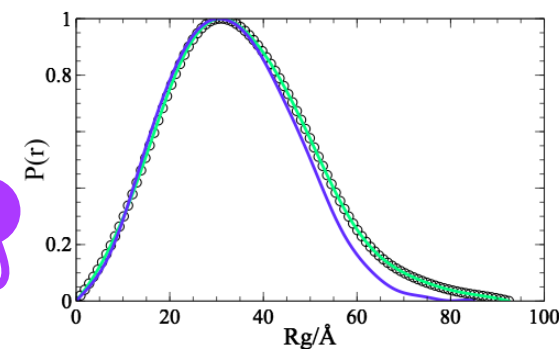
Kratky



Mass Density

- folded
- unfolded
- folded and unfolded

P(r)



1D-Structure
Maximum Distance

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

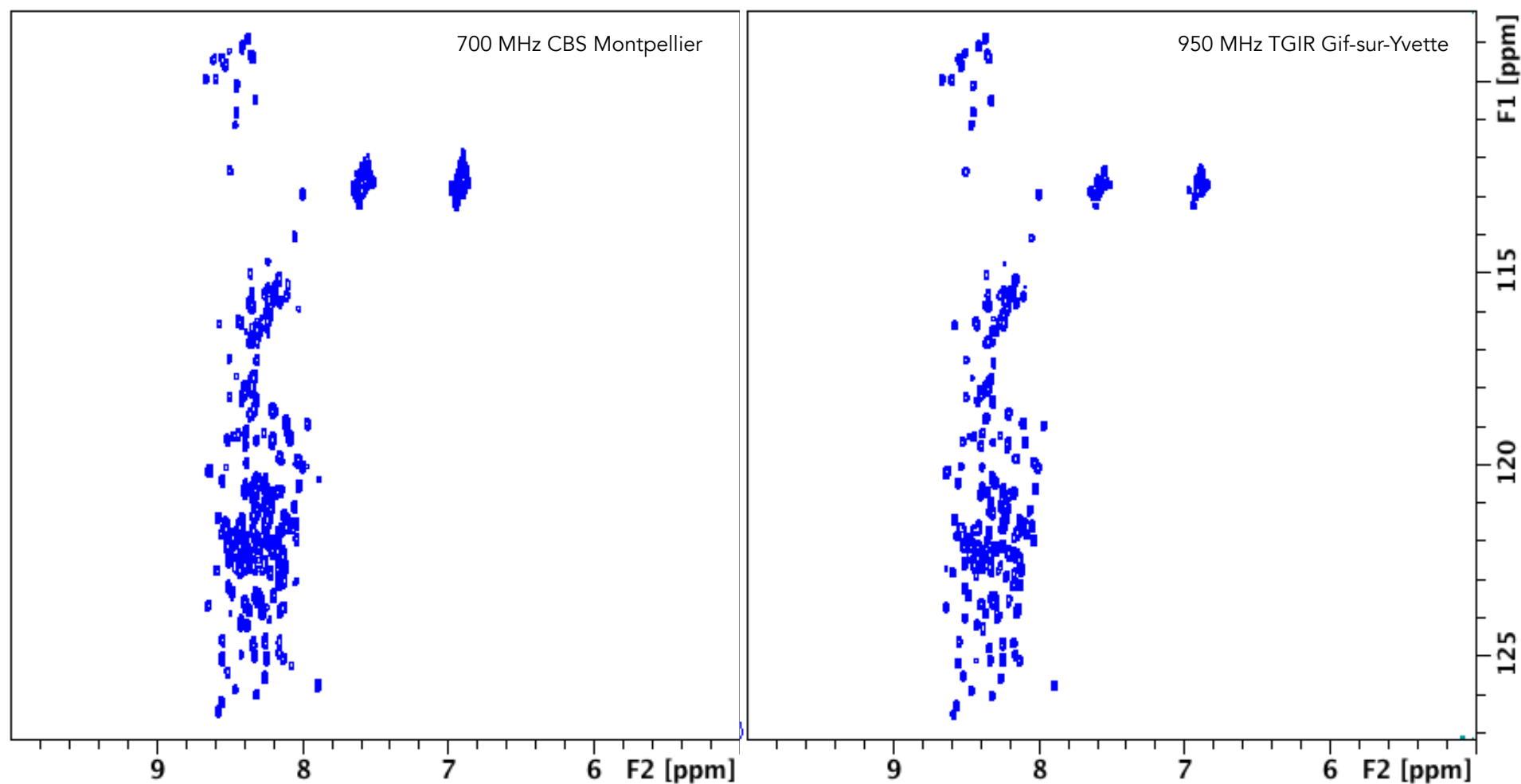
NMR

Sensitivity and Resolution

700 MHz

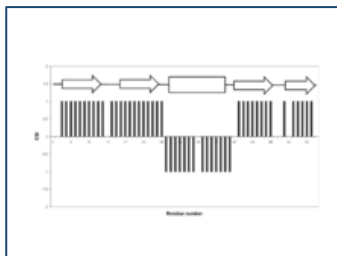
NCoR (269 aa)

950 MHz



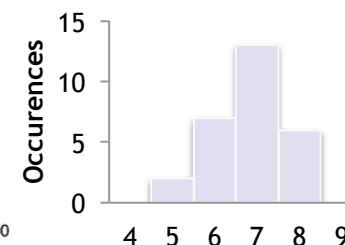
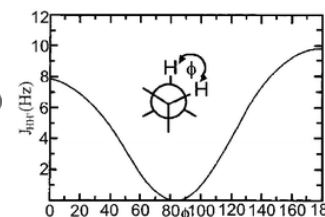
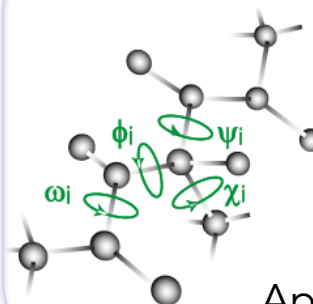
NMR is rich in structural information

Chemical Shifts



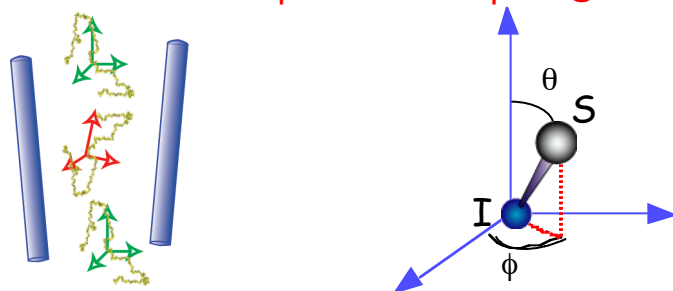
Secondary structure elements

J-Couplings



Approximate torsion angles $^3J_{HNHA}$ (Hz)

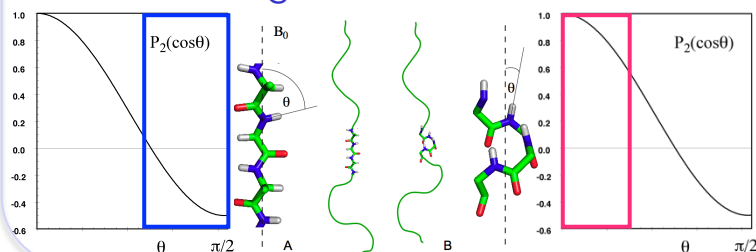
Residual Dipolar Couplings



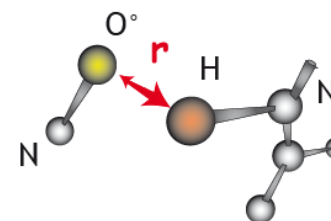
Relative Bond orientations

Extended Regions

α -Helical Regions

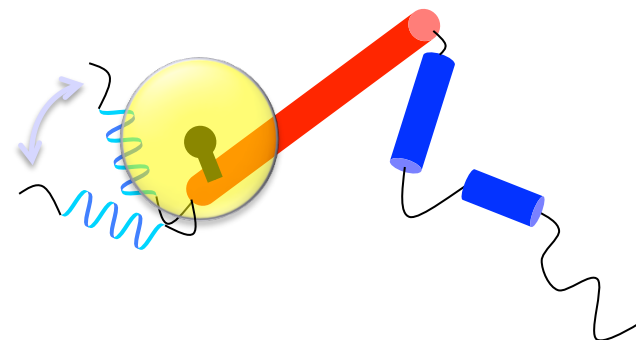


Paramagnetic Relaxation Enhancement



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

Transient long-range contacts



Chemical Shift Index

✧ Classical database (1994)

Wishart et al J Biomol NMR 5 (1995) 67-81 :

Based on **Ac-GGX(A/P)GG-NH2** peptides database

Proline correction factor on preceding residues of ~ 2 ppm on Ca

✧ Schwarzingler database (2001)

“Sequence-Dependent Correction of Random Coil NMR Chemical Shifts” by S.

Schwarzinger, et al. J. Am. Chem. Soc. 2001, 123, 2970-2978

Ac-G-G-X-G-G-NH2

Not representative for the other residue types & in 8 M urea pH 2.3

✧ Kjaergaard database (2011)

Random coil chemical shifts for intrinsically disordered proteins: Effects of temperature and pH. Kjaergaard, et al. (2011) J. Biomol. NMR 49(2):139-49

Ac-QQXQQ-NH2

Sample temperature (Celsius) and pH corrections, using GGXGG-based neighbor correction for glycines & corrections for perdeuterated protein

✧ Tamiola database : ncIDP (2010)

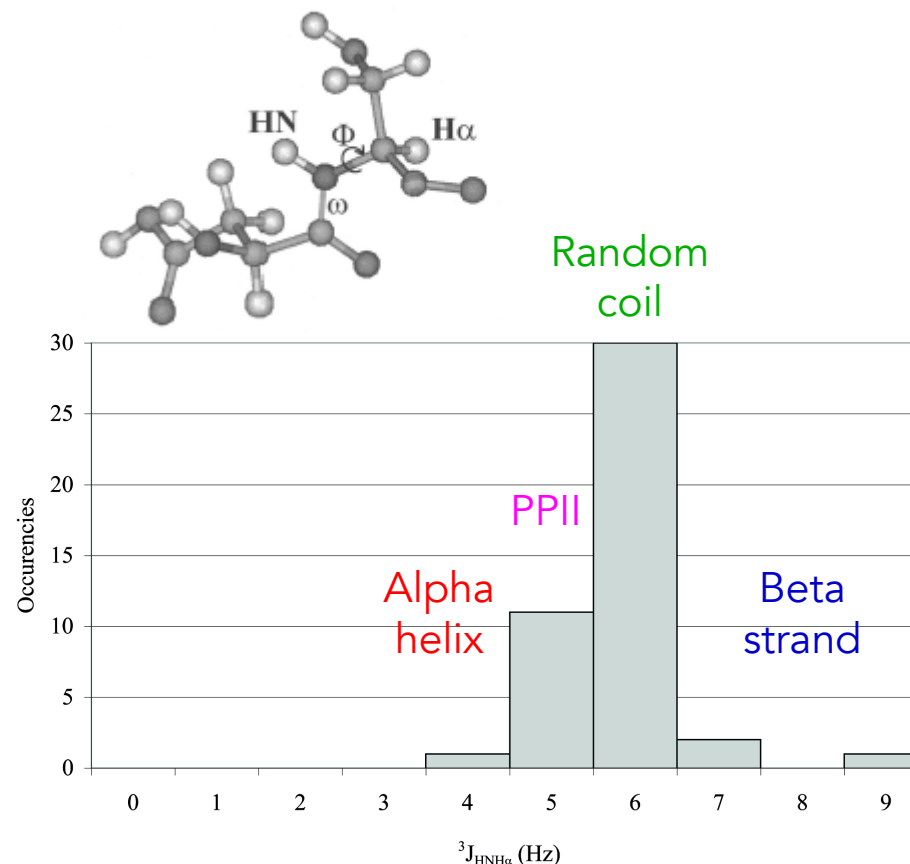
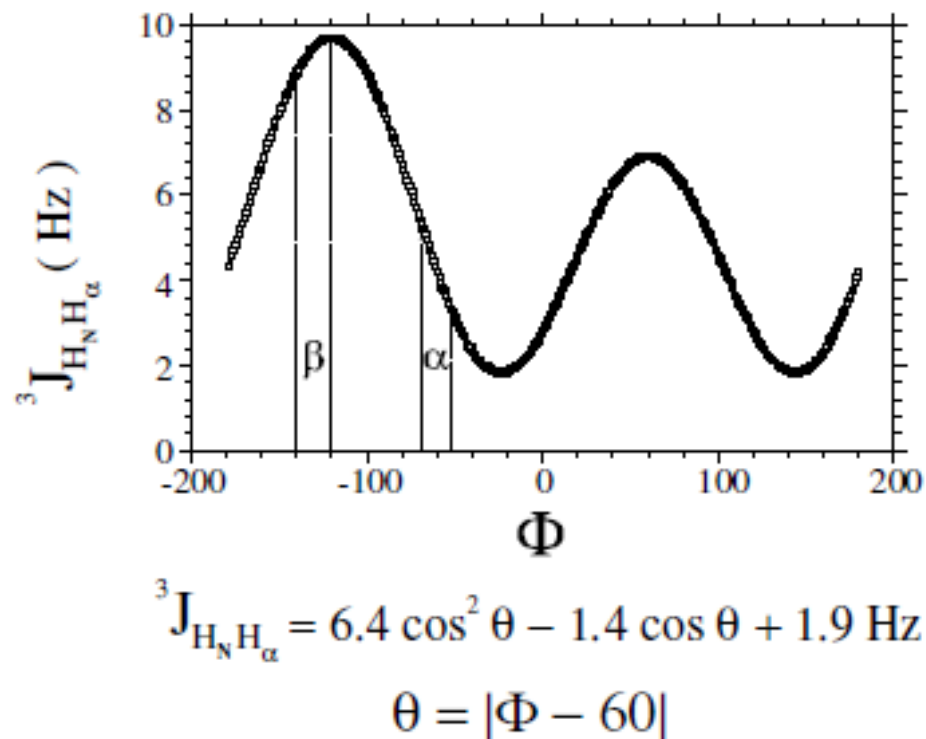
neighbor corrected Intrinsically Disordered Protein Library calculated according to

Tamiola, Acar, Mulder (2009) tables

IDPs, (POTENCI > 100 IDPs)

Scalar couplings

Karplus Law and Φ



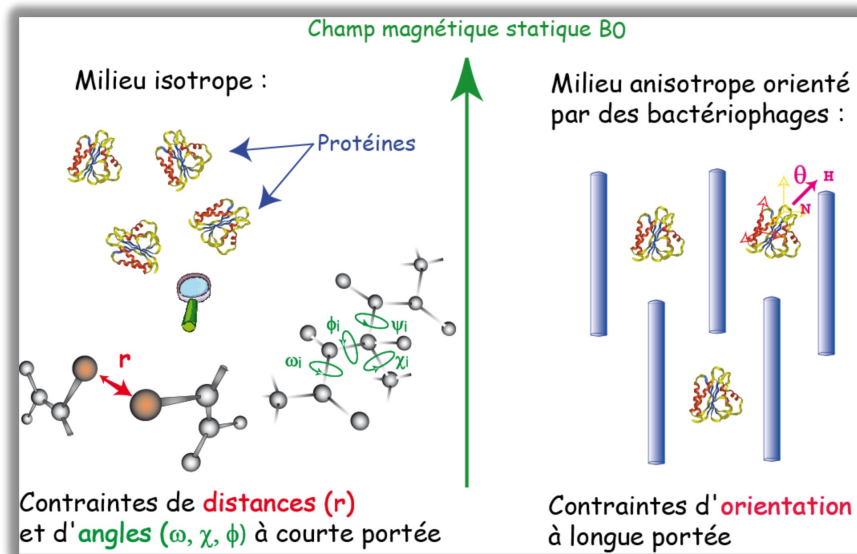
- $^3J_{H_N H_\alpha}$ Random Coil Values : Schwarzinger j biomol nmr 18 (2000) 43-48
- $^3J_{H_N H_\alpha}$ library : https://spin.niddk.nih.gov/bax/nmrserver/rc_3Jhnhalpha/

Shen, Y, Roche, J, Grishaev, A, Bax, A (2018) Prediction of nearest neighbor effects on backbone torsion angles and NMR scalar coupling constants in disordered proteins. Protein Sci 27(1):146-158

Residual Dipolar Couplings

Orientational information in an anisotropic medium

Liquid crystal Media



✧ Isotropic medium :

scalar coupling = J

✧ Anisotropic medium :

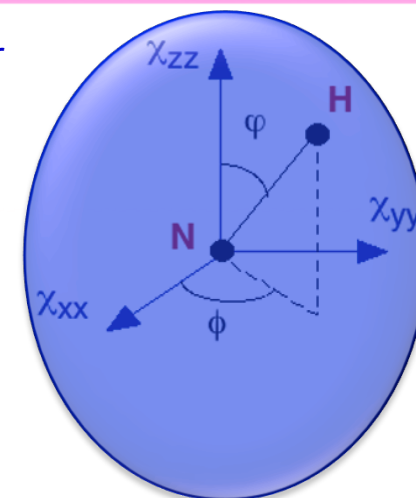
apparent splitting of the doublet = $J + D_{IS}$

=> Dipolar Interactions non averaged to zero

Geometrical dependency of RDCs :

$$D_{IS} = -S \frac{\mu_0 \hbar}{8\pi^2} \frac{\gamma_I \gamma_S}{\langle r_{IS}^3 \rangle} \left\{ A_a (3 \cos^2 \theta - 1) + \frac{3}{2} A_r \sin^2 \theta \cos 2\phi \right\}$$

Alignment tensor



Restricted
reorientation

Steric Interactions:

Alcohol mixture, gels, cellulose crystallites

Electrostatic Interactions:

Phage (Pf1), purple membranes, bicelles

Tjandra, N.; Bax, A. Science, 1997, 278, 1111-1114.

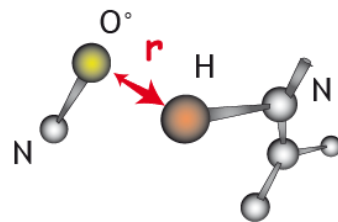


Paramagnetic Relaxation Enhancement

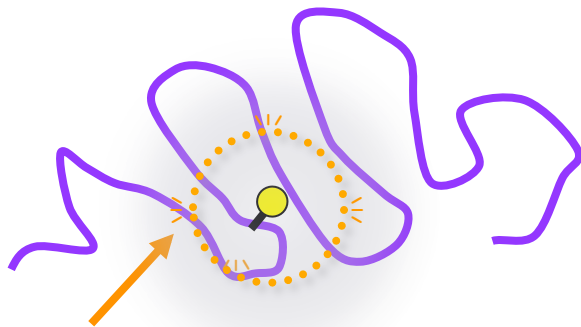
Dynamic dependant parameters

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

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



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



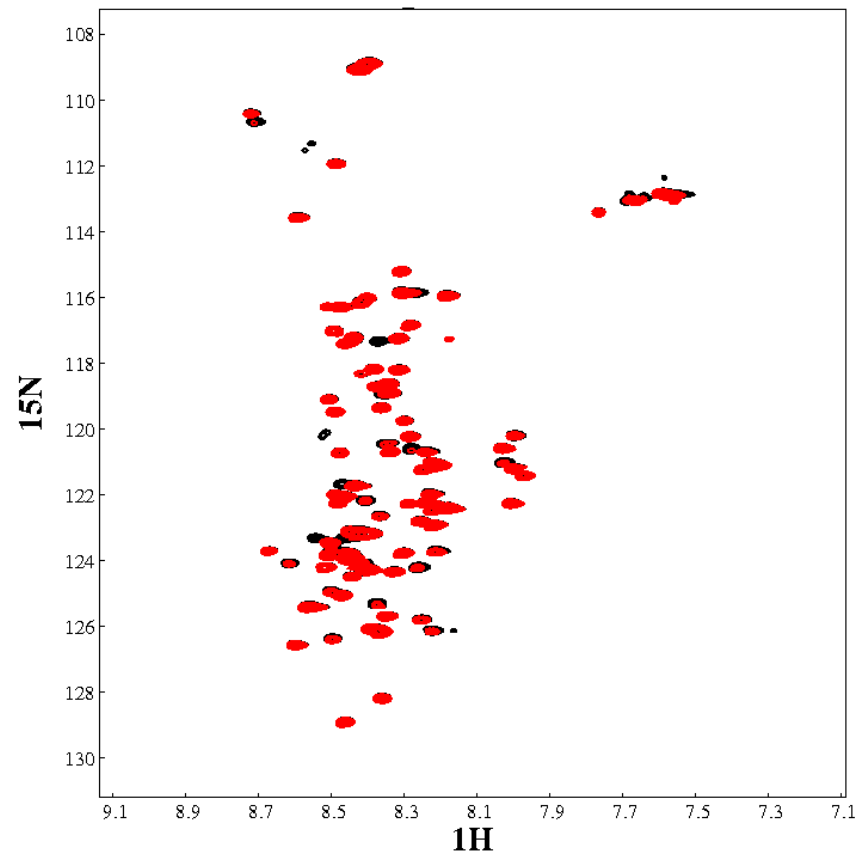
Effective radius $\sim 20 \text{ \AA}$

Forme para I :

NO(C291) 200 μM

Forme dia IO:

+ 5mM de vitamine C

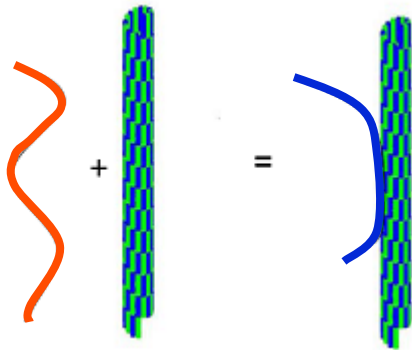


Is our protein an IDP ?

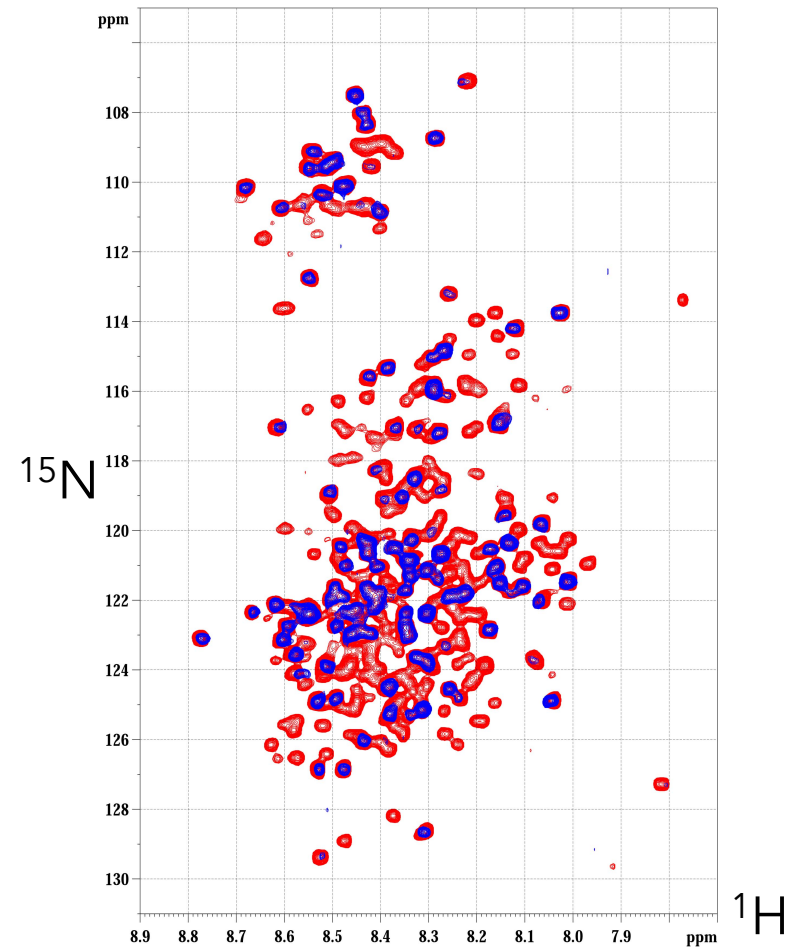
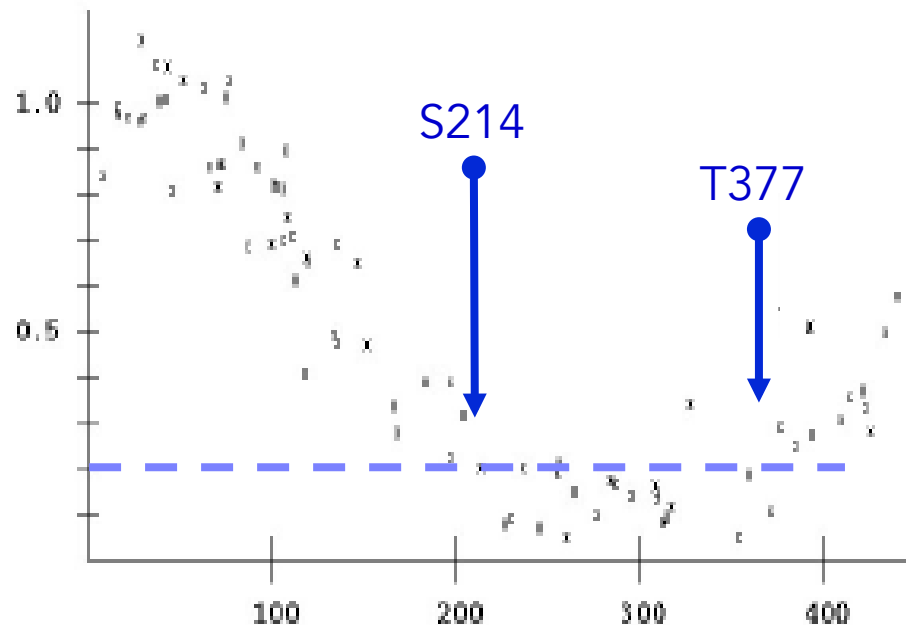
- ✧ Bioinformatics
Identification of disordered/ordered regions
- ✧ Biophysical characterization: CD, FTIR, FRET, hydrodynamics
Partial Information
- ✧ Small-Angle X-ray Scattering (SAXS)/Small-Angle neutron Scattering (SANS)
Averaged Intensity profiles...
Qualitative Interpretation of averaged R_g and 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

To X Ray cristallography

Structure determination of the bound form

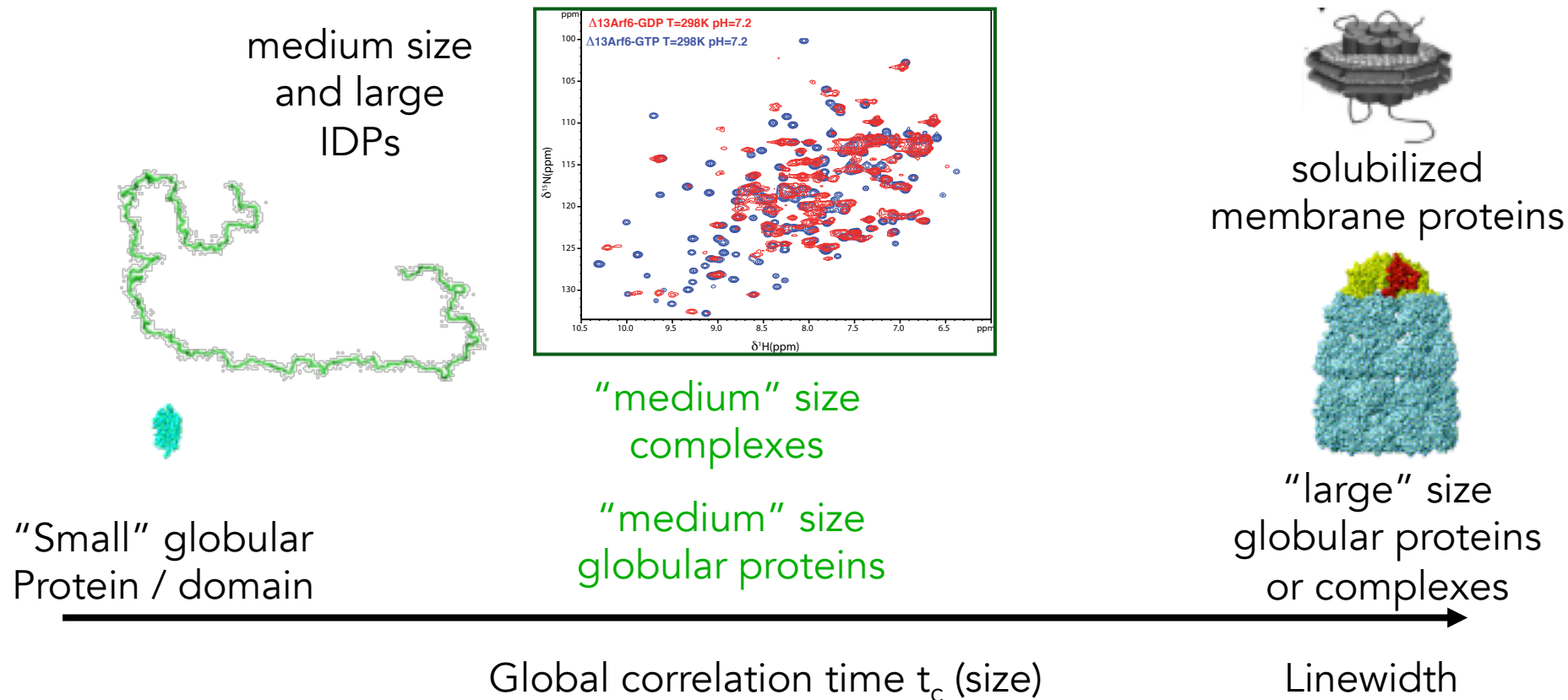


Delimitate the bound region to the partner



NMR spectra of **free form** (54 kDa) (5mM) and **bound form** (TeraDalton) (20mM)

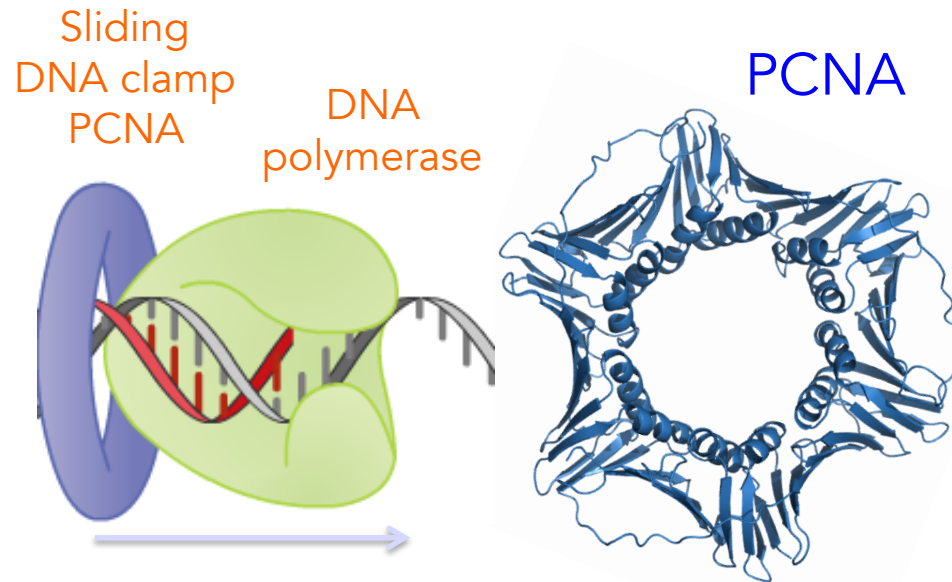
NMR - Spatial and Time scales are highly intermingled



+ Local Rotational fluctuation timescale
+ density of relaxation sources

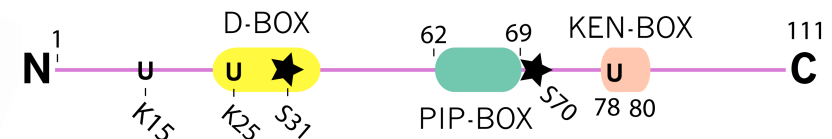
+ conformational/chemical exchange

p15^{PAF} a PCNA-Binding Protein



- ✓ 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

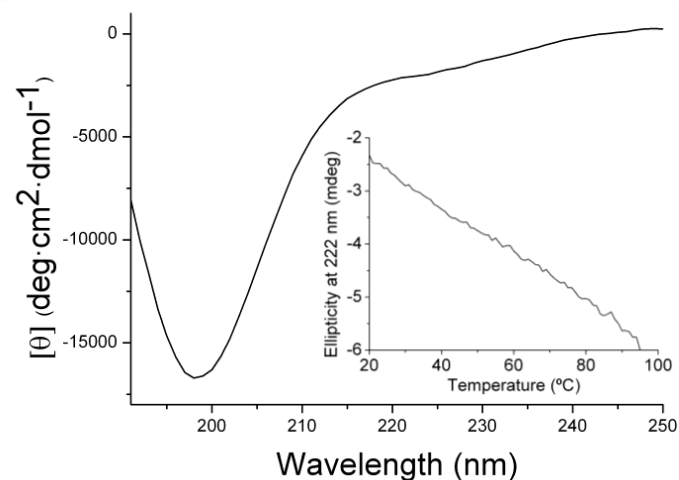
P15



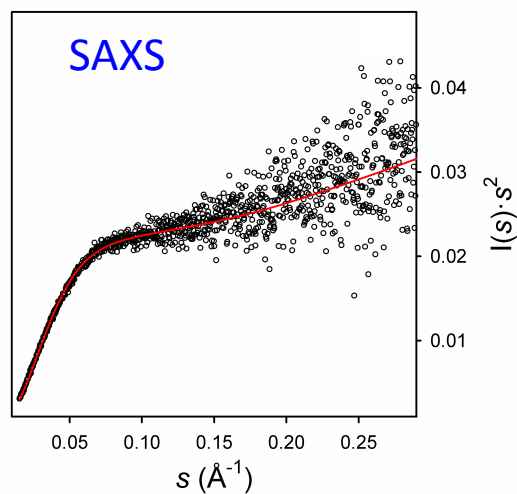
- ✓ P15 acts as a regulator of DNA repair
- ✓ is over-expressed in several types of human cancer
- ✓ It binds PCNA through its PIP-BOX
- ✓ Targeted for degradation by the ubiquitin ligase APC through its conserved KEN-BOX.

p15^{PAF} is an IDP ...

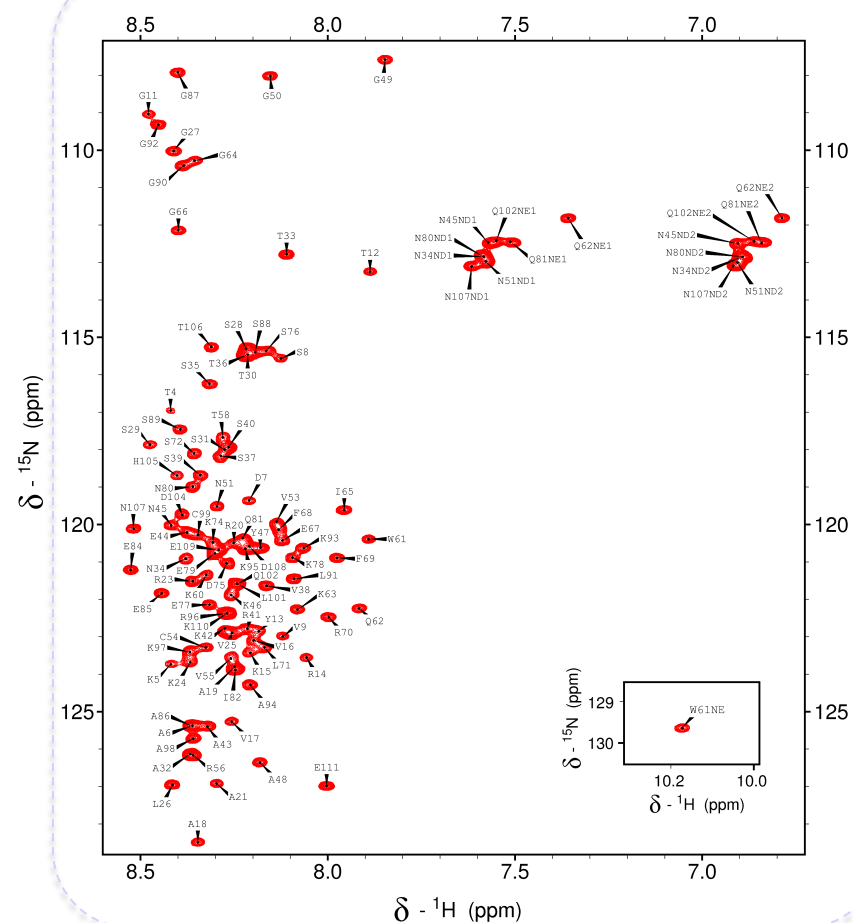
CD



SAXS

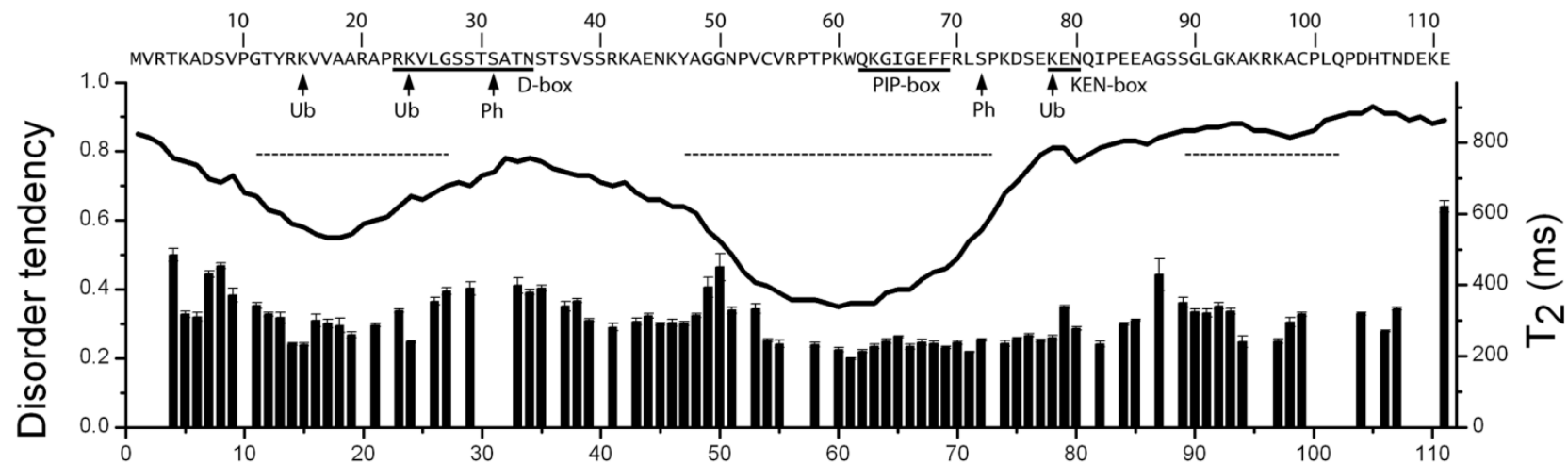


NMR

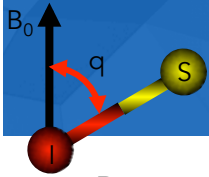


De Biasio *et al. Biophys J.* 2014

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 T₂ values, that correspond to sequences with reduced predicted disorder.



Transient structuration monitored by RDCs

Bernado et al. PNAS 2005

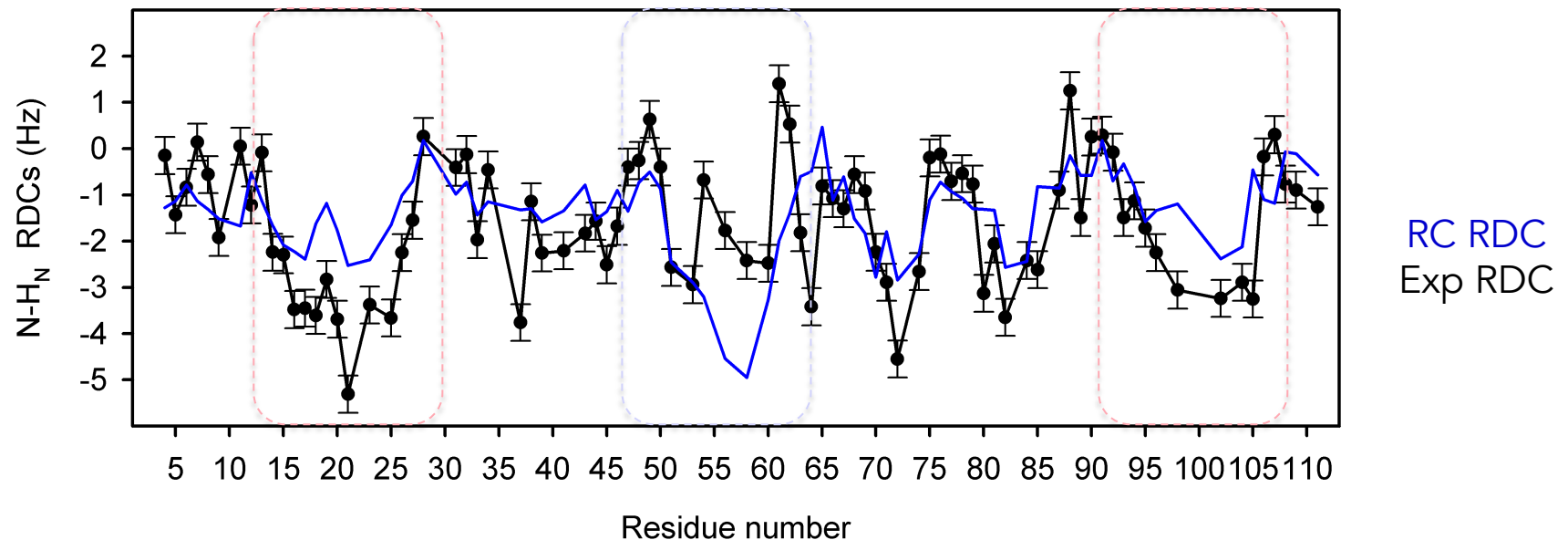
Disordered but, according to the RDCs, it presents several sites with partial structuration
FM Random Coil model helps in identifying the nature of the structural elements observed

Ensemble Description

Flexible-Meccano
100,000 RC structures

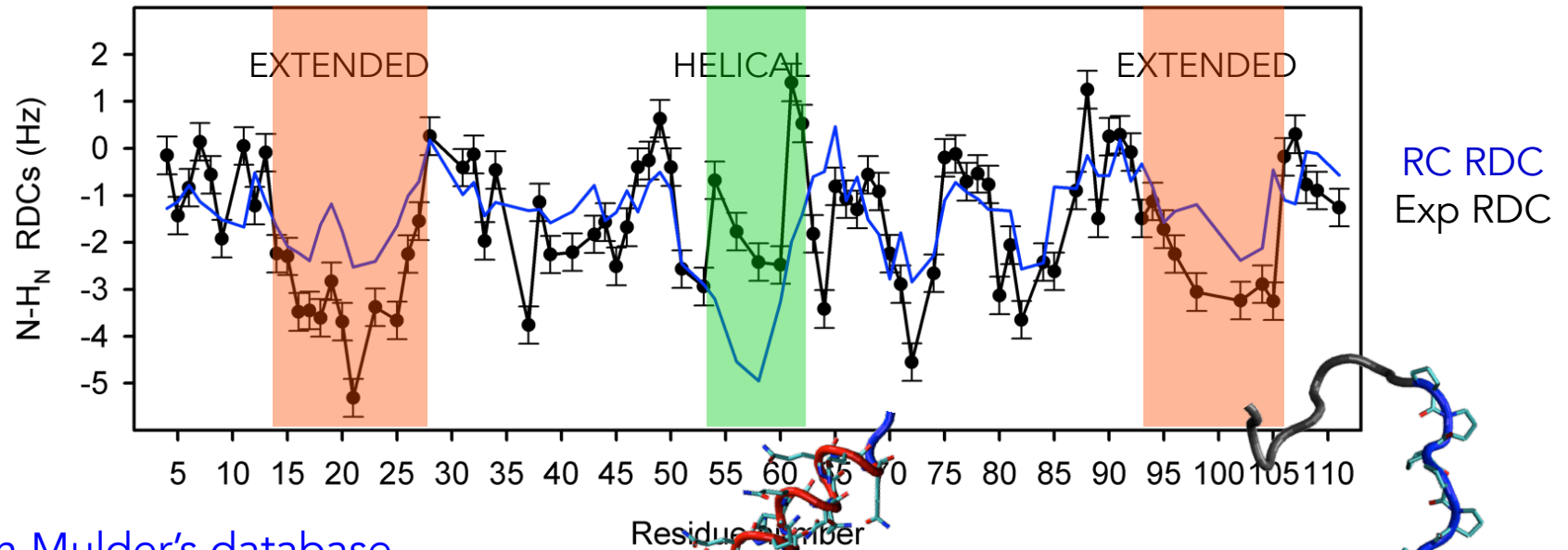
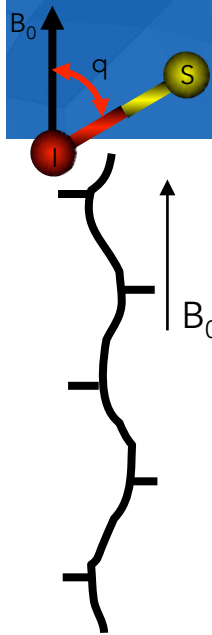
Conformational Preferences

Secondary Structural Element :
Position and Population

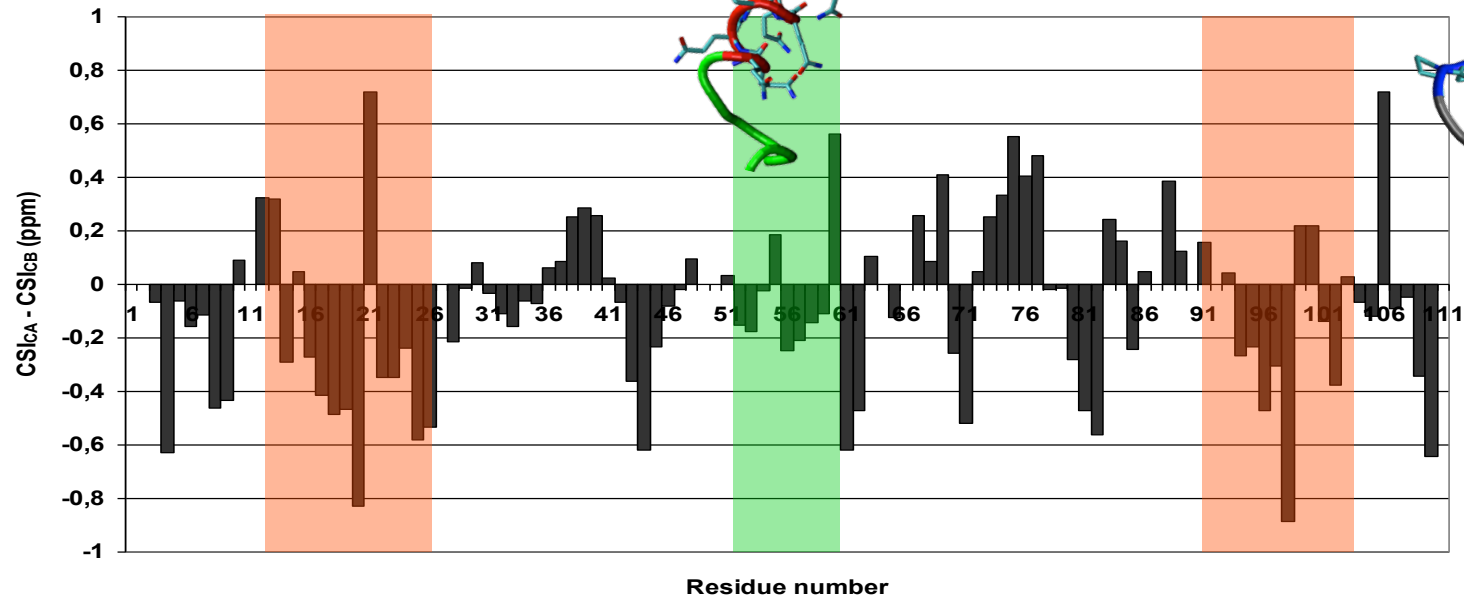


The RC model cannot fully reproduce the measured RDCs

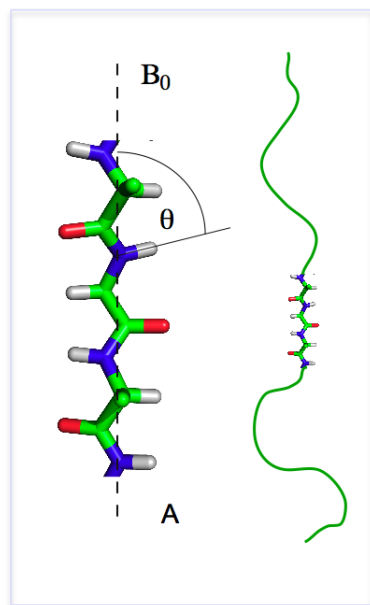
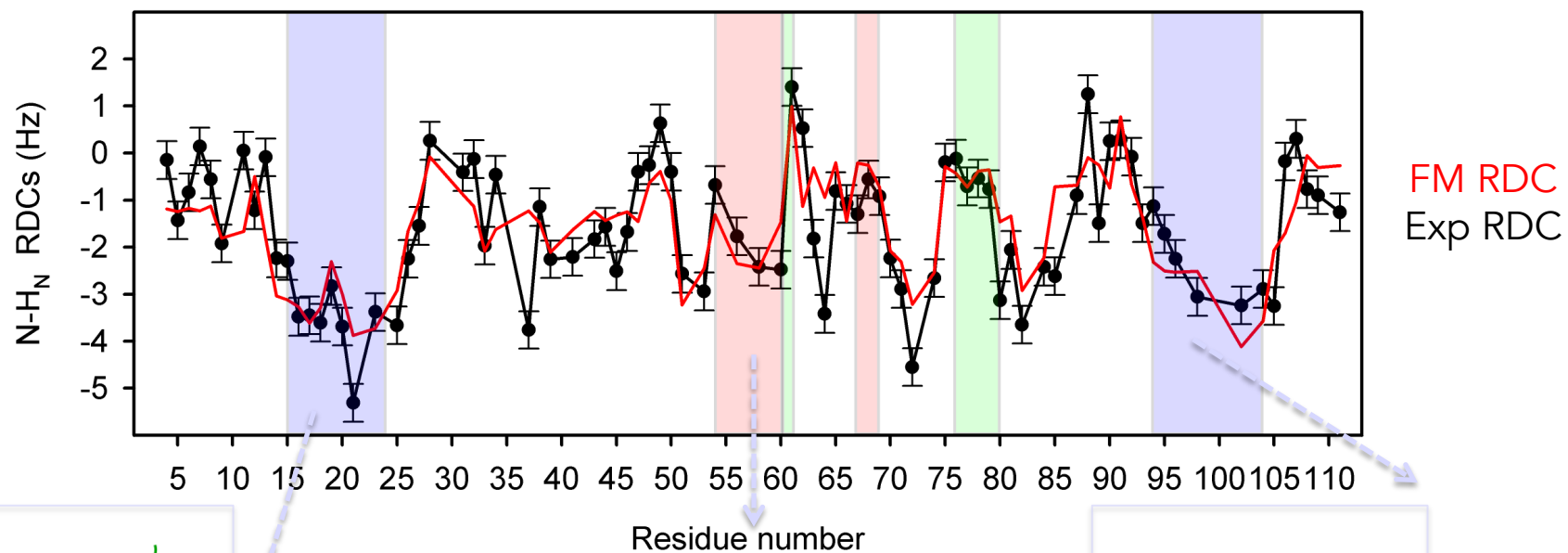
Transient structuration monitored by RDCs



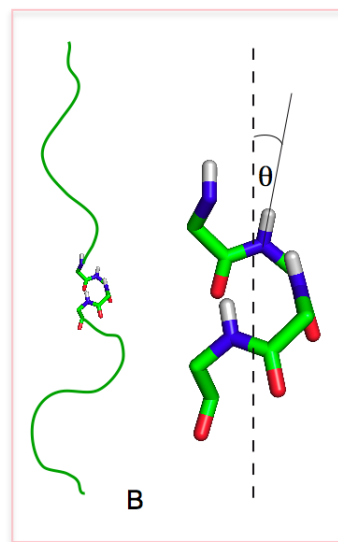
CSI with Mulder's database



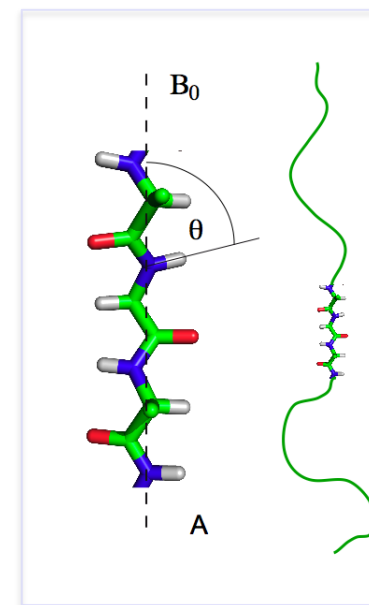
Structural Elements in p15



10%

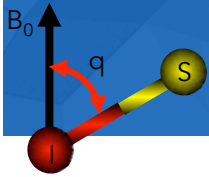


5%

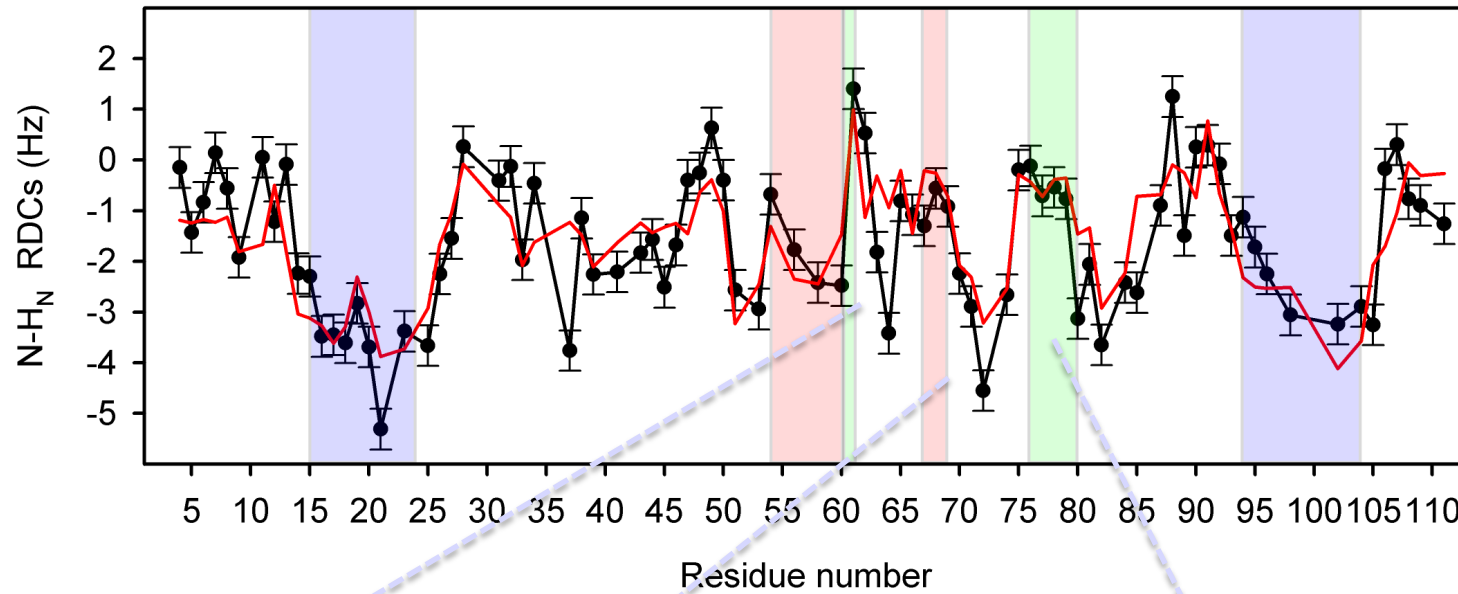


8%

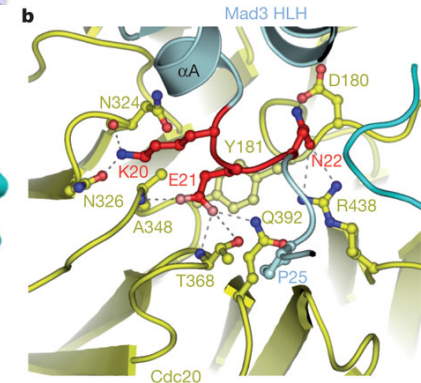
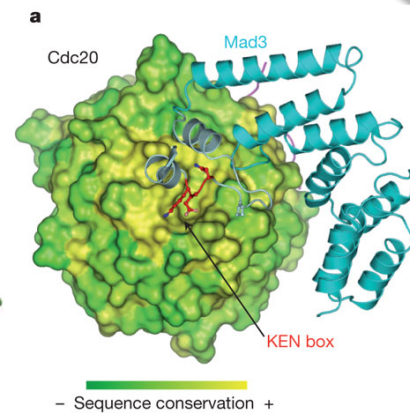
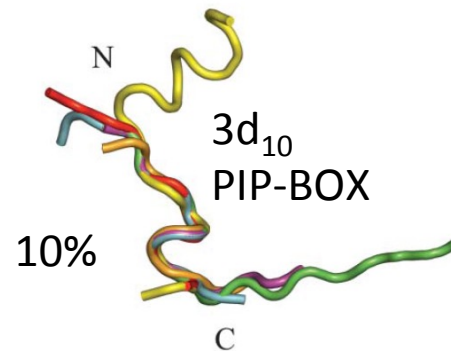
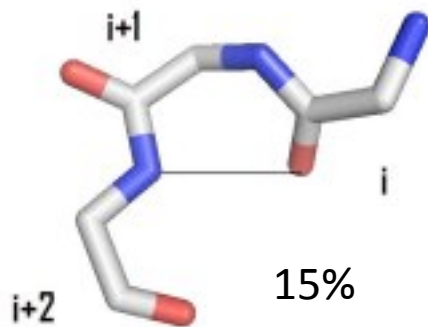
100,000 structures



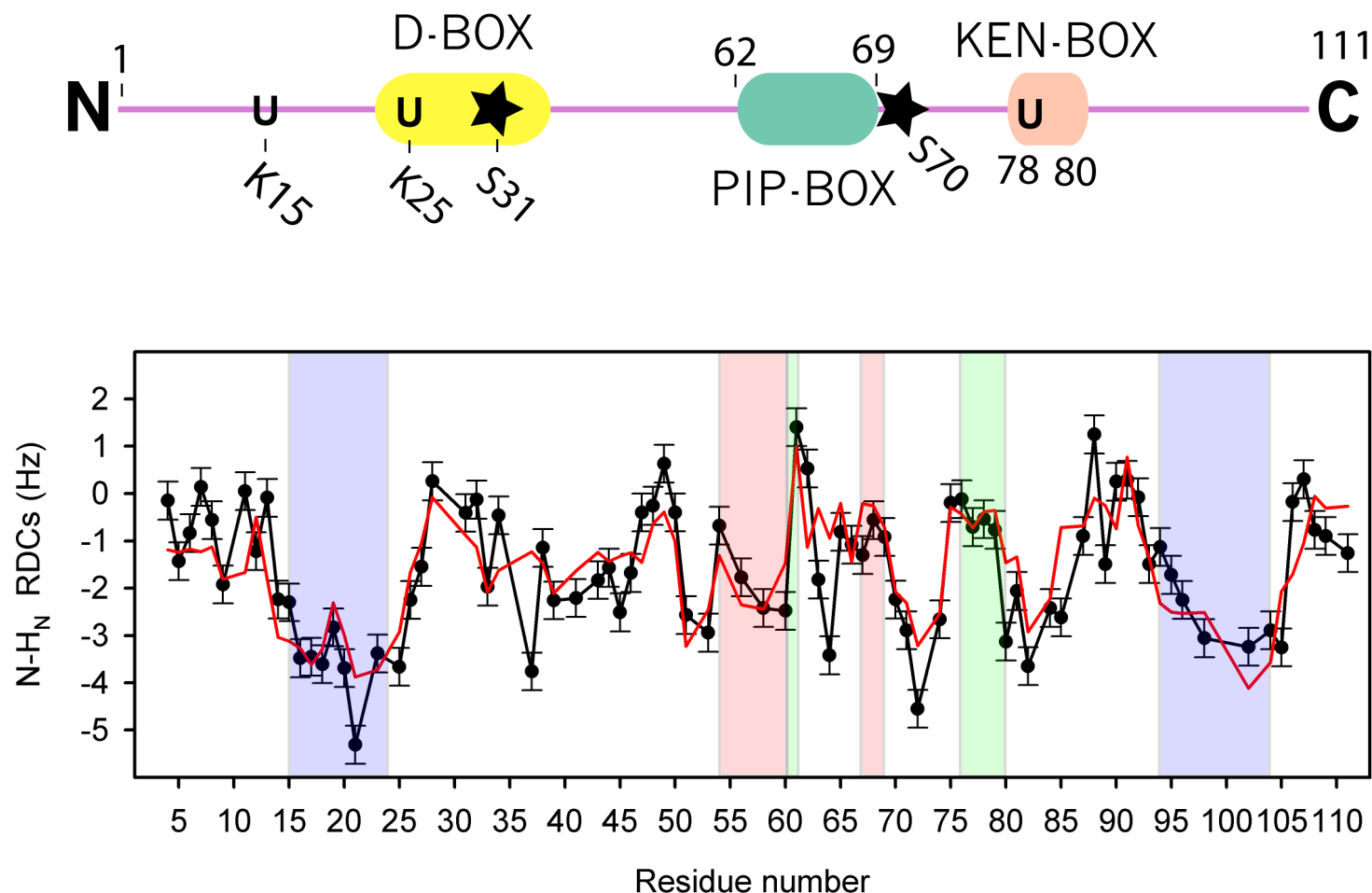
Structural Optimization



8%



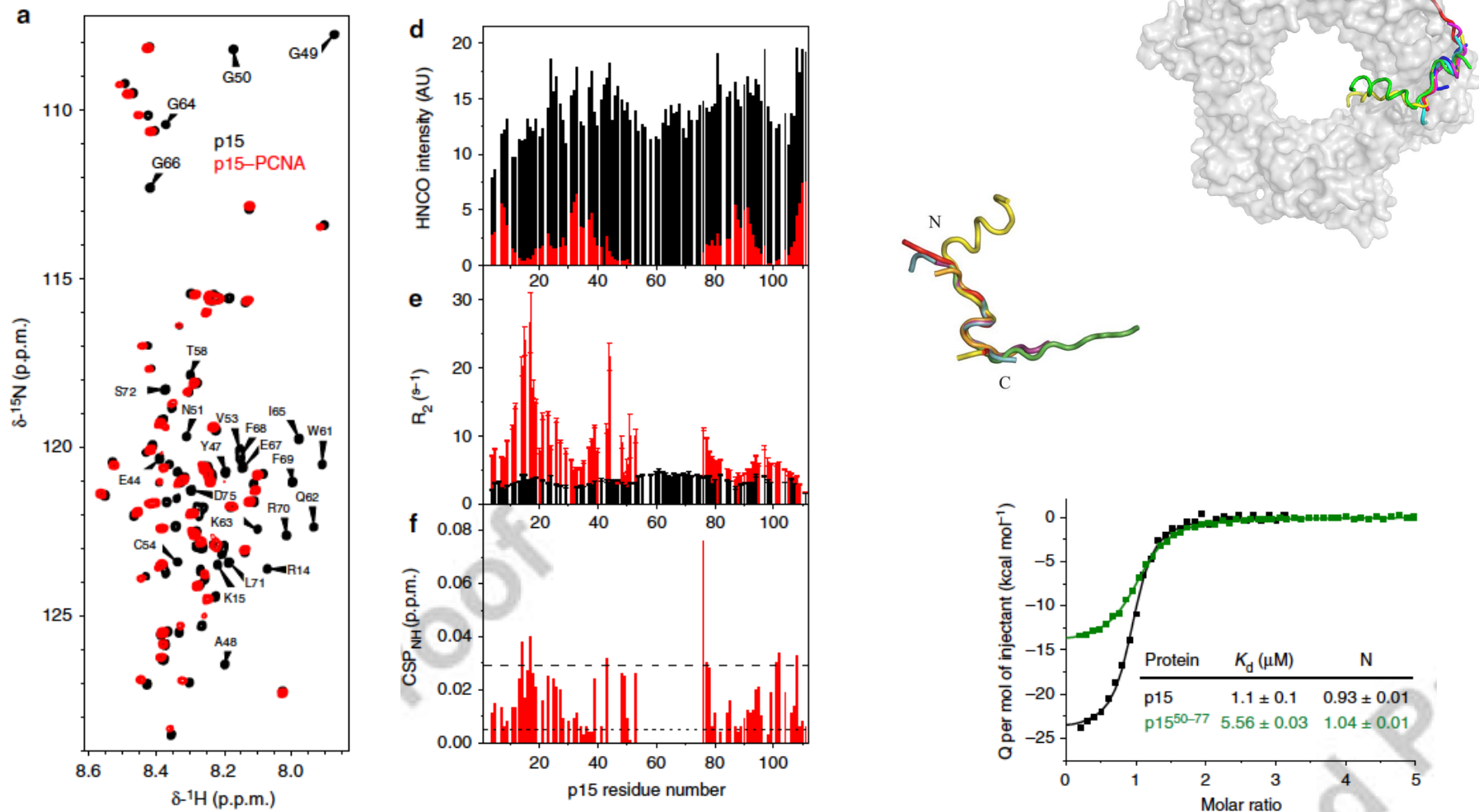
p15^{PAF} is an IDP with partial Structuration at Interaction Sites



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

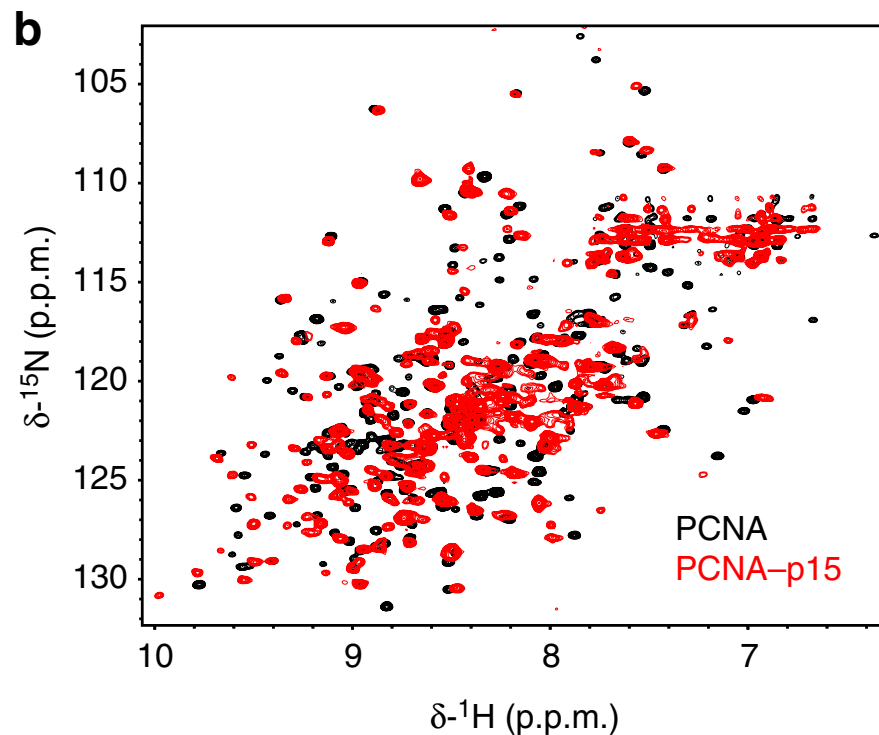
Interaction of p15^{PAF} with PCNA

- ✓ Mapping of PCNA binding on p15 (12 kDa) by NMR

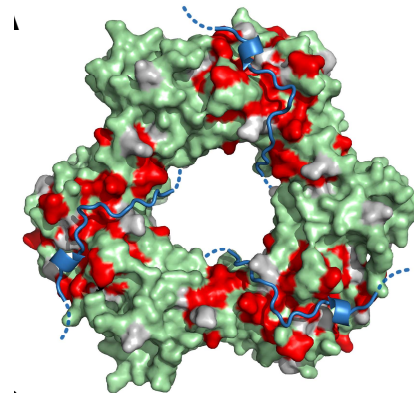


Interaction of p15^{PAF} on PCNA

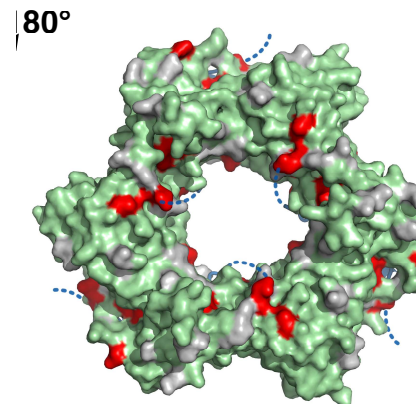
- ✓ Mapping of p15 binding on PCNA (86 kDa, ¹⁵N/²H) by NMR



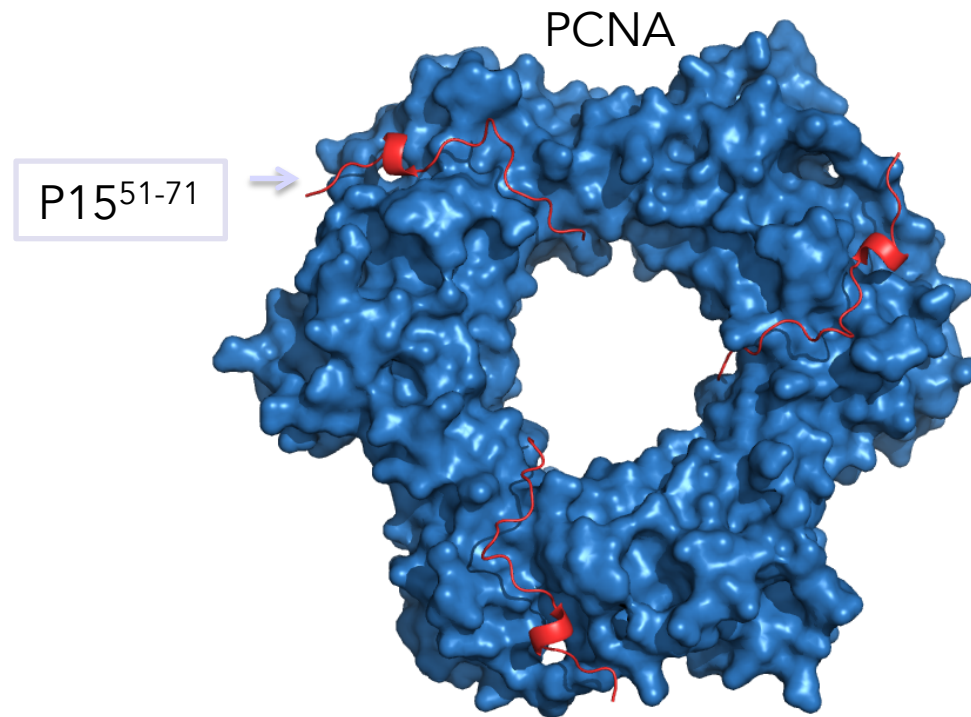
PCNA front-side



PCNA back-side

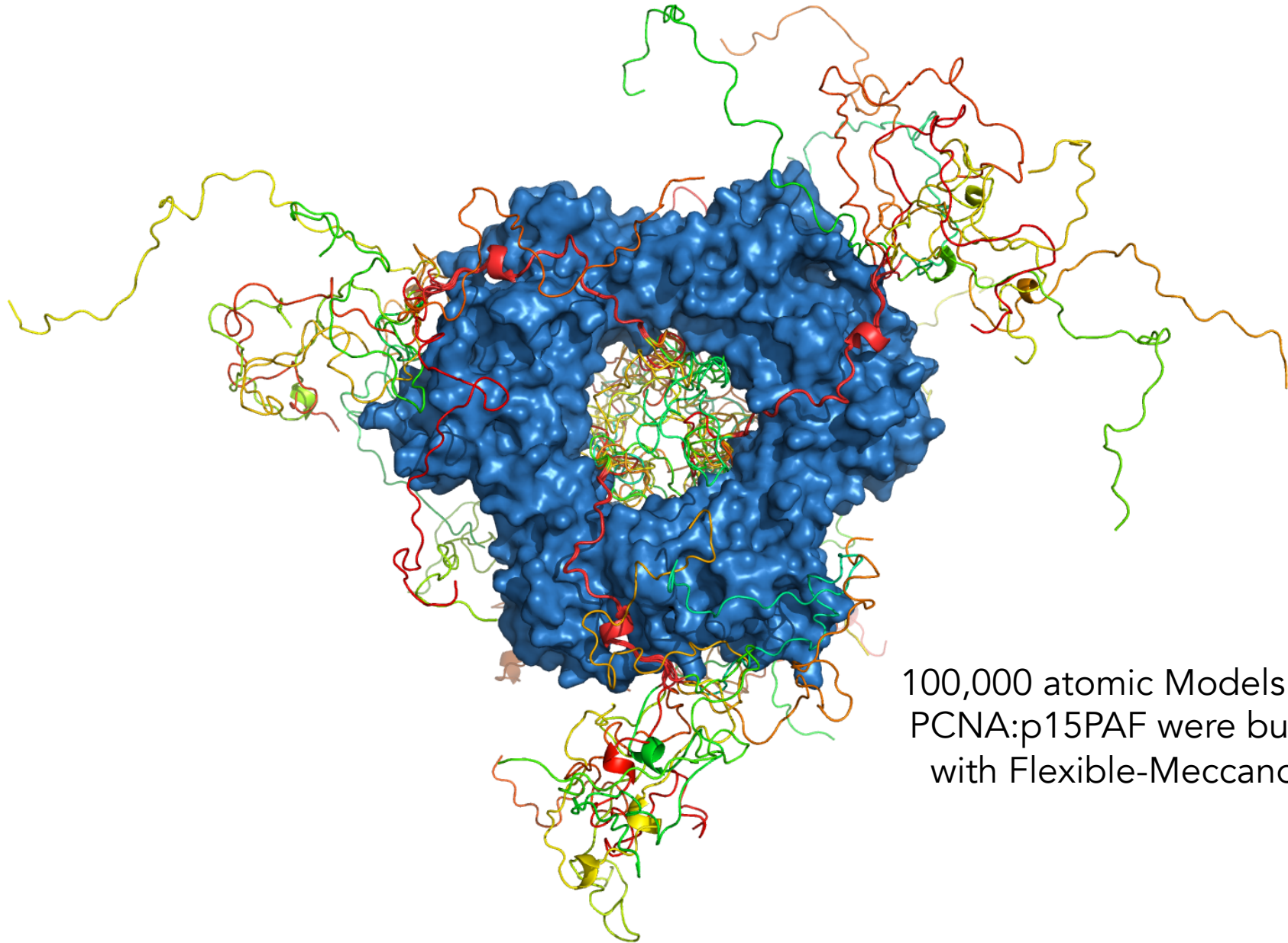


Crystallographic Structure of p15^{PAF} with PCNA



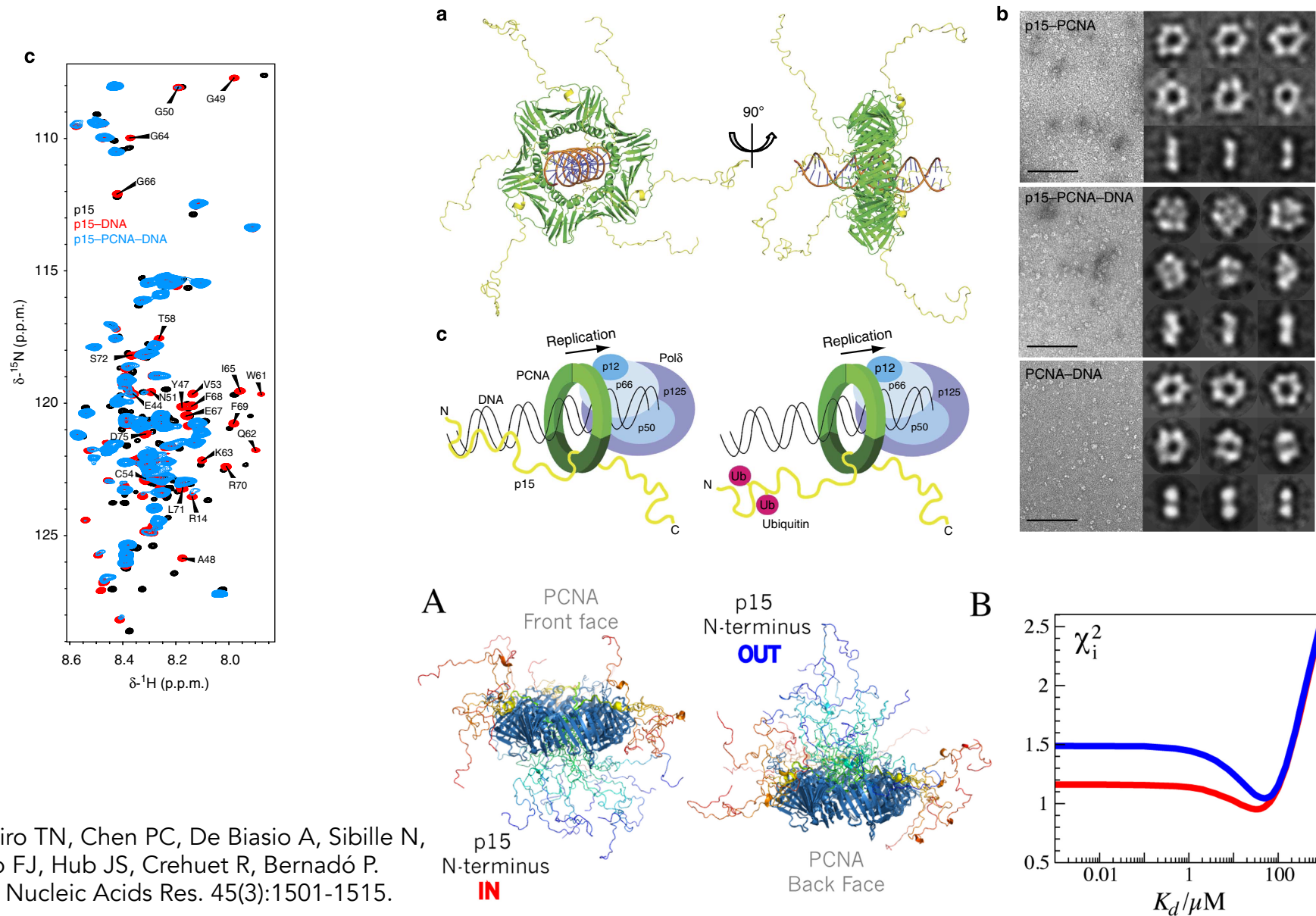
PDB:4D2G

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



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

Electron Microscopy

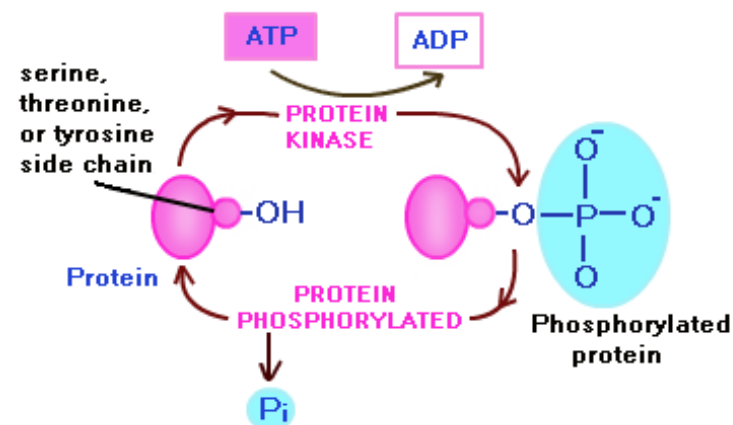


Cordeiro TN, Chen PC, De Biasio A, Sibille N, Blanco FJ, Hub JS, Crehuet R, Bernadó P. (2017) Nucleic Acids Res. 45(3):1501-1515.

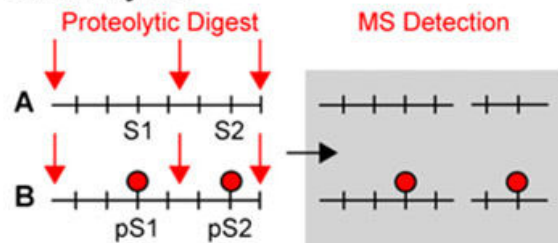
Phosphorylation

Effect of Post-translational Modifications

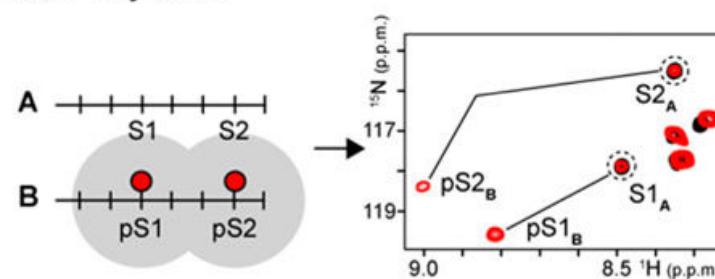
Even with recombinant kinases, mass spectrometry and immunodetection are not evident for determining the full phosphorylation pattern in a qualitative and quantitative manner.



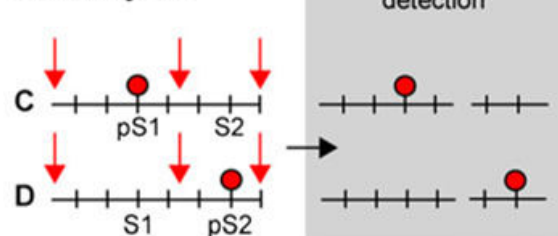
(a) Case 1 by MS



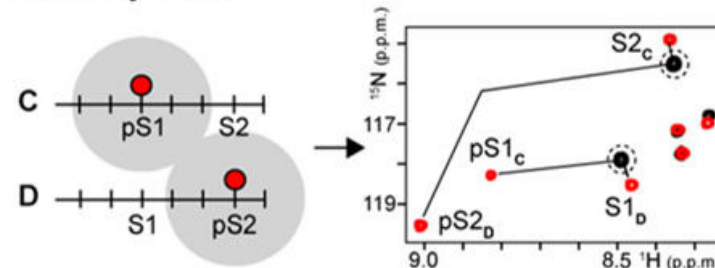
(b) Case 1 by NMR



Case 2 by MS



Case 2 by NMR

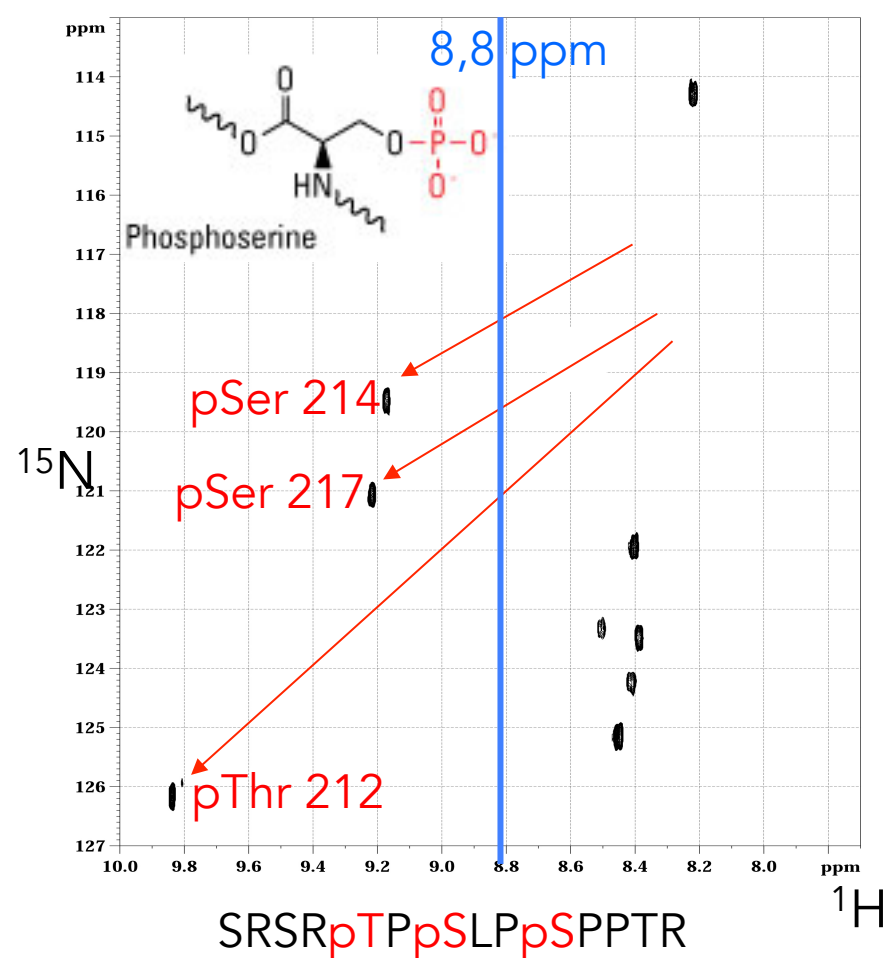
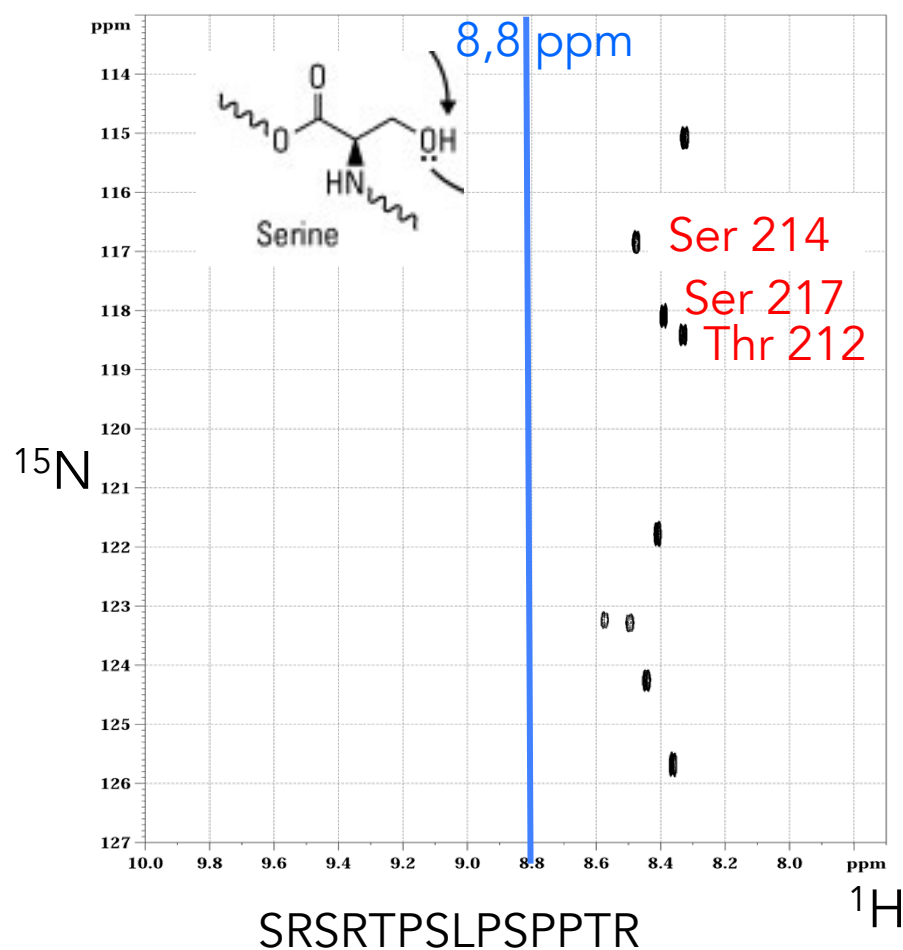


NMR is Qualitative and Quantitative



What sites? and to what extent?

→ In vitro phosphorylation by recombinant kinases: Identification and quantification

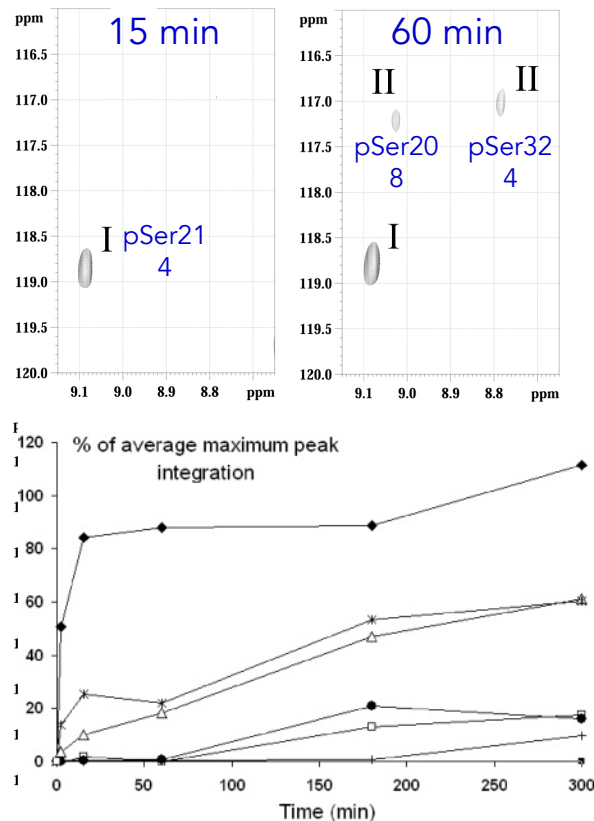


Kinetic of Phosphorylation

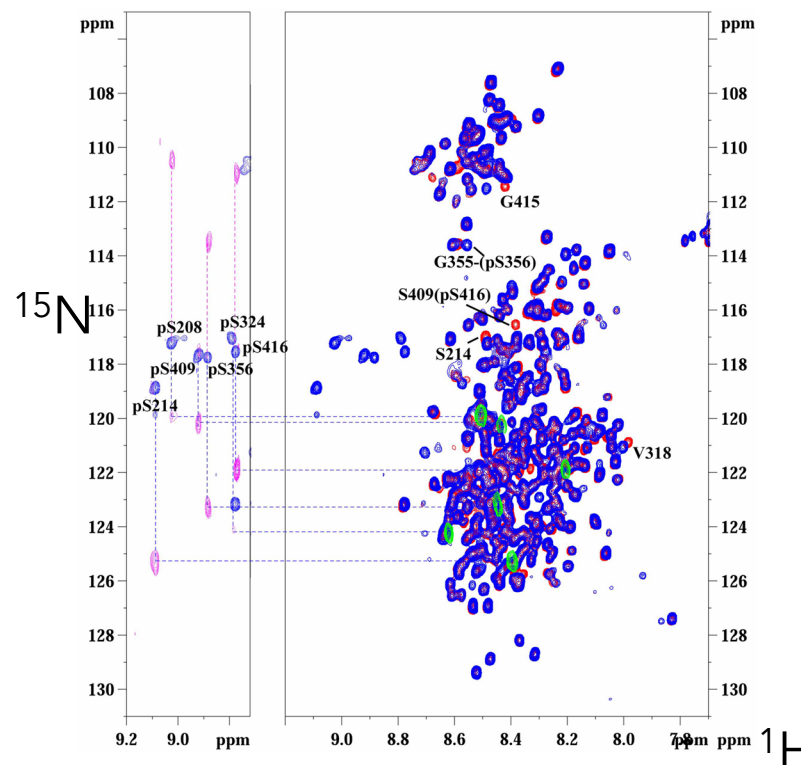


cAMP protein dependant kinase PKA

PKA from Prof. Langer/Schwalbe, Frankfurt, Germany



Sequential kinetic



Phosphorylated residues assignment
(triple resonance NMR experiments)

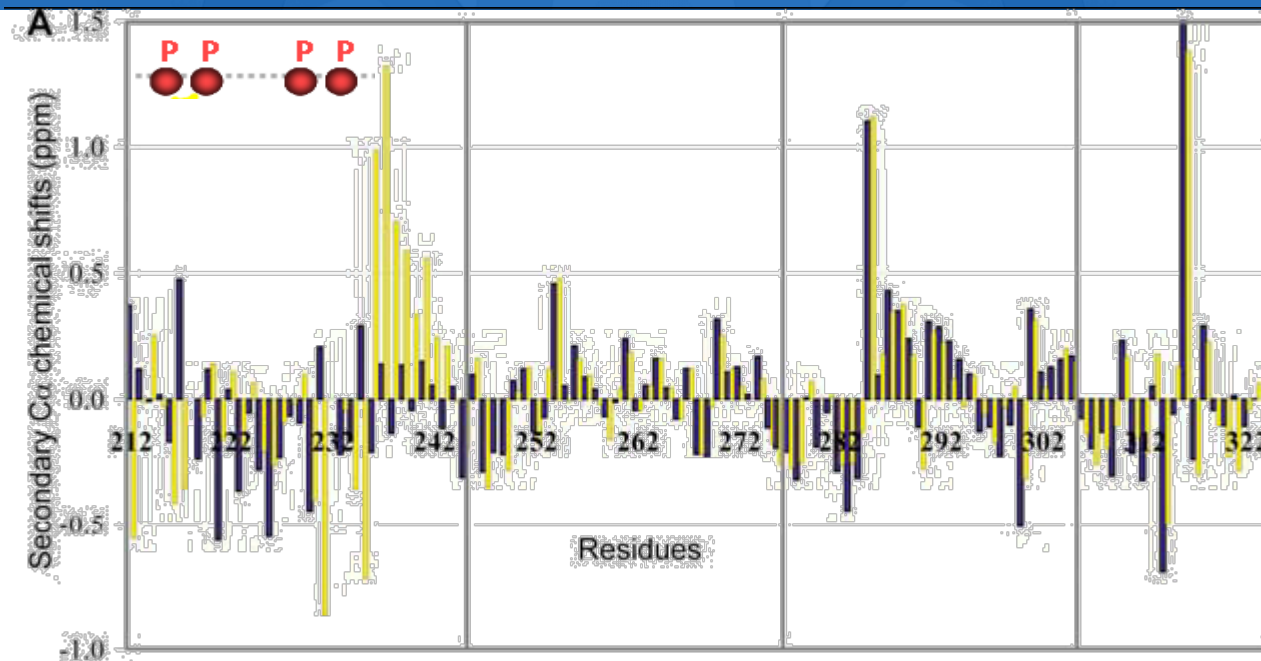
Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, N. Sibille, Schwalbe H, Saxena K, Langer T, Lippens G. NMR analysis of a Tau phosphorylation pattern. J Am Chem Soc 128:3575-83, 2006

Is there some degree of structuration ?

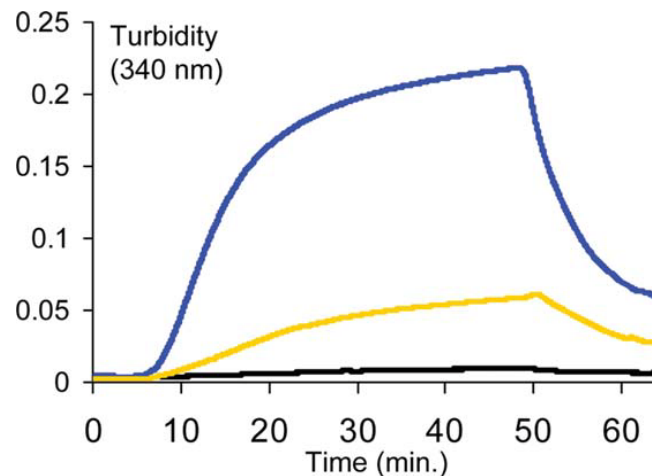
CDK2/CycA3
kinase on
Thr231 (AT180
epitope)
Thr235
Thr212
Thr217

Linked to
recognition
events

Here linked
to a loss of
function



Tubulin polymerization assay

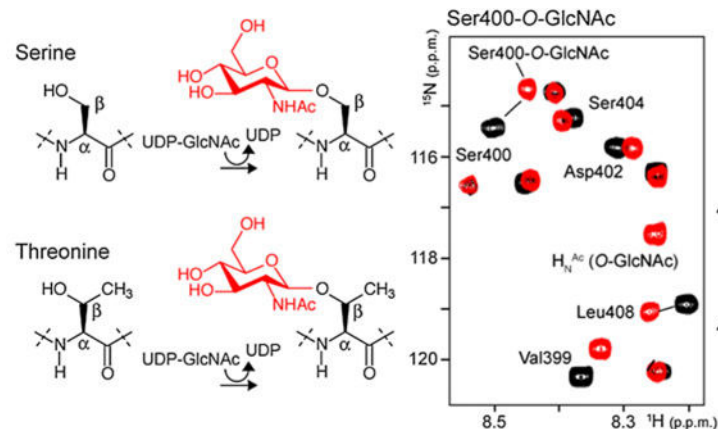


Where and which kind ?

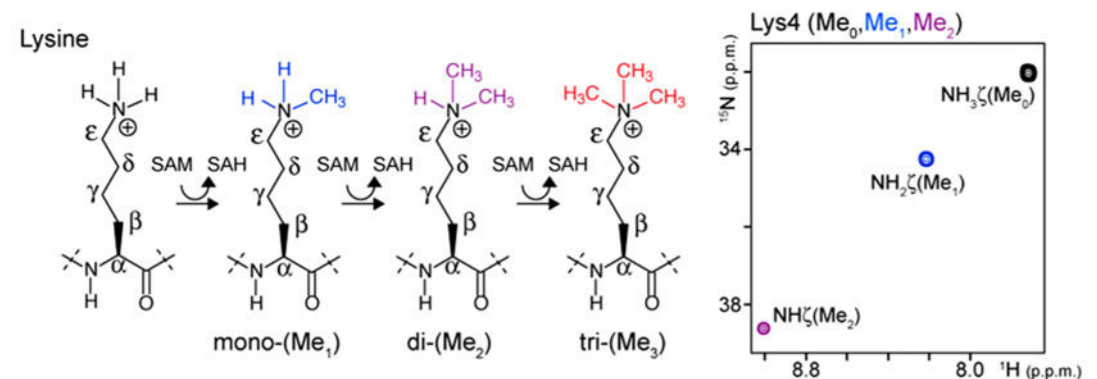
Other Post-translational Modification

Cell signaling, post-translational protein modifications and NMR spectroscopy
Francois-Xavier Theillet et al. J Biomol NMR. 2012 November ; 54(3): 217–236.

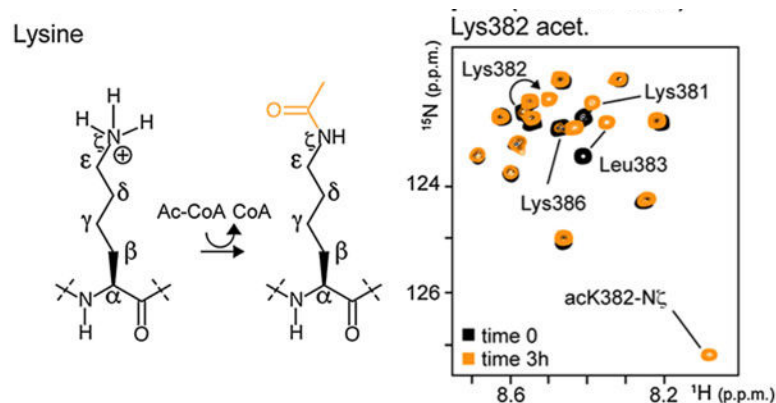
Serine, threonine glycosylation



Lysine and arginine alkylation



Lysine acylation



Serine, threonine, tyrosine and histidine phosphorylation

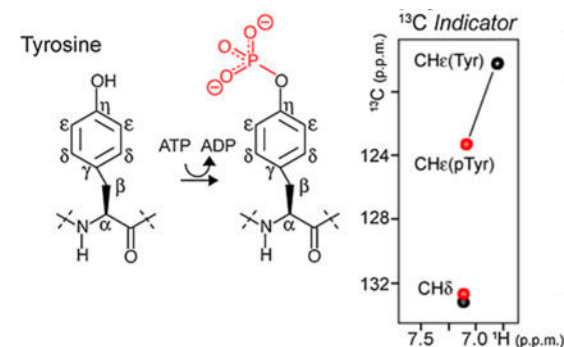


Illustration4science.com

Zhanna Santybayeva



Thanks for your attention

Questions

