Protein NMR Spectroscopy

Applications in Biological Sciences

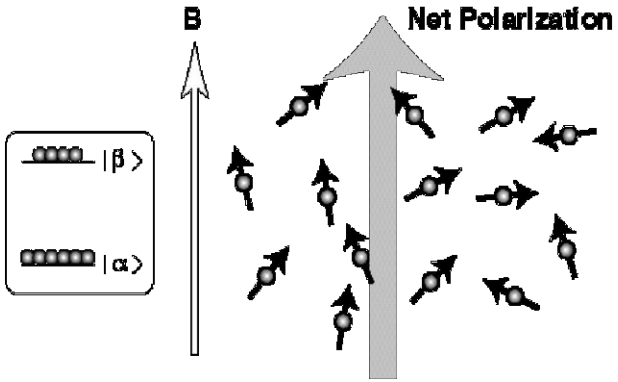
金长文

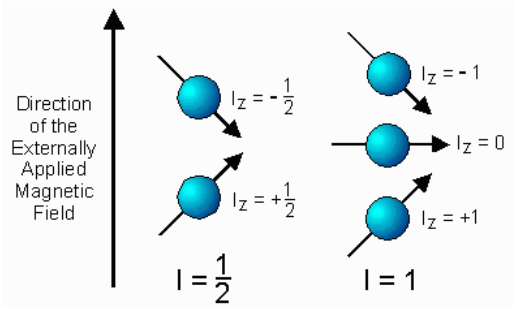
北京大学/北京核磁共振中心

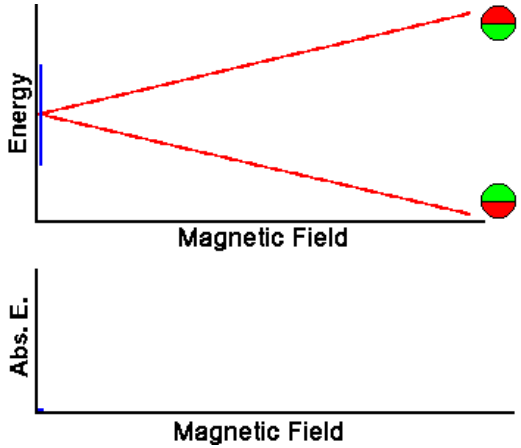
中科院微生物所

2011年7月

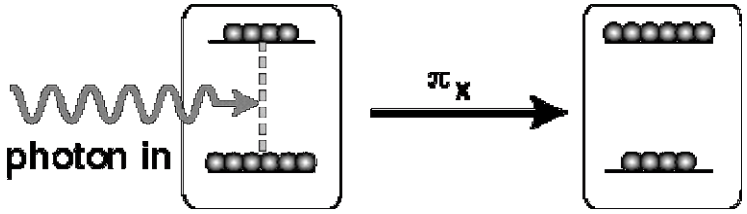
Nuclear Magnetic Resonance (NMR)

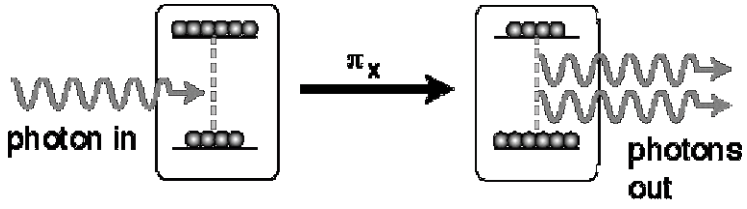






Boltzmann Distribution





- **Schrödinger equation: $E\Psi = H\Psi$**
- **Energy difference: $\Delta E = (h/2\pi)\,\gamma\, B = \hbar\omega$**
- **Larmor frequency: $\omega = \gamma B$**
- **Larmor frequency is the resonance frequency to create transitions between the energy levels.**
- **Larmor frequency ranges for ^{1}H encountered in modern NMR spectroscopy from 50 to 900 MHz.**

Bio-NMR active

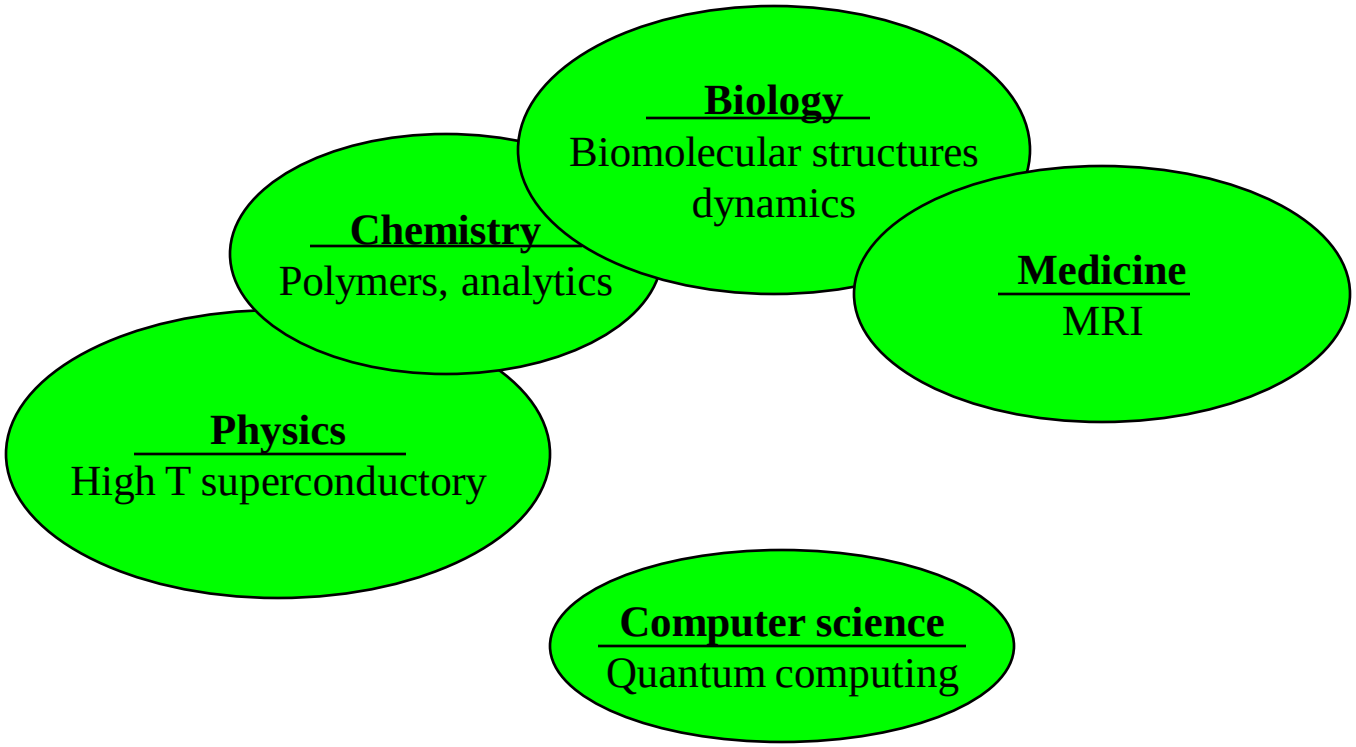
nucleus	I	γ (10^7/T·s)	abundance	sensitivity
^{1}H	1/2	26.752	99.98	1
^{2}H	1	4.107	0.02	0.00964
^{13}C	1/2	6.728	1.11	0.01559
^{15}N	1/2	-2.712	0.36	0.00104
^{31}P	1	1.0841	100	0.0664

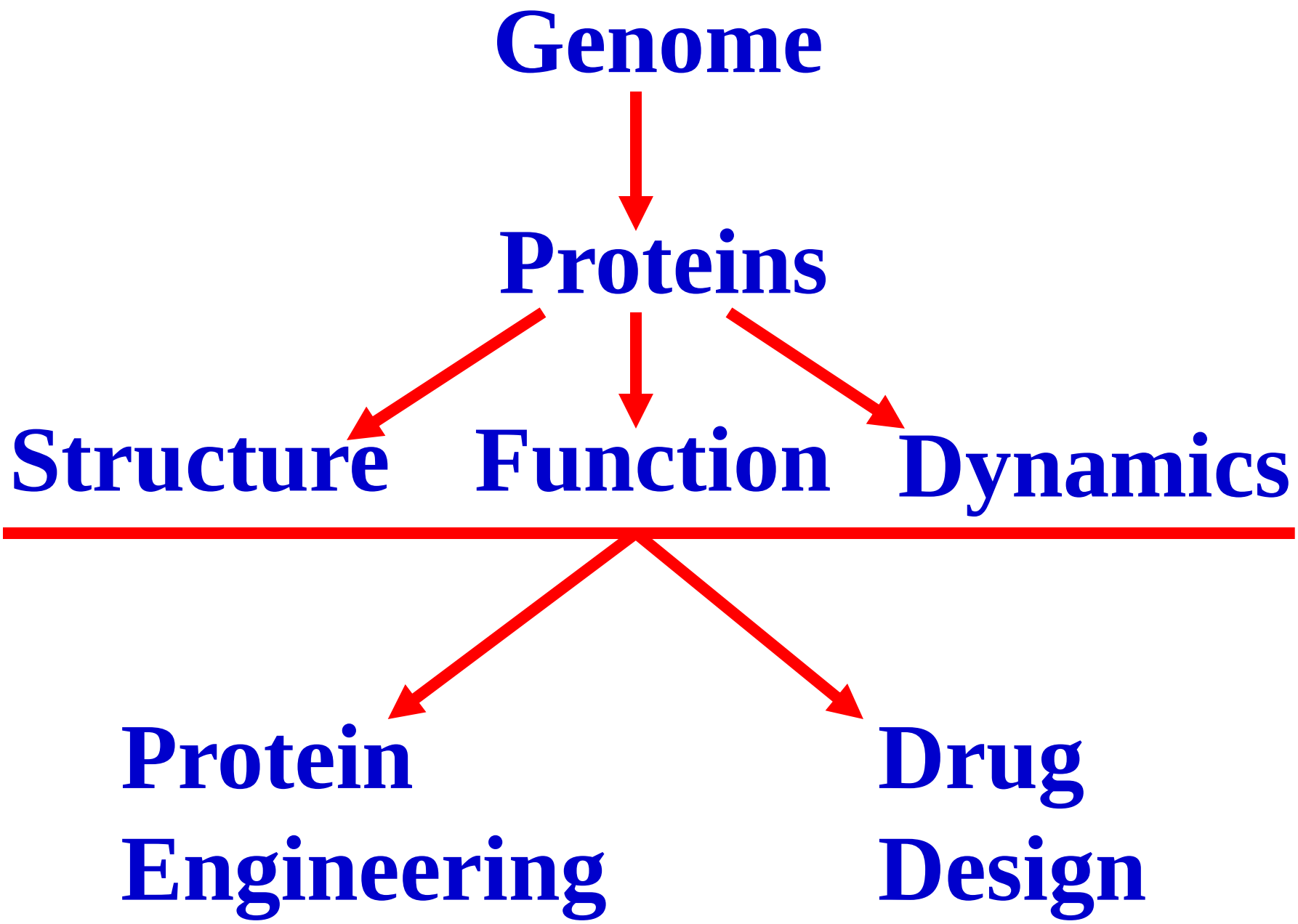

~ MHz
BRUKER
500
UltraShield™
BRUKER
CRYO PLATFORM
BRUKER
MHz – GHz
kHz

NMR SPECTROSCOPY

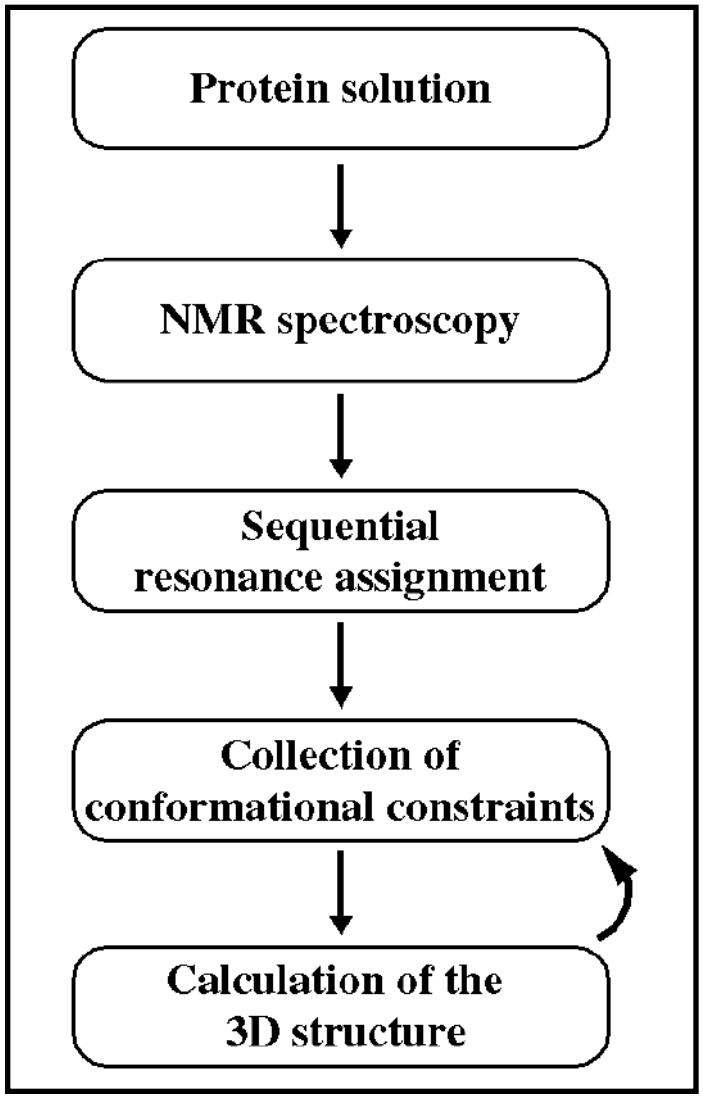
- **Key method for obtaining high resolution structure**
  -----in addition to X-ray Structure
- **Physiological temperature and condition**
  -----closer to native functional state
- **Structural information without solving structure**
  -----binding interaction, molecular motion, ect
- **Dynamics**
  -----motional properties
- **Protein size limitation: ~ 50 kDa**
  -----new methodology, higher magnetic field
- **Time consuming for data analysis**
  -----software for automated data analysis

The NMR World

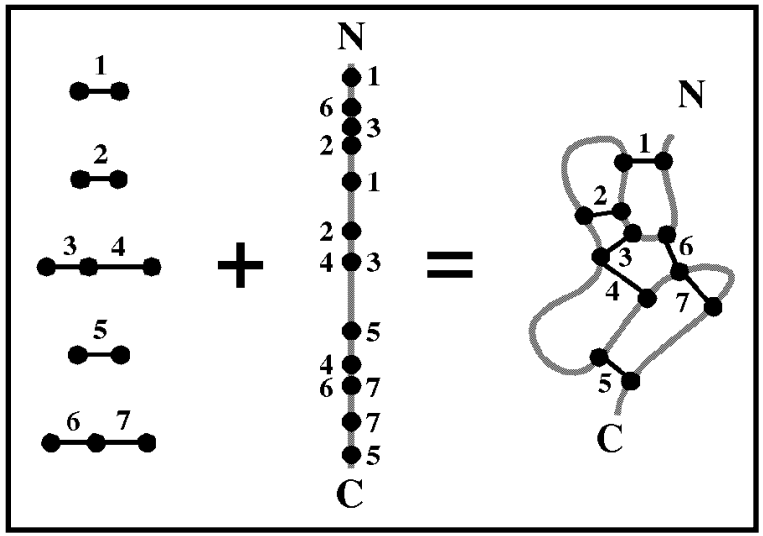


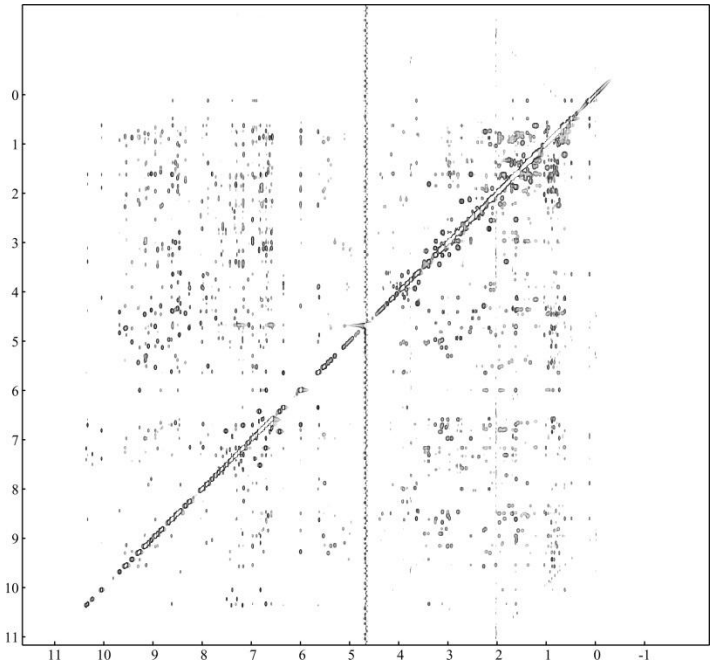

Genome
Proteins
Structure
Function
Dynamics
Protein Engineering
Drug Design

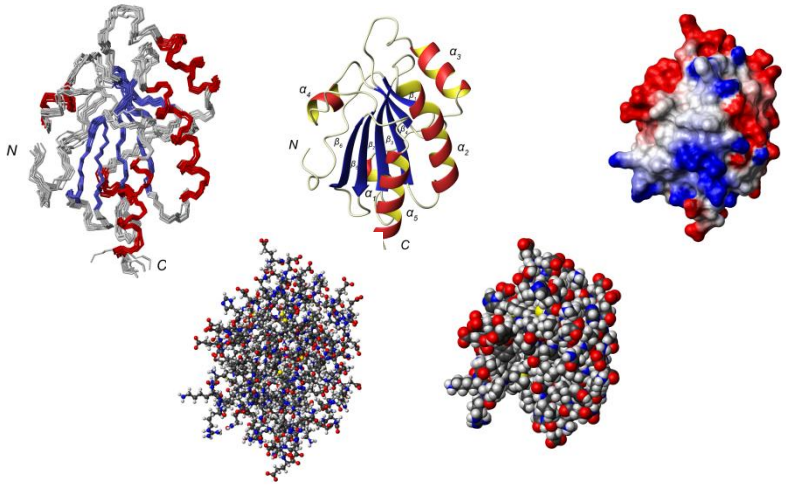
NMR Structure Determination



Protein Structure by NMR







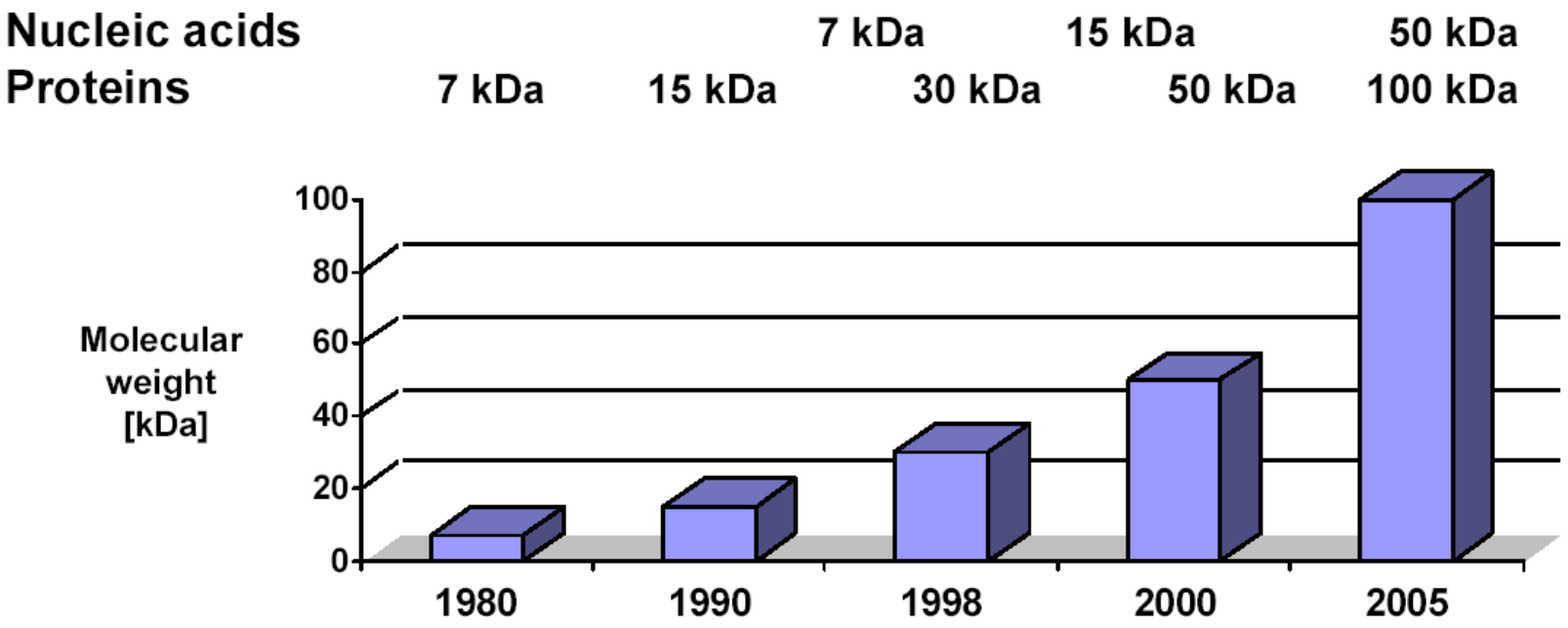
Progress in Structural Biology

PDB Current Holdings Breakdown

Exp.Method	Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
X-RAY	57898	1260	2784	17	61959
NMR	7657	934	168	7	8766
ELECTRON MICROSCOPY	245	22	86	0	353
HYBRID	28	2	1	1	32
other	132	4	5	13	154
Total	65960	2222	3044	38	71264

As of 2011

Progress in NMR

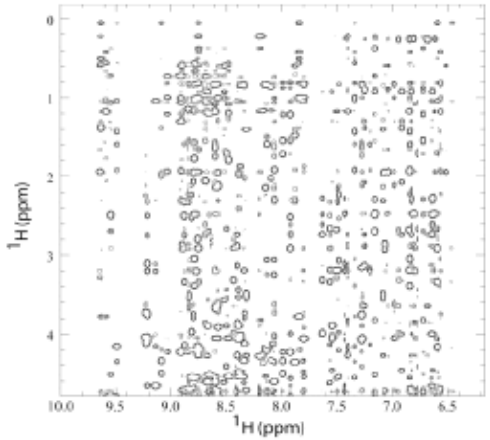


Recent Advances in Biomolecular NMR

- Large proteins
- Protein dynamics
- Residual dipolar couplings
- Protein-protein interactions and complexes
- Membrane proteins, In-Cell NMR, etc
- Automation & novel structure calculation methods
- Biomolecular NMR & drug discovery

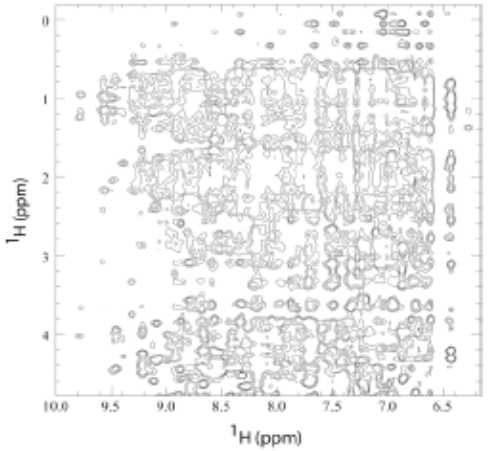
Large Proteins by NMR

- ^{1}H homonuclear 2D NMR → peptides and small proteins (< 10 kDa)
- ^{13}C/^{15}N labeling, 3D/4D NMR → proteins 10-30 kDa
- Main obstacles for larger proteins:
  1. Decrease in T_2 → line broadening, loss of sensitivity and resolution
  2. Overcrowded signals → loss of resolution



8 kDa (Tendamistat)

2D NOESY



21 kDa (Cdc42)

Large Proteins by NMR

- Hardware --- higher magnetic field, cryo-probe, ^{13}C detection, etc

Significant increasing of the cost!

Still, Other factors…

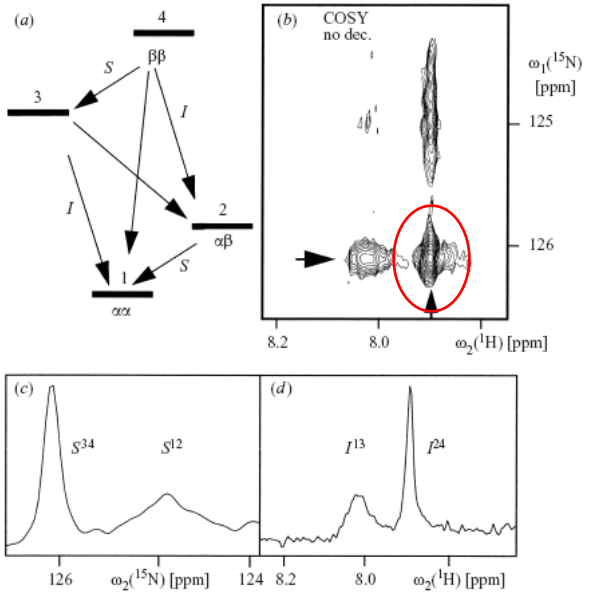
~$800,000 ~$2,000,000 ~$4,500,000 ~$8,000,000

Complicated…

Large Proteins by NMR

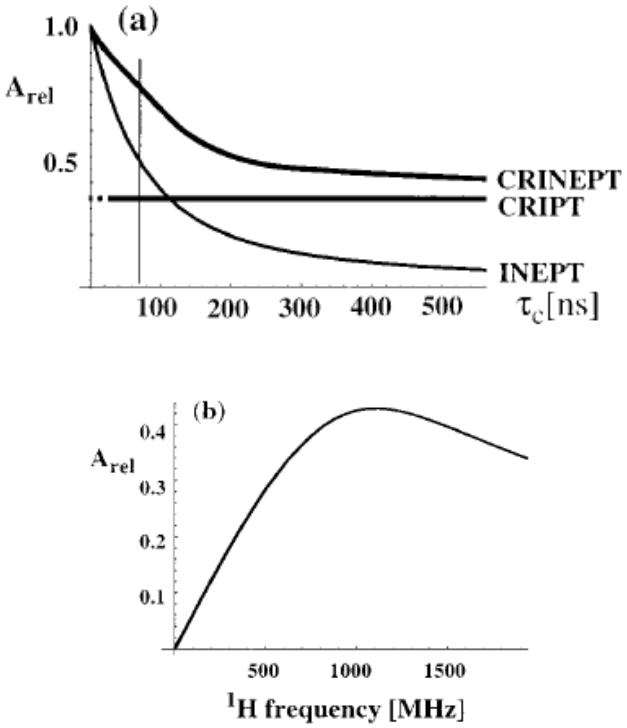
- NMR techniques --- TROSY, CRINEPT, etc

TROSY： Cancellation of DD and CSA, select the multiplet component with the narrowest linewidth



Pervushin (2000) *Q Rev Biophys*

CRINEPT： Cross-Relaxation Enhanced Polarization



Riek et al (1999) *PNAS*

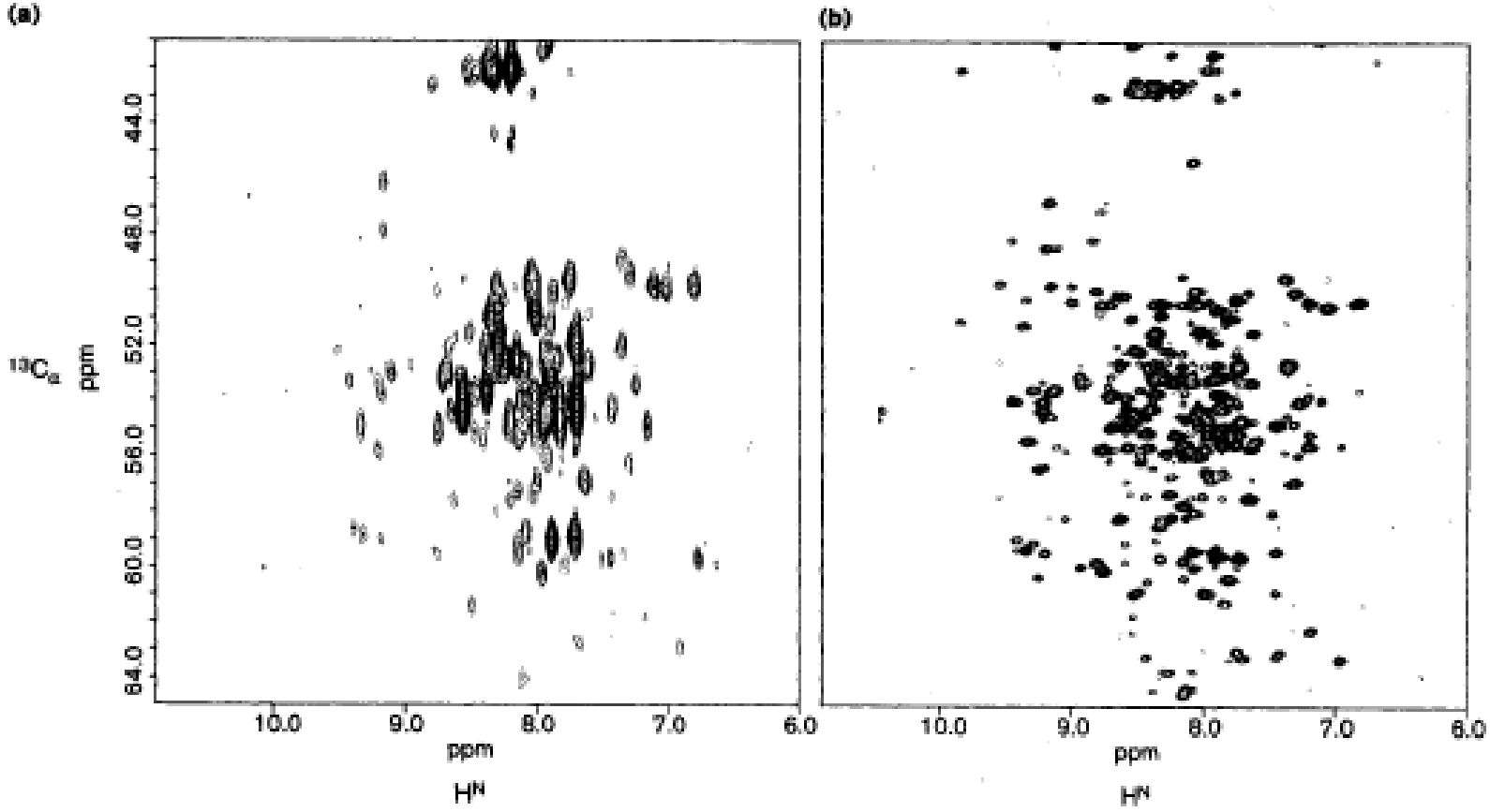


Figure 2. Comparison of 2D ($^{13}C_\alpha$, NH) projections from 3D HNCA spectra recorded of a 23-kDa Shc PTB domain/phosphotyrosine peptide complex. (a) ^{15}N, ^{13}C, ^{1}H uniformly labeled Shc PTB domain. (b) ^{15}N, ^{13}C, 75% ^{2}H uniformly labeled Shc PTB domain. Reprinted with permission from Sattler and Fesik (1996).

SAIL

- SAIL: stereo-array isotope labeling
- Reduced spectral complexity
- Improved sensitivity and resolution

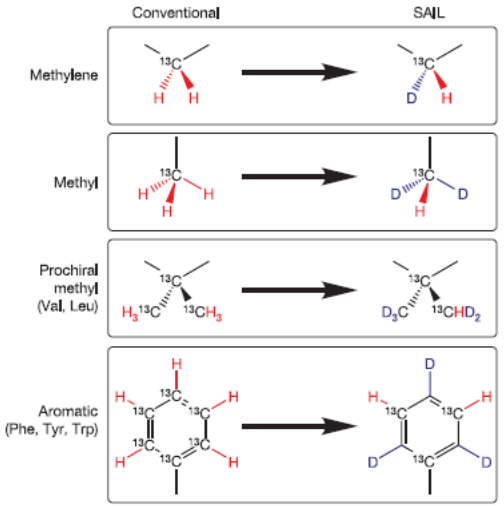
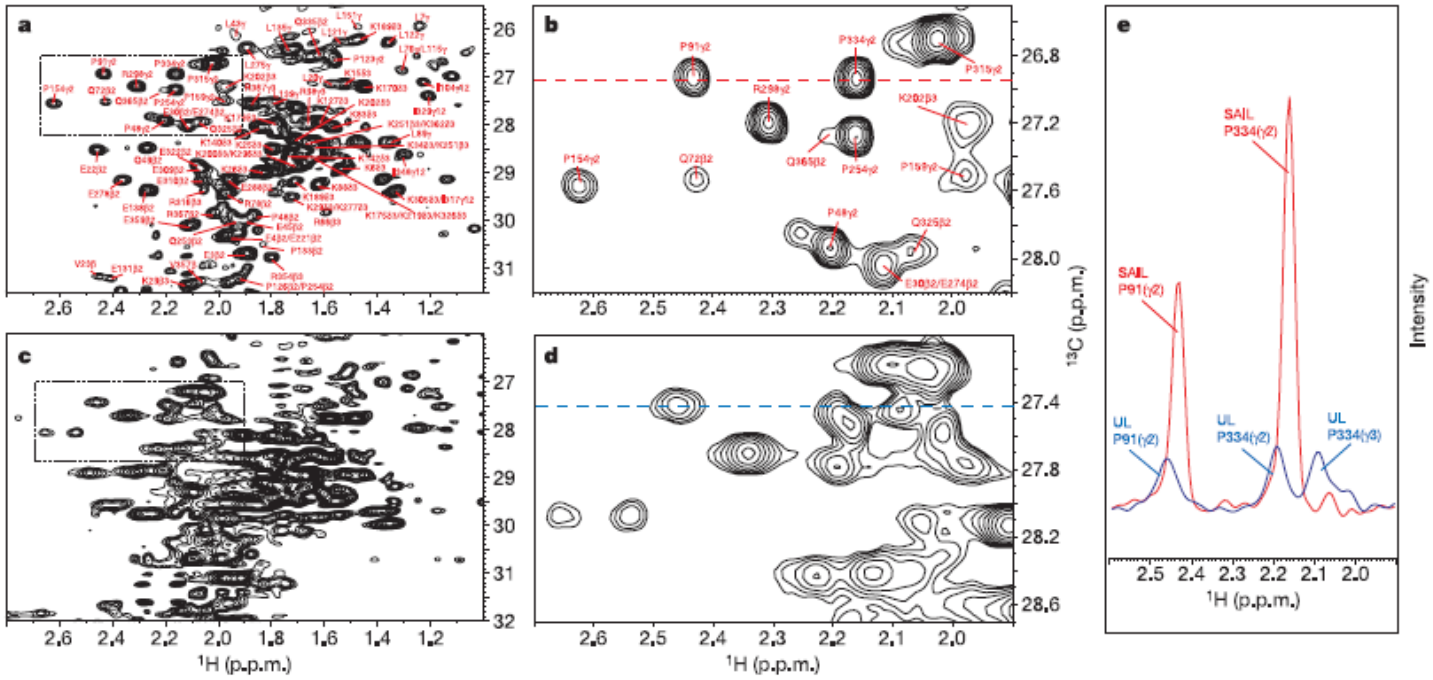


Figure 1 | SAIL amino acids. Design concepts embodied in the SAIL amino acids incorporated into CaM and MBP[13–15].

SAIL

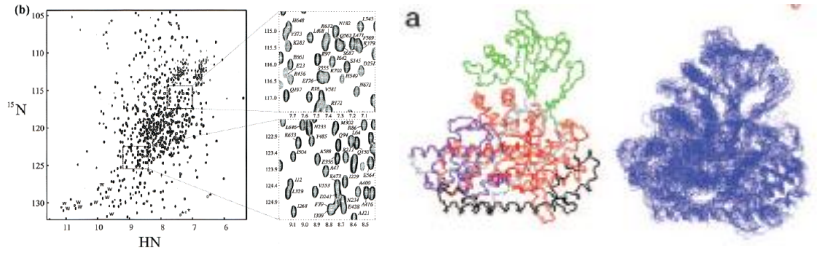
Conventional



Kainosho et al (2006) *Nature*

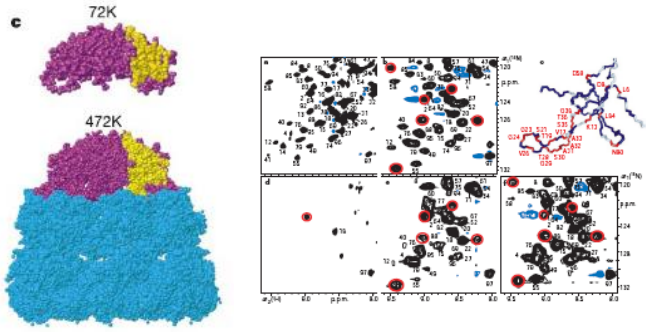
Large Proteins by NMR

1. Global fold of an 82-kDa enzyme malate synthase G



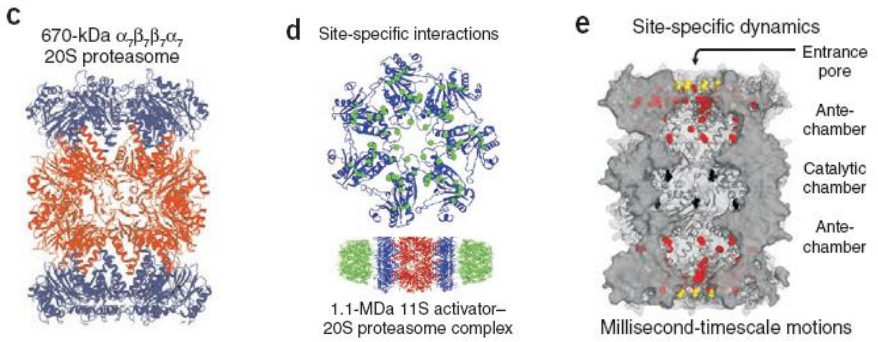
Tugarinov et al (2002) *JACS*
Tugarinov et al (2005) *PNAS*

2. Mapping binding interface 900-kDa GroEL-GroES complex



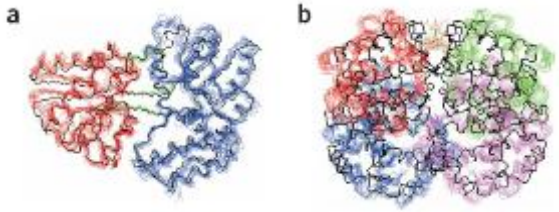
Fiaux et al (2002) *Nature*

3. Site-specific interactions and dynamics in 670-kDa 20S proteasome



Sprangers et al (2007) *Nature Methods*

4. NMR structures of 42-kDa MBP & 65-kDa hemoglobin using fully-protonated samples



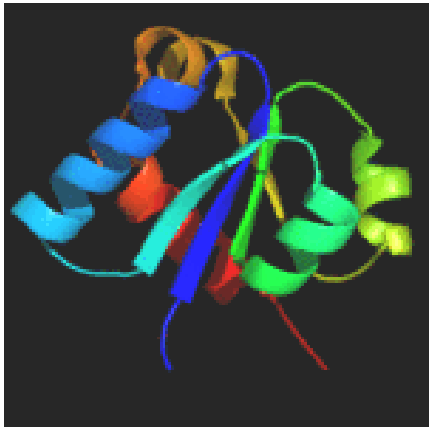
Xu et al (2006) *Nature Methods*

Protein Dynamics

Why Dynamics?

- Static three-dimensional structures alone often do not completely explain results from the biological assays, nor do they illuminate the path for protein engineering or rational drug design.

- The goal of dynamics studies is to bridge the gap between the static and dynamic pictures of molecular structure and to demonstrate how motion relates function.



Protein Dynamics by NMR

- Protein dynamics & enzyme catalysis
- Protein dynamics & folding
- Protein dynamics & protein-protein/protein-ligand interactions
- Protein dynamics in intrinsically disordered proteins
- …

Relaxation Dispersion

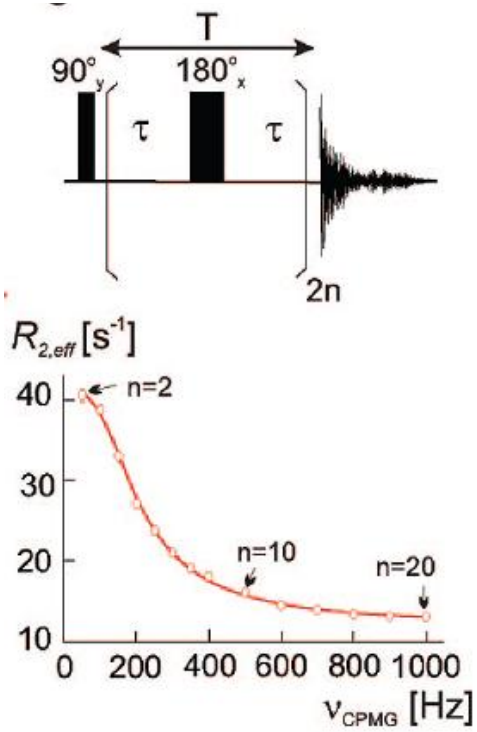
- Many of the important biological processes, (eg. protein interaction and signaling, enzyme catalysis, folding, allostery) involve conversion to low-populated, transient ‘excited’ states.

- These ‘invisible’ low-populated states are difficult to study using conventional methods, but can be probed using relaxation dispersion experiments by NMR.

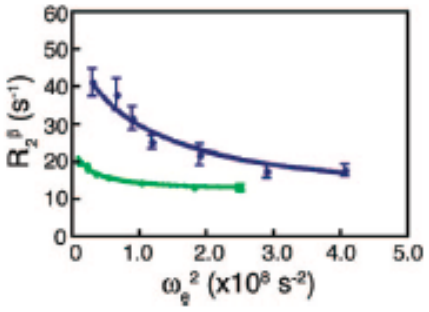
Relaxation Dispersion

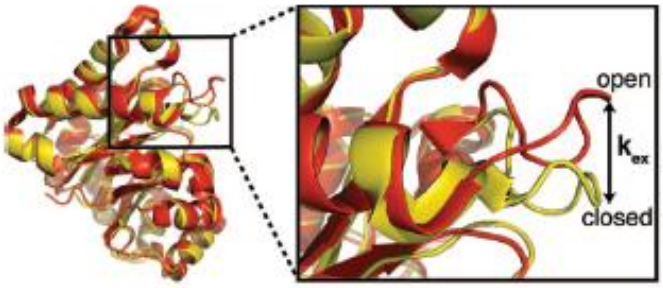
Quantifying micro- to milli-second timescale conformational exchanges.

Providing kinetic, thermodynamic and structural information of the conformational exchange process.



$$R_2(1/\tau_{cp}) = R_2^0 + \varphi_{ex}/k_{ex}[1 - 2\tanh(k_{ex}\tau_{cp}/2)/(k_{ex}\tau_{cp})]$$



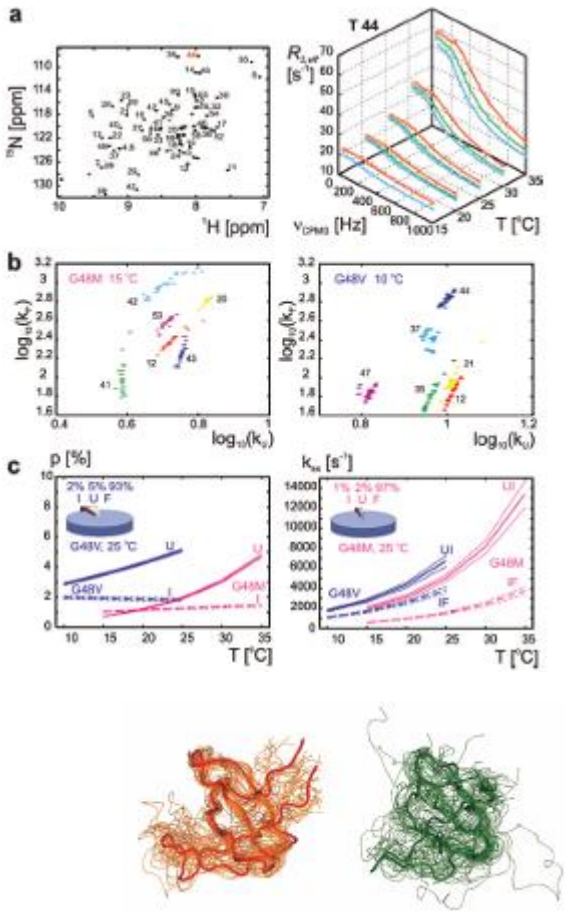


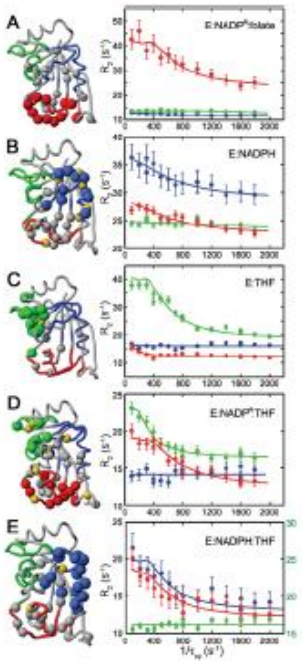
Korzhnev & Kay (2007) *Acc Chem Res*

Loria et al (2008) *Acc Chem Res*

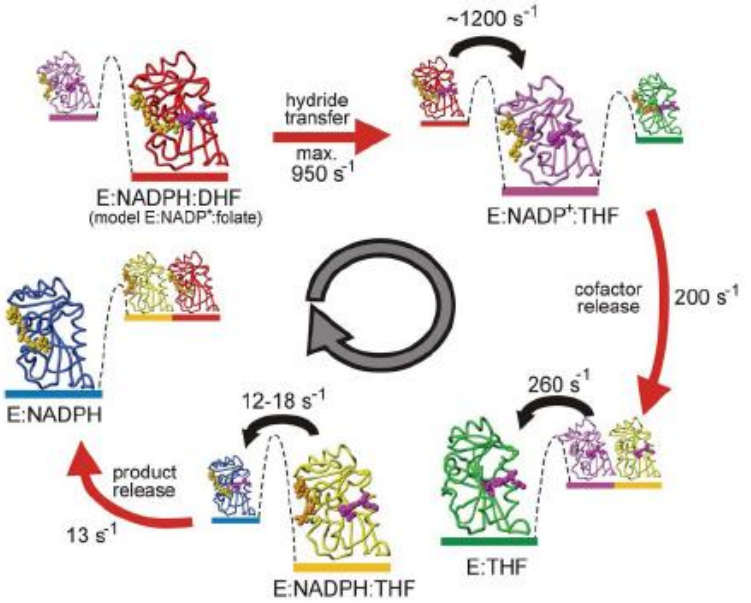
Relaxation Dispersion: Applications

Protein folding intermediates





Enzyme catalysis

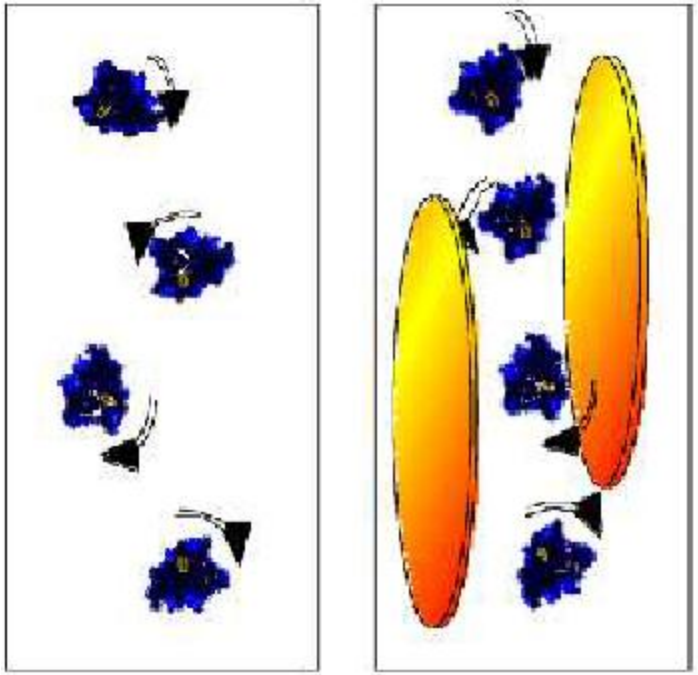


Korzhnev & Kay (2007) *Acc Chem Res*

Boehr et al (2006) *Science*

Residual Dipolar Couplings

Theoretical background



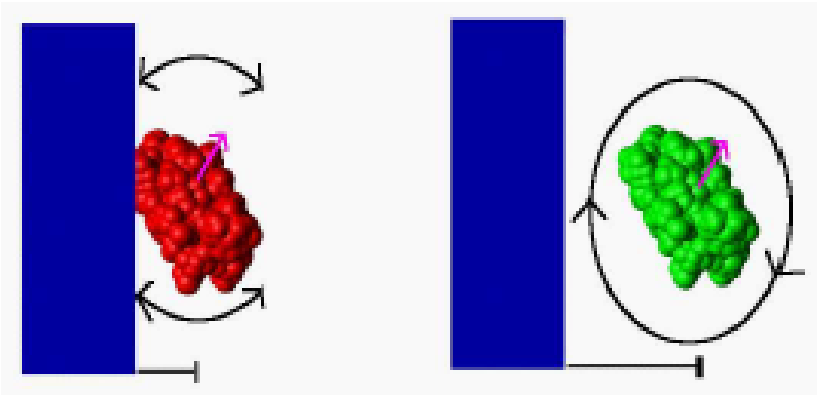
Isotropic tumbling

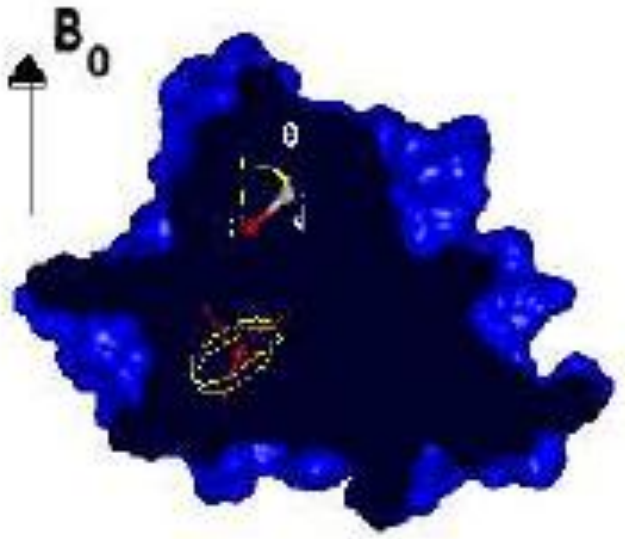
All orientations of the molecules are equivalent

Anisotropic tumbling

Some orientations are preferred to others

Alignment tensor





- The molecule prefers to be oriented in one direction of the space.
- Need of a mathematical expression (ex. according to magnetic field B_0).
- Not in laboratory frame but in molecular frame: easier to manipulate

What happens to the spectra?

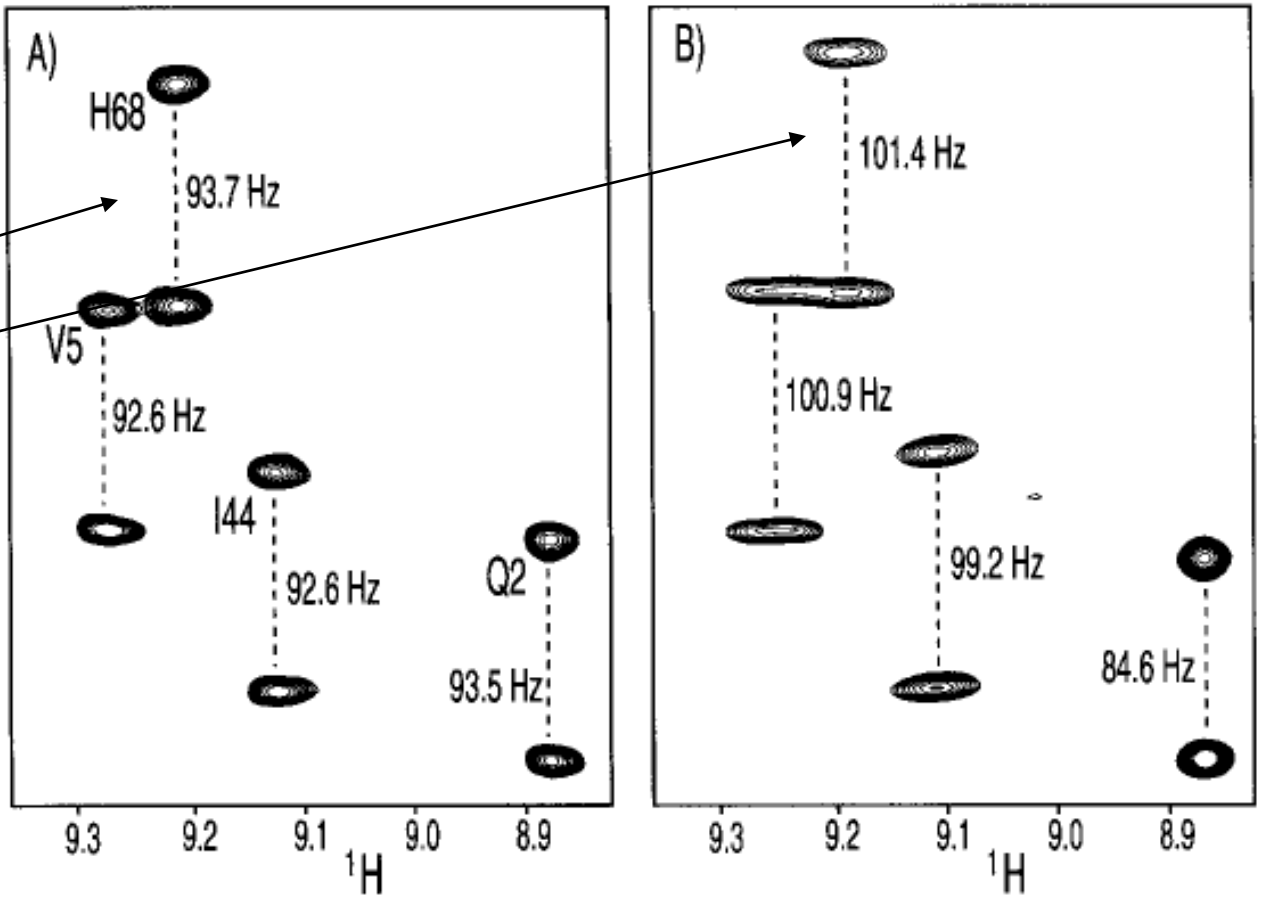
^{15}N-HSQC undecoupled in the indirect dimension

J_{HN}

$D+J_{HN}$

D=-7.7Hz

J_{HN}<0 !!!

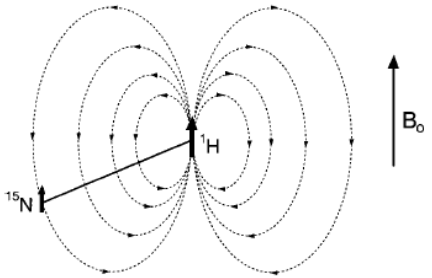


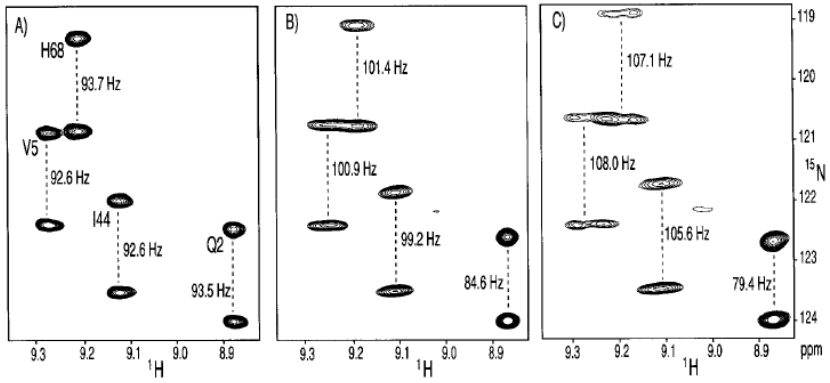
Isotropic
(in pure water)

Anisotropic
(in aligned media)

Residual Dipolar Couplings

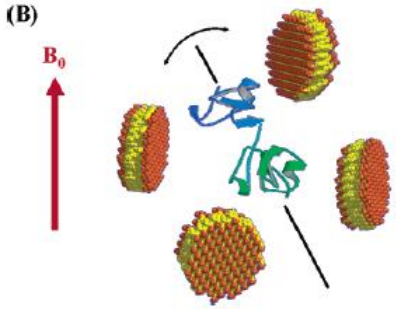
Magnetic dipole-dipole coupling





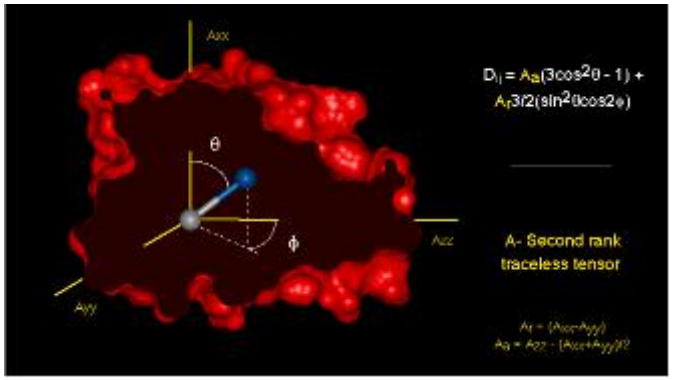
Isotropic **Anisotropic**

Bax (2003) *Protein Sci*



In slightly anisotropic solution, the proteins are weakly aligned.

The alignment tensor:



http://www.nmr.chem.uu.nl/education/module/module_doc/theory.html

Dosset et al (2001) *JBNMR*

What are RDCs for?

Potentially, everything related to:

- the orientation of an intramolecular vector (ex. HN)
    - structure validation
    - structure refinement (additional experimental restraints to NOE, dihedral angle, …)
    - relative domain orientation
    - motion of the vector
    - analysis of conformational equilibrium

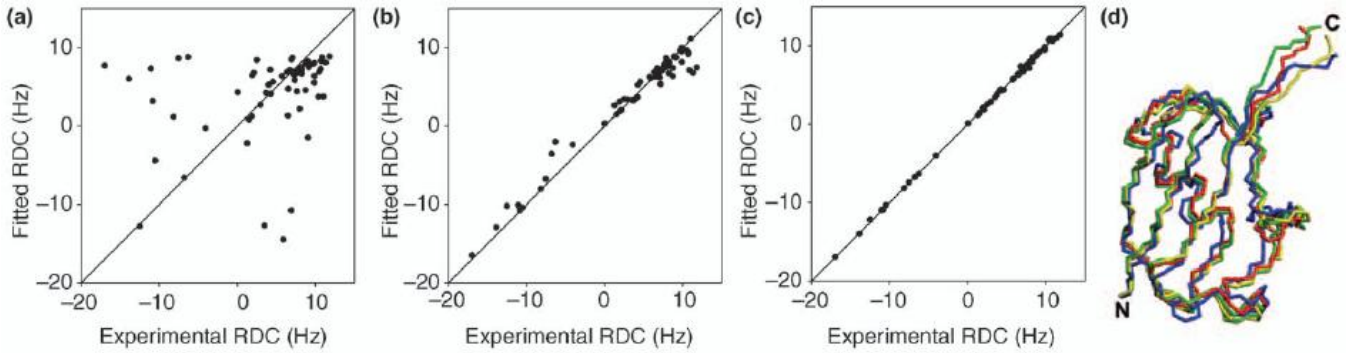
- the shape of the molecule
    - distinguish between a monomer, dimer….

Power of RDC: global information (contained in alignment tensor)

Residual Dipolar Couplings

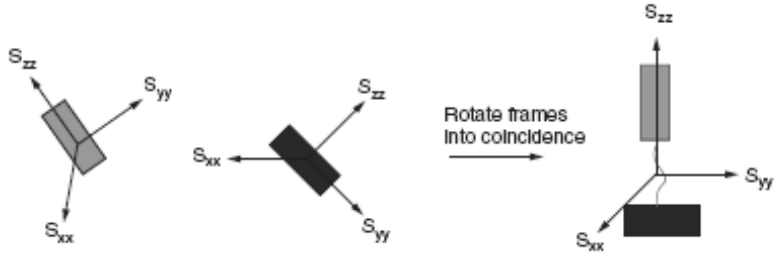
- Applications:

✓Protein structure refinement



Bax & Grishaev (2005) *COSB*

✓Intermolecular complexes and domain-domain orientations

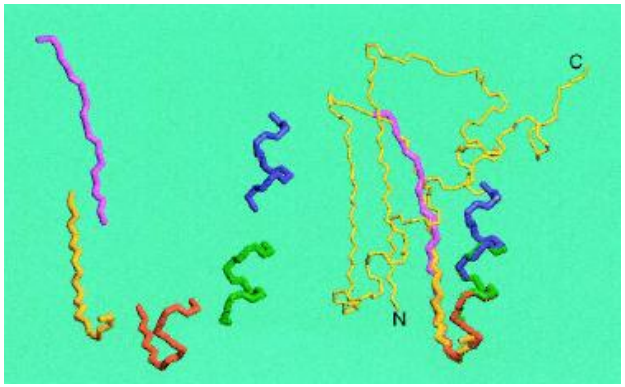


Tolman (2001) *COSB*

Residual Dipolar Couplings

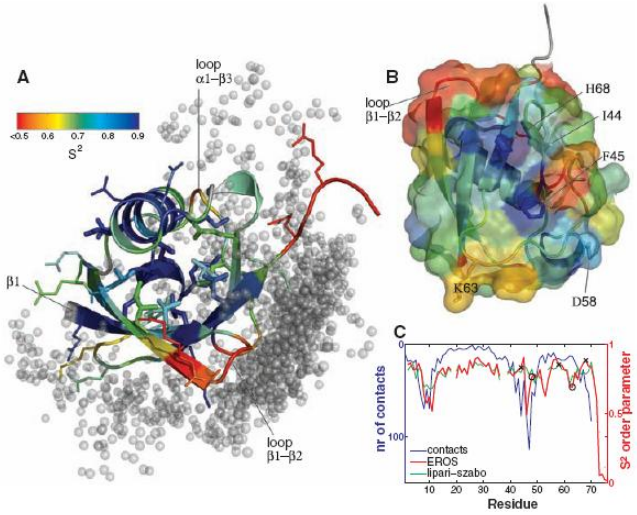
- Applications:

✓Protein structure by molecular fragment replacement (MFR)



Bax (2003) *Protein Sci*

✓Protein dynamics



Lange et al (2008) *Science*

Paramagnetic Relaxation Enhancement (PRE)

PRE

- Arise from unpaired electrons: paramagnetic centers
- Effective distance: up to 35 Å (depending on the paramagnetic group)

PRE through direct dipole-dipole interactions

Longitudinal PRE rate

$$\Gamma_1 = \frac{2}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1) J_{SB}(\omega_I)$$

Transverse PRE rate

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1)\{4J_{SB}(0) + 3J_{SB}(\omega_I)\}$$

$$J_{SB}(\omega) = r^{-6}\frac{\tau_c}{1 + (\omega\tau_c)^2}$$

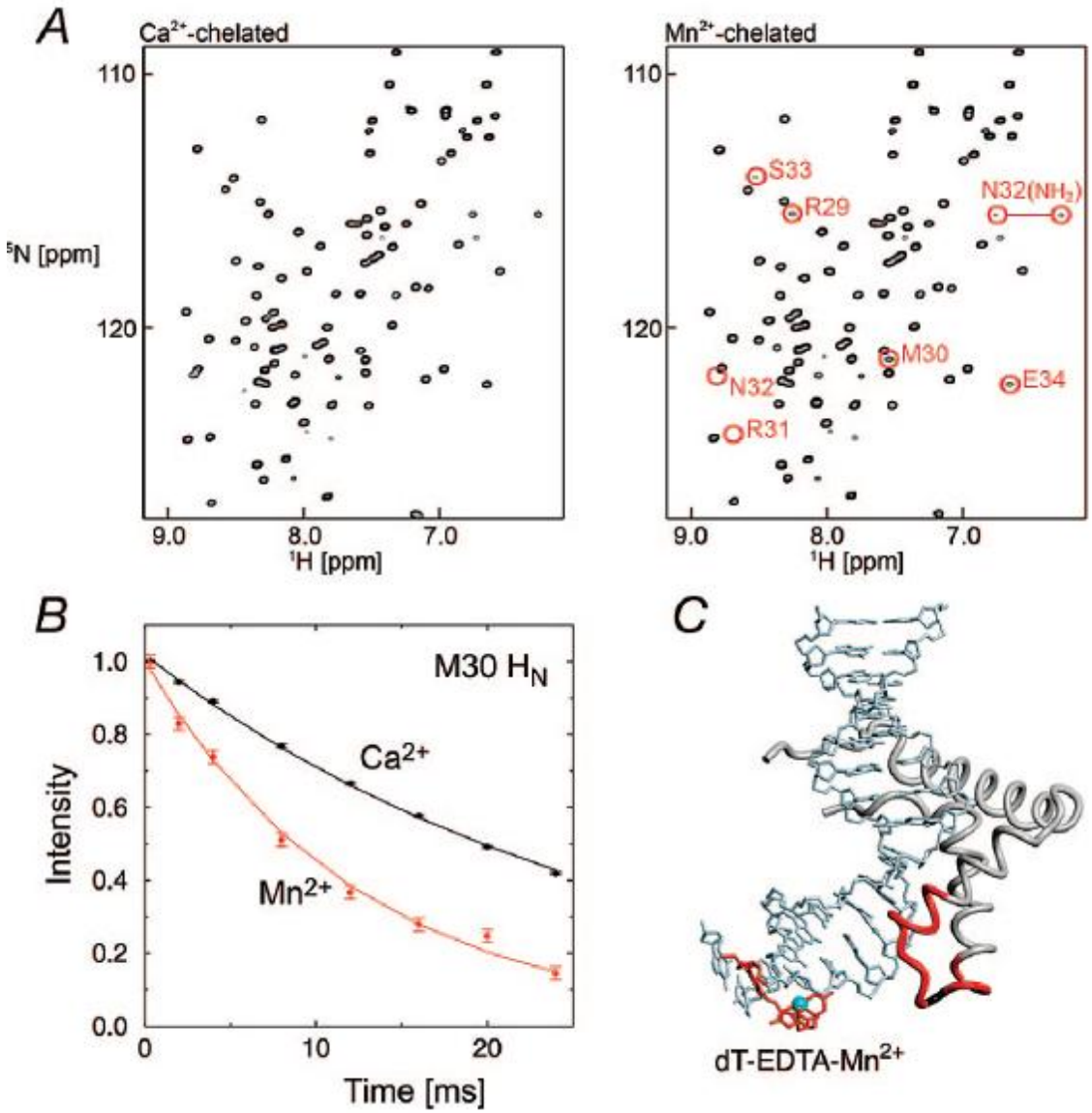
g: the electron *g*-factor

$\tau_c = (\tau_r^{-1} + \tau_s^{-1})^{-1}$

τ_r : the rotational correlation time of the macromolecule

τ_s : the effective electron relaxation time.

Diamagnetic Paramagnetic

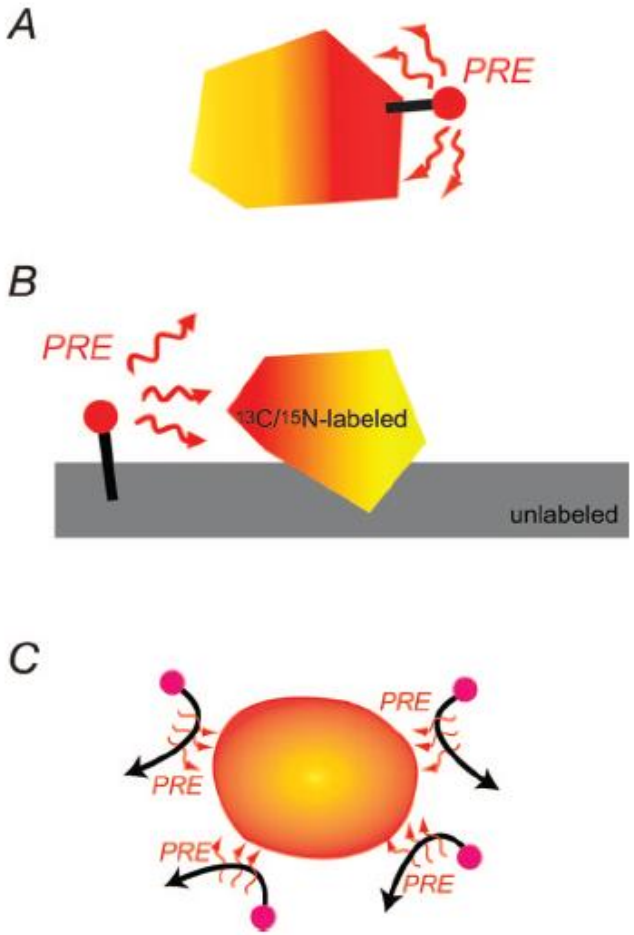


Different types of PRE

Intra-molecular PRE

Inter-molecular PRE

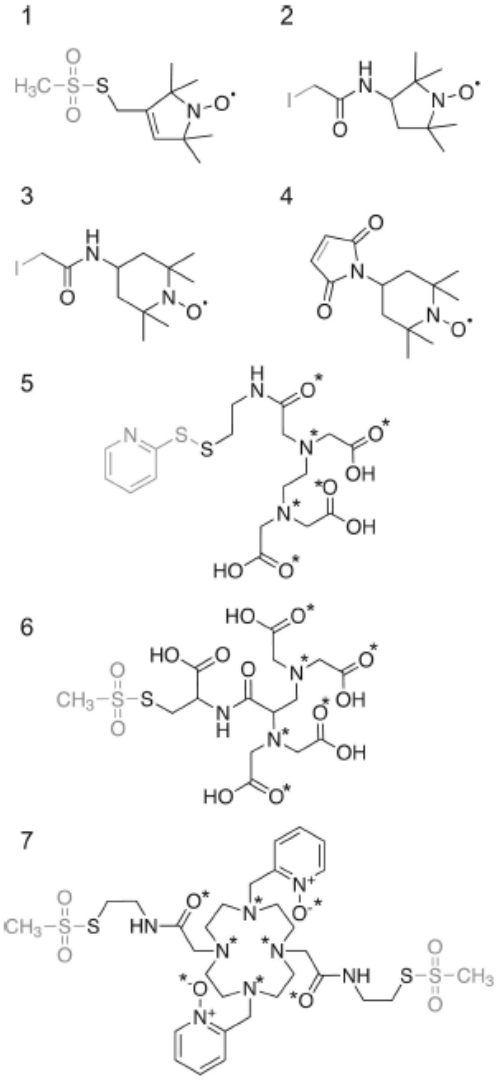
Solvent PRe



Paramagnetic Probes

(1) nitroxide stable radicals, >N-O•

(2) metal chelators (such as EDTA, DTPA, and metalbinding peptides) that bind paramagnetic metal ions with very high affinity.

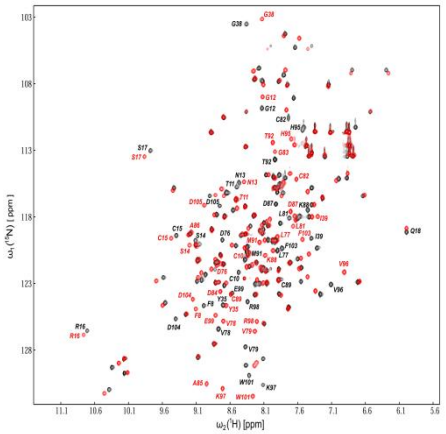


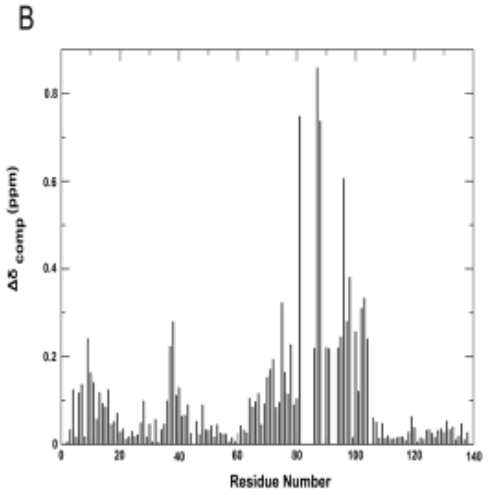
Applications of PRE

- New long-range structural restraints for protein structure determination
- Protein-Nucleic acid interactions
- Encounter complexes in protein-protein interaction / Transient, weak interactions
- Dynamics information, lowly populated states
- …

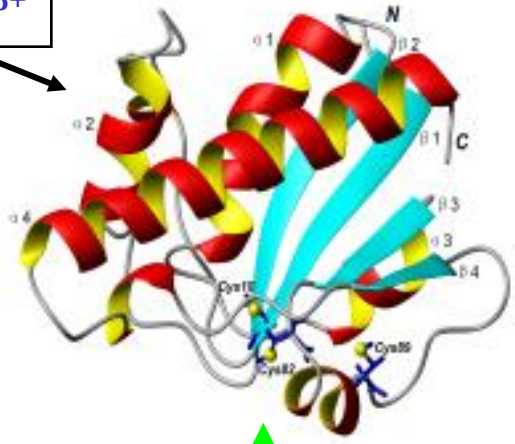
Protein Interactions and Complexes

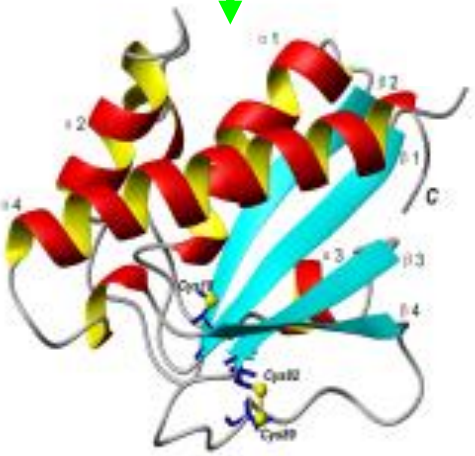
Conformational Switch



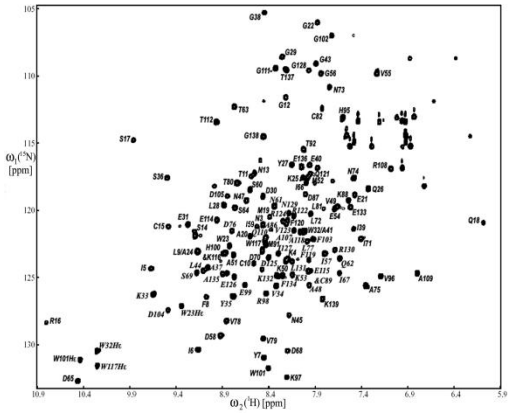


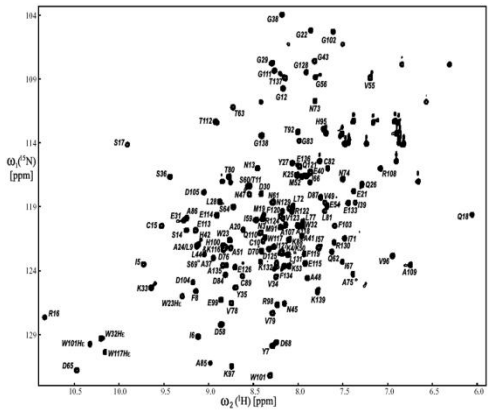
As^{5+}





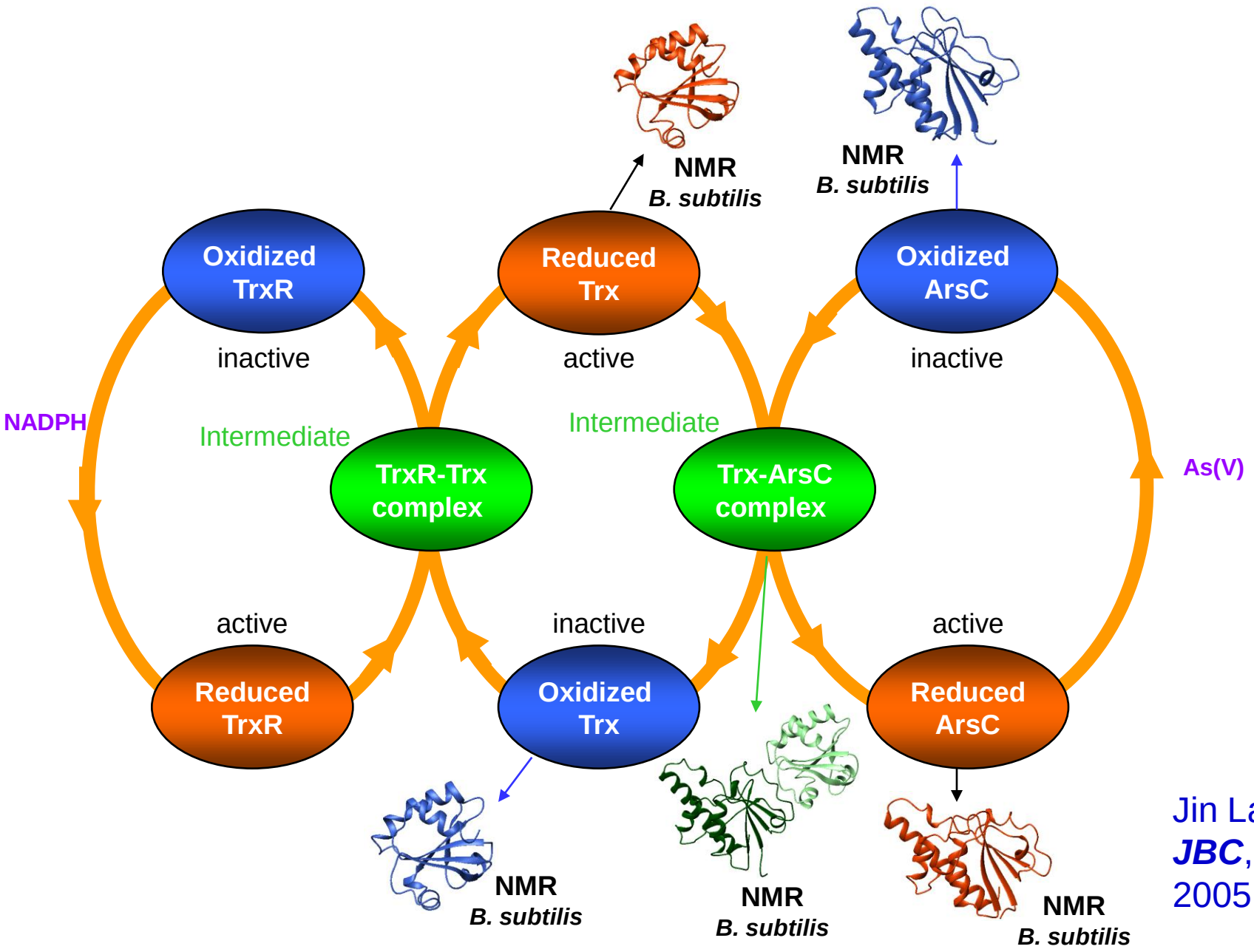
Thioredoxin





Jin Lab, ***JBC***, 2005

Structural snapshots along the reaction



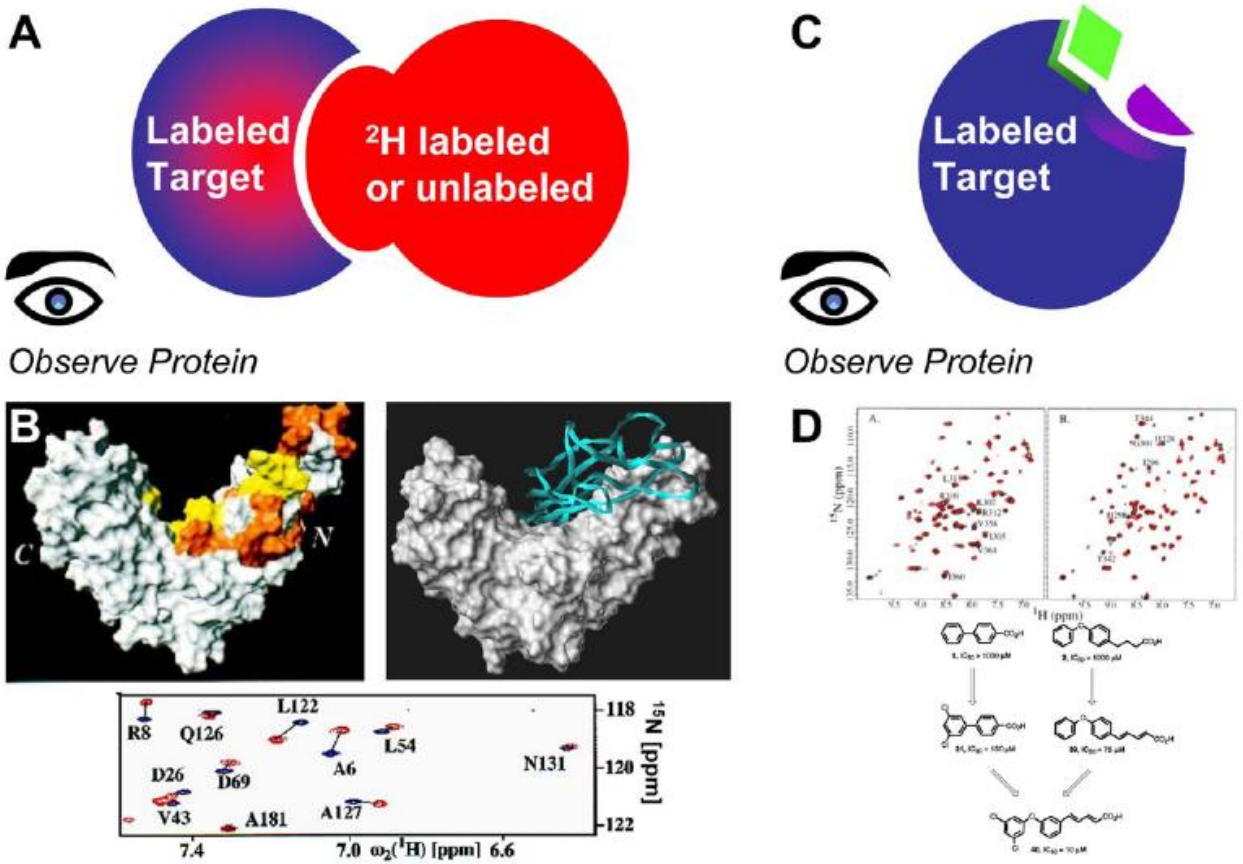
Jin Lab
JBC,
2005,2007

NMR Methods for Studying Protein Interactions

- **Chemical shift perturbation**
- **Chemical shift perturbation in combination with TROSY methods for larger proteins**

Chemical shift perturbation

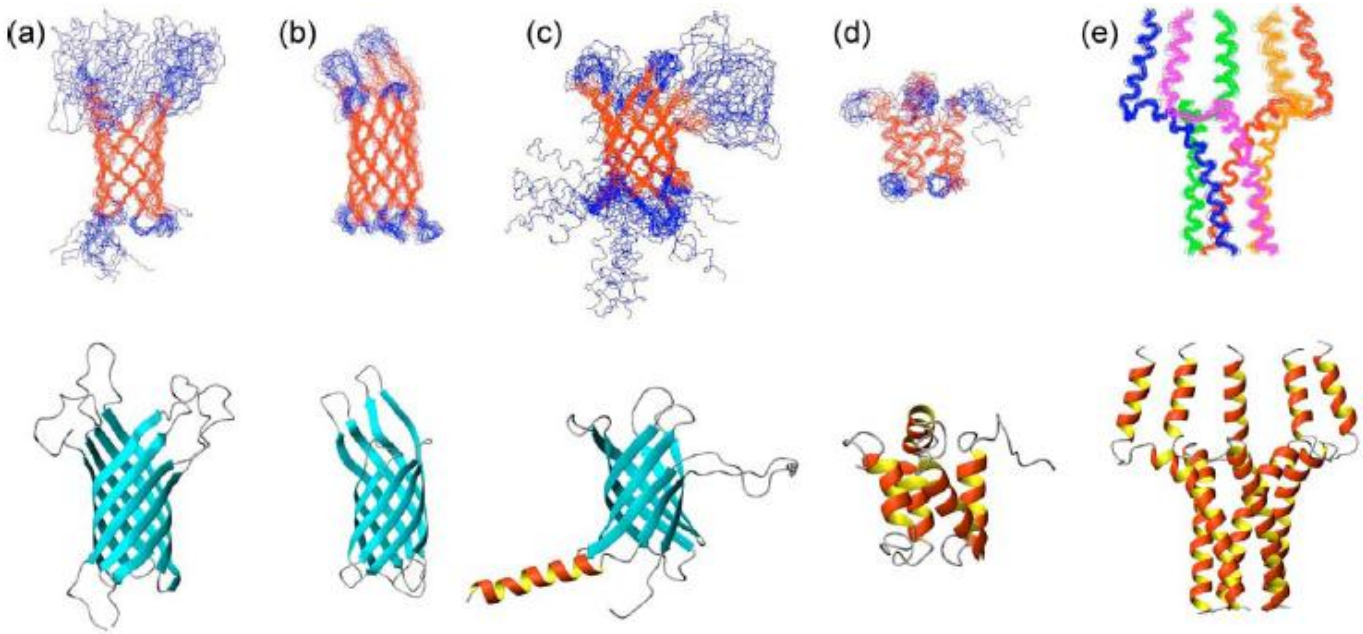
SAR by NMR



Pellecchia (2005) *Chem & Biol*

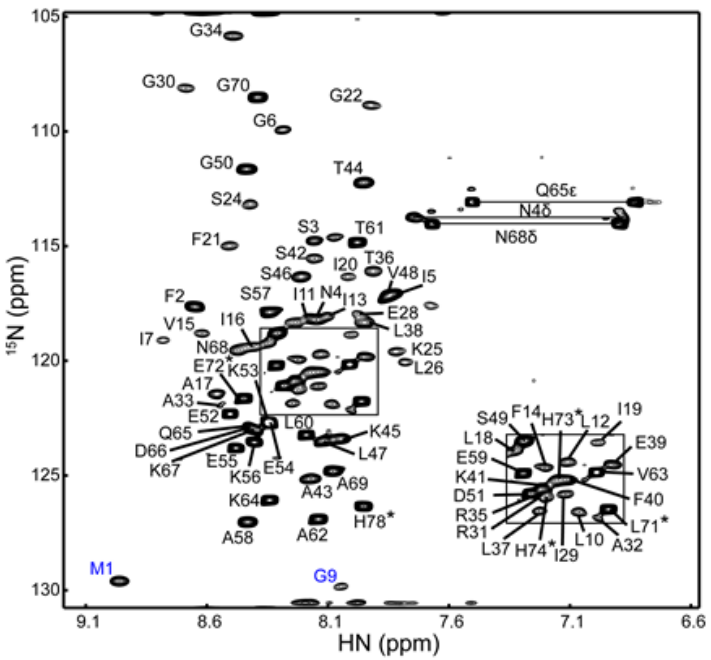
Membrane protein structures by solution NMR

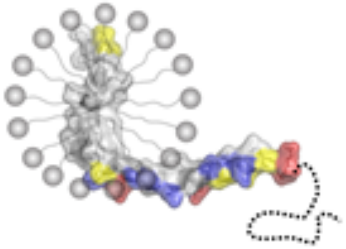
- **Protein expression, sample preparation in detergent micelles**
- **Using deuteration and refolding strategies**
- **Using a combination of TROSY, RDC and PRE (Paramagnetic Relaxation Enhancements) methods**

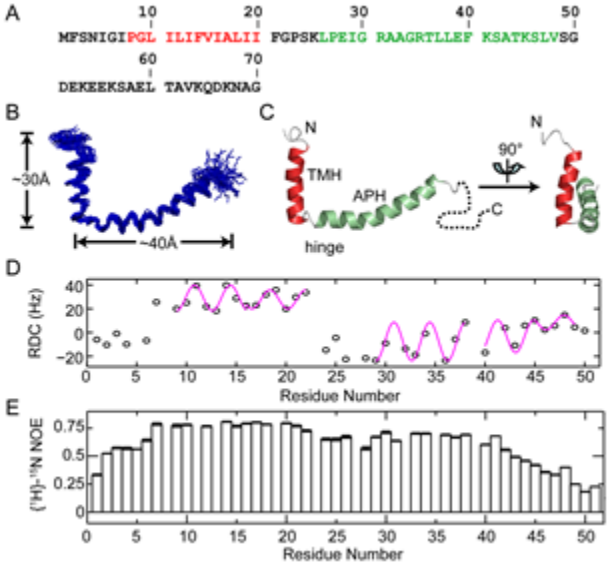


Tamm & Liang (2006) *PNMRS*

Solution NMR structure of the TatA component of the twin-arginine protein transport system from Gram-positive bacterium *Bacillus sutilis*



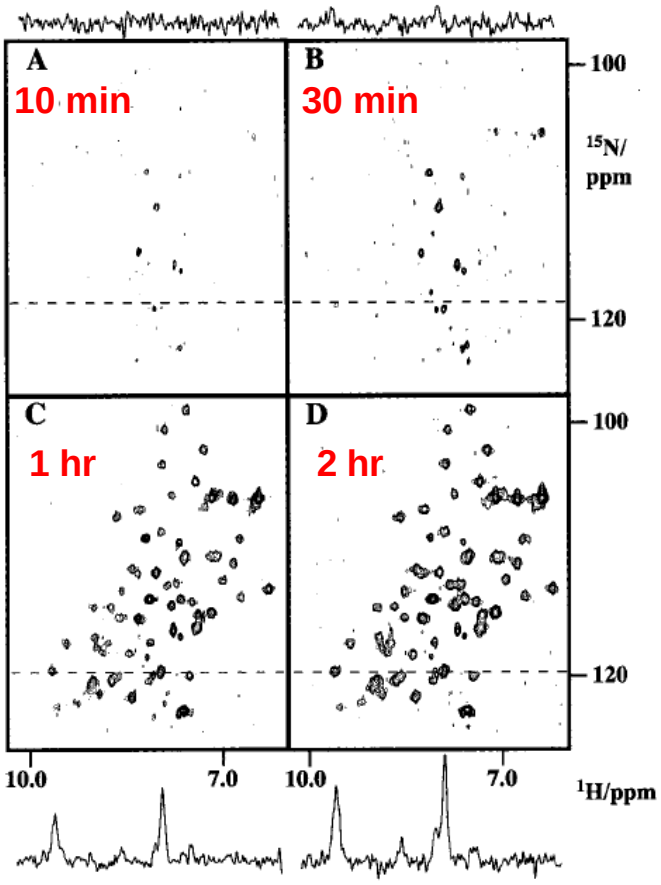




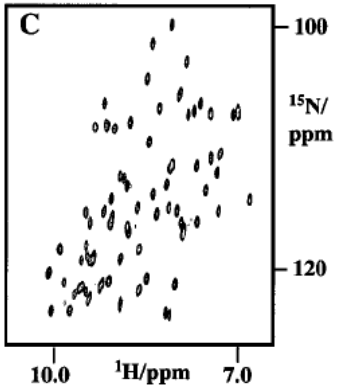
Hu *et al*, ***J. Am. Chem. Soc.* 132**, 15942, 2010

In-cell NMR

In vivo HSQC after IPTG induction



In vitro HSQC of purified sample

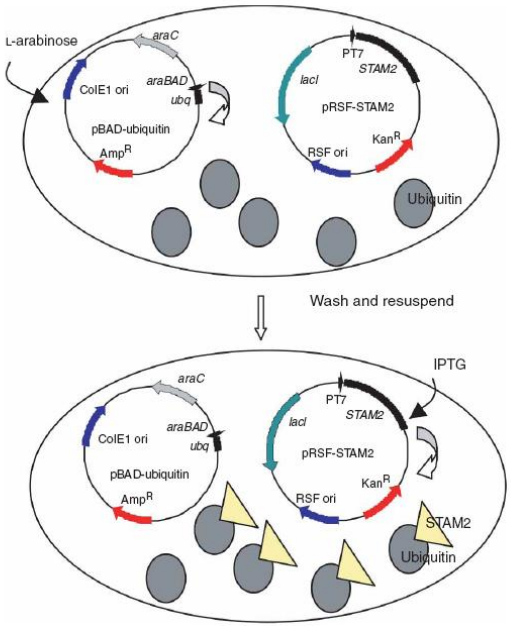


Serber et al. (2001) *JACS*

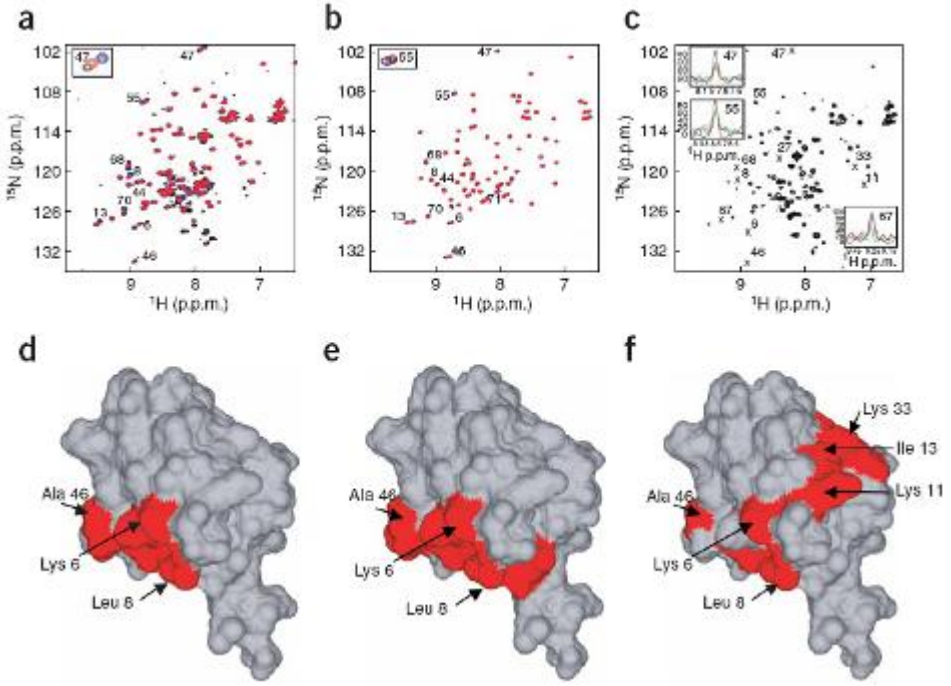
In-cell NMR: applications

- Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR)

Sequentially express two (or more) proteins within a single bacterial cell in a time-controlled manner and monitoring the protein interactions using in-cell NMR



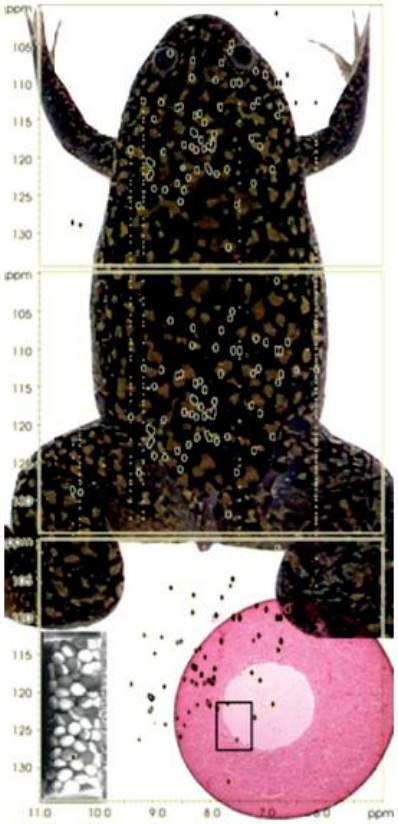
In-cell Ub + AUIM | In vitro Ub + AUIM | In-cell Ub + STAM2



Burz et al. (2006) *Nat Methods*

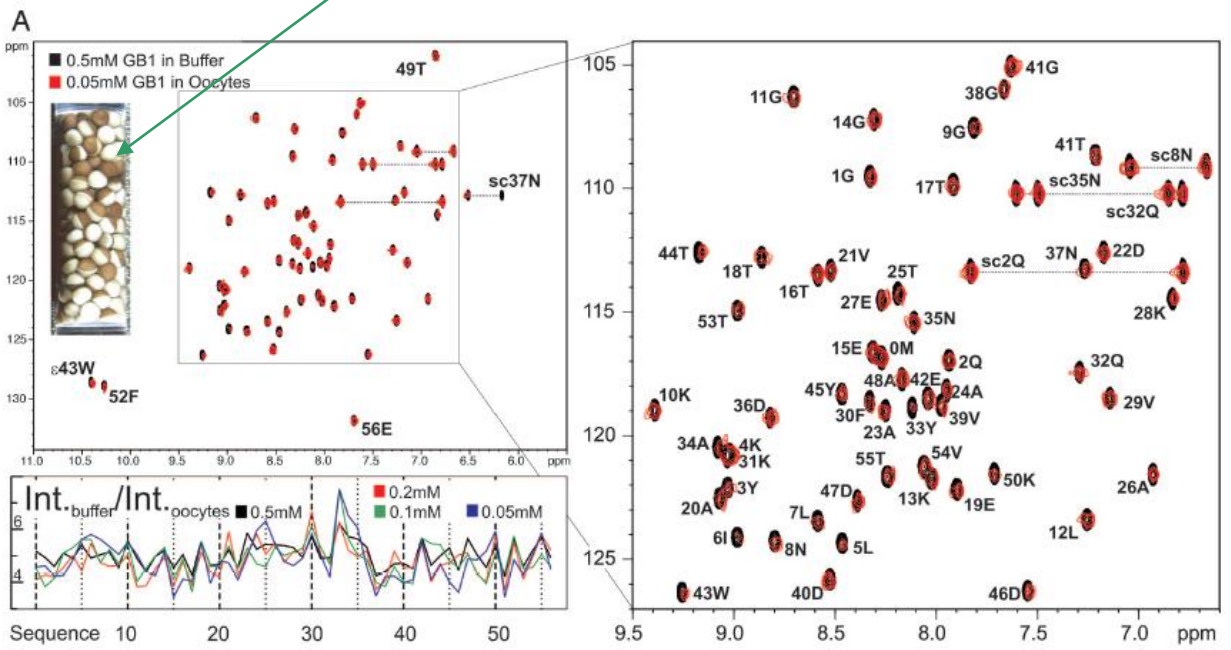
In-cell NMR: applications

- In-cell NMR in *Xenopus laevis* oocytes → mimicking eukaryotic cellular environment



Express and purify protein → **Microinject into oocytes**

200 oocytes in a Shigemi tube



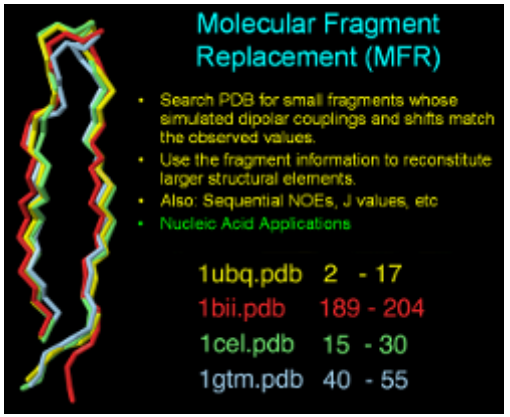
Charlton & Pielak (2006) ***PNAS***

Selenko et al (2006) ***PNAS***

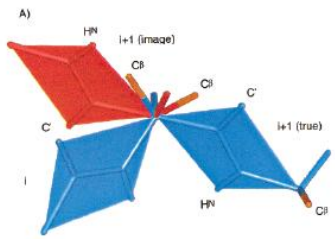
Automation & novel structure calculation methods

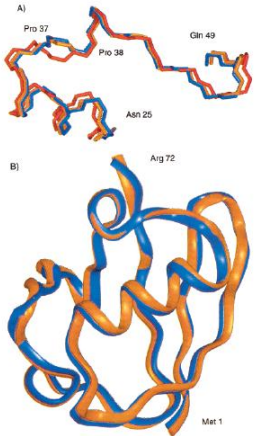
Novel structure calculation methods

- Molecular fragment replacement (MFR) using RDC data and database search
- Deriving molecular proton density from NOE cross peaks – the CLOUDS method → no assignments needed
- Structure calculation using only RDC data
- Simultaneously determination of backbone structure and dynamics from RDC

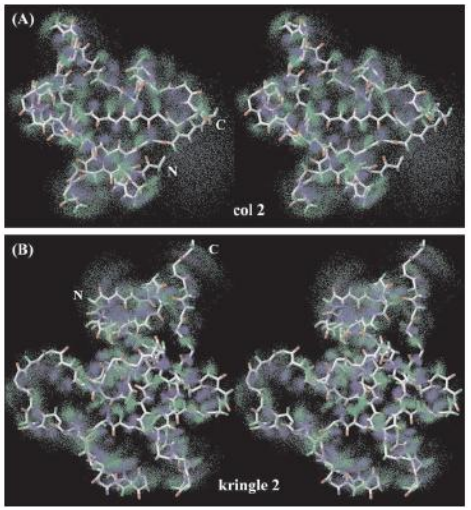


Delaglio: ***www.nmrscience.com***





Hus et al (2001) ***JACS***

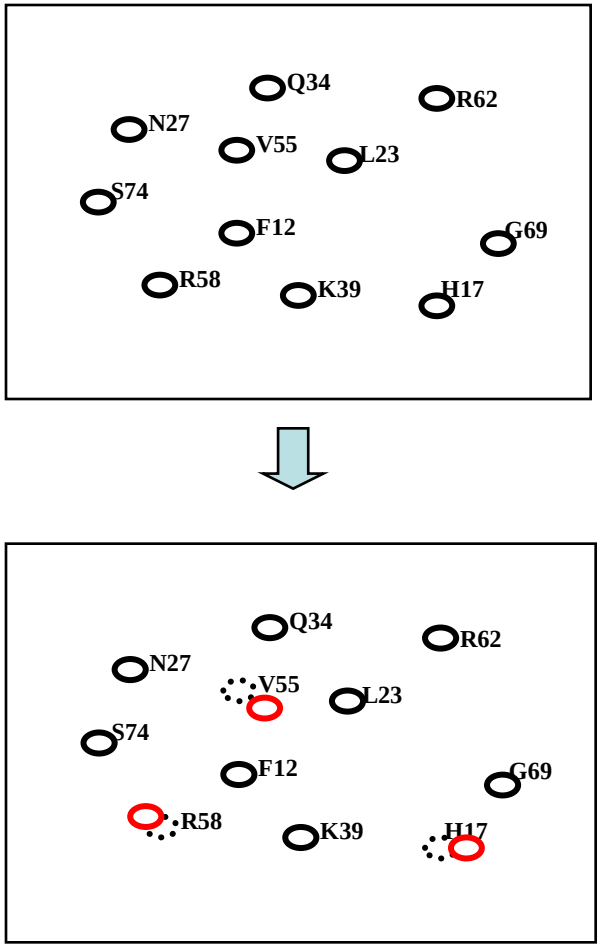


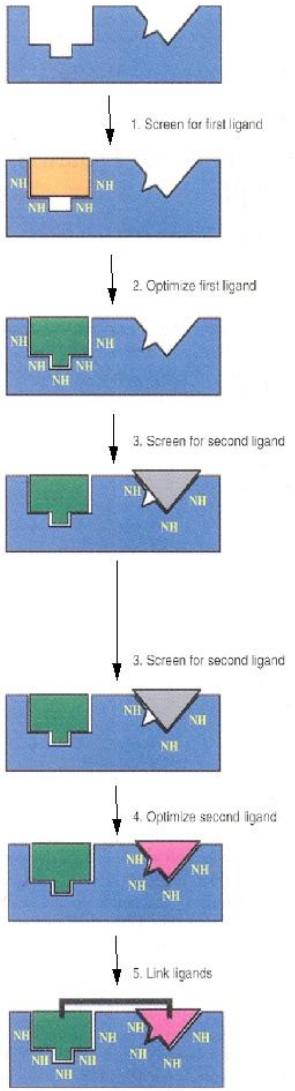
Grishaev & Llinas (2002) ***PNAS***

Biomolecular NMR & Drug Discovery

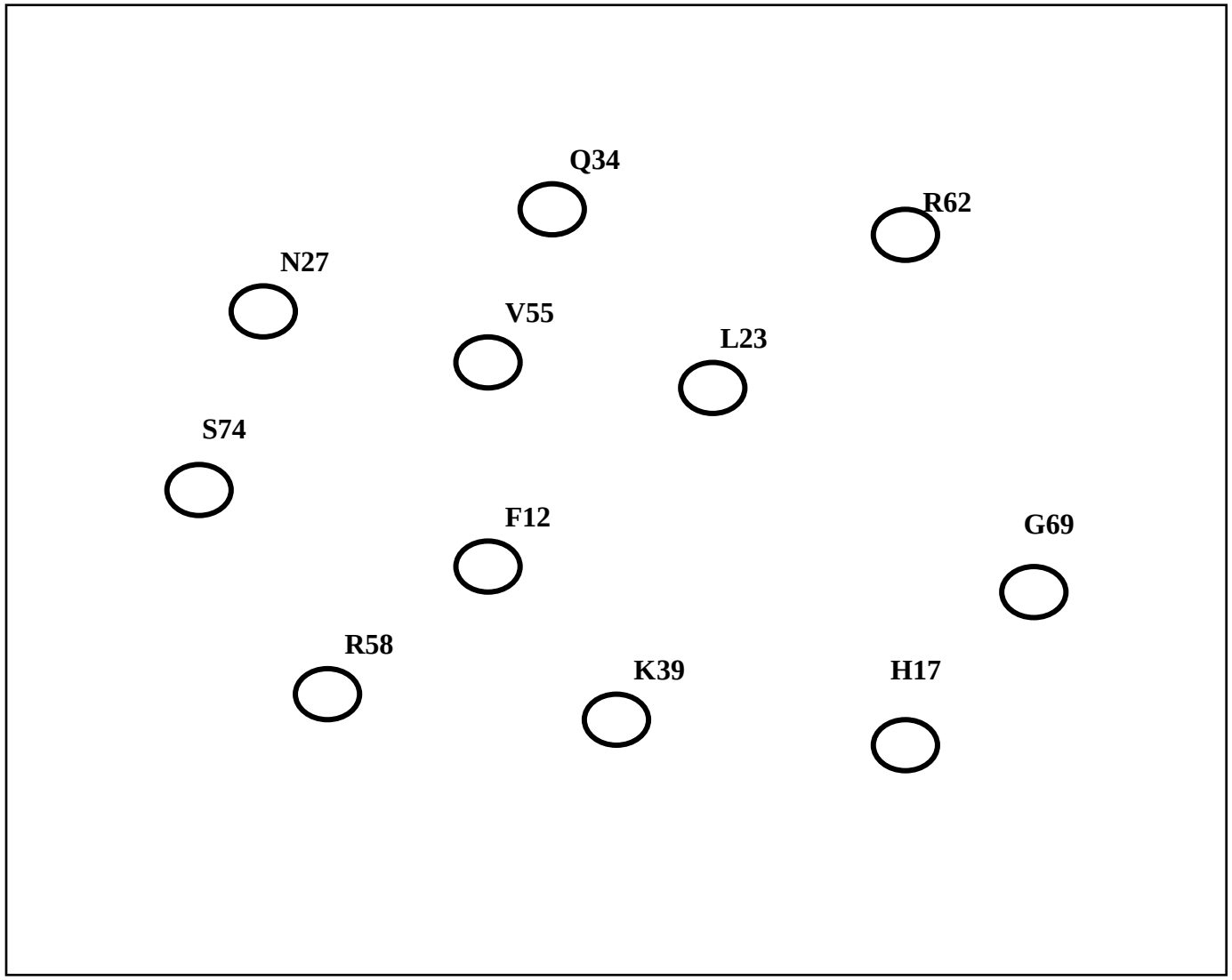
SAR by NMR

Proposed by the Abbott Laboratories (Stephen Fesik)
SAR: Structure-Activity Relationship

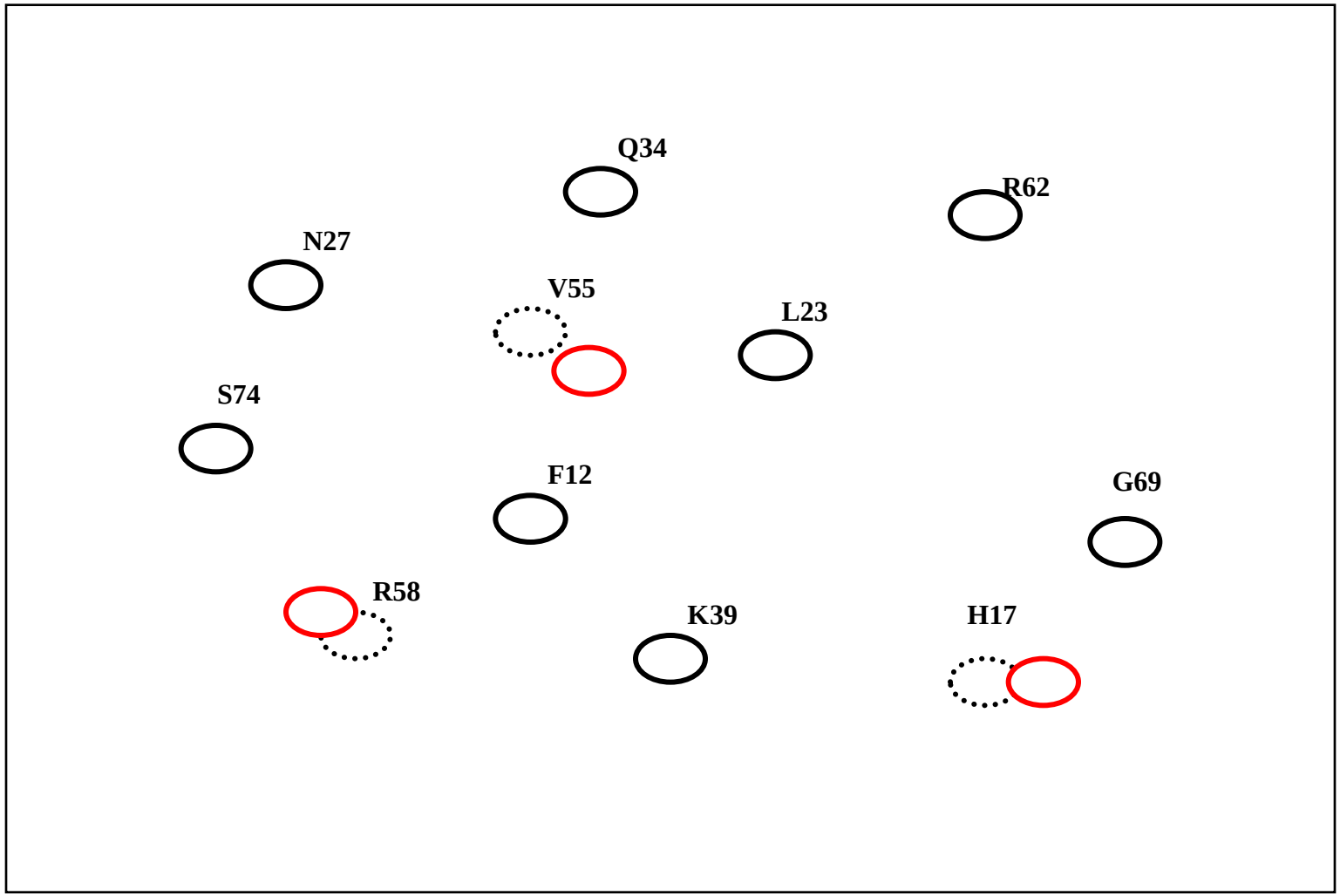




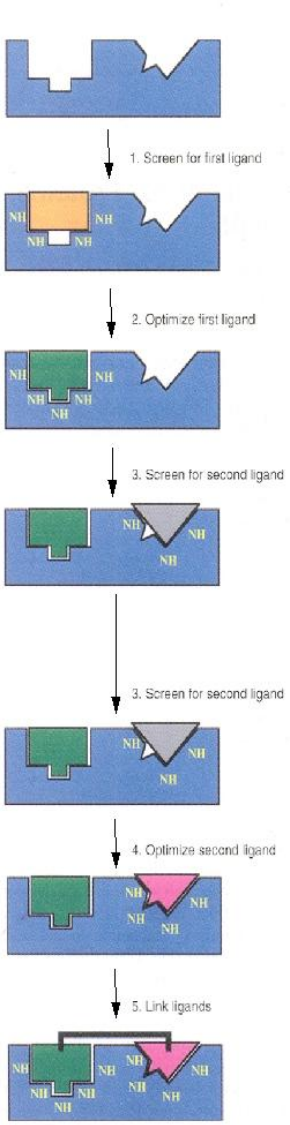
Chemical shift perturbation

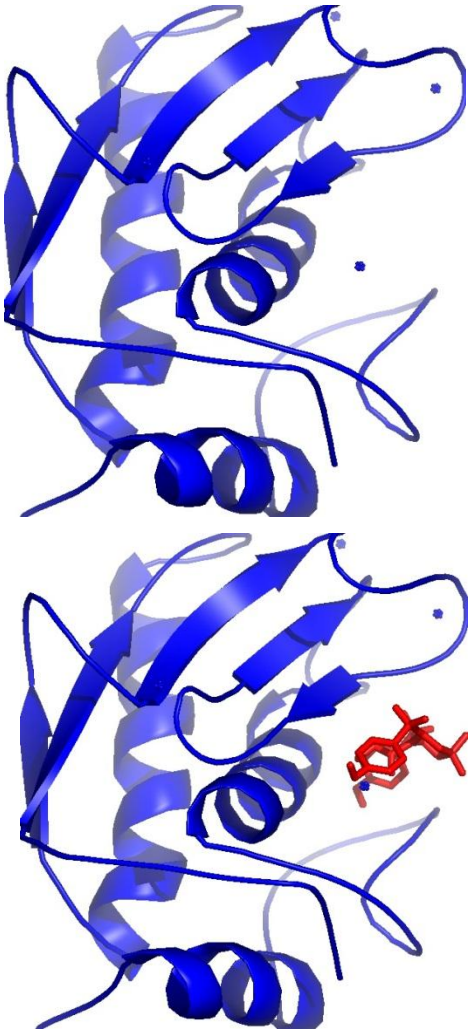


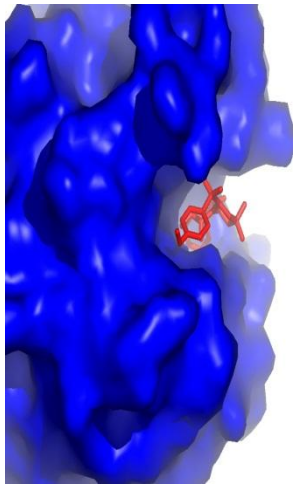
Chemical shift perturbation

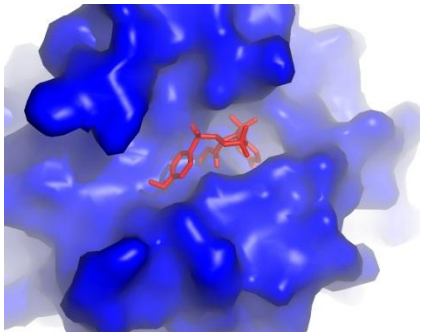


SAR by NMR

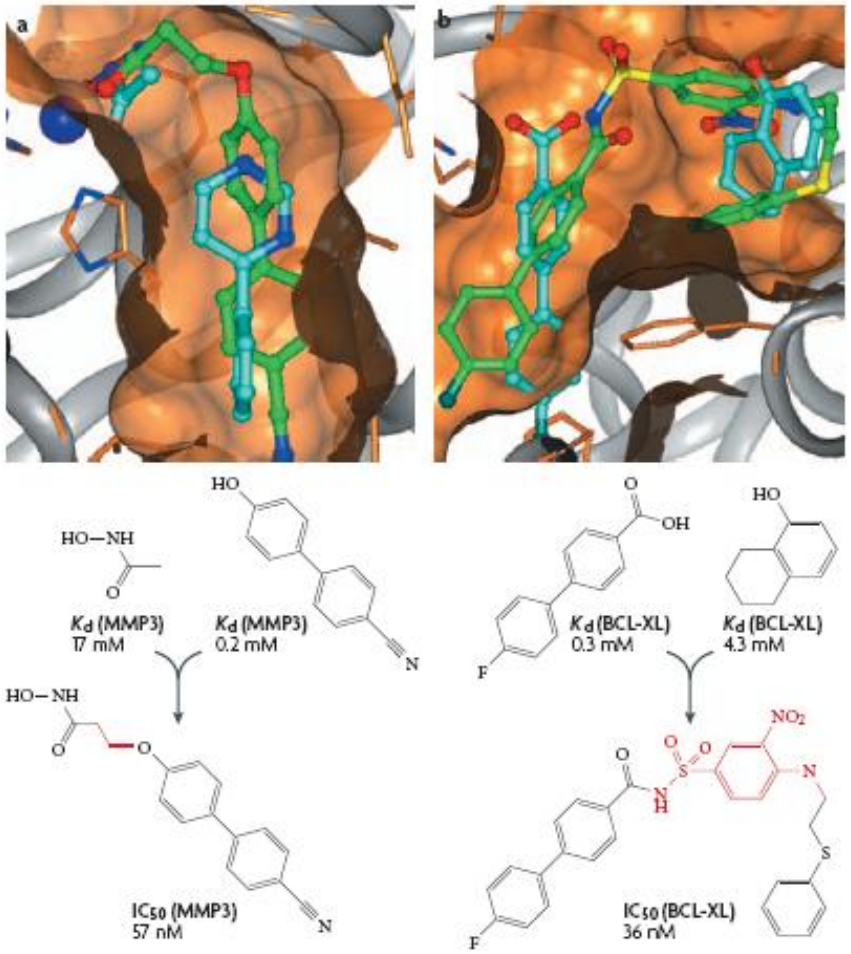

1. Screen for first ligand
2. Optimize first ligand
3. Screen for second ligand
3. Screen for second ligand
4. Optimize second ligand
5. Link ligands







SAR by NMR in fragment-based drug design (FBDD)



- Screening smaller numbers of compounds (typically several thousand) to find low-affinity fragments (with K_d values in the high uM to mM range)
- By proper optimization and tethering of the low-affinity fragments to produce high-affinity molecule

Cyan: low-affinity fragment leads
Green: high-affinity linked compounds

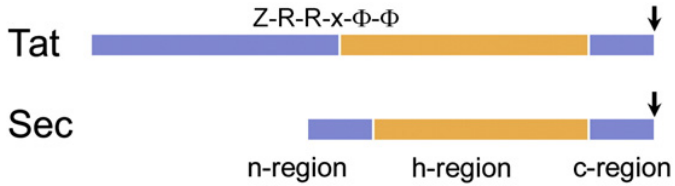
Hajduk & Greer (2007) *Nat Rev Drug Discov*

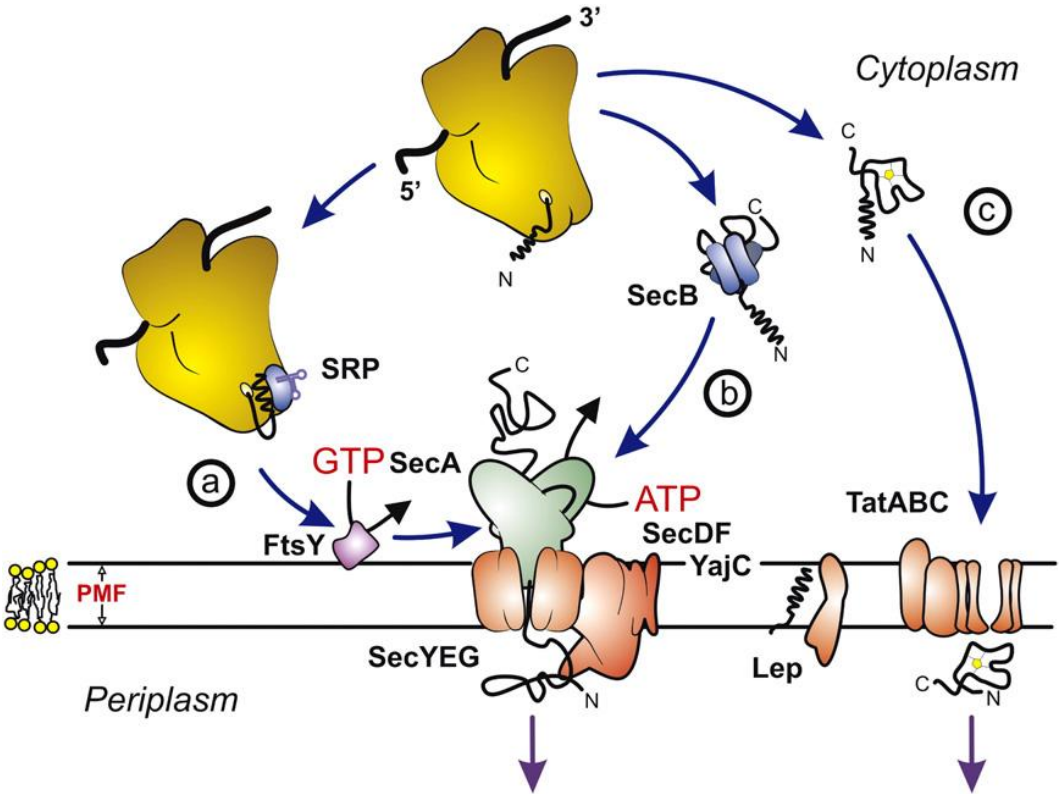
Solution structures of the TatA component of bacterial twin-arginine protein translocation system

Protein translocation across membrane

- Sec-pathway
- Tat-pathway

Signal peptide:





Natale P *et al* (2008) *Biochim Biophys Acta* 1778:1735-1756

Twin-arginine translocation (Tat) pathway

- Identified in bacteria, plant chloroplasts, archaea and some plant mitochondria
- Recognizes the twin-arginine motif of the signal peptide → **T**win-**A**rginine **T**ranslocation (Tat) system

Major differences between Sec and Tat systems

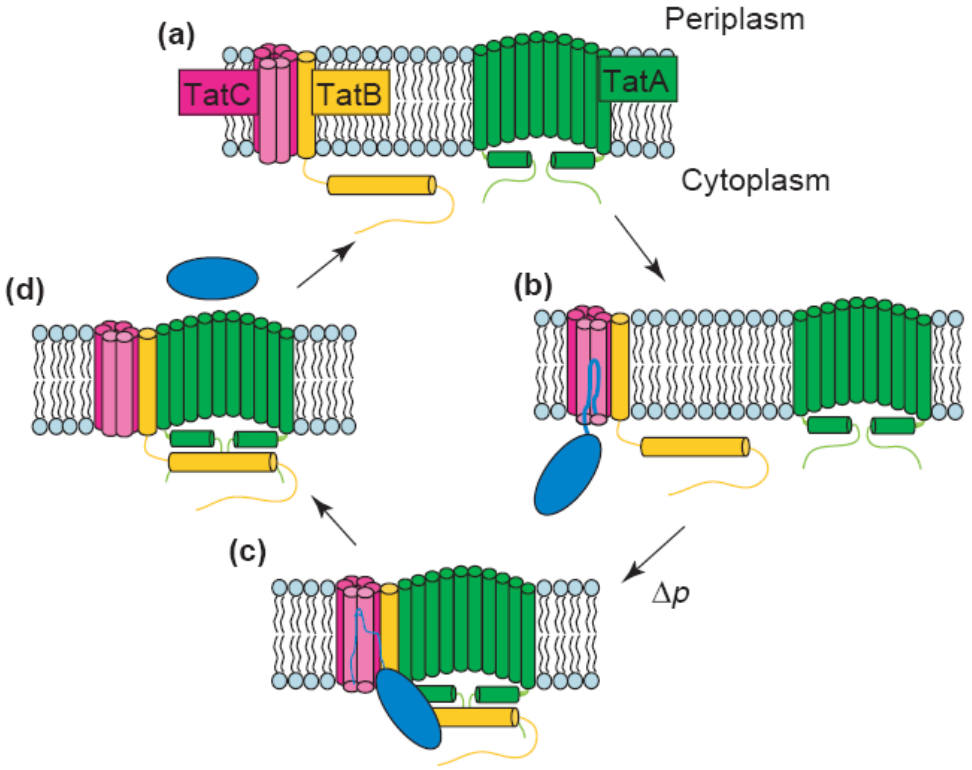
	Sec	Tat
RR motif in signal peptide	-	+
State of the target protein during translocation	unfolded	**folded**
Energy source	ATP	proton motive force

- Unique challenges for Tat system:
  - Hydrophilic environment for translocation of fully folded proteins
  - Variable protein translocation channel sizes for different substrates

Proposed mechanism of TatABC system

In *Escherichia coli* and plant chloroplasts

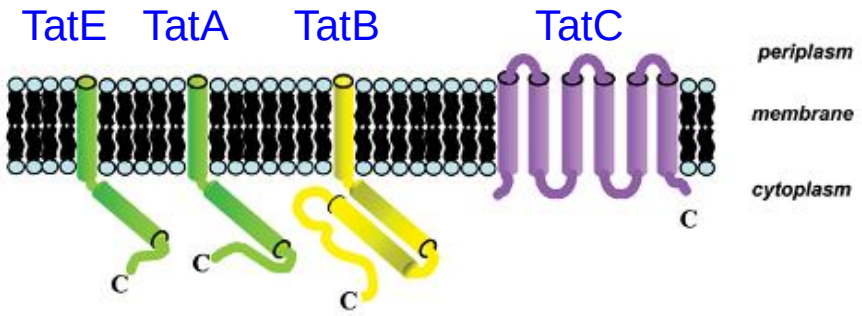
- Minimal functional components: TatA, TatB, TatC
  - TatA & TatB: single transmembrane protein
  - TatC: six transmembrane segments
- TatBC: recognition of signal peptide & recruitment of TatA component
- TatA: forms the protein translocation channel via self-oligomerization



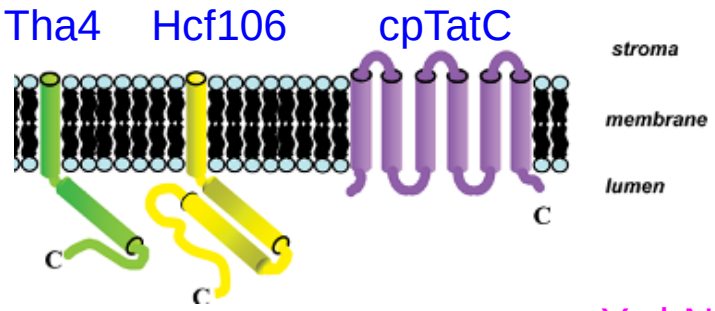
Palmer T *et al* (2005) *Trends Microbiol* 13:175-180

TatABC system *vs* TatAC system

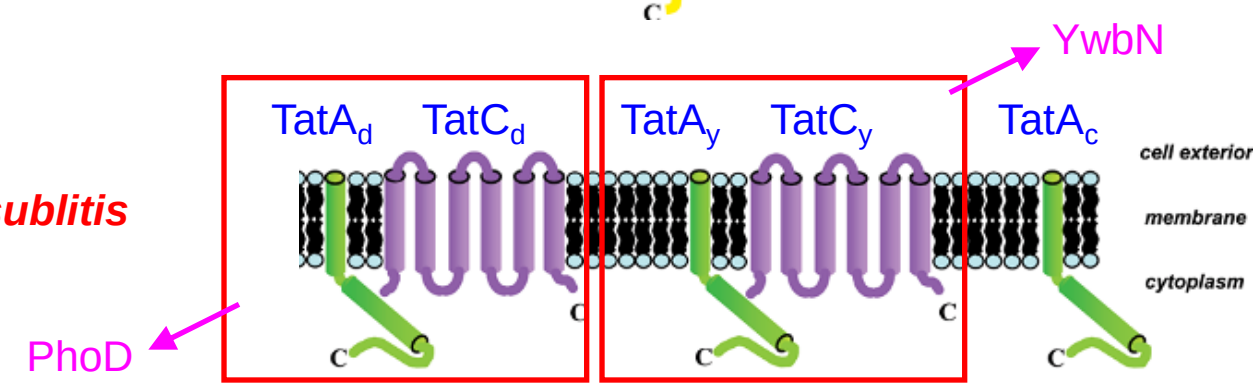
E. coli



Thylakoids of plant chloroplasts



Bacillus sublitis

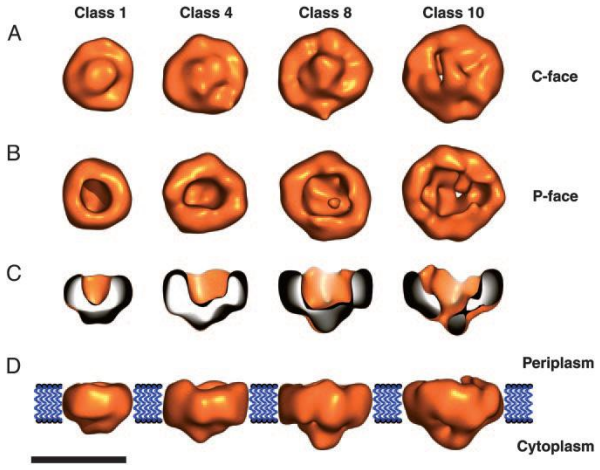


Sargent F (2007) *Biochem Soc Trans* 35:835-847

The channel-forming subunit: TatA

- Single transmembrane protein
- Two predicted helices:

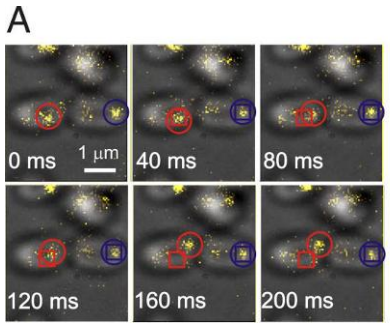
  TMH and APH

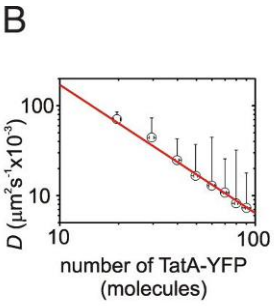


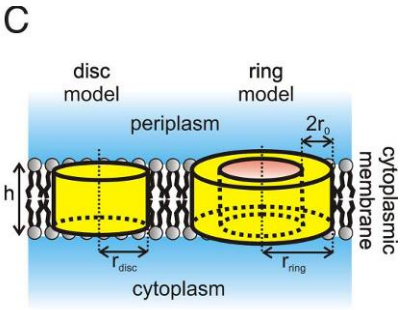
Gohlke U *et al* (2005) *PNAS* 102:10482-10486

Evidence for the channel-forming role of *E. coli* TatA:

- Self-oligomerization
- Ring structure with various diameters *in vitro* by EM
- Ring structure model by *in vivo* single-molecule imaging



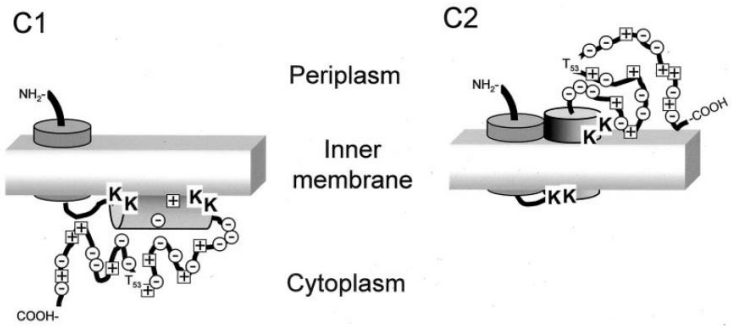


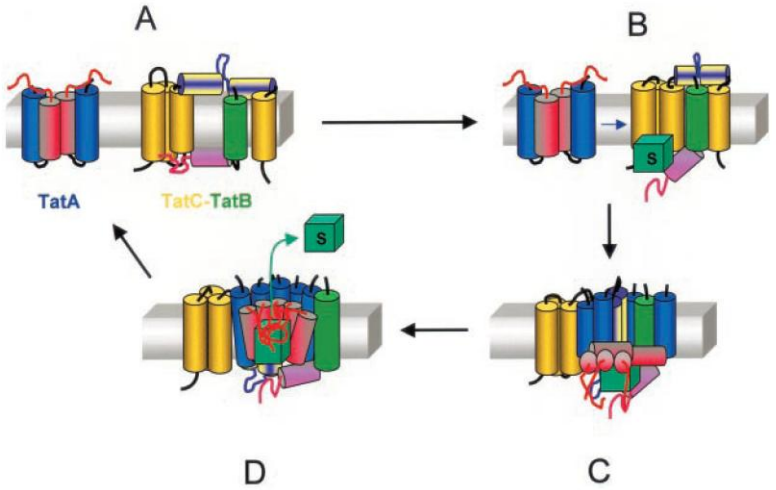


Leake MC *et al* (2008) *PNAS* 105:15376-15381

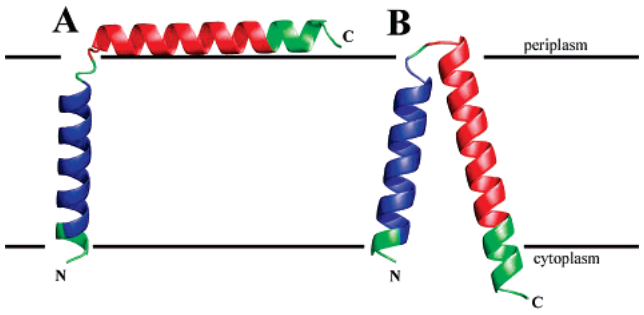
Dual topology of TatA

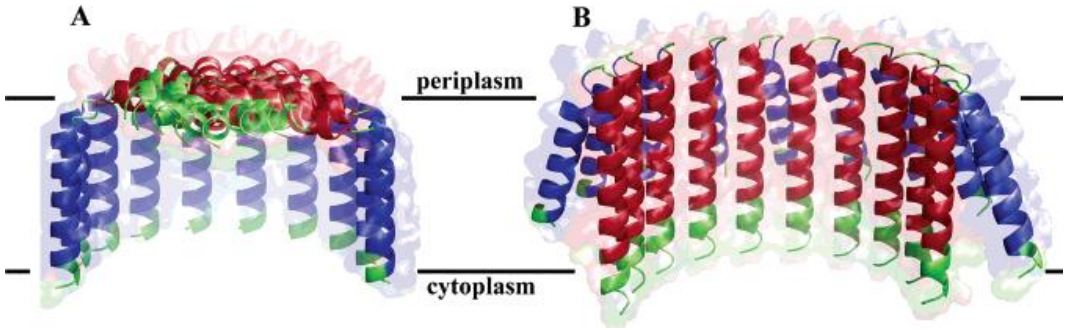
Proposed topology switch associated with protein transport





Gouffi K *et al* (2004) *J Biol Chem* 279:11608-11615



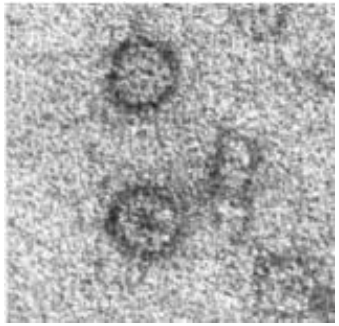


Chan CS *et al* (2007) *Biochemistry* 46:7396-7404

The channel-forming subunit: TatA

B. subtilis $TatA_d$:

- Soluble form: micelles
- Membrane-bound form: homo-multimeric complexes
  - homogeneous size
  - diameter ~10 nm



- Proposed mechanism
  - Soluble fraction of $TatA_d$ binds target proteins and recruits them to the cell membrane for translocation
  - Dual role of $TatA_d$: a combination of TatA and TatB in TatABC system

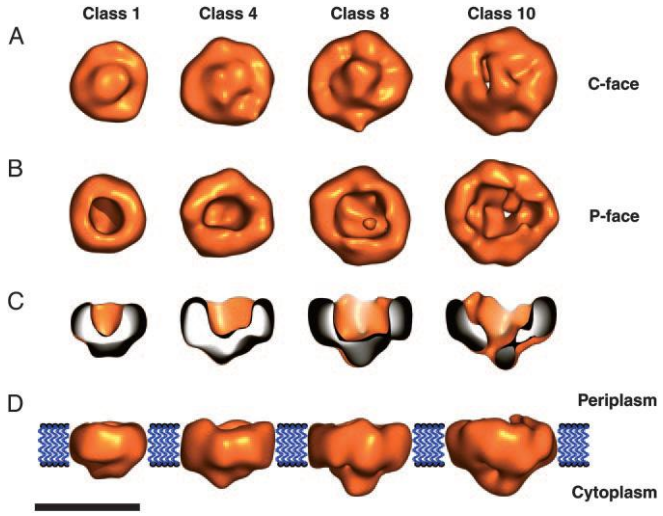
Westermann M *et al* (2006) *Biochim Biophys Acta* 1758:443-4451

Questions to be answered:

- Structural properties of TatA channels
- Mechanism of self-oligomerization (controlling pore size)
- Conformational changes during protein translocation
- Sequence and/or structural determinants of the different functions of TatA and TatB subunits

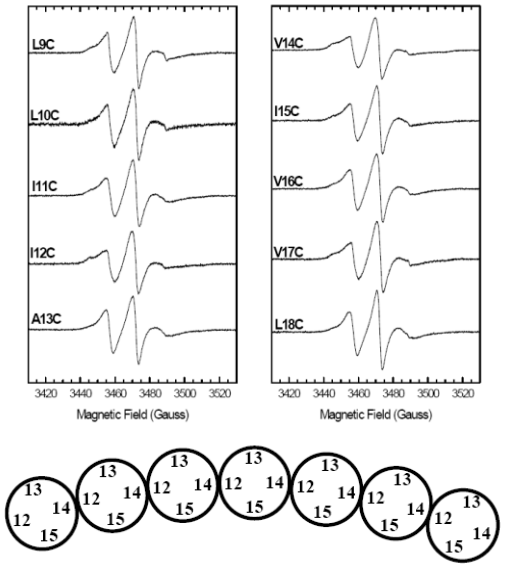
Structural information of TatA(B) in different functional states

Overall morphology (EM)



Gohlke U *et al* (2005) *PNAS* 102:10482-10486

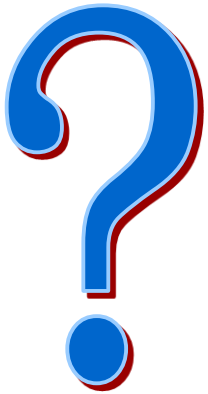
Inter-subunit contacts (EPR)



White GF *et al* (2010) *J Biol Chem* 285:2294-2301

High resolution structures

- sample inhomogeneity
- dynamic assembly
- intrinsic flexibility



Aims of the solution NMR study:

- High-resolution structures of TatA in monomeric (and low-oligomeric) state
- Self-oligomerization mechanism
- Conformational dynamics

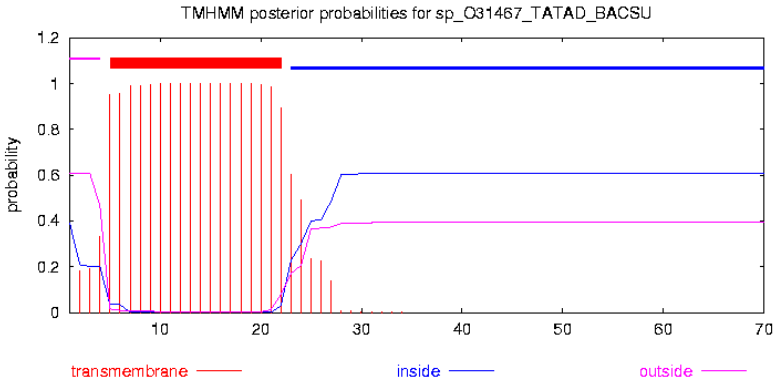
TatAC system: *B. subtilis* $TatA_d$ and $TatA_y$

TatABC system: *E. coli* TatA

B. subtilis $TatA_d$

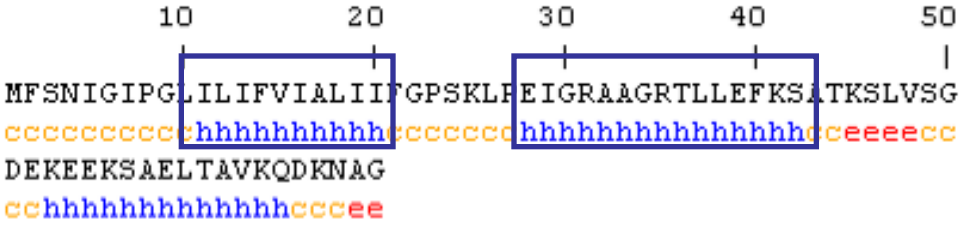
```
          10        20              30              40
MFSNIGIPGL ILIFVIALII FGPSKLPEIG RAAGRTLLEF
          50              60              70
KSATKSLVSG DEKEEKSAEL TAVKQDKNAG
```

Transmembrane segment prediction (TMHMM)



Amino acids: 70
Mw: 7430.7
Theoretical pI: 8.07

Secondary structure predictions (GOR)



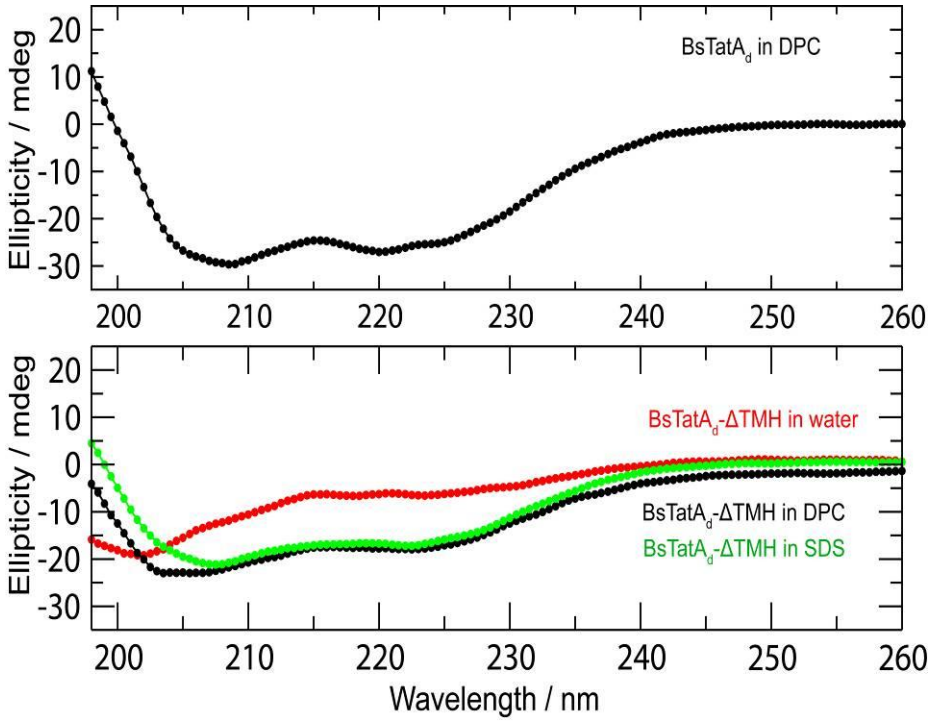
CD characterization of $BsTatA_d$

Full-length $BsTatA_d$: high helical content

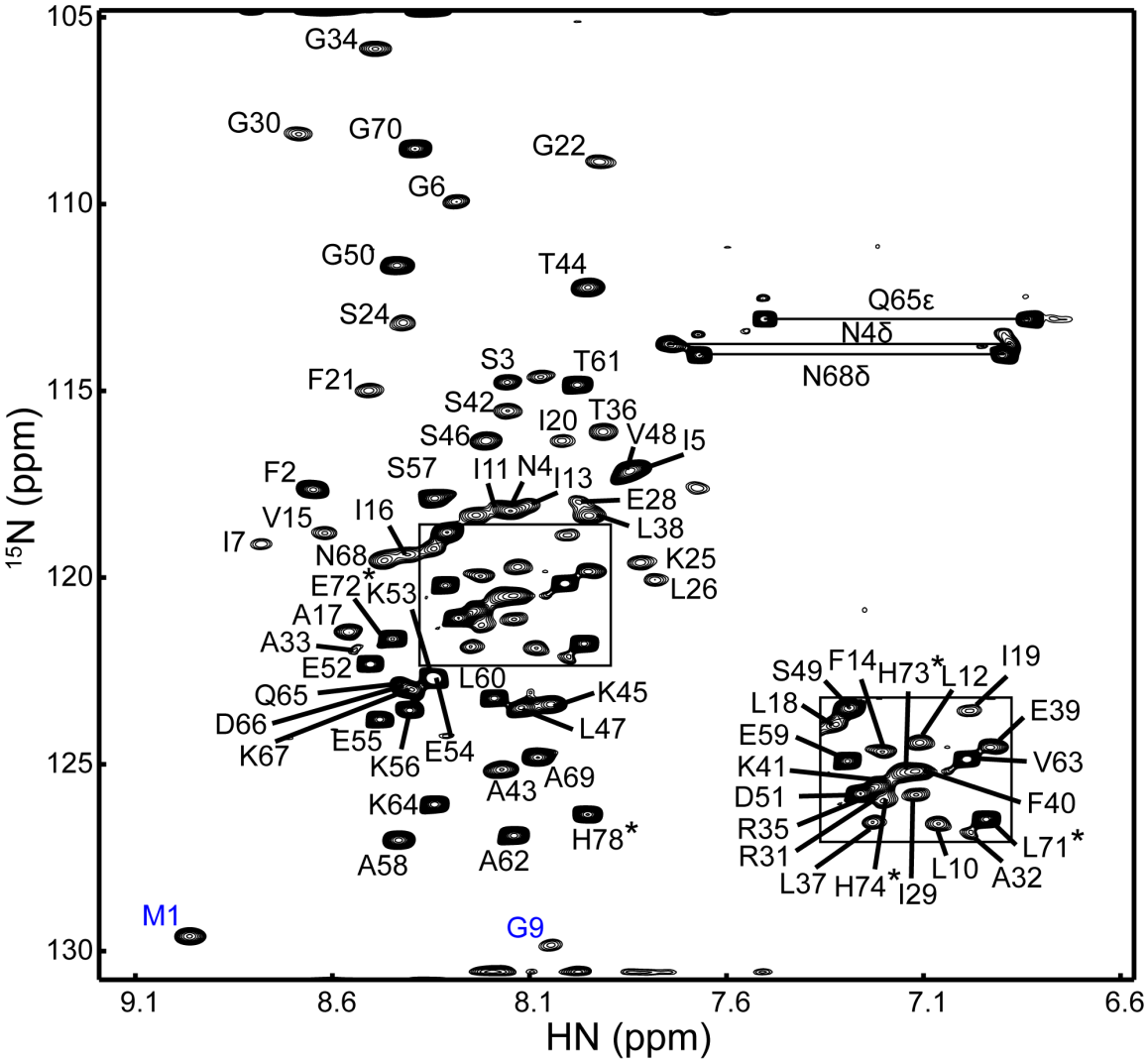
$BsTatA_d$-ΔTMH:

Unstructured in aqueous solution;

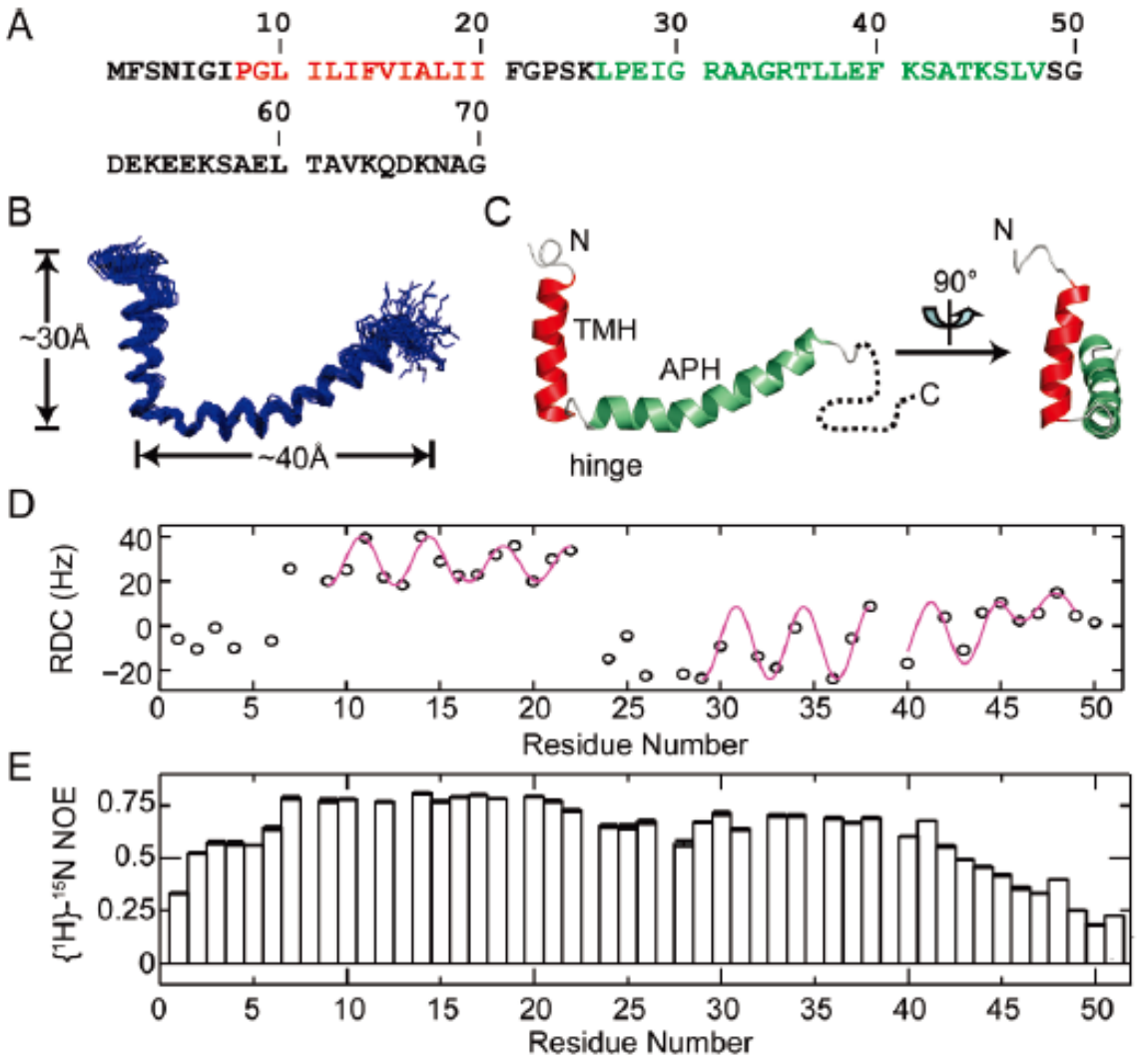
Induced helical structures in detergent micelles



2D HSQC of $BsTatA_d$ in DPC micelles



NMR structure of $BsTatA_d$ using NOE, dihedral angle and RDC restraints



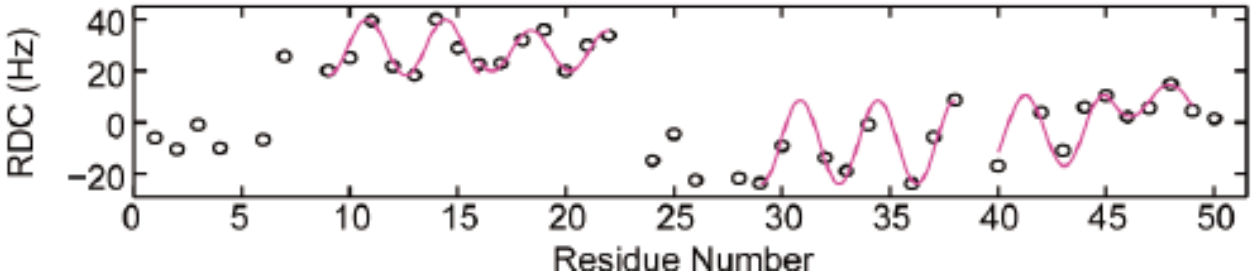
RDC measurements

- RDC medium: G-tetrad DNA (Lorieau J et al. *JACS* 2008)

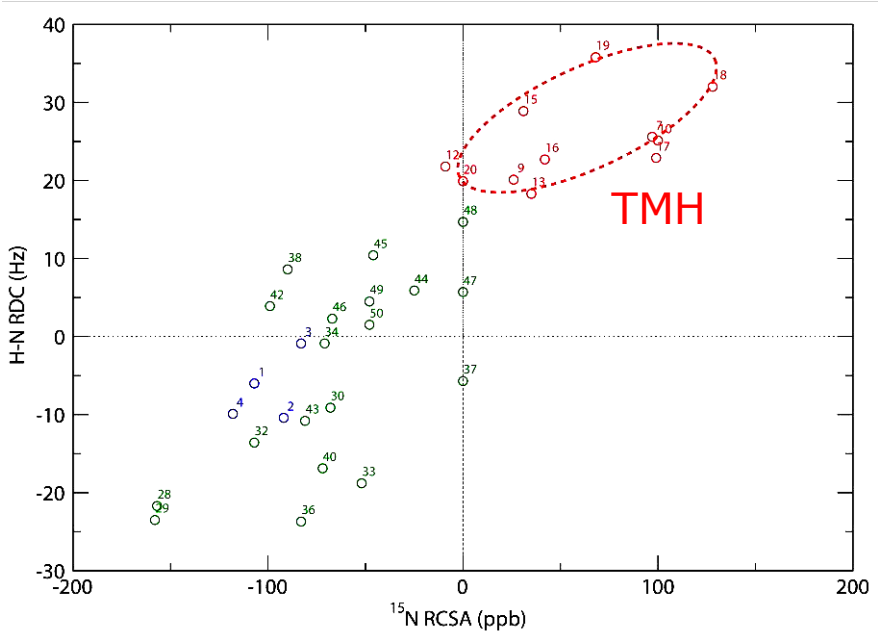
Dipolar wave:

Consistent with the L-shaped structure;

curvature in the APH segment

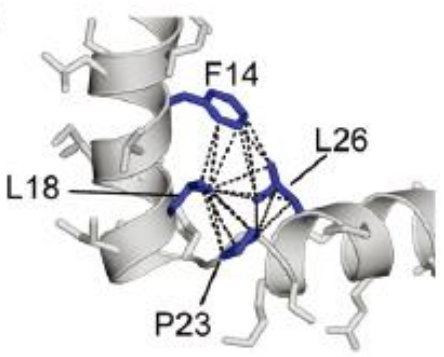


PISA wheel pattern of the TMH region

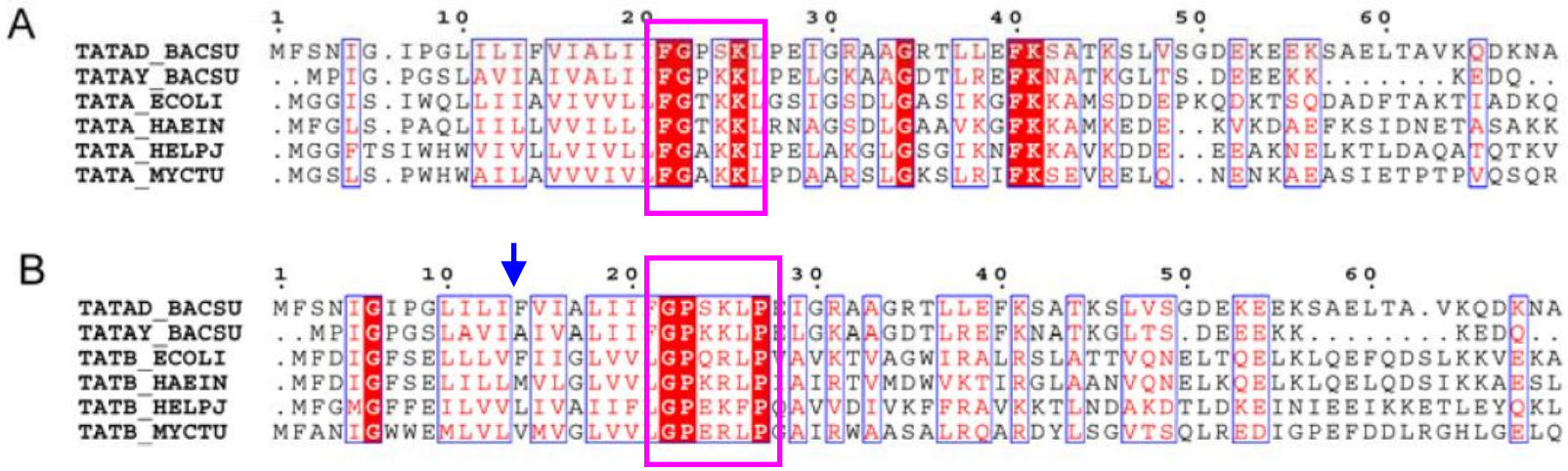


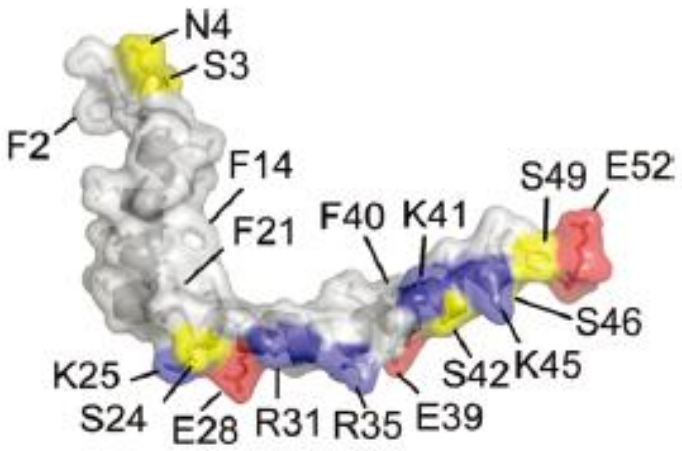
Local conformation at the hinge region

```
          10          20          30          40          50
          |           |           |           |           |
MFSNIGIPGL ILIFVIALII FGPSKLPEIG RAAGRTLLEF KSATKSLVSG
          60          70
          |           |
DEKEEKSAEL TAVKQDKNAG
```



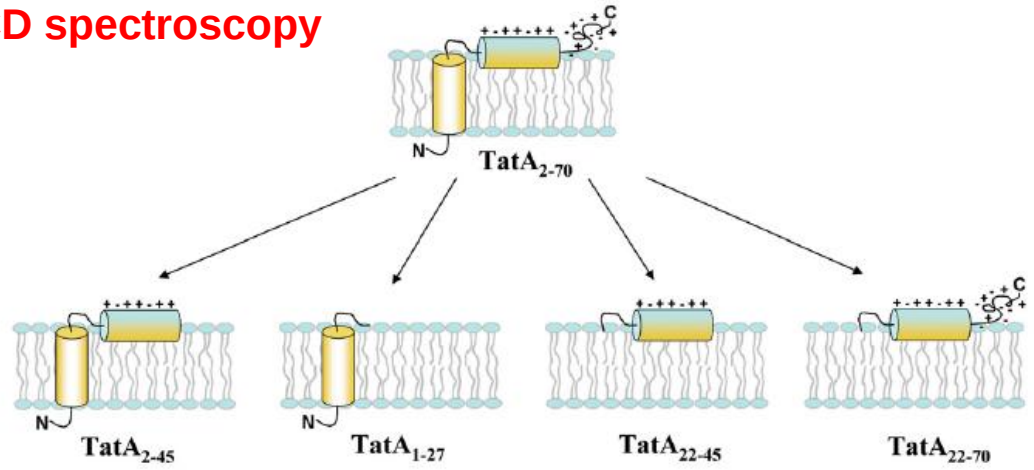
- Consensus sequence of TatA family:
  FG
- Consensus sequence of TatB family:
  G**P**xx**L**P





- TMH: fully buried
- N-terminal part of TMH (26-40): hydrophobic residues buried, hydrophilic residues partially exposed to water
- C-terminal part of TMH (41-48): mostly solvent exposed

Oriented CD spectroscopy



Lange C *et al* (2007) *Biochim Biophys Acta* 1768:2627-2634

Acknowledgements

Beijing NMR Center (BNMRC)

Peking University

Yunfei Hu

Yi Zhang

Hongwei Li

Enwei Zhao

Prof. Bin Xia (BNMRC)

$$$

863, 973 from MOST, China

Beijing NMR Center

Beijing NMR Center (BNMRC)

- A national NMR center
- Funded by the Ministry of Science and Technology, the Ministry of Education, the Chinese Academy of Science and Military Council of Medicine
- Operated by Peking University

- A research center, mainly focus on the structures, dynamics, and other applications of bio-macromolecules by high-field solution state NMR technique

NMR spectrometers at Beijing NMR Center

- 500 MHz, 600MHz, 700 MHz and 800 MHz spectrometers (with cryoprobes) : Bio-macromolecule structural and dynamic studies

- 600 MHz and 400 MHz spectrometers (with RT-probes): routine services for chemistry


BRUKER
500
UltraShield™
BRUKER
CRYO PLATFORM
BRUKER


BRUKER
600
UltraShield


BRUKER
800
UltraStabilized™

Wet lab at Beijing NMR Center

Computer Cluster

Members at Beijing NMR Center

- Principle Investigators:
  - Dr. Bin Xia (Professor)
  - Dr. Changwen Jin (Professor)
- Staff:
  - Dr. Jian Lin (Associate professor)
  - Dr. Hongwei Li (NMR manager)
  - Dr. Yunfei Hu & Dr. Supu Mi (Postdocs)
- Administration:
  - Mr. Jinxin Yang (Lab manager)
  - Ms. Xiangli Hu & Ms. Zhe Su (Technicians)

- Graduate students: ~20
- Undergraduate students: ~10

Workshops

- Jan. 2006: National NMR workshop
- Sept. 2008: EMBO World Practical Course

Research Interests

- Protein structures
- Protein-protein/protein-nucleic acid interactions and complex structures
- Protein dynamics
- Membrane proteins

International Colaborations

- Prof. Honggao Yan, Michigan State University, USA
- Prof. Zongchao Jia, Queen's University, Canada
- Prof. Jiyan Ma, Ohio State University, USA
- Prof. Jun Liu, University of Toronto, Canada
- Prof. JP Jacquot, Universite Henri Poincare, France
- ……

Thank you !