Recent Advances in Biomolecular NMR

NMR in Cellular Structural Biology: from Single Molecules to Pathways

Lucia Banci

Magnetic Resonance Center (CERM)
University of Florence

EMBO Global Exchange Course
Singapore 6-13 December 2017
Recent Advances in Biomolecular NMR

• Mechanistic Systems Biology
  *To describe and understand biological processes at molecular level*

• In cell NMR
  *For studying biomolecules in a cellular context*

• Structural Vaccinology
  *Rational vaccine design based on the structural knowledge of the antigen*
Living systems are complex: mixture of proteins, nucleic acids, other biomolecules, several cellular compartments, ...etc

A **Systems Biology** approach is needed. All the players involved in a given process have to be considered as well as their 3D structural and dynamical interactions determined.

**Proteins must be framed within their cellular context**
Integrating Atomic Resolution with the Cellular Context

Copper trafficking in human cells

No free copper ions in the cytoplasm

$E^\circ$ of cytosolic glutathione = -289 mV, corresponding to GSH and GSSG in vivo concentrations of 13 mM and 0.7 mM.
Let’s start with a single process

Maturation of Cu,Zn-SOD1

monomeric apo hSOD1$_{\text{SH-SH}}$

SOD1: present in cytoplasm, mitochondrial IMS, nucleus, peroxisomes
dimeric (Cu$_2$,Zn$_2$) hSOD1$_{\text{SS}}$
Active enzyme:

\[
2\text{O}_2^- + 2\text{H}^+ \rightarrow \text{O}_2 + \text{H}_2\text{O}_2
\]

These post translational modifications affect the fold properties and monomer/dimer equilibrium
In-cell NMR can monitor functional processes in live human cells

Understanding intracellular processes at the molecular level requires a high resolution description. In-cell NMR provides atomic-level information on a protein in the cellular environment.

Transfected HEK293T cells are used as a model system for human cells.

Isotopically labelled proteins are overexpressed and directly observed by hi-res NMR in living human cells.

Maturation processes such as protein folding, post translational modifications (i.e. metal binding, disulfide bond formation) are followed at atomic resolution.
In-cell NMR can monitor functional processes in live human cells

Understanding intracellular processes at the molecular level requires a high resolution description. In-cell NMR provides atomic-level information on a protein in the cellular environment.

Transfected HEK293T cells are used as a model system for human cells.

Isotopically labelled proteins are overexpressed and directly observed by hi-res NMR in living human cells.

Maturation processes such as protein folding, post translational modifications (i.e. metal binding, disulfide bond formation) are followed at atomic resolution.
In-cell NMR can monitor functional processes in live human cells

Understanding intracellular processes at the molecular level requires a high resolution description. In-cell NMR provides atomic-level information on a protein in the cellular environment. Transfected HEK293T cells are used as a model system for human cells. Isotopically labelled proteins are overexpressed and directly observed by hi-res NMR in living human cells.

Maturation processes such as protein folding, post translational modifications (i.e. metal binding, disulfide bond formation) are followed at atomic resolution.

Following SOD1 maturation steps in human cells

Incomplete maturation of SOD1 fALS mutants

- ALS: a motor neuron disease
- 20% of familial cases is related to mutations of SOD1.
- 165 mutations identified so far, scattered throughout the sequence.

- Mutations are thought to cause defects in SOD1 maturation, promoting aggregation of the apo protein

Maturation defects of fALS SOD1 mutants

Many SOD1 mutants do not bind zinc in the cell, and accumulate as an unstructured species, which does NOT evolve toward the native form


The mutations do not affect zinc binding in vitro

This unstructured species DOES NOT bind zinc

It could be a precursor of SOD1 aggregates

Copper trafficking in human cells

- Cu(I) to Cu(II)
- Ctr
- Red?
- SOD1
- CCS
- Sco1, Sco2
- Cox17, Cox19, Cox23
- Mitocondria
- Nucleus
- Regulators
- Golgi
- Amine Oxidase, Lysyl oxidase
- Ceruloplasmin
- MT
- No free copper ions in the cytoplasm

- Hah1
- MNK/WLN
- Ceruloplasmin
- Cox11
Towards systems biology of copper

The knowledge of the structures of the proteins and of their complexes allows the atomic level description of the transfer processes

System-wide understanding of biological processes requires an atomic-level description of the functional processes
Iron Sulfur Biogenesis in human cells

ANAMORSIN

MiA40

BOLA2

Cfd1

Nbp35

Nar1

CIAO1

MIP18

MMS19

Holo Cytosolic Proteins

Fe/S-X

ACONITASE

L IPOATE SYNTASE

Complex I

HSPA9

Frataxin

Nfs1-Isl11

GRX5

GSH

SSH

Holo

Fe2+

Fe2+

NAD(P)H

FDX

ISCU

IBC57

NFU1

IND1

NFU1

IBA57

ISCA1

ISCA2

2x

Apol

CIA Machinery
Structural properties of Anamorsin
An essential protein for FeS cluster biosynthesis

- **N-terminal domain**
- **Intrinsically Disordered Domain**
  - This motif binds a cluster in the cytoplasm
- **Folded domain**
  - “standard” approach
- **Linker**
- **[2Fe-2S] cluster**
  - Highly paramagnetic with fast relaxation and no PCS
  - The worst case!!

Mia40 forms two disulfide bonds when in mitochondria.

Mia40 recognition site
Playing For or Against Paramagnetic Relaxation?
Exploit the differences in longitudinal relaxation

The effect of $T_1$ relaxation

IR delays selected according to $T_1$ values

IR is combined with fast recycling times
Playing For or Against Paramagnetic Relaxation?
Limit the negative effects of fast transverse relaxation

**INEPT coherence transfer**  
Antiphase signal acquisition

The effect of $T_2$ relaxation

- No relax
- 100 ms
- 10 ms
- 5 ms
- 2 ms
- 1 ms
- 0.5 ms

We can detect peaks for signals with $\Delta \nu$ up to ca. 300 Hz

Remove the reverse INEPT Phasing AP in dispersion mode
Tailored $^{15}$N HSQC of the [2Fe-2S]-domain of anamorsin

13 HN peaks, missing in standard $^{15}$N HSQC, can be detected. $^1H$ $T_1$ values range from 5 to 30 ms.

Banci et al, PNAS 2013, 2014; Ciofi-Baffoni, Gallo, Piccioli, J. Biomol NMR 2014
Paramagnetic-tailored $^{13}$C-direct CACO of [2Fe-2S]-domain of anamorsin

$^{13}$C signals, absent in standard $^{13}$C-direct experiments, are also observed via $^{13}$C COSY and tailored CON

Overall, about 10 additional $^{13}$C resonances are detected
Structure of the [2Fe-2S] cluster binding region in anamorsin

The overall structure was then subjected to MD trajectory in explicit water

The “paramagnetic” $^{13}\text{C}$ and $^1\text{H}$ $T_1$'s provide structural info around the FeS cluster

Blue - residues detected in the “diamagnetic” experiments

Cyano - residues whose $^{13}\text{C}$ or $^{15}\text{N}$ signals were detected in paramagnetic-tailored $^{13}\text{C}$ or $^{15}\text{N}$ experiments
Anamorsin receives the [2Fe-2S] clusters from GRX3 in the cytoplasm.

The Trx domain of GRX3 is essential for protein-protein recognition.

Cluster site characterization complemented with EPR and Moessbauer data.

...but in oxidative conditions, the BolA2-GRX3 complex is more efficient in anamorsin maturation.

The GRX3/BolA2 heterocomplex is more stable in oxidative conditions.

An alternative route to anamorsin maturation.

Electron transfer between Ndor1 and anamorsin

- Anamorsin tightly interacts with Ndor1 through its flexible, unstructured linker.
- The [2Fe-2S]-CIAPIN1 domain transiently interacts with the FMN domain of Ndor1 transferring one electron from FMN to the [2Fe-2S] cluster.

Banci, Bertini, Calderone, Ciofi-Baffoni, Giachetti, Jaiswal, Mikolajczyk, Piccioli, Winkelmann. *PNAS, 2013*
Electron transfer between Ndor1 and anamorsin
An integrated approach to investigate electrons and FeS cluster transfer processes

ESI-MS

Para-NMR

Solution NMR, X-ray, Protein docking


Biophysics: EPR, UV-vis, Mössbauer CD …

FF parameters determination

Banci, Ciofi-Baffoni et al. PNAS 2013

Banci, Ciofi-Baffoni, Bill et al. JBIC 2013

Banci, Ciofi-Baffoni et al. PNAS 2014
Iron Sulfur Biogenesis in human cells

- NDOR1
- ANAMORSIN
- MiA40
- CIAO1
- MIP18
- MMS19
- Apo
- Holo Cytosolic Proteins

- GRX3
- Cfd1
- Nbp35
- Nar1
- CIA Machinery

- ATP/ADP
- HSPA9
- Frataxin
- Nfs1-Isl11
- SSH
- ISCU
- ISCA2
- ISCA1
- IBA57
- NFU1
- Lipoate Synthase
- Complex I

- Fe/S-X
- ACONITASE
- Complex III

- Fe^{2+}
- NAD(P)H

- e^-

- e^-
Structural Vaccinology: the structure-based rational vaccine design
STRUCTURAL VACCINOLOGY

Past

Conventional vaccinology

Cultivate microorganism

5-15 years

Antigen selection and validation

vaccine

Reverse vaccinology

Genome-based approaches

Pan-genome approach

1-2 years

Determination of antigens structure

Present

Rational design of target epitopes
Interaction between fHbpC and a fAb portion of the antibody mAb502 (as studied by NMR)

- 900 MHz Spectrometer (298K)

Chemical shift mapping

![Chemical shift mapping](image)

**1H-15N HSQC-TROSY of fHbpC alone**

**1H-15N HSQC-TROSY of fHbpC:mAb502 mixture**

**1H-15N HSQC-TROSY of fHbpC alone**

**1H-15N HSQC-TROSY-CRINEPT of fHbpC:mAb502 mixture**

Residues predicted by immunological data and confirmed by NMR

NMR data suggested the involvement of other amino acids

The data show that mAb502 recognizes a conformational epitope within a well-defined area of the immunodominant C-terminal domain of fHbp.
Model of the complex between fHbpC & fAb portion of mAb502

Residues of fHbpC involved in binding to mAb502 mapped onto the full length protein structure

1. They are still solvent accessibles!
2. fHbpC contains the major part of the epitope
Complex of a monoclonal antibody with a Meningococcus B antigen (Factor H binding protein)

fHbp is very effective in inducing protective immunity eliciting antibodies but has different sequence in different strains of MenB

Structure of antigen fHbp

Light chain of monoclonal antibody Mab502

Heavy chain of monoclonal antibody Mab502

Fab region of antibody

Structure-based design of a vaccine against *Meningocococcus B*

By knowing the *structural* properties of the antigen and of the epitopes in all the variants, a *chimera antigen* was produced which elicits *complete protective immunity*
CERM/CIRMMP - a core center of Instruct

1200 MHz (2018?)

- 11 NMR spectrometers + 1 Relaxometer,
- The largest available magnetic field range (0.01 – 950 MHz)

EPR, Pulsed EPR, e-Research Infrastructure, Biobank, MS X-ray, Genexpress, Cellular biology lab, Biolabs, Conference room, Library, He liquefier, Relaxometer,
Thank you for your attention!!

We may anticipate that the chemist of the future who is interested in biomolecules will come to rely upon a new structural chemistry, and that great progress will be made, through this technique, in biology and medicine.

From the Nobel lecture of Linus Pauling, 1954