Commendation International Young Chemistry Writer of the Year 2001

Studying how a Protein Folds

24 April 2002

Ever wondered how our body makes energy out of the food we eat? Have you ever thought who the workhorses are, when our brain demands some physical work or needs to sustain our life? How do the different chemical reactions occur in our body in a way that can't be performed easily in the laboratory? Asks M.M.G. Krishna.

Behind all the actions demanded by life, there are enzymes that catalyse different biochemical reactions. Enzymes are nothing but proteins. Proteins also carry out other major roles in the transport and processing of small molecules, in the generation and transmission of signals, and in controlling growth and differentiation. In short, proteins are an essential part of all life. Whether it is the binding of a small molecule or interaction with another macromolecule, the protein has to attain both the three-dimensional shape and structure required for its function.

Protein misfolding

How does a protein attain its three-dimensional structure? What happens if a protein misfolds? Protein misfolding causes diseases like Parkinson's, Alzheimer's, prion related diseases, and others that are just now being discovered. In all these diseases, critical examination of the affected individuals shows the presence of non-native non-functional protein aggregates. In order to understand the causes for these diseases and attempt to find a cure, and more fundamentally to understand how proteins attain their three-dimensional structures, many protein chemists, biologists, physicists and computer scientists are beginning to address the protein folding problem.

In order to study how a protein folds, we need a protein-specific signal that provides helpful information about a protein's folding status. It is difficult to do this *in vivo* due to interference from all of the other cell components. In general, what protein chemists try to do is to extract the protein from natural sources or overexpress the protein in some host organism, usually bacteria or yeast, by using well established molecular biology techniques. Given a sufficient quantity of the protein, protein folding reactions can be studied in a test tube in the laboratory.

Protein folding studies in a test tube

How are protein folding studies carried out in a test tube? The protein is first unfolded using different denaturing conditions, high or very low temperature, acidic or basic

related

stories from the Alchemist:

Alzheimer's Amyloid Proponents Pinpoint Precursors
news | research
8 April 2002

Unifying Pathway Points to Parkinson's
news | research
4 April 2002

Controlling foldamers
| hot papers
20 December 2001

Chemistry: the Queen and Servant of Biology
news | events
28 April 2000

links to ChemWeb.com:

International Young Chemistry Writer of the Year

Current Opinion in Structural Biology

services

EMAIL THIS STORY TO A FRIEND

Type e-mail address then hit return:

SEND

pH, or with chemical denaturants such as urea and guanidinium chloride. When the unfolded protein is returned to optimal conditions, it begins to fold back to its native state and may make that transition in a time that ranges from milliseconds to hours. Protein chemists try to gather information during this folding phase using special methods that can detect protein-specific signals. The traditional tools of the trade include a variety of spectroscopic methods, mainly fluorescence, absorbance, and circular dichroism. These methods provide localized information in the case of fluorescence and absorbance, or an averaged picture of the polypeptide chain in the case of circular dichroism. None of these spectroscopies can reveal what is happening to the polypeptide chain in any structural detail. The information one obtains from a single technique has been likened to what a blind man might learn about an elephant depending on what he happens to touch.

A different approach uses the technique of hydrogen exchange (HX) [1,2]. Although HX methods have been in use for almost 50 years, they have only very recently gained wide popularity among protein chemists to study protein folding. For example, a search of the Medline database with the keywords hydrogen exchange and protein folding provides the statistics for published papers, shown in Figure 1.

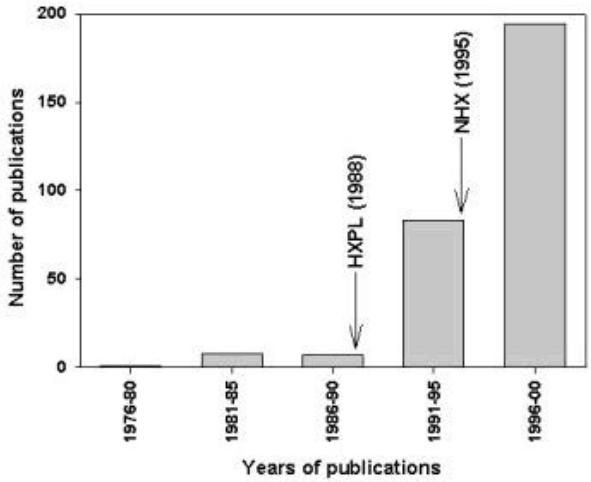

Figure 1: This graph shows the current literature indicating the usage of HX methods in studying protein folding. Most of the important contributions came from the past decade and the number of papers using HX methods is increasing rapidly. The arrows indicate the time periods during which the two major experiments that revolutionized the field were developed.

Hydrogen exchange

Let us understand the basic principles of HX. HX methods make use of the simple fact that the protons in polar bonds, such as amides (proton attached to nitrogen) or alcohols (proton attached to oxygen), exchange with solvent protons rapidly. Proteins are covalent polymers formed by amide (peptide) bonds between amino acids, and each of these carries an exchangeable proton (see Figure 2). When the protein is in an unstructured random chain, the amide protons can exchange freely with the solvent protons, for example, in about 1 second at pH 7 and 0 °C. But when the same amide proton is involved in some sort of hydrogen-bonding structure like a helix or sheet, the exchange is greatly slowed, and happens on a timescale of days or even months. Therefore, by measuring how fast or slow a particular amide proton exchanges with the water

protons, one can acquire information about structure formation *at that particular amide*.

Measured HX rates can yield information about how fast the structure opens (opening rate), or about the stability of the protecting structure, depending on conditions. Figure 2 illustrates these principles.

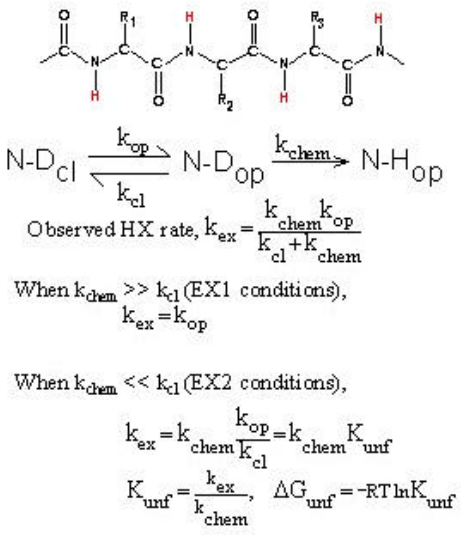

Figure 2: The random polypeptide chain with amide protons (colored red) distributed uniformly through out the sequence. The figure also shows the standard HX reaction scheme and the equations that define EX1 and EX2 limits. Under EX1 conditions, the exchange rate specifies the opening rate. Under EX2 conditions, it reveals the stability of the protecting structure. In the figure, k_{op}, k_{cl} and k_{chem} represent the opening and closing rates of the structure and the chemical exchange rate of the particular amide. K_{unf} and ?G_{unf} represent the equilibrium constant and the Gibb's free energy corresponding to the unfolding of structure that protects against exchange of that particular amide hydrogen.

HX is a chemical method rather than a spectroscopic method. As every amino acid has an amide proton (except for proline), one can acquire information about structure formation during the folding process at every amino acid along the protein sequence. That means, in the case of a protein with 100 amino acids, you have 100 signals (each from one amino acid) providing information about how that protein folds. Isn't it wonderful!

But how does one observe the exchange rate at each amide of a protein. Using two-dimensional nuclear magnetic resonance (2-D NMR) spectroscopy, one can resolve signals from each individual amide proton. As deuterium is not active in ^{1}H NMR, one can analyse how crosspeaks in 2-D NMR spectra change with exchange time to obtain the exchange rate of that particular proton.

HX pulse labeling

What are the different HX methods that are used by protein chemists? Here, I will explain mainly two HX experiments that have revolutionized the protein folding field in recent years. These were developed mainly in the laboratories of Professsor S.W. Englander at the University of Pennsylvania and Emeritus Professsor R.L. Baldwin at Stanford University , based on the earlier HX work by themselves and several other investigators [3]. The first method is the HX pulse

labeling (HXPL) technique [4].

In this method, the protein is initially unfolded in deuterated denaturant solution (D_2O). All the amide protons (NH) exchange with deuterium (ND). Rapid dilution into folding buffer solution using fast mixing stopped flow apparatus enables the protein to begin to fold. The folding buffer is at low pH so that no exchange happens during this phase. After some period of folding, a brief H-pulse is applied by increasing the pH of the solution to high pH which makes the exchange fast. Amide protons that have become protected by hydrogen-bonding structure formation remain as ND but amide protons that are still unprotected exchange to NH. A third mix into low pH terminates the HX labeling pulse. The protein folds into its native state, which freezes the labeling pattern. A 2-D NMR spectrum can then reveal which amide proton is D or H, that is which amino acids were protected in structure formed during the folding period. The brief HX-pulse essentially imprints structural information, on a millisecond timescale, that can be analysed later by 2-D NMR spectroscopy. Repetition of the experiment with different folding times can then yield information about the time course of structure formation at each amino acid as the protein folds.

How can we determine the stability of the structure formed? The intensity of the HX-pulse can be adjusted by varying the pulse pH. As the pH increases, the chemical exchange rate increases. In the regions where stable structure is formed, exchange will be slowed down by a factor that is proportional to the free energy of the structure, and stability can be quantified using the equations in Figure 2.

One of the first proteins to which this technique was applied is cytochrome *c*. It was observed that the N- and C-terminal helices form earlier than the other parts of the structure.

Native state HX

Another experiment which revolutionized the protein folding field is native state HX (NHX) [5]. The basic principles of NHX are as follows. The native state of a protein is the lowest free energy structure possible. All other intermediate structures and the unfolded state are at higher energy. Basic thermodynamics tells us that even when the protein is in its lowest energy native state, it has to cycle through these high energy states, as given by the Boltzmann distribution. This means that proteins must unfold and refold on a continuous basis, even under fully native conditions. When the protein visits its high energy intermediates transiently, the amide protons that become exposed exchange at a much faster rate than in the native state. So, if conditions are set in such a way to observe the exchange of the amide protons, then one can obtain the structures of the different intermediates through which the protein cycles.

When no denaturant is present, amide protons tend to exchange through small local unfolding reactions which obscure the large unfoldings that are characteristic of protein folding intermediates. By adding low levels of denaturant, the large unfoldings can be selectively promoted until they come to dominate the exchange that one measures. From these patterns, one can identify the otherwise invisible partially folded forms and obtain their stability.

For example, in the case of cytochrome *c*, the NHX

experiment reveals four unfolding units [6]. The free energy of all the amide protons in a structural unit joins into a HX isotherm as the denaturant concentration is increased. The unit (i) of lowest free energy unfolds first followed in order of energy by the second (ii) and third (iii) units. Finally, with the unfolding of the fourth (iv) unit, the protein unfolds completely. In other words, cytochrome *c* has three thermodynamically stable intermediates in between its native state (all four units folded) and the unfolded state (all four units unfolded): i open, i plus ii open, and i plus ii plus iii open. Wow!! No other spectroscopic/chemical technique has been able to identify folding intermediates so clearly and precisely.

A number of ensuing experiments carried out on cytochrome *c* mainly in Englander's lab confirm that the four structurally independent folding units detected by NHX do open sequentially in an unfolding pathway. In that case, the refolding pathway has to follow the same sequence in reverse. For example, the first unit seen to fold by HXPL, the N/C helical unit (iv), is the last unit to unfold, as seen by NHX.

The NHX method identifies the intermediates through which the protein cycles at infinitesimal and otherwise invisible levels under native conditions whereas the HXPL method identifies kinetic intermediates that accumulate transiently as the protein folds from the unfolded state. These two methods have been applied to a number of proteins in recent years. In some cases, the structures of the equilibrium intermediates determined by NHX were found to be the same as those of kinetic intermediates determined by HXPL.

What have we learnt so far about how a protein folds? All the information now available consistently supports the following picture of how a protein folds. The protein initially nucleates to a structure that has a native like topology and then rearranges itself through different intermediates that produce a pathway to the native state. The intermediates are robust in the sense that the same intermediates are observed under different conditions whether the protein is unfolded using temperature, pH or any chemical denaturant. The previously mysterious protein folding intermediates turn out to be partial replicas of the native protein. The intermediates are discrete and native like. The intermediates utilize the same secondary structural elements as in the native protein such as helices or sheets as the building blocks. In the kinetic folding process, interactions between the various secondary structures enable the initially formed intermediates to guide and stabilize the later ones.

As described above, HX methods have revolutionized how protein chemists study protein folding. The number of reported studies that use HX methods are increasing day by day. Although it will require many more proteins to study to come to a unique, undisputable picture of understanding how a protein folds, I am quite sure the HX methods that have been developed based on simple principles will take this branch of chemistry to new heights. So far we have seen only the initial phase of these advances!

M.M.G. Krishna is at the Department of Biochemistry and Biophysics, University Pennsylvania School of Medicine, Philadelphia, PA.

Acknowledgements: The author thanks Prof. S. Walter Englander for many helpful discussions regarding the subject in this manuscript.

References

1. S.W. Englander & N.R. Kallenbach. **Hydrogen exchange and structural dynamics of proteins and nucleic acids.** *Q Rev Biophys* 1984, **16:**521–655.

This review provides an excellent account of early HX studies.

2. S.W. Englander. **Protein folding intermediates and pathways studied by hydrogen exchange.** *Annu Rev Biophys Biomol Struct* 2000, **29:**213–238.

This review summarizes the application of HX methods to study protein folding and the conclusions obtained so far.

3. S.W. Englander & M.M.G. Krishna. **Hydrogen exchange.** *Nature Struct Biol* 2001, **8:**741–742.

This paper provides a brief history of hydrogen exchange.

4. H. Roder, G.A. Elove & S.W. Englander. **Structural characterization of folding intermediates in cytochrome *c* by H- exchange labeling and proton NMR.** *Nature* 1988, **335:**700–704.

This is one of the first papers where HX pulse labeling combined with NMR was used to identify the sequentiality of structure formation during the protein folding.

5 Y. Bai, T.R. Sosnick, L. Mayne & S.W. Englander. **Protein folding intermediates studied by native state hydrogen exchange.** *Science* 1995, **269:**192–197.

This is the first paper where native state HX was used to identify the thermodynamically stable folding intermediates.

6. S.W. Englander. **In pursuit of protein folding.** *Science* 1993, **262:**848–849.