Published in final edited form as:

Int J Mass Spectrom. 2012 December 15; 330-332: 63–70. doi:10.1016/j.ijms.2012.10.010.

Theory of the Protein Equilibrium Population Snapshot by H/D Exchange Electrospray Ionization Mass Spectrometry (PEPS-HDX-ESI-MS) Method used to obtain Protein Folding Energies/Rates and Selected Supporting Experimental Evidence

Rohana Liyanage, Nagarjuna Devarapalli, Derek B. Pyland, Latisha M. Puckett, N. H. Phan, Joel A. Starch, Mark R. Okimoto, Jennifer Gidden, Wesley E. Stites, and Jackson O. Lay Jr.
Department of Chemistry and Biochemistry University of Arkansas, Fayetteville, AR 72701

Abstract

Protein equilibrium snapshot by hydrogen/deuterium exchange electrospray ionization mass spectrometry (PEPS-HDX-ESI-MS or PEPS) is a method recently introduced for estimating protein folding energies and rates. Herein we describe the basis for this method using both theory and new experiments. Benchmark experiments were conducted using ubiquitin because of the availability of reference data for folding and unfolding rates from NMR studies. A second set of experiments was also conducted to illustrate the surprising resilience of the PEPS to changes in HDX time, using staphylococcal nuclease and time frames ranging from a few seconds to several minutes. Theory suggests that PEPS experiments should be conducted at relatively high denaturant concentrations, where the protein folding/unfolding rates are slow with respect to HDX and the life times of both the closed and open states are long enough to be sampled experimentally. Upon deliberate denaturation, changes in folding/unfolding are correlated with associated changes in the ESI-MS signal upon fast HDX. When experiments are done quickly, typically within a few seconds, ESI-MS signals, corresponding to the equilibrium population of the native (closed) and denatured (open) states can both be detected. The interior of folded proteins remains largely un-exchanged. Amongst MS methods, the simultaneous detection of both states in the spectrum is unique to PEPS and provides a “snapshot” of these populations. The associated ion intensities are used to estimate the protein folding equilibrium constant (or the free energy change, ΔG). Linear extrapolation method (LEM) plots of derived ΔG values for each denaturant concentration can then be used to calculate ΔG in the absence of denaturant, ΔG_{H_2O} . In accordance with the requirement for detection of signals for both the folded and unfolded states, this theoretical framework predicts that PEPS experiments work best at the middle of the denaturation curve where nated and denatured protein molecules are equilibrated at easily detectable ratios, namely 1:1. It also requires that closed and open states have lifetimes measurable in the time frame of the HDX experiment. Because both conditions are met by PEPS, these measurements can provide an accurate assessment of closed/open state populations and thus protein folding energies/rates.

© 2012 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Protein function within biological systems is highly dependent upon conformational state. Typically, a protein will maintain a folded state. Changes in the concentrations of denaturants such as urea or guanidine hydrochloride (GdHCl) or changes in pH and temperature can disrupt the forces stabilizing the folded state and denature the protein. Unfolding and changes in protein structure may contribute to disease as the efficacy of biologically important proteins is reduced [1]. Thus, various analytical methods to determine protein folding energies and protein folding/unfolding rates, described below, have been developed. These methods generally have some drawbacks and there is a need for the development of additional methods to supplement existing approaches. For example, methods requiring less time and protein and that are more tolerant of impurities may provide new tools for understanding how protein degradation or destabilization contributes to disease conditions.

Previous studies have often applied spectrophotometric methods to “probe” GdHCl/urea denaturation of proteins to extract protein thermodynamic stability. Even though spectrophotometric methods such as tryptophan fluorescence or circular dichroism (CD) are widely used to determine protein folding energies, there are several drawbacks [2,3]. Fluorescence studies provide relatively localized data for the region of the protein proximal to a tryptophan side chain. The usefulness of its fluorescence is also limited by the low abundance of tryptophan in most proteins. Although CD provides more complete, global data relevant to the protein as a whole, the signal obtained from this technique is weak. This technique also requires relatively large amounts of protein (milligrams) and also the protein must be pure. On the other hand, hydrogen deuterium exchange (HDX) kinetics studies of amide protons by NMR spectroscopy are very useful and provide highly accurate [4] information about specific regions of the molecule. NMR provides very precise information, including rates of protein folding and unfolding. NMR also requires specific hardware and skills for the interpretation of spectra from large proteins and more importantly also usually requires larger quantities of pure material.

Mass spectrometry is an alternative to NMR for monitoring HDX. It is a powerful tool for analyzing many different biologically relevant molecules including proteins. It has been widely used in protein identification [5,6], quantification [7,8], primary/secondary and tertiary structure elucidation [9], and in many other applications. Of particular importance here, the measurement requires very small amounts of material (micrograms on the high end and picograms on the low end have been reported) and works even for complex mixtures, thus overcoming many of the limitations associated with other techniques. Mass spectrometry-based methods used to determine protein folding energies typically utilize either hydrogen deuterium exchange (HDX) or controlled oxidation. HDX kinetics of backbone amide protons [10,11,12], or imidazole C₂ protons in histidine side chains [13] are typical examples of the first approach. The chemical modification kinetics of methionine oxidation reactions [14] is an example of the other general approach to estimating protein folding energies. The major advantage of the use of HDX with amide protons is that it is easily applied to almost any protein, as essentially all proteins have a hydrophobic region with globally protected amide hydrogen atoms which are exposed to solvent for exchange upon unfolding. The main disadvantage of this strategy is the sensitivity to back-exchange, which occurs rapidly and is difficult to control under typical LC-MS conditions. In contrast, studies using slowly exchanging imidazole C₂ protons in histidine [13] or methionine oxidation [14] to probe for unfolding are essentially unaffected by reversal of the reactions under LC-MS conditions. However, based on rarity, histidine and methionine are not the global probes of structure that the ubiquitous amide protons are.

Stability of unpurified proteins from rates of H/D exchange (SUPREX) [10, 11, 12] and stability of proteins from rates of oxidation (SPROX) [14] have been used to obtain both protein folding energies and affinity for binding proteins and ligands. Protein-ligand interactions by mass spectrometry, titration, and HDX (PLIMSTEX) [15,16] have also been used to obtain protein-ligand and protein-protein interaction data. A modified version of PLIMSTEX, self-association interactions using mass spectrometry, self-titration and HDX (SIMSTEX) [17], has been used to obtain affinity constants for self-association of proteins. All of these HDX methods are kinetic-based.

Recently, we introduced protein equilibrium snapshot by hydrogen/deuterium exchange electrospray ionization mass spectrometry (PEPS-HDX-ESI-MS or PEPS) to determine protein stabilities. The experimental method and the theoretical underpinning of PEPS are both much simpler than they are for these other approaches. Moreover, none of HDX-MS methods mentioned above can be simultaneously used to extract protein folding and unfolding rates. To our knowledge PEPS [18] is the only method which does not employ kinetics or time dependence of HDX to obtain protein thermodynamic stability and also the only method that can be used simultaneously to estimate protein folding/unfolding rates. The PEPS experiment obtains a “snapshot” of the relative equilibrium amounts of the open and closed populations of a protein at varying the concentration of GdHCl and using HDX during a short probing time “t”. When this probing time is long compared to the time required to exchange protons and short with respect to the lifetimes of the folded/unfolded states, PEPS will provide ion intensities reflective of the deuterated closed and open states of a protein from which protein folding energies and rates can be easily extracted. In contrast to our first publication [18], in this report, we outline the theoretical aspects of the PEPS method and explain in some detail when and why PEPS works. We do this using two model proteins, ubiquitin and staphylococcal nuclease. Ubiquitin was selected because it has known folding and unfolding rates [4]. Staphylococcal nuclease was selected because it is less stable than ubiquitin. PEPS experiments were conducted using lower denaturant concentrations because of the proteins lower stability. A wider range of exchange times compared to ubiquitin were also used to demonstrate that reasonably accurate protein folding energies could be extracted over a larger range of exchange times [19]. This work reports new experimental data targeted at (a) providing a benchmark comparison of our data with accepted NMR embedded in the description of PEPS theory (with ubiquitin) and (b) demonstrating the surprising insensitivity of the method to changes in timeframe of H/D exchange (using staphylococcal nuclease). Also, for the first time, in this manuscript, we show how PEPS experiments can estimate protein folding and unfolding rates. .

2. Experimental

2.1 Chemicals

Bovine ubiquitin, sodium hydrogen phosphate, and acetonitrile (ACN) were purchased from Sigma-Aldrich. Trifluoroacetic acid (TFA) was purchased from Halocarbon. Guanidine hydrochloride (GdHCl) omni pure was purchased from VWR. Sodium phosphate and sodium chloride were purchased from Mallinckrodt. Deuterium oxide (D₂O; 99.8% atom D) was purchased from Cambridge Isotope Laboratories. The expression and purification of the staphylococcal nuclease used in this study has been described elsewhere [19, 20]. H₂O and D₂O phosphate buffers were prepared to have final concentrations of 100 mM NaCl, and 25 mM sodium phosphate, at pH 7.0. A stock solution of 6M GdHCl was prepared in 100 mM NaCl with 25 mM sodium phosphate at pH 7.0. Deuterated guanidine hydrochloride (GdDCI) was prepared by ten cycles of evaporation and reconstitution to the same volume with pure D₂O to maintain the buffer condition and pH. All protein samples were dissolved in the H₂O buffer to achieve a final protein concentration of about 2 mg/ml. The pH

measurements were not adjusted for deuterium. The HPLC mobile phases were 0.1% TFA (A) and 0.1% TFA in ACN (B).

2.2 Instrumentation

A Bruker Esquire 2000 (Billerica, MA) LC ion trap equipped with an electrospray ionization source was used for LC-MS. It was operated in the positive ion mode with a nebulizing gas pressure (N_2) of 32 Psi and a drying gas flow of 12 ml/min maintained at 250° C. The mass spectrometer was optimized at m/z 1000 with low skimmer voltage (instrument default for this mass). The HPLC was a Hewlett Packard (Palo Alto, CA) 1100 series instrument equipped with an autosampler. HPLC (desalting) was accomplished using a Supelco C₁₈ column (4.5 mm × 50 mm, 5 μ m) at a flow rate of 0.8 ml/min using a rapid gradient from 5%B-100B% over 3 minutes.

2.3 PEPS HDX-ESI-MS method

The equilibrium protein folded (closed) and unfolded (open) populations were first established by mixing 5-10 μ l of protein stock solution with specific volumes of 6 M GdHCl stock solutions and H₂O phosphate buffer described above to achieve specific final concentrations of GdHCl in the test solution. For staphylococcal nuclease the denaturant concentration ranged between 0 to 2.5 M. For the more stable protein, ubiquitin, it ranged between 0 to 5 M. In all cases 100 μ l was the total volume. The protein concentration was selected to give good S/N by mass spectrometry. Because of the small volumes used in PEPS experiments, low microgram quantities of protein are required for each experiment. Equilibrium for all solutions was established at 25° C using a water bath. Sufficient time, on the order of 5-10 minutes, was allowed to insure proper equilibration of the native closed and denatured open states after the addition of denaturant and before the HDX. After presumed establishment of the denaturant-induced equilibrium between the open and closed states HDX was initiated by dilution of 5 μ l of protein solution into 45 μ l of exchange buffer made of the same concentration of GdDCI in D₂O and D₂O phosphate buffer. After quickly mixing the two solutions, HDX quenching was accomplished at the desired times by 1:1 dilution into 50 μ l of 0.2% TFA incubated in an ice cold bath. For ubiquitin, 1s HDX experiments were performed (at approximately 0.25M increments) between 0 to 5 M GdHCl. For staphylococcal nuclease 1s, 5s, 15s, 1.5 min, 11 min, and 19 min HDX experiments were performed between 0 to 2.5 M GdHCl with denaturant concentrations selected at about 0.06M increments. Half of this (25 μ l) sample was immediately injected using a Rheodyne injector for HPLC/MS analysis. Both HPLC solvents, A and B, were kept in an ice bath continuously. The HPLC column was also kept at 0° C.

3. Theory and Data Analysis

3.1 Theory

The PEPS method is an equilibrium-based method where the populations of the closed and open states of a protein, at a given denaturant (GdHCl) concentration, are measured directly from the intensities of two m/z peaks observed by ESI-MS. Knowledge of these populations is then used, with linear extrapolation to denaturant free conditions, to calculate protein folding energies and protein folding rates. We recently applied PEPS to ubiquitin and both wild-type and mutant staphylococcal nucleases. Based on benchmark data from other methods, PEPS was demonstrated to be an accurate and easy to use method to obtain protein folding energies [18]. We presumed that this approach worked because both the open and closed state populations had sufficiently long lifetimes that they could be sampled by mass changes after brief HDX. The implication is that for the short times of these studies the exchange and measurement processes are completed before any significant changes occur in the folded and unfolded populations. A typical example of an unprocessed and deconvoluted

spectrum, where these two states are easily resolved can be seen in Figure 1(a), (b) for ubiquitin.

The HDX reaction, as suggested in Figure 2, can probe the entire protein backbone. When the protein is folded (closed), part of the peptide backbone is buried in the hydrophobic core and the amide protons in that region are protected from solvent exposure. This hinders deuterium from exchanging with these protons. In other words, when protein has intact H-bonding (is in its folded or closed state), amide protons have less availability or accessibility to exchange. When the protein unfolds (open), this globally protected region becomes accessible and can be exchanged with deuterium in the solvent in conditions with faster HDX rates compared to the folding and unfolding rates. As a result, the closed and open states can be differentiated by mass, yielding two separate and assignable signals by ESI-MS as shown in Figure 1.

To a first approximation, the folded and unfolded states are fixed over the timeframe of the experiment. Folding and unfolding of the proteins would cause loss of information. Proteins that unfolded during the measurement process would be deuterated in the “protected” region and retain their increased mass, even if the protein refolded, and “counted” as unfolded. Thus, for PEPS to work, the populations must be relatively stable over the timeframe of the experiment. To minimize problems with folding/unfolding, the typical PEPS experiment employs short fixed exchange times and varies only [GdHCl]. Some obvious questions are: (1) Why does this work? (2) What are reasonable time frames for PEPS experiments? and (3) How hard is it to tell when PEPS is not working? A classical “breathing model” [21,22,23] involving a two state folding mechanism is used to model the HDX of the interior region in the PEPS experiment and better explain how and why PEPS works.

In broad terms, there are two accepted HDX mechanisms, named EX1 and EX2, referring to the rate of exchange of the globally protected amide protons of a protein [24]. EX1 is defined as the condition in which a single unfolding event results in a majority, or the entire molecule's, globally protected amide protons exchanging. Under EX2 conditions, a single unfolding event results in little or no exchange. PEPS works accurately in EX1 region where the folding rate (k_{close}) and unfolding rate (k_{open}) are significantly less than the intrinsic exchange rate ($k_{\text{int}}^{\text{H/D}}$). The EX1 region is shown in figure 3(a) for ubiquitin based on data from denaturation studied using NMR [4]. Figure 3(b) shows the same data as 3(a) except that the vertical axis shows the “life time” associated with folding states rather than rates. Examination of the magnitude of these values provides some insight into how PEPS works and when. From this figure it is possible to compare the time frames for which the folded and unfolded molecules survive (1-10s). The figure also depicts the timeframe needed to exchange an exposed amide proton ($\ll 1$ s). For ubiquitin, it is easy to see that for a GdHCl concentration of around 3.5 - 4 M and an exchange time of about 1 sec the ESI mass spectra should show population correlated signals for both the deuterated-folded and deuterated-unfolded proteins. This is because the life times of the folded (closed) and unfolded (open) states are sufficiently large with respect to the exchange time that it is possible to sample the equilibrium before a significant number of folding/unfolding events take place.

Using this graph and theory, it is a simple matter to predict what errors should occur for exchange times approaching (or exceeding) the life time of the folding/unfolding state and at higher or lower denaturant concentrations. For example, at the low end of the concentration scale in Figure 3(a) where EX2 is the predominant HDX mechanism, the life time of open state is much shorter and only partial or no deuterium exchange will occur during a single protein unfolding event. Under this situation, it is unlikely that two discrete peaks would be observed. At the high end of the concentration scale in Figure 3(a), the open state is persistent and the closed state has a short life time. In this case also, only one ESI-MS signal

is observed, for the deuterated-unfolded state. The deuterated-folded state could not be measured (because of its low intensity) or easily interpreted if detected (because of folded-state turn over during the experiment). In less extreme situations, where the EX1 mechanism is predominant as shown in Figure 3(a), 3(b) the open state will be over represented because of the inclusion of proteins that were closed at the onset of HDX exchange, but which opened during the measurement process and were counted as open because of the mass change. Fortunately this situation can be predicted mathematically and compared to experimental data (see below) and also can be avoided by employing short HDX times.

3.2 Calculation of the observed closed and open state signals in PEPS

For the simplest case with PEPS, the equilibrium constant at given GdHCl concentration and temperature, K_{app} can be obtained directly from intensity values of the closed and open states observed in the mass spectra after HDX: ($K_{app}=I_{closed}/I_{open}$).

How can we explain the validity of this simplification? This manuscript is mainly intended to address this question. A logical starting point would be to relate the rate of exchange of the globally protected region to the folding and unfolding rates of a protein as described below. The standard equation for the exchange rate of the globally protected amide protons (k_{ex}) is given by Equation (1) using the classical HDX model (breathing model) [21, 22, 23, 25].

$$k_{ex} = (k_{open}k_{int}^{H/D}) / (k_{open} + k_{close} + k_{int}^{H/D}) \quad (1)$$

Where, k_{open} , k_{close} , and $k_{int}^{H/D}$ refer to the rate of opening, rate of closing, and the intrinsic exchange rate of unprotected amide protons respectively at 25° C. It is worth noticing here that in the EX1 regime, when $k_{int}^{H/D} \gg k_{close}$ and k_{open} , k_{ex} approximates to k_{open} , as shown in Figure 3(a), 3(b). The intrinsic HDX rate ($\sim 5 \text{ s}^{-1}$ at pH=7, at 25° C) [10] is depicted by the solid horizontal line in Figure 3(a). If the normalized closed state population is assumed to be $N_{closed}^{H,H}$, then the normalized open state population would be given by $N_{open}^{H,H} = 1 - N_{closed}^{H,H}$ (H,H refer to two types of amide protons, the first “H” is for the already exposed proton and the second “H” is for the globally protected amide protons. This second “H” is exposed only in the open state). Hypothetically, if no folding/unfolding takes place during HDX time “t”, the population of the closed and open states that will be exchanged is the same as the normalized equilibrium population of the closed and open states initially established based on the GdHCl concentration ($N_{closed}^{H,H}$, and $N_{open}^{H,H}$). Hence without any HDX of the globally protected region due to unfolding during HDX time “t”, $N_{closed}^{H,H}$ and $N_{open}^{H,H}$ (initial values) should be equal to $N_{closed}^{D,H}$ and $N_{open}^{D,D}$ (after exchange) in EX1 region where $k_{int}^{H/D}$ is $\gg k_{open}$ and k_{close} . The correspondence between these populations and the observed ion intensities is the basis for PEPS. Amongst other things, the accuracy of the population measurement will depend on how quickly the quenching can be performed, after initiation of exchange, to minimize the number of folding/unfolding events occurring during HDX time “t”.

The actual expected equilibrium population (and thus the expected ESI-MS signals) at $t=0$, $N_{closed}^{D,H}(t=0)$ and $N_{open}^{D,D}(t=0)$, can be calculated from k_{closed} and k_{open} using values derived from NMR data for GdHCl concentrations between (0-6)M [4].

$$\begin{aligned} N_{closed}^{D,H}(t=0) &= k_{closed} / (k_{closed} + k_{open}), \\ N_{open}^{D,D}(t=0) &= 1 - N_{closed}^{D,H}(t=0) \end{aligned} \quad (2)$$

In the real-world case for a PEPS experiment performed for HDX time “t”, the populations of the closed state remaining globally protected (observed ESI-MS signal for closed state) and exchanged are given by Equations (3) and (4).

$$I_{\text{closed}}(\text{PEPS, theory}) = N_{\text{closed}}^{D,H}(t=t) = N_{\text{closed}}^{D,H}(t=0) e^{-k_{\text{ex}} \cdot t} \quad (3)$$

$$N_{\text{closed}}^{D,D}(t=t) = N_{\text{closed}}^{D,H}(t=0) (1 - e^{-k_{\text{ex}} \cdot t}) \quad (4)$$

$N_{\text{closed}}^{D,D}(t=t)$ is the population of closed state at $t=0$ that unfolded during the HDX time “t” and resulted in exchange of the globally protected region (denoted by the second “H”). Hence the open state ESI-MS signal, $I_{\text{open}}(\text{PEPS, theory}) = N_{\text{open}}^{D,D}(t=t)$, is given by the sum of the initial open state population $N_{\text{open}}^{D,D}(t=0)$ (before HDX or at $t=0$) and $N_{\text{closed}}^{D,D}(t=t)$ as shown in Equation (5).

$$I_{\text{open}}(\text{PEPS, theory}) = N_{\text{open}}^{D,D}(t=t) = N_{\text{closed}}^{D,D}(t=t) + N_{\text{open}}^{D,D}(t=0) = N_{\text{closed}}^{D,H}(t=0) (1 - e^{-k_{\text{ex}} \cdot t}) + N_{\text{open}}^{D,D}(t=0) \quad (5)$$

Using equations 1, 2, 3, 4, and 5, the apparent PEPS folding equilibrium constant at [GdHCl] (M) and temperature $T^\circ \text{C}$ (typically at 25°C) is given by,

$$\begin{aligned} K_{\text{app}}^{\text{PEPS}(\text{theory})} &= I_{\text{closed}}(\text{PEPS, theory}) / I_{\text{open}}(\text{PEPS, theory}) \\ &= (k_{\text{close}} e^{-k_{\text{ex}} \cdot t}) / (k_{\text{open}} + k_{\text{close}} (1 - e^{-k_{\text{ex}} \cdot t})) \end{aligned} \quad (6)$$

Once the equilibrium folding constant was calculated for each concentration of GdHCl, the linear extrapolation method (LEM) was used to calculate the physiologically relevant folding energy by the equation

$$\Delta G_{\text{app}}^{\text{PEPS}(\text{theory})} = \Delta G_{\text{H}_2\text{O}}^{\text{PEPS}(\text{theory})} + mC \quad (7)$$

where $\Delta G_{\text{app}}^{\text{PEPS}(\text{theory})} = -RT \ln K_{\text{app}}^{\text{PEPS}(\text{theory})}$, $\Delta G_{\text{H}_2\text{O}}^{\text{PEPS}(\text{theory})}$ is the folding energy at 0M denaturant, C is the molar concentration of the denaturant, and m is $\delta \Delta G_{\text{H}_2\text{O}} / \delta C$. Equation (7) is originated from the assumption that the protein denatures linearly with the denaturant or in other words, free energy change per unit change in denaturant concentration ($\delta \Delta G_{\text{H}_2\text{O}} / \delta C$) is a constant.

3.3 PEPS-HDX-ESI-MS Experimental Data Analysis

The PEPS method is an equilibrium-based method where the population of the closed and open state for a given GdHCl concentration is inferred from the intensities of the two m/z peaks in the ESI mass spectra. As shown in figure 1, two clear peaks are readily detected, with very little (or no) change in the m/z (directly or after the deconvolution), and whose intensity varies as a function of GdHCl. Hence, $K_{\text{app}}^{\text{PEPS}(\text{exp})}$ is obtained simply from the ratio of these two signals at any given GdHCl concentration as shown in Equation 8.

$$K_{\text{app}}^{\text{PEPS}(\text{exp})} = I_{\text{closed}}(\text{PEPS, exp}) / I_{\text{open}}(\text{PEPS, exp}) \quad (8)$$

Data presented here were analyzed both before and after charge state deconvolution using both intensity and area of the peaks for the closed and open states. Average $K_{\text{app}}^{\text{PEPS}(\text{exp})}$ values at varying concentrations of GdHCl were used to generate $\Delta G_{\text{app}}^{\text{PEPS}(\text{exp})}$, $N_{\text{closed}}^{\text{PEPS}(\text{exp})}$, $N_{\text{open}}^{\text{PEPS}(\text{exp})}$ and LEM plots as described above.

3.4 Estimation of protein folding and unfolding rates by PEPS

Equations 1 and 6 can be further simplified by using reasonable approximations and assumptions at the middle of the GdHCl denaturation curve, where PEPS works best. In this region, both k_{close} and k_{open} are much smaller than $k_{\text{int}}^{\text{H/D}}$. This simplifies the rate of exchange of the globally protected region in the folded protein, k_{ex} , in Equation 1 to k_{op} and Equation 6 can be rearranged to

$$K_{\text{app}}^{\text{PEPS}} = (K e^{-k_{\text{open}} \cdot t}) / (1 + K (1 - e^{-k_{\text{open}} \cdot t})) \quad (9)$$

Rearrangement of equation 9 leads to estimation for k_{open} and k_{close} .

$$k_{\text{open}} = -\ln \left[K_{\text{app}}^{\text{PEPS}} (1 + K_{\text{app}}) / K_{\text{app}} (1 + K_{\text{app}}^{\text{PEPS}}) \right] / t \quad (10a)$$

$$k_{\text{close}} = K_{\text{app}} \cdot k_{\text{open}} \quad (10b)$$

Here K_{app} , the protein folding equilibrium constant, can be taken from fluorescence data, calorimetric data, CD data, or PEPS experiments recorded for the shortest time “t”. $K_{\text{app}}^{\text{PEPS}}$ from multiple time points can be used to obtain an average for rates of folding (k_{close}) and unfolding (k_{open}). Because rates obtained this way are at high (relative to physiological) denaturant (GdHCl) concentrations and in a much narrower range, linear extrapolation is used to obtain the rates at other concentrations. This is again consistent with the assumption of a two state folding mechanism where GdHCl linearly denatures closed native state to single open denatured state.

4. Results and Discussion

Data obtained from PEPS experiments for ubiquitin and staphylococcal nuclease are presented here. Theoretical PEPS populations for the closed and open states as well as $\Delta G_{\text{app}}^{\text{PEPS(theory)}}$ are calculated according to the equations above and are considered as a function of both HDX time “t” and GdHCl concentration for ubiquitin. The staphylococcal nuclease represents a less stable protein than ubiquitin [4, 12, 18, 20] and it is used to demonstrate how PEPS can be employed to extract protein folding and unfolding rates in addition to equilibrium folding constants and folding energies.

4.1 Ubiquitin Data

Figure 4 shows the theoretical closed (native), $N^{\text{H,H}}_{\text{closed}}$, and open (denatured), $N^{\text{H,H}}_{\text{open}}$, state populations for ubiquitin, calculated using k_{close} and k_{open} obtained from NMR data [4], and the calculated population movement from closed to open and open to closed as a function of time “t” and GdHCl concentration. The movement in population from one state to the other was calculated using simple first order rate equations. In one second (a typical PEPS HDX time) not much folding/unfolding takes place. The closed and open populations are very close to those at $t=0$. While the denaturant concentration controls the relative populations of the open and closed states, the short time frame for the PEPS experiment limits the changes of individual proteins as well. Relatively few proteins change from closed to open within 1 second. On the other hand, about 20% of the population shifts within 5 seconds. This movement of closed to open is clearly the reason for alteration in the PEPS signals observed by MS and the PEPS derived closed and open state populations calculated using equations 5 and 6.

Figure 5(a) shows a similar plot for open state compared to the plot in Figure 4. The only difference is inclusion of the extent of the globally protected region of the closed state which

exchanged during the unfolding of the protein within the HDX time as calculated by the equation 4, $N^{D,D}_{closed}(t=t)$. As seen in Figure 4, within a second, closed state unfolded to open state and resulted in slight decrease in the closed state population corresponding to $N^{H,H}_{closed}(1-e^{-k_{open}})$ as shown by solid triangle (Δ). The population of the closed state that unfolded to open state accompanies exchange of the globally protected region and is reflected as slight increase in open state population by $N^{D,D}_{closed}(1s)$ in PEPS experiments as shown in Figure 5 (a), open square ($\square$). Similar trends can be seen for 5 seconds, open triangle (Δ), and 15 seconds, open diamond ($\diamond$). At these times, PEPS clearly overestimates the open state population as shown in Figure 5(a) and Figure 5(b). This is because the closed state unfolds and eventually the globally protected region gets exchanged/labeled to a larger extent than expected based on the protein population as reflected in the ESI-MS peak for the open state.

It is clearly evident from Figure 5 that if the HDX in PEPS is carried out at physiological pH ~ 7 , the best HDX time (of the three) for PEPS experiment is close to $\sim 1s$ as it is very close to the NMR derived populations and LEM plot [open square ($\square$)]. It is also clearly evident from Figure 5 that PEPS work best in the middle of the denaturation curve as all the theoretical curves for 1s (open square ($\square$)), 5s (open triangle (Δ)), and 15s (open diamond ($\diamond$)) merge close to the NMR data. A HDX time shorter than 1s is not practical if the HDX is performed at pH ~ 7 . Controlling times using shorter timeframes is also difficult to implement experimentally. Short HDX time with faster HDX reaction rate could conceivably achieve more accurate PEPS population measurements as suggested in theoretical plots for 50 milliseconds HDX time at pH ~ 10 , open circles (o), in Figure 5. The implications of this are discussed again below in the section on improving PEPS.

It should also be mentioned that the NMR derived open state populations denoted by solid line and PEPS ($\sim 1s$ HDX time) derived populations denoted by solid circles (o) are slightly different from each other due to the temperature at which the NMR study was performed as seen in both population plots in Figure 5(a) and the LEM plots in Figure 5(b). Our PEPS was performed at 25° C whereas NMR study was performed at 10° C. But ΔG_{H_2O} at [0] M GdHCl and m_{GdHCl} values obtained by standard linear extrapolation of the PEPS data are in good agreement with the values obtained from CD, SUPREX, and SPROX [10,11,12,14] as shown by dash trend line in the Figure 5(b). Overall, Figure 5 clearly makes the point that using short $\sim 1s$ HDX time PEPS at the middle of the middle of the denaturation curve, 3-5 M GdHCl, provides accurate closed and open state population information and in turn provide accurate folding energy for ubiquitin.

4.2 Wild type Staphylococcal Nuclease Data

PEPS derived ΔG_{H_2O} and m_{GdHCl} values for staphylococcal nuclease wild type for $\sim 1s$, $\sim 5s$, and $\sim 15s$ HDX times (Figure 6) are all in reasonable agreement with fluorescence values. However, the values for 1.5 min, 11 min and 19 min HDX times are systematically higher than the fluorescence data. This trend in general is consistent with the fact that a folding and unfolding event that occurs during the HDX time leads to an over-representation of the open states. Once the protein opens, protons in the previously “protected” region of the protein deuterate and the refolded proteins do not lose those deuteriums and revert to the lower mass corresponding to the closed state. Perhaps the most interesting observation here is the fact that all three HDX times (corresponding to 1s, 5s, and 15s) gave accurate (± 0.3 kcal/mol) protein folding energies for the staphylococcal nuclease wild type. This is also true for the two mutants, 23I/25V/66L/72L, and 117G/124L/128A/41I/59A/21K/21N [18]. In contrast, the data for ubiquitin showed the best accuracy, only with very short HDX times. The only possible explanation for this observation is that both k_{close} and k_{open} for the staphylococcal nucleases are about half an order of magnitude lower in the middle of the GdHCl denaturation curve than the corresponding rate constants for ubiquitin as explained in

the section 4.3 and can be seen in Figure 7. Thus, when the staphylococcal nuclease proteins fold and unfold about three times more slowly than ubiquitin, and population states can be sampled using a wider range (up to three times longer) of exchange times without introducing significant error. As noted below, the PEPS measurement allows a calculation of a numerical value for both the opening and closing rates (as well as a relative value) and it is useful to compare the calculated values with the real world observations. The plots clearly predict the observed behaviour and provide good evidence that this method of calculating folding and unfolding rates is reliable, at least under the conditions presented here.

4.3 Estimation of numerical values of protein folding and unfolding rates by PEPS

Equations 9 and 10 can be used to extract protein folding and unfolding rates simultaneously with the protein folding constant and the folding energy. k_{close} and k_{open} values were estimated using equation (10) with the apparent folding equilibrium constants, $K_{\text{app}}^{\text{PEPS}}$, at 25° C for GdHCl concentrations between 0.5 – 1M. K , in turn, was obtained by using either PEPS (shortest time frame, as suggested by equation (9), experimentally determined MS intensity ratio values) or fluorescence (experimental values from the literature). For example, $K_{\text{app}}^{\text{PEPS}}$ values for 1s, 5s, and 15s are shown in Figure 6 at varying concentrations of GdHCl (0.5 M-0.9 M). Under these conditions, equations 9 and 10 result in three different (denaturant dependent) values for k_{close} and k_{open} . For each point three measurements were used to obtain an average to produce the data shown on the graph as open circles (o) for k_{open} , and solid circles (o) for k_{close} (Figure 7). Since explicit data for rates of folding and unfolding are available only in the region where the PEPS is performed, linear extrapolation is required. Standard linear regression is the accepted approach and allows predictions at physiological conditions and other denaturant concentrations. Extrapolation of the experimental data in Figure 7 for staphylococcal nuclease produces trend lines with correlation coefficients of 0.96 for k_{close} (solid line) and 0.85 for k_{open} (dash-dot-dash, -.-). The trend line predicts k_{close} and k_{open} at physiological conditions (0M GdHCl, pH~7, at 25° C) to be $10 \pm 1 \text{ s}^{-1}$ and $0.00158 \pm 0.0005 \text{ s}^{-1}$ respectively. Prediction of the trend line is more reliable using shorter HDX exchange times. Very long times deplete the undeuterated population sufficiently that measurement becomes difficult. Small errors in measurement have the least impact when measuring peaks of similar intensity as opposed to a very large and very small peak.

Numerical evaluation gives the same result suggested by graphical evaluation, namely that both k_{close} and k_{open} values for ubiquitin are half an order of magnitude larger in the region where the PEPS experiments were carried out (gray box in the right corner) compared to staphylococcal nuclease wild type (transparent box in the left corner). This is the reason why all three PEPS experiments performed at 1s, 5s, and 15s HDX times resulted in reasonably accurate protein folding energies in contrast to ubiquitin where errors became noticeable for longer exchange times.

How can we improve PEPS experiments?—The Most important factor that predicts the accuracy of the PEPS derived open and closed population estimations (the correlation of the two ESI-MS signals with the closed and open state equilibrium populations) is the product of k_{ex} and time “t” ($k_{\text{ex}} \cdot t$), as seen in Equation 6. Of course the stability of the ESI-MS signal itself is a key factor, but it is an instrument dependent factor that is primarily protein concentration dependent. The mass spectral signal will only be a limiting factor when the smallest amount of protein is available. But even considering this restriction, this method requires less protein than non-MS methods, namely micrograms rather than milligrams. As shown in Equation 6, when both “t”~ 0 and $k_{\text{ex}} \sim 0$, K^{PEPS} approximates to K (true equilibrium folding constant). Clearly when $k_{\text{ex}} = 0$, the outcome is meaningless because nothing is being probed. Ideally, PEPS experiments should be performed using

short HDX times where “t” is shorter than the life time of both the closed and open states. Normally this will occur when the denaturant concentration produces reasonable and measurable percentages of both closed and open states. Ultimately, the use of shorter HDX times reaches a limit set by the HDX rate. Speeding the rate of H/D exchange would clearly extend the range over which PEPS could efficiently measure various folded/unfolded state populations at very short exchange times. Theoretically, chemical reactions could also be considered as an alternative to HDX to monitor the globally protected region of the closed state of the protein, but these reactions would also have to be fast. The HDX reaction rate in physiological pH for an exposed amide proton is $\sim 5 \text{ s}^{-1}$ but the $k_{\text{int}}^{\text{H/D}}$ value at pH ~ 10 is about 2000 s^{-1} [10]. This could extend the ability to estimate populations accurately between 1M to 6M GdHCl compared to 2.5M to 5M (figure 3) at physiological pH for ubiquitin. This is evident from the calculated PEPS population and the LEM plot for 50ms HDX time at pH ~ 10 indicated as open circles in Figure 5. To accomplish this, the PEPS experimental design would need to be modified to include robotic liquid handling devices, a quench flow apparatus, or a micro fluidic platform to achieve millisecond time scale HDX [26,27].

5. Conclusions

Standard theoretical models can be used to explain PEPS by applying a two state protein folding mechanism and the “breathing model” of HDX. The theoretical and experimental data for bovine ubiquitin are in good agreement with the best results for 1s exchange times. PEPS works because the life times of the open state and the closed state are long compared to the life time of the exposed amide proton under exchange (or the intrinsic rate of HDX, $k_{\text{int}}^{\text{H/D}}$) and also long compared to the time used to implement the exchange process experimentally. This condition is generally best satisfied in the middle of the denaturation curve. Hence it is important to use the PEPS measurements obtained only in the middle of this curve. PEPS using $\sim 1\text{s}$ HDX worked well both theoretically and experimentally for ubiquitin. However, theory predicts the observed degradation of the data for longer exchange times (15s). PEPS can be used to show that the folding and unfolding rates are about 3 times longer for staphylococcal nuclease. Hence better accuracy and precision for exchange times up to 15 seconds was observed with this protein. Theory predicts the ability to sample folded state populations with PEPS over an extended region of denaturant concentrations even for difficult proteins (like ubiquitin) at higher pH by speeding up the intrinsic exchange rate, leading to millisecond HDX times. In summary PEPS currently provides useful data consistent with the PEPS theory that can be extrapolated to physiological pH. PEPS should provide better data using high pH HDX measurements and millisecond HDX time. This improvement to PEPS is the focus of our plans for continuing work in this field.

This study also showed that PEPS can be used to estimate protein folding and unfolding rates.

Acknowledgments

“This publication was supported by Grant Number P30 GM103450 from the National Institute of General Medical Sciences of the National Institutes of Health (NIH).”

References

- [1]. Dobson CM. Trends Biochem. Sci. 1999; 24:329. [PubMed: 10470028]
- [2]. Brandts JF, Lin LN. Biochemistry. 1990; 29:6927. [PubMed: 2204424]
- [3]. Hill JJ, Royer CA. Methods Enzymol. 1997; 278:390. [PubMed: 9170324]

- [4]. Sivaraman T, Arrington CB, Robertson AD. Nat. Struct. Biol. 2001; 8:331. [PubMed: 11276253]
- [5]. Calhoun LN, Liyanage R, Lay JO, Kwon Y. BMC Microbiology. 2010; 10:249. [PubMed: 20920181]
- [6]. Kannan L, Rath NC, Liyanage R, Lay JO Jr. Ann N Y Acad Sci. 2007; 1112:425. [PubMed: 17947593]
- [7]. Kannan L, Liyanage R, Lay JO Jr, Rath NC. Mol Immunol. 2009; 46(15):3151. [PubMed: 19665233]
- [8]. Kannan L, Rath NC, Liyanage R, Lay JO Jr. Mol Cell Biochem. 2010; 344(1-2):55. [PubMed: 20614231]
- [9]. Sides CR, Liyanage R, Lay JO Jr, Philominathan ST, Matsushita O, Sakon J. J Am Soc Mass Spectrom. 2012; 23(3):505. [PubMed: 22207568]
- [10]. Ghaemmaghani S, Fitzgerald MC, Oas TG. Proc. Natl. Acad. Sci. U.S.A. 2000; 97:8296. [PubMed: 10890887]
- [11]. Powell KD, Fitzgerald MC. Biochemistry. 2003; 42:4962. [PubMed: 12718538]
- [12]. Dai SY, Fitzgerald MC. J. Am. Soc. Mass Spectrom. 2006; 17:1535. [PubMed: 16872837]
- [13]. Tran DT, Banerjee S, Alayash AI, Crumbliss AL, Fitzgerald MC. Anal. Chem. 2012; 84:1653. [PubMed: 22185579]
- [14]. Graham MW, Liangjie T, Fitzgerald MC. Anal. Chem. 2008; 80:4175. [PubMed: 18457414]
- [15]. Sperry JB, Shi X, Rempel DL, Nishimura Y, Akashi S, Gross ML. Biochemistry. 2008; 47:1797. [PubMed: 18197706]
- [16]. Zhu MM, Rempel DL, Du Z, Gross ML. J. Am. Chem. Soc. 2003; 125:5252. [PubMed: 12720418]
- [17]. Chitta RK, Rempel DL, Grayson MA, Remsen EE, Gross ML. J. Am. Soc. Mass Spectrom. 2006; 17:1526. [PubMed: 16952461]
- [18]. Liyanage R, Devarapalli N, Puckett LM, Phan NH, Gidden J, Stites WE, Lay JO. International Journal of Mass Spectrometry. 2009; 287:96. [PubMed: 22427739]
- [19]. Chen J, Stites WE. Biochemistry. 2001; 40:14012. [PubMed: 11705393]
- [20]. Chen J, Stites WE. Biochemistry. 2001; 40:14015.
- [21]. Hvidt A, Nielsen S. Ado. Protein Chem. 1966; 21:287.
- [22]. Englander SW, Downer NW, Teitelbaum H. Annu. Rev. Biochem. 1972; 41:903. [PubMed: 4563445]
- [23]. Englander SW, Kallenbach NR. Q. Rev. Bio- phys. 1984; 16:521.
- [24]. Weiss DD, Wales TE, Engen JR, Hotchko M, Ten Eyck LF. J.Am. Soc. Mass Spectrom. 2006; 17:1498. [PubMed: 16875839]
- [25]. Roder H, Wagner G, Wiithrich K. Biochemistry. 1985; 24:7397.
- [26]. Englander SW, Mayne L. Annu. Rev. Biophys. Biomol. Struct. 1992; 21:243. [PubMed: 1525469]
- [27]. Kaltashov IA, Eyles SJ. Mass Spectrometry Reviews. 2002; 21:37. [PubMed: 12210613]

- !! ~1s HDX PEPS intensities of deuterated closed and open state forms of ubiquitin represents a true “snapshot” of the protein equilibrium.
- !! Theoretical model which used a two state protein folding mechanism and the “breathing model” of HDX explain how short 1s HDX PEPS experiments work in estimating protein folding energies.
- !! PEPS can also be used to estimate protein folding/unfolding rates.

Figure 1.

a) Raw ESI and **b)** deconvoluted mass spectra; of ubiquitin, for one second HDX PEPS taken close to the midpoint of the linear denaturation curve. Clearly resolved peaks corresponding to deuterated closed and open states can be seen in every charge state and at all four GdHCl concentrations.

Figure 2.

Classical "breathing model" of HDX with a two state folding mechanisms [21, 22, 23]. Only globally protected amide protons are shown in its exposed denatured (open) state. These protected amide protons can exchange only when the protein opens up using the classical HDX model. k_{close} , and k_{open} are given by rates of folding and unfolding of the protein. $k_{\text{int}}^{\text{H/D}}$ is the rate of H/D exchange of an unprotected amide proton.

Figure 3.

a) Folding and unfolding rates of ubiquitin with the intrinsic H/D exchange rate of an exposed amide proton. Folding/unfolding rates for ubiquitin were obtained from H/D exchange studies of globally protected amide protons at 10° C by NMR [4]. --- (Dash), -.- (Dash double dot), and — (Solid) lines represent folding rate, unfolding rate, and H/D exchange rate of an exposed amide proton [10] as a function of GdHCl concentration.

b) Life times of closed (folded) and open (unfolded) states of ubiquitin with the life time of an exposed amide proton. Similar to **a)**, --- (dash), -.- (dash double dot), and — (solid) lines represent the life time of the closed, open, and exposed amide proton as a function of GdHCl concentration.

Figure 4.

Shows the movement between closed and open state as a function of both time and the GdHCl concentration for the ubiquitin calculated using known k_{close} (folding) and k_{open} (unfolding) rates at 10° C. -- line is the fraction of the closed state at $t=0$, and -.-.- line is the fraction of open state population at $t=0$. Solid square ($\blacksquare$) represents fraction of the open state remained in the open state while the solid $\blacktriangle$ represents the fraction of the closed state remained in closed configuration after one second. Likewise, open square ($\square$) represents the fraction of the open state remain in the open state while open triangle ($\triangle$) represents the fraction of the closed state remained in the closed configuration after five seconds.

Figure 5.

a) Compares open state population obtained using PEPS-HX-ESI-MS theory described in the theory section as a function of GdHCl concentration for 1s (open square, $\square$), 5s (open triangle, Δ), 15s (open diamond, $\diamond$), and 50ms (at pH~10, open circle, $\circ$) HDX time to the NMR, 10° C, population (solid line) and population estimated using 1s PEPS (solid circle, $\circ$) experiments at 25° C.

b) Compares linear extrapolation plots (LEM plots) generated by the equation 7 and (8) described in the PEPS theory section for 1s (open square, $\square$), 5s (open triangle, Δ), 15s (open diamond, $\diamond$), and 50ms (at pH~10, open circle, $\circ$) HDX time to the data from NMR,

10° C experiments (solid line), and data from one second PEPS experiments at 25° C (solid circle, o).

Figure 6.

Comparison of linear extrapolation plots (LEM plots) for 1s (open circle, ○), 5s (open triangle, ▽), 15s (open square, □), 1.5 min (solid square, ■), 11 min (solid triangle, ▲), and 19 min (solid circle, ●) HDX times PEPS experiments to the data from fluorescence, 25° C, (solid line).

Figure 7.

PEPS estimated staphylococcal nuclease wild type folding (k_{close} , solid circle, o) and unfolding (k_{open} , open circle, o) rates for varying concentration of GdHCl concentrations using equations 9 and 10 and 1,5 and 15s experiments (averaged). The solid trend line for k_{close} is given by the equation $-2.8x + 1$ with a 0.96 correlation coefficient and the dash-dot-dash (-.-) line for k_{open} is given by the equation $1.7x - 2.8$ with a 0.85 correlation coefficient. The graph also compares the staphylococcal nuclease folding/unfolding rates with that of ubiquitin. The dashed line denotes k_{close} , and the dash-dot-dot-dash (-.-.-) line denotes k_{open} .