Fast multidimensional NMR spectroscopy of proteins in solution

Development of new methods and application to the study of structure and kinetics

Paul Schanda
Directeur de thèse : Bernhard Brutscher

Laboratoire de RMN
Institut de Biologie Structurale
Grenoble
Major challenges in structural biology

Structures

X-ray diffraction
NMR (20% of PDB)

β-catenin in complex with the intrinsically unfolded protein Tcf4
Major challenges in structural biology

Structures

- X-ray diffraction
- NMR (20% of PDB)

Dynamics

- Transient Interactions
- Folding
- Unfolded proteins
- Coupled binding/folding
- Catalysis
- Misfolding and fibrils
- Ensembles of structures changing with time

Spectroscopies

- IR
- UV/Vis
- NMR
- Scattering
- Diffraction

Static
Introduction

Nuclear spins report on the environment at hundreds of sites within the molecule.

Each individual NMR-active nucleus ($^1\text{H}, ^2\text{H}, ^{13}\text{C}, ^{15}\text{N}, ...$) gives local information:

- Chemical environment
- Structure
- Dynamics

[Diagram showing molecular structure with arrows indicating nuclear spins reporting on the environment.]
Introduction

Requirements to observe individual atomic sites

1. Sensitivity

Only 1 out of ~10000 molecules contributes to NMR signal

Solutions

High $B_0$-fields

400 $\rightarrow$ 900MHz

factor $\approx 3.0$

Cryogenic probes

factor $\approx 2-4$

Optimized pulse sequences

Sensitivity has advanced a lot in recent years making possible a significant reduction of experimental durations
Introduction

Requirements to observe individual atomic sites

2. Resolution

Observation of signals of individual nuclei is hindered by signal overlap

Solution

Multidimensional NMR
Introduction

Multidimensional NMR

... correlates the frequencies of several interacting nuclear spins and spreads the signals over a multidimensional space.
Time requirements of nD NMR

1D: seconds
2D: minutes
3D: hours
4D: weeks

In many cases, experimental durations are dictated by the sampling requirement, not by the sensitivity.
Limitations of nD NMR related to long experimental durations

* Sample stability and long acquisition times?

* High throughput capabilities for structural genomics projects?

* High-dimensional data sets as required for proteins with strong overlap (unfolded proteins) require unrealistically long times

* Spectral changes during a reaction occurring on a seconds time scale cannot be followed with atomic resolution
Outline

Part 1: Methods for fast multidimensional NMR

Concepts

Sensitive fast-pulsing experiments: SOFAST HMQC, BEST experiments

Part 2: Applications

Application to the study of protein folding and unfolding

HET-SOFAST NMR: a fast tool for characterizing protein structure
Accelerating nD NMR

**Concepts**

1. Reduce the number of mandatory scans
Accelerating nD NMR Concepts

1. Reduce the number of mandatory scans

Alternative ways of data sampling:

- Non-linear data sampling of time domain
- Hadamard NMR spectroscopy (frequency domain)
- “Single scan” NMR (spatial/orientational encoding)

...
Accelerating nD NMR

Concepts

1. Reduce the number of mandatory scans

Alternative ways of data sampling:

- Non-linear data sampling of time domain
- Hadamard NMR spectroscopy (frequency domain)
- "Single scan" NMR (spatial/orientational encoding)

2. Reduce the duration of each scan

---

**Pulse sequence**

**Recovery delay**

---

0.1 sec 1-2 sec
### Accelerating nD NMR Concepts

1. **Reduce the number of mandatory scans**

   Alternative ways of data sampling:
   - Non-linear data sampling of time domain
   - Hadamard NMR spectroscopy (frequency domain)
   - “Single scan” NMR (spatial/orientational encoding)

2. **Reduce the duration of each scan**

   - **0.1 sec**
   - **<0.1 sec**
Considerations about sensitivity and interscan delay

- **Sensitivity**
  - As a function of recovery delay, sensitivity increases initially and plateaus after a certain delay.

- **Available z-magnetization $I_z$**
  - Longitudinal relaxation increases with recovery delay.

- **S/N**
  - $S/N \sim \sqrt{\text{number of scans}}$
  - $S/N \sim 1/\sqrt{\text{scan time}}$

- **Accumulation of S/N**
  - The accumulation of S/N decreases with increasing recovery delay.
Considerations about sensitivity and interscan delay

- **Sensitivity**
  - Available z-magnetization $I_z$
  - Longitudinal relaxation
  - $S/N \sim \sqrt{\text{(number of scans)}}$
  - $\sim 1/\sqrt{\text{(scan time)}}$

- **Recovery delay**
  - $t_{acq}$
  - ~1 [sec]
  - ~2 [sec]
  - ~3 [sec]
Considerations about sensitivity and interscan delay

Challenge: retain high sensitivity for high repetition rates of the experiment

S/N $\sim \sqrt{\text{number of scans}}$
$\sim 1/\sqrt{\text{scan time}}$

Longitudinal relaxation

Available z-magnetization $I_z$
Considerations about sensitivity and interscan delay

Available z-magnetization $I_z$

$T_1=0.5$ sec

$T_1=1$ sec

Accelerated longitudinal relaxation allows fast repetition of scans while retaining high sensitivity
Proton longitudinal relaxation in proteins

Most NMR experiments in liquids excite and detect proton spins. The $^1$H longitudinal relaxation determines the sensitivity.
Longitudinal relaxation in a two spin system

\[- \frac{d}{dt} \begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \end{pmatrix} = \begin{pmatrix} \rho & \sigma \\ \sigma & \rho \end{pmatrix} \begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \end{pmatrix} \]

\( \rho \)  Auto-relaxation rate constant
\( \sigma \)  Cross-relaxation rate constant

Non-selective excitation

Faster relaxing component due to cross-relaxation

Selective excitation
Proton longitudinal relaxation in proteins

The relaxation in a multi-spin system can be considered as a sum of two-spin interactions.

\[ -\frac{d}{dt} \begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \\ \vdots \\ I_{nz} - I_{nz}^0 \end{pmatrix} = \begin{pmatrix} \sum_j \rho_{1j} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sum_j \rho_{2j} & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sum_j \rho_{nj} \end{pmatrix} \begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \\ \vdots \\ I_{nz} - I_{nz}^0 \end{pmatrix} \]

\( \rho \)  Auto-relaxation rate constant

\( \sigma \)  Cross-relaxation rate constant
Amide $^1$H relaxation after non-selective/selective excitation

Non-selective excitation

Selective excitation of $H^N$
Outline

Part 1: Methods for fast multidimensional NMR

Concepts

Sensitive fast-pulsing experiments: SOFAST HMQC, BEST experiments

Part 2: Applications

Application to the study of protein folding and unfolding

HET-SOFAST NMR: a fast tool for characterizing protein structure
2D $^1$H-$^{15}$N correlation experiment: **SOFAST-HMQC**

**Band-Selective Optimized-Flip-Angle Short-Transient**

**Properties:**
* Selective pulses
* Small number of rf pulses: reduces loss due to $B_1$ inhomogeneity
* Good water suppression in only one scan

$^{1}$$H$-$^{15}$N correlation experiment: **SOFAST-HMQC**

**Band-Selective Optimized-Flip-Angle Short-Transient**

Properties:
* Selective pulses
* Small number of rf pulses: reduces loss due to $B_1$ inhomogeneity
* Good water suppression in only one scan

$^1$H-$^{15}$N correlation experiment: **SOFAST-HMQC**

**Band-Selective Optimized-Flip-Angle Short-Transient**

Properties:
* Selective pulses
* Small number of rf pulses: reduces loss due to $B_1$ inhomogeneity
* Good water suppression in only one scan
* Partial excitation (Ernst angle)

15N-ubiquitin (8.6 kDa, 25°C)

800 MHz, coldprobe

concentration: ~ 0.2 mM

Exp. time: 4 seconds

Fast-pulsing regime

Scan time [sec]

standard se-HSQC

SOFAST 90°
SOFAST 120°
SOFAST 150°
Optimal-sensitivity regime

SOFAST-HMQC performs best in terms of absolute sensitivity

\[ \alpha\text{-lactalbumin (14 kDa, 25°C)} \]

800 MHz, coldprobe

no \(^{15}\text{N}\) enrichment, \(~ 4.0 \text{ mM}\)

\(~ 10 \mu\text{M} \, ^{15}\text{N}-\text{labeled protein} \)

Exp. time: 1 hour

\[ t_{\text{rec}} \approx 0.2-0.3 \text{ sec} \]
Can we go even faster?

Combination with alternative encoding schemes

standard $^{15}$N labeling

Exp 1

Exp 2

Exp N

$n \Delta t_1$

$\Delta t_1$

$\cdots$

$N \approx 100$
Can we go even faster?
Combination with alternative encoding schemes

standard $^{15}$N labeling

Hadamard $^{15}$N labeling

Spatial $^{15}$N labeling

Exp 1

Exp 2

Exp N

$\Delta t_1$

$n \Delta t_1$

$N = 4, 8, 12, \ldots$

$N = 1 - 4$

$N \approx 100$
Hadamard encoded SOFAST HMQC

$^{15}\text{N-ubiquitin}$ (8.6 kDa, 25°C)

600 MHz concentration: ~2 mM

8 scans

Exp. time: 1 second

Spatially-encoded SOFAST-HMQC

3.4 mM ubiquitin
25°C, pH 6.0
800 MHz,
standard probe

Total acquisition time:
(8 interleaved scans)
1.4 seconds

Gal, Schanda, Brutscher & Frydman,
Extension to 3D triple resonance experiments:
Band-selective Excitation Short Transient Experiments
Example: 3D HNCA
3D HNCA experiment for resonance assignment in a few minutes

H-CA strips extracted from a 3D-BEST HNCA spectrum recorded in 15 min.
Outline

Part 1: Methods for fast multidimensional NMR

Concepts

Sensitive fast-pulsing experiments: SOFAST HMQC, BEST experiments

Part 2: Applications

Application to the study of protein folding and unfolding

HET-SOFAST NMR: a fast tool for characterizing protein structure
Dynamics in proteins studied with atomic resolution by NMR

- ps  ns  µs  ms  seconds  minutes

Steady-state NMR

Real-time NMR

Time (minutes)
Dynamics in proteins studied with atomic resolution by NMR

| ps | ns | µs | ms | seconds | minutes |

Steady-state NMR

Standard real-time NMR

SOFAST real-time 2D NMR

Protein (un)folding, H/D exchange,...
SOFAST real-time 2D NMR

Required tools

Fast data acquisition

SOFAST HMQC
SOFAST real-time 2D NMR

Required tools

Fast data acquisition

SOFAST HMQC

Initiation of the kinetic event

Fast mixing inside the spectrometer within ca. 100 ms
Application I

Folding of $\alpha$-lactalbumin followed in real time

$\alpha$-lactalbumin 14 kDa
Native state pH 8
Application I

Folding of apo $\alpha$-lactalbumin followed in real time

“Molten Globule”

pH 2

$\alpha$-lactalbumin 14 kDa

Native state pH 8
Application I

Folding of apo α-lactalbumin followed in real time

2D snapshots of the folding process every 10 seconds

fast pH jump

pH2

+10 sec

+120 sec
Application I

Folding of apo α-lactalbumin followed in real time

2D snapshots of the folding process every 10 seconds
Application I

Folding of apo α-lactalbumin followed in real time

No intermediate peaks observed

Same rate of disappearance of molten globule state and appearance of native state

Conclusion:
A single transition state ensemble controls folding
more folding...

Folding of an amyloidogenic protein

Unfolded → Intermediate(s) ? → Folded

? Fibrils

Work in progress

SOFAST-HMQC spectra recorded during refolding every 15 sec

Unfolded, pH 2 → pH jump → t=0 seconds

Final spectrum
more folding...

Folding of an amyloidogenic protein

Disappearance of I-state peaks

Appearance of native state peaks

Intermediate peaks

Multiexponential intensity changes

Site-specific analysis of kinetic behavior will reveal nature of these states.
Application II

(Un)folding under native conditions studied by H/D exchange

folded

$\text{k}_{\text{unfold}}$

$\text{k}_{\text{fold}}$

Unfolded
minor population
"NMR-invisible"
(Un)folding under native conditions studied by H/D exchange

H/D exchange gives access to the “invisible” manifold of partially unfolded conformations under native conditions
Fast H/D exchange gives insight into unfolding kinetics.

At high pH, the observed exchange rates directly reflect (partial) unfolding rates.

Application II
Global unfolding rate determined from fluorescence Chevron plots:

\[ k_{\text{NU}} = 0.0084 \text{ s}^{-1} \quad \tau_{\text{NU}} = 119 \text{ s} \]

Bofill, R., ..., Searle, M., *J. Mol. Biol.* (2005) 349, 205

SOFAST HMQC-based H/D exchange experiments on 0.7mM ubiquitin at pH 12, 298 K
Unfolding of ubiquitin studied by EX1 H/D exchange

* Unfolding in ubiquitin is a heterogeneous process
* Partial unfolding of the C-terminal face of the β-strand is faster
* Gradual decrease of unfolding rates
* The globular unfolding observed by fluorescence appears to be the unfolding of the N-terminal “folding nucleus”
Heterogeneous unfolding behavior of ubiquitin

**EX1 H/D exchange**

Cold denaturation

Ubiquitin in reverse micelles, -20°C

**Most thermostable H-bond:** I3 $\rightarrow$ L15

**Biological implication?**

Engineering a disulfide bridge between 4 and 66: 70-80% decrease of activity in signaling proteolysis
Ecker et al. (1989) J Biol Chem 264, 1887
Outline

Part 1: Methods for fast multidimensional NMR

Concepts

Sensitive fast-pulsing experiments: SOFAST HMQC, BEST experiments

Part 2: Applications

Application to the study of protein folding and unfolding

HET-SOFAST NMR: a fast tool for characterizing protein structure
HET-SOFAST NMR

Amide proton relaxation exploited for structural characterisation

Towards a quantitative measure of structure:
what distinguishes **structured** and **unstructured** polypeptide chains?

- high proton density
- low water accessibility
- low proton density
- high water accessibility
Amide proton relaxation exploited for structural characterisation

Amide proton relaxation depends on the structural context:
* number of aliphatic protons close in space
* local dynamics
* water exchange of $H^N$

Selective
Aliphatics saturated

$$\lambda = \frac{|\text{sat}|}{|\text{ref}|}$$
HET-SOFAST NMR

HET(eroogeneity) SOFAST NMR *
Quantification of proton density and water accessibility along polypeptide chains

$\text{H}_{\text{water/aliph}}$ $180^\circ$ $t_c$

$\text{H}_{\text{amide}}$ SOFAST HMQC

Ubiquitin (2 mM, 25°C, 600 MHz)
Acquisition time: 10 s / spectrum

2D HET-SOFAST NMR: characterizing structural compactness for assigned protein samples

Similar information to heteronuclear NOE but obtained in significantly reduced experimental time.
Structural and dynamic heterogeneities along polypeptide chains

prior to resonance assignment
HET-SOFAST NMR

Application: N-tailC, an unstructured protein with a small helical propensity

HET-SOFAST NMR allows to detect even partial structural preferences

Houben et al
HET-SOFAST NMR

Following subtle structural changes upon metal binding

\[ \alpha\text{-lactalbumin} \]

The structure of apo and holo forms are very similar.

Metal binding was reported to induce some “structure rigidification”

HET-SOFAST NMR experiments can be used to follow small changes of protein “compactness”
CONCLUSIONS

Longitudinal relaxation optimized experiments (SOFAST, BEST) allow significant acceleration of nD NMR data acquisition.

These approaches are fully compatible with alternative sampling methods (Hadamard, spatial encoding, reduced dimensionality,...)

Kinetic reactions in proteins can be probed simultaneously residue by residue on a time scale of seconds.

$^1H$ longitudinal relaxation can be used as a very sensitive probe for local structure and dynamics (HET-SOFAST NMR).
<table>
<thead>
<tr>
<th>Merci</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bernhard Brutscher</strong></td>
</tr>
<tr>
<td>Ewen Lescop</td>
</tr>
<tr>
<td>Mirjam Falge</td>
</tr>
<tr>
<td>Beate Bersch</td>
</tr>
<tr>
<td>Isabel Ayala</td>
</tr>
<tr>
<td>Cécile Giustini</td>
</tr>
</tbody>
</table>
Thank you

Folding
Vincent Forge

Pulses, technical things
Eriks Kupce
Peter Sandor

Ultrafast NMR
Lucio Frydman
Maayan Gal
Longitudinal $^1H$ relaxation enhancement

Magnetic field dependence

Longitudinal relaxation optimized (selective) experiments become important especially for high field strength.

Simulation

Experiment
Amide $^1$H relaxation in proteins: Insight from simulation

**Size of the molecule (tumbling rate)**

- **Effective $T_1$ [sec]**
  - Non-selective
  - Selective

**Magnetic field strength**

- **Effective $T_1$ [sec]**
  - Non-selective
  - Selective

$\tau_c \text{ [ns]}$

$^1$H resonance frequency
**Fast pulsing**

**Aliphatic $^1H$ polarization recovery**

- **Graph 1:**
  - Y-axis: Recovered polarization [%] vs. Effective amide $^1H$ $T_1$ [s]
  - The graph shows a decrease in polarization as $T_1$ increases.

- **Graph 2:**
  - X-axis: $t_{rec}$ [s]
  - Y-axis: Recovered polarization [%]
  - Data points for SE-AFB-HSQC, BEST-HSQC, and SOFAST-HMQC methods.

- **Diagram:**
  - SE-AFB-HSQC sequence with timing parameters $t_{rec}$, $\Delta$, $\tau$, and $\tau/2$.
  - EBURP-2 sequence with aliphatic flip back.

- **Formula:**
  - Effective amide $^1H$ $T_1$ [s]

- **Table:**
  - Recovered polarization [%]
  - Time $t_{rec}$ [s]
  - Methods: SE-AFB-HSQC, BEST-HSQC, SOFAST-HMQC
Choice of selective pulses

**Broadband Inversion Pulses**
(Shaka & co, JMR 2001, 151, 269)

**Time-optimized INEPT transfer**
using PC9 (Kupce & Freeman, JMR 1993, 102A, 122)

**Planar mixing without general rotation pulses**
using EBURP2 (Geen & Freeman, JMR 1991, 93, 93)
Hadamard encoded SOFAST HMQC
Hadamard encoded SOFAST HMQC

Retrieving the 2D information is achieved in only 2 scans!