Protein Folding and Dynamics based on Chemical Exchange/Hydrogen Exchange

- Protein folding and misfolding
- Amide exchange and applications
- Monitoring folding and dynamics by detecting intermediates and excited states
  - EXSY, CPMG, RDC/PRE

Jenny J. Yang
Jenny@gsu.edu
Chemistry Department
Center for Diagnostics & Therapeutics
Georgia State University
http://www.gsuyanglab.com/research
Protein Folding

• Protein folding considers the question of how the process of protein folding occurs, i.e. how the unfolded protein adopts the native state.

• This has proved to be a very challenging problem. It has aptly been described as the second half of the genetic code, and as the three-dimensional code, as opposed to the one-dimensional code involved in nucleotide/amino acid sequence.
  – Predict 3D structure from primary sequence
  – Avoid misfolding related to human diseases
  – Design proteins with novel functions
Unfolded State

• The unfolded state is an ensemble of a large number of molecules with different conformations.
A new view of protein folding suggested that there is no single route, but a large ensemble of structures follow a many dimensional funnel to its native structure.

Progress from the top to the bottom of the funnel is accompanied by an increase in the native-like structure as folding proceeds.
MG is a Key Kinetic Intermediate
Molten Globule State (MG)

- It is an intermediate of the folding transition U→MG→F
- It is a compact globule, yet expanded over a native radius
- Native-like secondary structure, can be measured by CD and NMR proton exchange rate
- It has a slowly fluctuating tertiary structure which gives no detectable near UV CD signal and gives quenched fluorescence signal with broadened NMR chemical peaks
- Non-specific assembly of secondary structure and hydrophobic interactions, which allows ANS to bind and gives an enhanced ANS fluorescence
- MG is about a 10% increase in size than the native state
MG for folding and misfolding

Unfolding/Folding and Misfolding

**FIGURE 6-4.** Five hundred-MHz $^1$H-NMR spectra of guinea pig $\alpha$-lactalbumin in the native (pH 5.4), acid (pH 2.0), and unfolded (in 9-M urea) states recorded at 52°C. (Adapted from Baum et al., 1989)
Classic Models of Protein Folding

- **The Framework model** - Local elements of native local secondary structure could form independently of tertiary structure (Kim and Baldwin).

- **Diffusion-collision model** - These preformed 2nd elements would diffuse until they collided, successfully adhering and coalescing to give the tertiary structure (Karplus & Weaver).

- **The classic nucleation model** - Some neighboring residues in the sequence would form native secondary structure that would act as a nucleus from which the native structure would propagate, in a stepwise manner. Thus, the tertiary structure would form as a necessary consequence of the secondary structure (Wetlaufer).

- **The hydrophobic-collapse model** - A protein would collapse rapidly around its hydrophobic sidechains and then rearrange from restricted conformational space occupied by the intermediate. 2nd structure would be directed by native-like tertiary structure (Ptitsyn & Kuwajima).
Kinetic Folding Pathways

- U → I → II → N
- Not all steps have the same rate constants.
- Intermediates accumulate to relatively low concentrations, and always present as a mixture.
- Identify kinetic intermediates
- Measuring the rate constants
- Figure out the pathways
- Slow folding
  - Formation of disulfide bond
  - Pro isomerization
Equilibrium Unfolding

- Using many probes to investigate the number of transitions during unfolding and folding
- For 2-state unfolding, all probes give the same transition curves. Single domains or small proteins usually have two-state folding behavior.
- For 3-state unfolding, there are more than one transitions or different probes have different transition curves
Initial NMR dynamics experiments in 1970s.

Rapid advancements due to ability to label specific positions in bio-molecules and methodologies development

Magnetization exchange spectroscopy (EXSY)-slow exchange $0.5 \, \text{s}^{-1}$ to over $50 \, \text{s}^{-1}$

CPMG relaxation dispersion: chemical shifts $100 \sim 2000 \, \text{s}^{-1}$

R1rho can extend to more rapid exchange (dot)

Residual dipolar coupling (RDC) and PRE

Spin relaxation for ns-ps

CPMG and PRE are sensitive to low-lying excited states with populations $> 0.5\%$

H/D exchange can detect high energy excited states with much lower population

A. Mittermaier, L. Kay (2006) Science. 312, 224
Hydrogen exchange (HX) techniques is described for measuring the approximate exchange rates of the more labile amide protons in a macromolecule. The exchangeable amides in proteins are:
Exchangeable Nucleotides
Hydrogen-Exchange Chemistry

- Hx rate is catalyzed by OH⁻ and H⁺
  \[ k_{\text{ex}} = k_{oH} [\text{OH}^-] + k_{H}[\text{H}^+] + k_w \]
- All exchange rates are referenced to random coil polyAla at 0°C.
- HX rates are sensitive to pH, local chemical environment, solvent, sidechain type, neighboring amino acids and temperature
- \( k_{\text{intrinsic}} \) for each amino acid is different
- \( pD = \text{pH}^* + 0.4 \)

A minimum ~ pH 3.5
> 1hr at pH 3
< 1ms at pH 10

Bai. And Englander. (1993) Proteins, 17, 75;
Simulated Exchange Rates for Labile Protons of Polypeptides

In \( H_2O \) solution at 25 °C.

Im stands for imidazole ring NH, Gua for guanidinium NH, bb for backbone.

The amide protons have a large range of possible exchange rates under physiological pH (pH 6.5–7.5).

Wuthrich & Wagner JMB 1979
HX vs. Protein Structure

In proteins, HX rates can be altered:
- H-bonding
- Shielding in the center of protein
- Shielding by binding another molecules
- pH and temperature

Extremely slow exchange can be months, yrs
Protection factor \( \theta_p = \frac{k_{\text{intrinsic}}}{k_{\text{obs}}} \)

\( \theta_p > 10^6-10^7 \) for slow exchange

Amide exchange rate contains information about secondary structural elements
Hx Mechanism (Ex1/EX2)

Close $\xrightarrow{pK_{cl}}$ Open $\xrightarrow{k_{intrinc}}$ Exchanged

- Hvidt & Nielsen, 1966

- Solvent penetrates protein secondary structure
- A protected amide hydrogen is ‘closed’ to exchange and becomes accessible to exchange through an ‘open’ state at the exchange rate for an unstructured peptide.

Ex1: $k_{intrinc} \gg k_{cl}$ $k_{obs} = k_{op}$ independent of pH
Ex2: $k_{intrinc} \ll k_{cl}$ $k_{obs} = k_{op}k_{intrinc}$ pH dependent

Ex2 is typically encountered in proteins under conditions where folded state is stable and intrinsic exchange is relative slow
HX is an excellent way to look at the stability of proteins

• The rates of amide proton exchange for individual protons can be related to equilibrium constants for opening of individual hydrogen bonds. Knowing the equilibrium constants, one can calculate the free energy for the conformational transition which allows exchange to occur.

• When certain protons are only exposed in the completely unfolded form then the equilibrium constants and ΔGs correspond to the global unfolding reaction. These protons are usually the slowest exchanging protons in the molecule.

\[
\Delta G_{HX} = -RT \ln \left( \frac{k_{obs}}{k_{intrinc}} \right)
\]

• For mutation, the change of stability:

\[
\Delta \Delta G_{HX} = (\Delta G_{HX})_{wt} - (\Delta G_{HX})_{mut} = -RT \ln \left( \frac{k_{ex, wt}}{k_{ex, mut}} \right)
\]
Adding D\textsubscript{2}O to our H\textsubscript{2}O solution and take spectra at different times, signals from different amide protons will decrease in size at different rates. We look at the NH to H\textalpha fingerprint at different times in DQF-COSY or HSQC.
Amide exchange rates in *Escherichia coli* acyl carrier protein: Correlation with protein structure and dynamics

MICHAEL ANDREC, R. BLAKE HILL, AND JAMES H. PRESTEGARD
Department of Chemistry, Yale University, New Haven, Connecticut 06511

**Fig. 1.** Selected spectra from the amide H$_2$O/D$_2$O exchange time-course. A: HMQC spectrum obtained in H$_2$O. B, C, D: Spectra obtained 9, 63, and 133 min after addition of D$_2$O (see the Materials and methods). Peak corresponding to Ser 27 (10.07, 116.2 ppm) is visible in the spectra corresponding to A and B but is not shown in the plot limits of this figure. This signal does not appear in the spectra of C and D.

**Fig. 3.** Examples of signal loss due to D$_2$O exchange as a function of time after addition of D$_2$O for three selected residues. Peak height is in arbitrary units.
Amide Exchange Rates in ACP

- Residues at the center of helices and hydrophobic core have slow exchange rates.

The overall protection factors ($< 10^{4.5}$) are smaller than other proteins suggesting that ACP has high mobility.

Helix II exchanges faster than helix I and helix III suggesting that Helix II is highly dynamic.

The refolding experiment involved dilution of droplets of protein denatured in 6 M GuHCl in H$_2$O solution into a denaturant free solution of D$_2$O to initiate refolding and hydrogen exchange simultaneously.

After folding completed, HX is quenched by lowering the pH.

Comparing 2D NMR spectrum of the refolded protein with that which was not denatured. Residues protected early in refolding can be detected using NMR.
Competition HX of lysozyme

- Using 65 slowly exchanging amide hydrogen as probes.

The majority of residues in β-domain have exchange >30%. The majority of residues in α-domain have exchange <30% suggesting that two structural domains of lysozyme are folding domains that differ significantly in the extent to which protected structure accumulates early in the folding process.

After an adjustable refolding time, $t_f$, the protein is subjected to a short high pH pulse, where exchange of the unprotected NHs is very fast. NHs protected by structure within the folding time does not exchange during the short pulse.

After a pulse time $t_p$. The D to H exchange is quenched by rapidly lowering the pH.

After folding completed, the pattern of NH and ND labels in the refolded protein is analyzed by 2D NMR.

Increasing $t_f$ time, proton occupancies measured in the NMR spectrum decreases. Plotting proton occupancy vs. folding time $t_f$. 
Identifying Folding Pathway by HX Pulse-Labeling

(a) pure 2-state
All probes achieve 100% protection at the same rate in a single kinetic step.

(b) U -> I -> N sequential, I has A&B H-bonds with the same HX constant

(c) U1 -> N(30%) U2 -> N(70%) two heterogeneous parallel paths

(d) U1-> N U2->I->N contribution of intermediate and heterogeneous folding
Parallel Folding Pathway of Lysozyme

- All probes (50% of 126 amides) have one fast phase and one slow phase.
- Fast phase rates for both $\alpha$- and $\beta$ domains are 10 ms.
- Slow phase rates for $\alpha$ is 65 ms and $\beta$ is 350 ms, respectively.
- $\alpha$ domains fold before $\beta$ domain and different populations of molecules folded by kinetically distinct pathways.
Additional Methods for Amide-Water Exchange


CLEANEX-PM has the ability to specifically monitor water-proton exchange without 1) exchange relayed NOE/ROE from rapidly exchanging protons (hydroxyl or amide groups) in the macromolecules, 2) intra-molecular NOE/ROE peaks from protein $\text{CaH}$ protons which has chemical shifts coincident with water, or TOCSY-type interactions.
The FHSQC indicates proton signal that remain at the amide resonance through out the pulse sequence. The CLEANEX-PM indicates $^1$H signals that initiate in the $^1$H$_2$O resonance and then transfer to $^1$H amide resonance during the mixing period of the pulse sequence.

Staphylococcal nuclease – Hwang et al., 1998
Quantitation of rapid proton-deuteron amide exchange using hadamard spectroscopy

Catherine Bougault\textsuperscript{a}, Lianmei Feng\textsuperscript{b}, John Glushka\textsuperscript{b}, Eriks Kupče\textsuperscript{c} & J.H. Prestegard\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a}Structural Biology Institute, CNRS, Grenoble, France; \textsuperscript{b}Complex Carbohydrate Research Center, University of Georgia, Athens, GA, U.S.A. \textsuperscript{c}Varian Associates, Palo Alto, CA, U.S.A.

Received 14 October 2003; Accepted 20 October 2003

\textit{Key words:} HSQC, $^{15}$N, NMR, protein structure, ubiquitin

\textbf{Abstract}

The rates of amide proton exchange in protein backbones are very useful reporters of accessibility and structural stability of specific residues and secondary structure elements. Measurement by monitoring changes in intensity of cross-peaks in standard $^{15}$N-$^1$H HSQC spectra as protons are replaced by solvent deuterons has become widely accepted. However, these methods are limited to relatively slow rates due to time limitations of the conventional 2D HSQC experiment. Here we show that a Hadamard encoded version of the HSQC, which relies on a multiplexed, frequency selective, excitation in the $^{15}$N dimension, extends application to rates that are as much as an order of magnitude faster than those previously accessible.
Conventional [1H, 15N]-HSQC

A 0.5 mM 15N-labeled sample was prepared in 50 mM potassium phosphate buffer in H2O/D2O 5/95, pH=6.2 600 MHz with cryoprobe. It required approximately 21 min using 128 t1 time increments and 4 scans per increment.
Reconstructed Hadamard [1H,15N]-HSQC Spectra for Ubiquitin

(A) Data in $^1$H$_2$O collected with 128 t1 increments in 20 min. The sample was then lyophilized overnight and brought back to its initial volume with 99.9% $^2$H$_2$O and immediately returned to the spectrometer for rapid collection of a series of Hadamard spectra.
(B) First point after 1 min in $^2$H$_2$O collected with 4 scans in 42 s.
Cross-peak intensities as a function of time

Lines are best fits to

\[ I(t) = I_0 \exp(-kt) + \text{const}. \]

The precision of the data is quite high with the estimated errors for rates in the range of \(1 \times 10^{-3} \, \text{min}^{-1}\) being on the order of 5%. Rates derived also show reasonable agreement with previously published rates.
Exchange Spectroscopy (EXSY)

ClpP, an oligomeric protease comprising 14 subunits with a total molecular mass of 300 kDa. (a) Surface representation of ClpP with two monomers shown as yellow and blue ribbons. Locations of dynamic isoleucine residues are identified by green and red circles. Substrate entry pores are indicated with blue arrows. (b) The $^1$H/$^{13}$C methyl TROSY correlation spectrum collected for a uniformly [$^{15}$N, $^2$H], Ile $\delta_1$ [$^{13}$C, $^1$H] labeled ClpP sample. I149 and I151 are each associated with two $\delta_1$ methyl peaks, designated F and S, reflecting slow exchange between two distinct, functionally important, conformations. ($50^0$C, rotational correlational time >0.4 us)
Carr–Purcell–Meiboom–Gill (CPMG) Relaxation Dispersion

Exchange between ground state and excited state is in the millisecond time scale

\[ A \xrightleftharpoons[k_{BA}][k_{AB}] B \]

\[ k_{ex} = k_{AB} + k_{BA} \]

\[ k_{BA} > k_{AB} \]

Shape of the dispersion depend on:

- populations of the two states
- chemical shift difference
- the rate of exchange

In a typical series of experiments, variable numbers of refocusing pulses are applied to magnetization as it evolves under the influence of a chemical shift that varies stochastically due to the exchange process

Allosteric Transmission along a Loosely Structured Backbone Allows a Cardiac Troponin C Mutant to Function with Only One Ca^2+ Ion.

Quantitative analysis of exchange dynamics in cTnC D145E using CPMG


©2017 by American Society for Biochemistry and Molecular Biology
Probing Conformational Exchange Dynamics in a Short-Lived Protein Folding Intermediate by Real-Time Relaxation–Dispersion NMR

- Using relaxation–dispersion NMR and real-time NMR to reveal the conformational exchange dynamics present in short-lived excited protein states, such as those transiently accumulated during protein folding. Amyloidogenic protein \( \beta_2 \)-microglobulin folds via an intermediate state which is believed to be responsible for the onset of the aggregation process leading to amyloid formation.

- Rémi Franco, et al JACS, 2017
Residual Dipolar Coupling (RDC)

Spin Relaxation and Paramagnetic Relaxation Enhancement

PRE Study of Maltose Binding Protein

Grey: N-terminal domain
Blue: C-terminal domain, apo
Red: C-terminal domain, holo

- PRE of holo form agree with X-ray structure
- PRE of apo form is larger than X-ray structure
- The apo form is in a transient close form (5%)