Method Validation and Standards in Hydrogen/Deuterium Exchange Mass Spectrometry

Jeffrey W. Hudgens, Richard Y.-C. Huang, and Emma D'Ambro

Institute for Bioscience and Biotechnology Research, BioProcess Measurements Group, Biomolecular Measurement Division, National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD 20850

Table of Contents

4.0 Method Validation and Standards in Hydrogen/Deuterium Exchange Mass Spectrometry..... 3

4.1 Introduction 3

4.2 Rationale for a Reference Measurement System for HDX-MS 4

4.3 General Metrological Terminology..... 7

4.4 Method Validation: General Conditions 7

4.5 Method Validation: Precision..... 9

4.6 Method Validation: Bias 12

4.7 Method Validation: Accuracy Improvements 14

4.8 Method Validation: HDX-MS and HDX-NMR Cross Comparisons..... 16

4.9 Standards: Reference Materials..... 16

4.10 Summary: Maintaining Standards and Monitoring Performance 18

4.11 References 18

4.0 Method Validation and Standards in Hydrogen/Deuterium Exchange Mass Spectrometry

Jeffrey W. Hudgens, Richard Y.-C. Huang, and Emma D'Ambro

Institute for Bioscience and Biotechnology Research, BioProcess Measurements Group, Biomolecular Measurement Division, National Institute of Standards and Technology (NIST), 9600 Gudelsky Drive, Rockville, MD 20850

4.1 Introduction

Over the past twenty years scientific journals have documented the development of standard operating procedures for hydrogen deuterium exchange mass spectrometry (HDX-MS) that are based on proteolytic fragmentation. These procedures include criteria for the assignment of measurement uncertainties. To assure that other investigators can obtain the same results, studies usually set their HDX-MS measurements to a specific pH, temperature, and ionic strength. To further reduce measurement variability, laboratories commonly conduct relative measurements between a test system and a protein standard. [1, 2] Using epitope mapping as an example, the HDX-MS D uptake patterns of the uncomplexed (apo) and ligand-complexed (holo) test protein are obtained using the same solutions, buffers, digestion column, chromatographic analysis column, and chromatographic conditions. The commonality of the apparatus, solution manipulations, and solution properties minimizes the measurement uncertainties and enables direct comparison of two or more data sets. Using this protocol during measurement campaigns of short duration, small measurement errors of pH, temperature, pressure, ionic strength, and other variables have little or no effect upon the test conclusions. [3] Such practices have proved sufficient for proteolytic fragmentation HDX-MS studies conducted within a single laboratory, where methods under development rely heavily upon the judgment of the investigator. [1, 4, 5] This local model of traceability relies on the integrity and technical mastery of individual scientists, and it is a fundamental underpinning of the archival scientific literature.

More recently, successes in the basic research laboratory have emboldened the pharmaceutical industry to employ HDX-MS for the characterization of biopharmaceutical products that will enter commerce. To assure the quality of the product, the offeror might present HDX-MS data that establish:

- The temperature and pH stability range of a biotherapeutic drug; [6]
- The invariance of higher order structure among manufacturing lots of a biotherapeutic drug; **(Part 3: Chapter 3)**
- The comparability of therapeutic proteins following manufacturing changes;[7, 8]
- The fidelity of a candidate biosimilar to the name brand innovator biotherapeutic drug; **(Ref. Part 3: Chapter 3)**. [9, 10]

In such cases the HDX-MS data may originate from one or more laboratories, and the period of data collection may span years. Customers may rightly ask: "Just how good is that measurement? Has the measurement process changed over time? Is the HDX-MS measurement conducted in one lab the same as in others anywhere in the world?" In the commercial sphere the customer relies on the international model of measurement traceability comprising an unbroken chain of measurement calibrations from the testing laboratory to fundamental standards defined by international agreement. Skeptical customers and regulatory institutions will require assurance that properties claimed for the biotherapeutic are in accord with those determined from HDX-MS data. This assurance becomes possible when all measurements are traceable to a common set of reference materials and validated measurement procedures. The standardization of procedures reduces the need for expert judgment and facilitates comparisons of data across different laboratories. Standards can also incorporate accumulated knowledge of the measurement procedure.

A recent report by the International Bureau of Weights and Measures (BIPM) has identified measurement services, research, and development needed to underpin the comparability of bio-measurements for the health industry. [11] The report identified HDX-MS as a key physicochemical technique that can improve higher order structure characterizations of single purified proteins. The report also identified the need for metrological traceability with respect to HDX-MS measurements of proteins. Such traceability would significantly broaden the acceptance of HDX-MS measurements by conferring universal understanding to results obtained in different laboratories at different locations and times. [12] The foundation of traceability is a set of well-tested, reliable laboratory protocols that can reproduce measurements of reference standards with known uncertainty. These protocols are refined and confirmed through tests and calibrations against selected biological material standards. Thus, method validation and standards development are key metrological activities that are essential to scientific understanding and the promotion of commerce.

This chapter is directed to the process of method validation and use of standards that can help investigators obtain reliable rate coefficients for H/D exchange by proteolytic fragmentation HDX-MS. We base this chapter on our observations of the excellent metrological HDX-MS science that is reported in journal articles and with reference to reports issued by the international standards organizations. Its intended audience is investigators who need to achieve consistent higher resolution measurements of protein dynamics by HDX-MS over the course of time and, perhaps, at multiple locations.

4.2 Rationale for a Reference Measurement System for HDX-MS

The apparatus and samples examined by HDX-MS measurements present a great degree of metrological complexity, which raises the question: "Can we formulate proteolytic fragmentation HDX-MS protocols so that they assure traceability?" The Vocabulary in Metrology (VIM3) defines metrological traceability as the *"property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty"* [13] This definition specifies that traceability begins with the measurement result and not with the instrument, [14] *i.e.*, the results obtained from two calibrated, competently constructed apparatuses will yield measurement results that are in accord within their combined uncertainties. This definition is tied to the *Guide to the Expression of Uncertainty in Measurement* (GUM) [15], ISO 5725 [16], and ISO 17025 [17]. These documents require the use of a calibration hierarchy that traces and

propagates the measurement uncertainty from the reference materials (RM) to the measured sample. This procedure yields a measure of the proximity of the measurement to the (unknowable) actual value. [16-18] The GUM, VIM3, and other metrological documents require that measurements and their associated uncertainties are expressed in conformance with the International System of Units (SI).

The calibration hierarchy is the sequence of calibrations from a reference to the final measuring system, where the outcome of each calibration depends on the outcome of the previous calibration. [13] This hierarchy requires that for measurements incorporating more than one input quantity in the measurement model (*e.g.*, pH, *T*), each input quantity must itself be metrologically traceable. In addition, each measurement and derived quantity is listed with an evaluated uncertainty that captures the uncertainties of the measurements and of the calibration hierarchy. Also, because the propagation of variances is additive, measurement uncertainty increases throughout the calibration hierarchy from the reference material (which is ideally a certified reference material aka CRM) to the sample. A statement describing the uncertainty is essential, as a measured quantity value unaccompanied by a measurement uncertainty is not only useless, but it is potentially dangerous because the measured value may be misinterpreted or misused.

The ISO and GUM documents were written to encompass most common quantitative measurands. For biological entities some extension is necessary, but these documents provide sufficient guidance to place HDX-MS upon the strong foundation of the SI. For HDX-MS the relevant fundamental units are kilogram (kg), second (s), ampere (A), kelvin (K), and mole (mol). The measurand of HDX-MS is always the *m/z* spectrum of the deuterium shifted analyte or one of its proteolytic peptide ions. [19] Computational evaluation of the ion current vs. *m/z* spectrum yields the centroid of the isotopic envelope, $\langle m \rangle$, which is defined as:

$$\langle m \rangle = z \cdot \frac{\sum_{i=1}^{n} \left(m/z \right)_i \cdot I_i}{\sum_{i=1}^{n} I_i} \tag{4.1}$$

where z is the ion charge, n is the number of isotopic *m/z* features in mass spectrum of the ion, and $(m/z)_i$ and I_i are, respectively, the measured mass to charge ratio and intensity of the i^{th} ion feature. The centroid of the isotopic envelope is a derived measurand that reflects the average mass of an isotopic envelope originating from the same peptide. It contains no information about the measurement system or about the properties of the analyzed ion. This derived measurand is appropriate only for EX2 kinetics data. [20] After the subject protein is incubated for time t_{ex} in a solution rich in D_2O, the deuterium content of each proteolytic peptide, $D_0(t_{ex})$, is computed with the equation:

$$D_0(t_{ex}) = \frac{\langle m \rangle - \langle m_{0\%} \rangle}{\langle m_{100\%} \rangle - \langle m_{0\%} \rangle} \cdot N \tag{4.2}$$

where $\langle m_{0\%} \rangle$ and $\langle m_{100\%} \rangle$ are the centroid values of the undeuterated peptide and fully deuterated peptide, respectively, and N is the total number of amide hydrogens in the peptide. [21] Since the change in deuterium content is measured as a function of time and the measurand

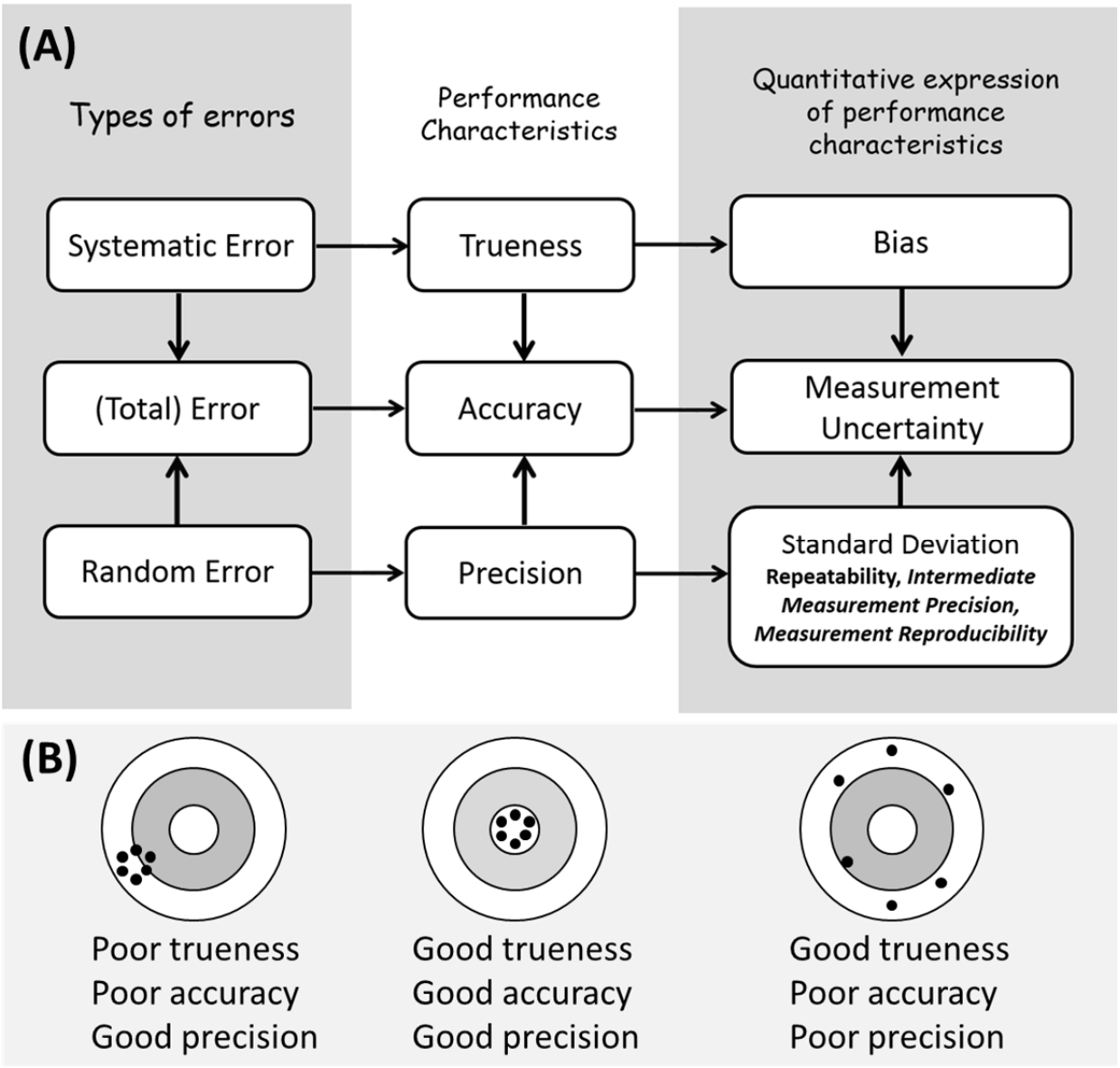

Figure 4.1. Diagrams depicting relationships of trueness, accuracy, and precision. A) Connection map among the error, performance characteristics during measurements, and the quantitative performance characteristics. The measurement is more accurate when it presents a smaller measurement error, more true when the bias is small, and more precise when the random error is small. [23] Figure is adapted from Ref. [24]. B) Bullseye target representations of shooting patterns manifesting good and poor trueness, accuracy, and precision.

of a pure protein can be expressed directly in kilogram (or mole), a rate coefficient (s^{-1}) for deuterium exchange can be computed for the analyte (*e.g.*, protein, peptide, or amide site) under defined conditions, *i.e.*, pH, temperature, ionic strength, etc. [22]

When establishing a traceable calibration hierarchy, both the measurement procedure and the measured substance must be in accord with the SI. VIM3 defines two types of analyte.[14]

Type A analytes are chemically well-defined compounds that are available in pure form (*e.g.* electrolytes, urea, many proteins, etc.). These analytes can be expressed in molar units. Since Type A materials are expressed in SI units, Type A materials can be certified as primary reference materials. Type B analytes are heterogeneous mixtures of chemicals. The components of the mixture may vary from sample to sample, as occurs for glycoproteins, cardiac troponin, etc. Hence, Type B quantities cannot be expressed in SI units, and thus, Type B materials are not suitable reference materials.

In summary, HDX-MS measurements of Type A analytes can be firmly based in the SI and the HDX-MS instrument and analysis method can be traced through the calibration hierarchy to fundamental standards. Assured that this rigorous foundation can exist, the development of validated measurement methods and reference materials for use in proteolytic fragmentation HDX-MS is sensible.

4.3 General Metrological Terminology.

The greater portion of a validation exercise is devoted to the establishment of the precision and accuracy of the method. For this discussion we define terms in accordance with the GUM [15] and VIM3 [13], as these documents are the foundation of ISO standards. Figure 4.1A lays out their relationships, which differ somewhat from common usage, particularly, with respect to terminology for error**.** [24]

Measurement accuracy is the closeness of agreement between the measured quantity value and the true quantity value of a measurand. VIM3 regards accuracy and measurement error as idealized quantities that we may not know exactly. [14] VIM3 terminology uses measurement *trueness,* which is the closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value, as measured using a certified reference material (CRM). [13] In the absence of a certified reference material that shares sufficient similarity to the protein of interest, any locally prepared material can only provide insight into precision, not trueness.

As diagrammed in Figure 4.1A, accuracy may be regarded as the combination of trueness and precision. Figure 4.1B depicts the relationship among these terms in terms of target shooting or a game of darts. Good accuracy requires both good trueness and good precision. Trueness is inversely related to systematic measurement error or bias, which itself is the component of the measurement difference that in replicate measurements remains constant or varies in a predictable manner. Note that bias does not contain contributions from mistakes, such as using an uncalibrated pH meter or collecting data with a Fourier transform instrument set to an inappropriate resolution. [25] Once evaluated, the measurement bias can be used to correct the measurement quantities to truer values. The evaluated precision contains only the measurement uncertainty from random sources intrinsic to the measurement procedure and the calibration hierarchy. Each accurate measurement value is always reported with its estimated uncertainty. [13-15]

4.4 Method Validation: General Conditions

Method validation is the process of proving that the HDX-MS method is acceptable for its intended purpose. [23] In addition, the investigators must document the experiments conducted during the validation exercise that assure the accuracy and uncertainty of the method.

[26] Before beginning a method validation exercise, investigators should set the minimum performance metrics required of the HDX-MS measurement platform. Performing a thorough method validation can be tedious, but executing it in a second-rate manner is simply wasted time and resources. Even when carefully executed, a thorough method validation exercise cannot rule out all potential problems, but the method validation process should identify the more common ones. Method validation becomes particularly important for organizations that extend their testing across several laboratories, additional laboratory personnel, and mass spectrometer platforms.

Analytical chemistry method validation exercises require the collection of extensive data that are used to determine the specificity, linearity, accuracy, precision, range, quantization limit, and robustness of the method. Extensive data not only establish reliable statistical measures, but the statistically reliable results will serve during later times as a tool for evaluating questionable data.

Some elements of method validation for proteolytic fragmentation HDX-MS are relatively straightforward to document. Method specificity for HDX-MS is determined by the chosen peptide map, the proteolytic enzyme, the chromatographic column and apparatus, and the mass spectrometer.[27] The peptide map is the filter through which mass spectral data are selected for analyses. The enzyme efficiency largely determines whether the peptide abundance is sufficient for detection. The resolution of the chromatographic apparatus determines whether eluted peaks are adequately resolved run after run to allow accurate measurement of the deuterium uptake by each peptide. Method linearity and range of HDX-MS can be demonstrated from plots of deuterium uptake versus time for the peptide set, showing smooth lines that intersect with the corresponding $\langle m_{0\%} \rangle$ and $\langle m_{100\%} \rangle$ points. The method quantization limit for HDX-MS is the minimum signal level for which the mass shift of the peptide isotopic envelope is properly measured with good precision. This limit must be considered when examining the effects of baseline shifts and inhomogeneous noise on data for each peptide ion. The method validation exercise should explore the protein concentration necessary to support proper measurement of the isotopic shift for each member of the proteolytic peptide set. Here, a well-understood reference material can assure consistent proteolytic activity and chromatographic fitness.

Method robustness is accessed by varying parameters that are not expected to disturb the HDX-MS measurements. For example, HDX-MS measurements of stable proteins should be unaffected by changes in the injection volume, injection velocity, column wash duration, number and range of the time points collected, etc. Although nearly every change of solution environment (pH, *T*, ionic strength) affects the exchange kinetics, the evaluation of robustness should include collection of measurements as a function of these parameters, followed by back corrections via thermodynamic relationships to a set of chosen conditions, *e.g.*, 25 ^{o}C, pH 7.4, etc. If the back corrected values are inconsistent with runs collected at standard conditions, the method is not robust and further characterization of the measurement is warranted. In addition, the D uptake vs. time for each peptide should be examined to determine whether EX1 or EX2 behavior is conserved as a function of solution conditions. A change in kinetics behavior may indicate a change in the physical properties of the protein, or it may indicate a change in the protein ensemble, *e.g.*, aggregate formation, shifts of equilibrium among protein conformations. Robustness studies of a sample protein are best accompanied by a parallel study that uses a reference material (*e.g.*, a CRM). [26] For HDX-MS the reference material should manifest

complexity and physiochemical properties akin to the sample, so that the barriers to obtaining good analytical measurements are similar. For example, the sample protein and reference material might share similar proteolytic digestion efficiency difficulties, and their proteolytic peptide sets might present similar analytical problems.

4.5 Method Validation: Precision

Precision is the closeness of agreement among measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. It is specified by a standard deviation of the mean. [13] Precision is defined by three components:

- *Repeatability*, which is the measurement precision for replicate measurements on the same (or similar) objects over a short period of time. When evaluating repeatability, the measurement conditions require the same measurement procedure, same operators, same measuring instruments, same operating conditions, and same location;

- *Intermediate Measurement Precision* (*IMP*), which is the set of conditions that includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time. This specification allows for other conditions (*e.g.*, different reagent lots) involving changes;

- *Measurement Reproducibility*, which is the precision of the analytical protocol after taking into account its application across different laboratories during an inter-laboratory collaborative study.

The graphs in Figure 4.2 illustrate the terms *Repeatability* and *IMP*. [Insert figure 4.2 here] These data were obtained by the authors during an evaluation of their robotic HDX-MS system. (See ref. [28, 29] for details.) The experiments measured the deuterium-uptake by bovine ubiquitin after immersion in D_2O for 30 seconds. This short exchange period was chosen because such data are very sensitive to fluctuations in sample manipulations. The evaluation of such data sets can give evidence for variations of sample mixing, syringe injections, chromatograph performance, and the sample environment. Figure 4.2A maps the 39 analyzed peptide ions onto the amino acid sequence of bovine ubiquitin (100% protein coverage). In Figure 4.2B each dot, square, and triangle represents the average fractional D uptake of each peptide for twenty replicants, as observed on days 1, 2, and 37 of the campaign. The bars extending from symbols represent the standard deviation (SD) of twenty measurements. On Figure 4.2C, the daily *Repeatability* computed from the standard deviations (Type A only) of 780 measurements is about s_r = 0.9 %. The dispersion among all measurements is less than 2 % (Table 4.1). The measurement campaign extended to 37 days, which suggests a s_I = 1.0 %. In practice, our laboratory would continue to collect additional replicants and maintain a running record of *Within Lab Reproducibility*, as a quality assurance measure.

Automation has proven essential for the collection of the large datasets needed for establishing *Reproducibility* and *IMP*. The statistics provided by large datasets allow investigators to explore subtle environmental and operational variables that may affect measurement precision and bias.

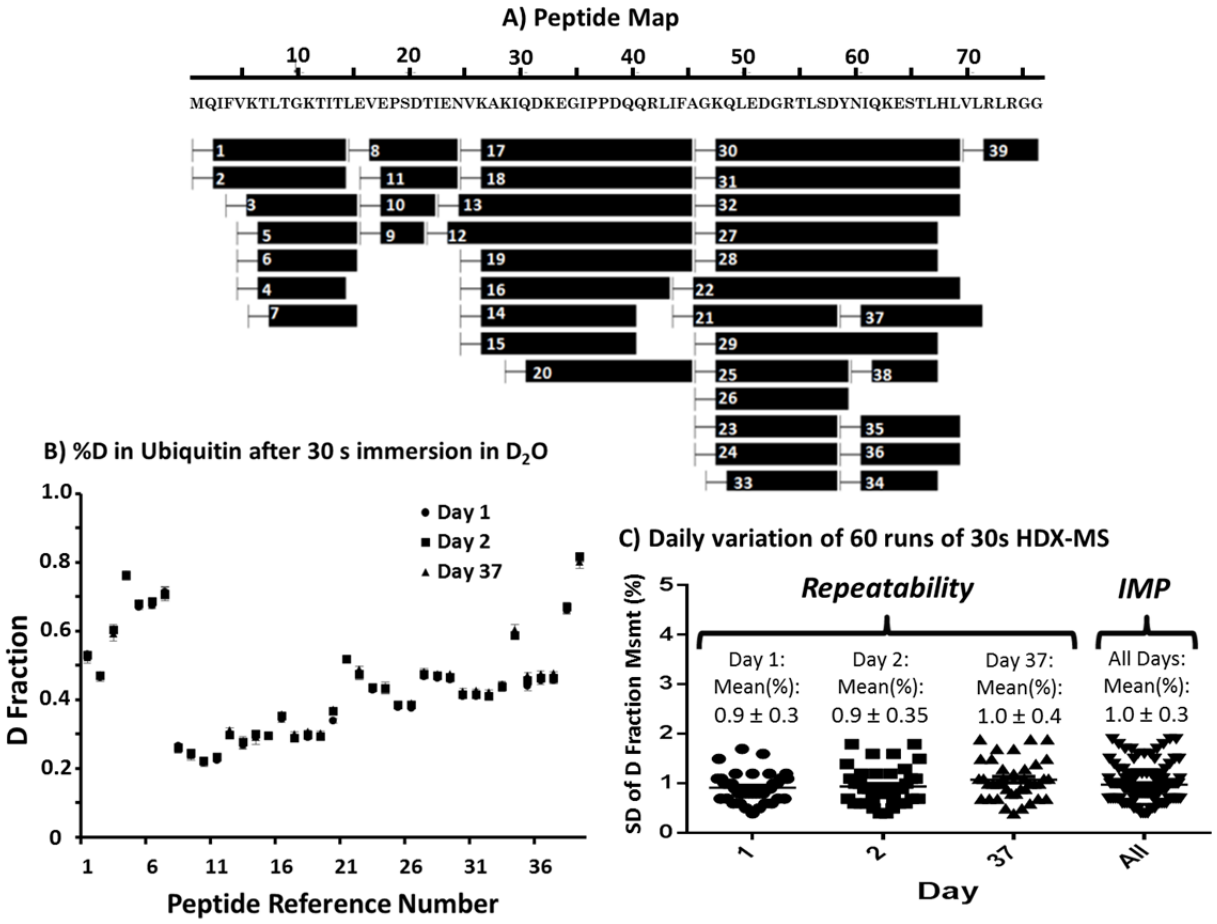

Figure 4.2. The *Repeatability and Intermediate Measurement Precision (IMP)* of a robotic HDX-MS apparatus. The study is based on replicant measurements of bovine ubiquitin deuterium-uptake after 30 seconds immersion in D_2O. A) The peptide map of ubiquitin, comprising 29 peptides (100% protein coverage) observed in 39 analyzed peptide ions, used for the HDX-MS analyses. See refs. [28-30]. B) The average fractional D uptake observed for each of the 39 analyzed peptide ions. As shown on the key, each symbol presents the average of 20 measurements conducted on Day 1, 2, and 37. Bars on each symbol indicate the standard deviation (SD) when it exceeds the symbol size. C) *Repeatability* and *IMP* for each peptide set, plotted as the standard deviation of the D fraction measurements obtained for each peptide. The horizontal solid line in each group indicates the mean SD of the peptide set (780 measurements).

Table 4.1 lists the *Repeatability* and *Intermediate Measurement Precision* derived from five studies spanning protein samples of 8.6 kDa to 58.2 kDa. These laboratories conducted the experiments using the same automated sample robot. The robots manipulated samples uniformly between temperature controlled sample drawers and the chromatograph. The robot allowed investigators to customize sample manipulation attributes, *e.g.*, syringe injection rates, volumes, exchange and quench temperatures, etc. The system scheduler could arrange experiments so that two or more experiments of differing H/D exchange time (t_{ex}) overlapped. It could randomize the order of t_{ex} H/D exchange measurements. To reduce fluctuations of back exchange due to

Table 4.1 *Repeatability* and *Intermediate Measurement Precision (IMP)* observed during robotic, proteolytic fragmentation HDX-MS measurement campaigns.

Subject Molecule	Expt. Conditions (t_{ex}, pH, T)	Daily Repeatability (s_r)	*IMP* (s_I)	No. %D values, No. of peptides	Study Length	Ref.
Vitamin D receptor, 29.8 kDa (PDB: 1DB1)	30 s @ pH 7.5, 4 °C		9 %	4191, 33	8 mo	[1]
Equine cytochrome c, 12.4 kDa (PDB: 3O1Y)	1 s to 180 s @ pH 7.9, 25 °C	2.9 % (0.1 to 0.2) Da		156, 52		[31]
Equine cytochrome c, 12.4 kDa (PDB: 3O1Y)	130 ms @ pH 7.9, 25 °C	0.7 %		185, 37		[31]
bovine carbonic anhydrase (BCA), 58.2 kDa (PDB: 1v9e)	30 s @ pH 6.8, 20 °C	0.3 % to 1.5 %	1.9 % to 3.8 %	140, 5	2 mo	[32]
interferon-β-1a, 20 kDa	30 s to 240 min @ pH 7.2, 20 °C	0.28 Da[a]		> 1000, 67		[9]
Bovine ubiquitin, 8.6 kDa (PDB: 1UD7)	30 s @ pD 7.6, 3 °C	0.9 %	1.0 %	2340, 39	37 d	This work

variation of ambient temperature during transfers between drawers, the entire robotic system could be placed in a refrigerator. [1] Mass spectra obtained during wash cycles could verify that the desalting trap and analytical column were free of protein and peptide residues. As long as good system hygiene was maintained, the robot would tirelessly collect useful data until the sample backlog was depleted.

The study of interferon-β-1a variants by Houde *et al.* found that centroid shifts of greater than 0.5 Da for any digested peptide had a 98 % likelihood of being significant for determinations of comparability. [9] Statistical analysis of their data shows that this deviation criterion has weak or no dependence on peptide size. The deviation criterion is independent of the H/D exchange time (t_{ex}). It is also independent of the magnitude of the mass difference between the deuterated and undeuterated peptides.

Examination of Table 4.1 reveals that the precision of proteolytic fragmentation HDX-MS is quite good. The *Reproducibility* and *IMP* figures of merit all reside below 3 % and 10 %, respectively. Because the mass spectrometers employed for HDX-MS can measure the mass of a peptide with tolerance of ± 0.01 Da or better, the mass spectrometers contributed little to measurement variability. [2] Because the underlying data in Table 4.1 originated from nearly identical robotic HDX-MS systems, the different figures of merit found during these investigations most likely reflect the measurement complexities that arose from the operational issues of processing the protein (*e.g.*, sample concentration, digestion efficiency, sample injection rates, aggregates, and column properties). Variability can arise from inhomogeneous labeling of the protein due to the presence of different conformation folding forms, where each manifests a different set of H/D exchange rates. [2, 21] Moreover, any small fluctuation in

conditions that changes the Gibbs energy of the proteins prior to, or during, the exchange measurement (*e.g.*, pH, temperature, ionic strength, solvent composition, etc.) can introduce variability by changing the equilibrium distribution of conformation forms. In order to obtain the most reproducible results, each variable must be understood, controlled, and made as nearly identical as possible. Again, such investigations will generally require large data sets.

Finally, we consider *Measurement Reproducibility*, which is determined through a statistical analysis of HDX-MS measurements conducted on the same protein sample in many laboratories. [13] The determination of the consensus *measurement reproducibility* for HDX-MS is an integral part of method validation, as its derived uncertainty supports the estimation of precision under reproducibility conditions. *Measurement reproducibility* is determined through a statistical analysis of HDX-MS measurements conducted on the same protein sample in many laboratories. The results from this analysis can help investigators detect measurement variance due to different realizations of the HDX-MS technique. For studies of unknowns *measurement reproducibility* defines where measurements of D uptake differ statistically. It is expected that rigorously evaluated *measurement reproducibility* can foster a broader understanding and acceptance of HDX-MS data. To date, no study leading to the determination of the *measurement reproducibility* of HDX-MS has been reported.

4.6 Method Validation: Bias

Data never lie, but data are bent by the astigmatism of bias. Much bias can be prevented during the design of experiments. When devising a measurement campaign, each process step must be reviewed for potential problems that might corrupt the sample or the measurement of its HDX-MS data. This review of potential sources of bias is a key step of method validation.

Bias can originate from any component of the measurement apparatus or even from the sample itself. As examples, during sample preparation bubbles can change protein conformations, [33, 34] and ice can unfold proteins. [35] The selected sample pH may promote protein agglomeration or aggregation or conformation changes. The chromatographic data may contain undetected sample carryover, which may cause the signal to exhibit false EX1 signatures. [36] The composition of the beads in the proteolysis column can induce extensive back-exchange. [37] Interactions between the peptide and analytical column can also induce excessive or unaccounted back-exchange variations across the peptide set. [38]

Design of the experiment is equally important. [2] Heavy water can change the flexibility of proteins, obscuring protein details. [39] Here, the experiments may need to employ lower D_2O concentrations during the exchange step, a step that can often be implemented with no significant loss of precision. [40] For samples exhibiting conformer-specific properties, "top-down" HDX-MS measurement strategies may be required. [41] For glycoproteins additional experiments that determine the deuteron content of glycan groups may be required. [42, 43] To increase the coverage of rapidly exchanging portions of the protein, the investigator may choose to expand the H/D exchange time window with pH adjustments. [44]

Systematic measurement errors may reside implicitly in the customary reporting of pH and temperature. H/D exchange rates can vary by a factor of ten for each pH unit. [45] When pH is recorded to the nearest tenth unit, the expanded uncertainty of the exchange rate coefficient, k_{ex}, is no less than ≈ 10 %. This uncertainty can be reduced substantially. Modern, commercially-available pH probes that are calibrated with two point bracketing have an expanded uncertainty of $U_{95}(\text{pH}) \approx 0.02$. When pH is determined at this accuracy, the contribution of $U_{95}(\text{pH})$ to the

expanded uncertainty of k_{ex} is ≈ 1 %. This uncertainty contribution is halved when the probe is calibrated with five secondary standard solutions. [46] In practice, the stated uncertainty may be slightly higher due to junction effects induced by the protein and solution; however, probe designs that mitigate these effects are commercially available. Temperature recorded with one degree accuracy also increases the uncertainty slightly. Inexpensive probes can measure temperature to $U_{95}(T) = 0.02\ ^{o}C$ accuracy. As higher HDX-MS measurement quality is attained, the supporting measurements will need smaller reported uncertainties.

A fundamental limitation of the proteolytic fragmentation HDX-MS method arises from the loss of the deuterium label (back-exchange) during immersion in the quench solution and throughout chromatographic analysis. When averaged across all analyzed peptides, the typical proteolytic fragmentation HDX-MS study reports deuterium recovery in the range of 70 %. [45] The fractional losses of the peptides vary widely, due to differences among intrinsic amino acid D/H exchange rates. The eluting peptides exhibit large variability in D-recovery from one residue to another, as well as one peptide to another.

To address back-exchange, the Englander group has reported experiments that optimized each step of the proteolytic fragmentation HDX-MS analysis. Based on their studies, they recommended method revisions that can improve deuterium recovery into the range of 90 % to 100 %. [45] (Their quench and chromatography steps were conducted at ≈ 0 ^{o}C.) This study found that the back exchange rate varies with ionic strength of the solution. For higher ionic strength solutions the average intrinsic back exchange rate is at a minimum at pH 2.5, which is also the customary acidity of the quench buffer. However, their measurements found that low ionic strength solutions exhibit minimum back exchange rates at pH 2.25. Based on these measurements, they recommended using pH 2.5 quench solutions and then flowing solvents adjusted to pH 2.25 through the desalting and analytical columns. They recommended measuring back exchange as a function of sheath gas desolvation temperature within the electrospray source of the mass spectrometer, as the optimum is likely ≈ 100 ^{o}C instead of the 200 ^{o}C to 300 ^{o}C used during proteomic studies. Their study supports the use of large chromatographic flow rates that effectively reduce system volume.

The study by the Englander group reported column interactions that result in a negative correlation between back exchange rates and peptide elution time. [45] Knowledge of the back exchange profiles of peptides can afford the investigator freedom to choose a set of peptides that manifest exceptionally low back exchange by sacrificing the faster exchanging peptide fraction. By choosing this peptide set judiciously, excellent protein coverage can be maintained. [45]

Brock's laboratory has reported a method that limits losses from D-for-H back-exchange by using subzero temperature reversed-phase chromatography. [47] Their method employs buffer modifiers that prevent freezing. They obtained good results when using ethylene glycol because it has good electrospray ionization source compatibility. Their study demonstrated that the use of buffer modifiers allows separations to be accomplished at temperatures as low as −30 °C with negligible (< 5 %) loss of the deuterium label even during long chromatographic separations. Much of this study examined the loss of deuterium from whole, fully deuterated proteins. At −20 °C the observed loss of deuterium from fibrinopeptide A was only 10 % after 20 minutes, as compared to 38 % loss at 0 °C. [47] This approach allows chromatography gradients of 30 minutes or longer, which improves separations and can increase the number of resolved peptides.

In liquid chromatography-mass spectrometry (LC-MS), variations in environmental conditions and samples can cause retention times to fluctuate among experiments. As described

in Zhang *et. al.*, the presence of internal standards, such as a mixture of unstructured peptides, in each exchange solution can be used to monitor the extent of back exchange from run to run. (44) When comparing features among several datasets, an accurate time alignment among the datasets can improve the quality of data analysis and validation. In HDX experiments, the non-deuterated and deuterated isotope clusters may exhibit mass shifts of more than 10 Da and the monoisotopic peak in many deuterated isotopic clusters may not be observed. The position of the natural abundance *m/z* peak must be inferred. Venable *et al.* have reported a procedure by which the systematic increase of temporal uncertainty for the entire set of chromatographic runs can be minimized by fitting the collection of possible peak pairs across the entire set of chromatographic runs. [48] This fitting process reliably corrects for retention time variation, and it overcomes the uncertainty in feature pairings. The method takes into account charge, retention time and mass of the monoisotopic peak. The quality of the alignment is improved by the use of an iterative approach that preferentially eliminates incorrectly assigned features. This procedure seems mathematically robust. In HDExaminer the operator may invoke this bias correction by checking an options box. [29, 49]

4.7 Method Validation: Accuracy Improvements

The choice of the centroid as the derived measurand can introduce bias, particularly, within automated analysis programs because the centroid cannot evidence the presence of I_i vs. $(m/z)_i$ patterns that are characteristic of EX1 kinetics data. [20] Although the centroid measurand is adequate for the analysis of EX2 kinetics data, it ignores much useful data contained in the mass spectra. [50, 51]

Guttman *et al.* [50] has addressed this problem by adopting a derived measurand comprising the binomial distribution function:

$$I_m^{calc} = A \cdot \frac{n!}{m!(n-m)!} \cdot p^x (1-p)^{n-m} \tag{4.3}$$

where the isotopic peaks of the calculated envelope, I_m^{calc}, have been converted to integers of index *m*, which are referenced against the undeuterated spectrum and multiplied by the charge state, *z*. The variable, *A,* is an adjustable weighting term initially set to the base peak intensity. The value of *n* is adjustable, but it is approximated by the number of slow exchanging amide hydrogens in the peptide sequence minus one to account for the N-terminal residue and additional decrements to account for each contained proline. The parameter, *p*, is the binomial probability that signifies the deuterium content of the peptide. The peptide natural abundance isotopic distribution is either read directly from the undeuterated spectra or computed from the amino acid composition. The defined measurand incorporates the instrumental line function:

$$f(x) = a_0 + \sum_{\mu=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.4}$$

where a_0 is the baseline offset constant (or function), x is the running *m/z* variable, μ is the *m/z* peak position, and σ is the Gaussian peak width estimated from the instrument resolution. Using an asymmetric, linear least squares regression, the observed mass spectrum envelope for each peptide, I_m^{obs}, is fit using the function:

$$\chi^2 = \lambda \cdot \sum_{m=0}^{n} (I_m^{calc} - I_m^{obs})^2 \quad (4.5)$$

where the spectral weighting function λ =1 if $I_m^{calc} < I_m^{obs}$; λ >1 if $I_m^{calc} > I_m^{obs}$; and χ^2 is the quantity minimized.

The centroid shift in each peptide spectrum relative to the undeuterated profile serves as an initial estimate of the binomial distribution probability, *p*. The mass spectrum envelope, I_m^{calc}, is described by applying the natural abundance profile to each peak in the binomial distribution. Scaling the highest point of I_m^{calc} so that it matches I_m^{obs} provides the initial estimate of the weighting coefficient *A*. Recursive minimization of χ^2 finds the optimum coefficients. The optimum *p* is the *m/z* value of the peptide profile, and the degree of deuteration is calculated by comparing the *p* values with those of the undeuterated and fully deuterated peptides (*cf*. eq. 4.2, and let $p_k \triangleq m_k$).

For high quality mass spectra resulting from EX2 exchange kinetics, Guttman *et al.'s* derived measurand yields the same degree of deuteration as obtained by using the centroid. Unlike the centroid, Guttman *et al.*'s derived measurand, comprising equations 4.3 and 4.4, incorporates information about the instrument and the isotopic spectrum of the natural abundance peptide. Although computation of the binomial measurand consumes a factor of 10 more time than computation of a centroid, its computational expense affords significant advantages over the centroid. The value of the weighting coefficient, λ, can diminish the sensitivity to overlapping mass peaks that are unrelated to the analyzed peptide. [50] For cases where one binomial function fails to yield a good fit, the measurand can be augmented with two or three binomial functions. When the measurand for a peptide comprises multiple functions, analyses of the fitting coefficients (*i.e.*, A_i, p_i) can reveal the presence of multiple conformers. In addition, analyses of these coefficients can untangle comingled EX1 and EX2 kinetics data, yielding powerful new insights into protein dynamics.

Englander's group has offered an alternate derived measurand that includes the isotopic envelope shape and instrument contributions to mass spectra, Englander's group found that amide resolution was improved and that back exchange can be properly corrected. [51] The method is best implemented on data sets containing many sequentially overlapping peptide fragments, so that the subject protein is covered several times over. Their analysis method fits the observed envelopes of peptide mass spectra with trial envelopes comprising a convolution of the residue-level distribution of carried deuterons with the natural abundances of other atom isotopes, ^{13}C, ^{15}N, etc. For each cycle of fitting, the D-occupancies of the amide sites within each peptide are adjusted for the peptide elution time. After applying appropriate site-specific, intrinsic, back-exchange rates, new deuterium site occupancies are re-estimated from the fitting error. The converged calculation provides the D-occupancy of each amide site for each given t_{ex} time point. Then, the temporal D-occupancies for each amide site are fit to a single exponential rate equation yielding the k_{ex} rate coefficient and the associated protection factor.

Where the centroid of each proteolytic peptide contributes only one constraint to fit, the *m/z* envelope originating from a peptide ion contains ten times as many constraints. This abundance of constraints allows the fitting process to converge to a unique solution that also accounts for D-for-H back-exchange. [51] In contrast, computational attempts to fit centroid data sets to solve for D uptake and H-for-D back-change have found multiple sets of similar but not unique solutions. [51, 52] Both band envelope fitting methods benefit from H/D exchange data obtained in solutions that contain the greatest practical deuterium content (> 80%). [50, 51] In

addition, the decomposition of protein ensembles into multiple conformations exhibiting EX1 and EX2 behaviors will likely be improved by measurement of more time points, including those obtained at greater exchange duration.

4.8 Method Validation: HDX-MS and HDX-NMR Cross Comparisons

An important element of method validation is the comparison of HDX-MS measurements with those determined using a different technology; particularly, technology known to yield reliable measurement data. The amide hydrogen exchange rates derived from HDX-NMR measurements provide such orthogonal reference measurements. The methodology of HDX-NMR differs substantially from HDX-MS. Two dimensional NMR can interrogate an intact protein and measure the exchange rate coefficient of an amide site individually. [54] By changing the sample pH, the exchange rates of individual amide sites are adjusted into a range accessible to the NMR instrument. In this way a large fraction of the amide sites of an intact protein can be determined.

Due to the differences between their respective measurement protocols, the principal sources of error for the HDX-NMR measurement differ significantly from those affecting HDX-MS. Where derived rate coefficient determinations from HDX-NMR and HDX-MS overlap, the data sets can be compared directly. The degree of accord between the two sets of k_{ex} values derived from HDX-NMR and HDX-MS provides an independent measure of the true values of k_{ex} and of the reliability of both measurement methods.

In a study of HDX rates by two dimensional NMR, Englander's group determined the exchange rates for individual amide sites of staphylococcal nuclease, SNase (Protein Data Bank: 1SNO) [53]. [54, 55] The rate coefficients for the amide hydrogen exchange are precise within 14 % (u_{cA}). [54] Using HDX-MS, they also measured the D uptake by SNase after its immersion in D_2O. Data sets at four pH's were collected for exchange intervals ranging between $t_{ex} \approx 10$ seconds to $t_{ex} \approx 4$ weeks. The measured amide hydrogen exchange rates varied by a factor of 10^7. Each data set comprises ≈ 300 proteolytic peptide fragment ions. The peptide sequences overlapped extensively. [51] Using the fitting method described above, the solutions of the HDX-MS data converged to yield the back-exchange corrected, D-occupancy at each amide site. The time-dependent D-occupancy at each amide site was then fit to a single exponential decay function, yielding the site-resolved HDX rate coefficient and its protection factor. Figure 4.3A shows the quality of such fits for three residues.

The protection factors of SNase amide sites determined from the HDX-NMR and HDX-MS data have 99 sites in common (Figure 4.3B). Direct comparison of the results reveals that 81 amino acids lie within threefold of the measured NMR protection factors, 16 lie within 10-fold, and 2 are outliers (Figure 4.3C). This comparison is direct. It required no scaling factors. Although it is not reported in this study, global comparability value for the HDX-MS and HDX-NMR sets can be computed from the rate coefficients and associated uncertainties. The global comparability value may prove to be a useful criterion for assessing accuracy during optimization of HDX-MS methodologies.

4.9 Standards: Reference Materials

Periodic measurements of a reference material can provide a base of values useful for assuring consistent measurements in laboratory settings. When the measurement campaign involves two or more laboratories, reference materials are an essential tool for maintaining

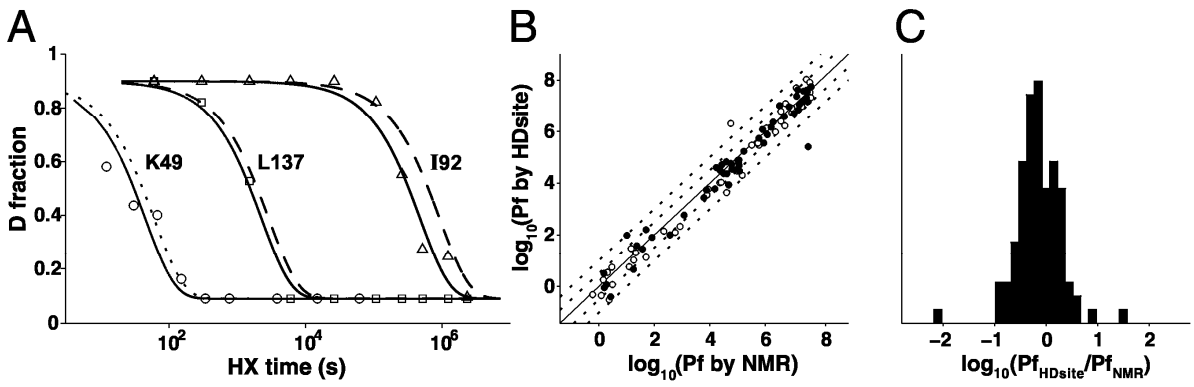

Figure 4.3. Toward single amino acid resolution. (A) Some residue-resolved HDX-MS results for SNase (PDB: 1SNO) [53] compared with HDX-NMR results (dashed curves) or with the calculated rate for an unprotected amide not measured by NMR (dotted curve). (B) Comparison of HDX-NMR and HDX-MS data for SNase plotted in terms of HDX protection factor (P_f = measured HDX rate/expected unprotected rate). Filled symbols indicate directly determined HDX-MS D-occupancy. Open symbols, switchable sites due to incomplete MS peptide overlap, are paired with their apparent NMR identities (this does not alter the fit quality in B and C). Dotted lines show deviations of threefold and 10-fold from the identity line. (C) Population distribution of site-resolved protection factors computed from HDX-MS data versus measured by NMR. Figure from ref. [51] is used with permission of PNAS.

comparable interlaboratory performance. Furthermore, reference materials are critically important for maintaining consistent performance during investigations lasting months or years.

Each future reference material for hydrogen exchange measurements must be a Type A protein (ref. section 4.2), so that connection to the SI is maintained. The protein ensemble should comprise a single conformation. Due to their microheterogeneity, most glycoproteins are Type B materials, which disqualify their use as primary reference materials. Glycoproteins sharing a single glycan structure can serve as primary reference materials.

The quality and utility of each reference material will be greatly enhanced by the availability of a set of amide exchange rate coefficients derived from measurements by an orthogonal method, such as HDX-NMR. The physical properties of proteins and their behaviours during proteolysis vary considerably. The optimum set of reference proteins should reflect these varied behaviours. This will enable researchers to pair the measurement difficulties of their commercial biotherapeutic product with a primary reference material. The chosen primary reference material can then be used to measure HDX-MS laboratory performance over time and location.

Certified reference materials, or CRM's, issued by a national laboratory differ from the "reference materials" specified in the ICH Q6B document issued by the Food and Drug Administration (FDA). ICH Q6B describes requirements for approvals for new drugs and for changes in manufacturing. It stipulates that: "*… the manufacturer should have established an appropriately characterized in-house primary reference material, prepared from lot(s) representative of production and clinical materials. In-house working reference material(s) used*

in the testing of production lots should be calibrated against this primary reference material." The document continues: "*Where an international or national standard is available and appropriate, reference materials should be calibrated against it.*" [56] Thus, a pharmaceutical company uses its in-house reference materials to demonstrate the comparability of each biotherapeutic lot. Ideally, the company would use certified reference materials, issued by a national laboratory or other standards organization, to assure consistent performance of the laboratory instrumentation. Alternately, the research community may propose and validate an interim consensus standard, which can cover the immediate need for a common standard for evaluating laboratory performance. [14, 57, 58]

Thus far, no national standards laboratory (*e.g.*, NIST) or private standards organization has issued, or recommended, suitable proteins for use in a HDX-MS calibration hierarchy. This absence is not surprising, as investigators using HDX-MS have yet to settle on a set of consensus standard reference materials. In view of the rapid expansion in the use of HDX-MS for research applications and its likely future applications as a quality control for biopharmaceuticals, perhaps it is time for the HDX-MS community, the national standards laboratories, and reference material producers to devote resources to identifying and characterizing proteins that can serve as suitable reference proteins.

4.10 Summary: Maintaining Standards and Monitoring Performance

The HDX-MS laboratory must expend considerable effort validating their measurement methods, developing internal reference materials and procedures, and constructing the traceable calibration hierarchy that demonstrates laboratory consistency. However, attaining this state of affairs is only a beginning. After a method has been validated and implemented, the laboratory organization must continue monitoring laboratory measurement quality to assure that the HDX-MS method is performing within specifications.

This monitoring program involves on-going quality control of HDX-MS laboratory performance through the use of reference materials. Clear performance limits for each HDX-MS system component should be specified, and a standard operating procedure (SOP) document should be developed. This document should include corrective actions that should be applied when a limit is exceeded. These actions include component recalibration and then revalidation of the HDX-MS method. Collaborative exercises and proficiency testing schemes (usually round robins) [59] can be used to monitor and compare measurement performance with previous determinations and against the performance of other laboratories. These studies are essential for establishing the reliability and comparability of data that will be shared with the scientific and commercial communities.

4.11 References

1. M. J. Chalmers, B. D. Pascal, S. Willis, J. Zhang, S. J. Iturria, J. A. Dodge and P. R. Griffin, Methods for the Analysis of High Precision Differential Hydrogen–Deuterium Exchange Data, *International Journal of Mass Spectrometry*, **302**(1–3), 59-68 (2011).
2. R. E. Iacob and J. R. Engen, Hydrogen Exchange Mass Spectrometry: Are We out of the Quicksand?, *Journal of the American Society for Mass Spectrometry*, **23**(6), 1003-1010 (2012).
3. Y. Hamuro, S. J. Coales, M. R. Southern, J. F. Nemeth-Cawley, D. D. Stranz and P. R. Griffin, Rapid Analysis of Protein Structure and Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry, *Journal of Biomolecular Techniques*, **14**(3), 171–182 (2003).

4. M. J. Chalmers, S. A. Busby, B. D. Pascal, M. R. Southern and P. R. Griffin, A Two-Stage Differential Hydrogen Deuterium Exchange Method for the Rapid Characterization of Protein/Ligand Interactions, *Journal of Biomolecular Techniques*, **18**(4), 194–204 (2007).
5. T. Liu, D. Pantazatos, S. Li, Y. Hamuro, V. J. Hilser and V. L. Woods, Jr., Quantitative Assessment of Protein Structural Models by Comparison of H/D Exchange MS Data with Exchange Behavior Accurately Predicted by DXCOREX, *Journal of the American Society for Mass Spectrometry*, **23**(1), 43-56 (2012).
6. L. Swint-Kruse and A. D. Robertson, Temperature and pH Dependences of Hydrogen Exchange and Global Stability for Ovomucoid Third Domain, *Biochemistry*, **35**(1), 171-180 (1996).
7. A. J. Chirino and A. Mire-Sluis, Characterizing Biological Products and Assessing Comparability Following Manufacturing Changes, *Nature Biotechnology*, **22**(11), 1383–1391 (2004).
8. A. J. Chirino and M.-S. A. R., State of the Art Analytical Comparability: A Review, *Developments in biologicals*, **122**, 3–26 (2005).
9. D. Houde, S. A. Berkowitz and J. R. Engen, The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies, *Journal of Pharmaceutical Sciences*, **100**(6), 2071-2086 (2011).
10. M. Federici, A. Lubiniecki, P. Manikwar and D. B. Volkin, Analytical Lessons Learned from Selected Therapeutic Protein Drug Comparability Studies, *Biologicals*, **41**(3), 131-147 (2013).
11. J. Marriott , G. O'Connor and H. Parkes. Study of Measurement Service and Comparison Needs for an International Measurement Infrastructure for the Biosciences and Biotechnology: Input for the BIPM Work Programme Paris, France: International Bureau of Weights and Measures (BIPM), 2011, (Available from: http://www.bipm.org/.
12. M. Panteghini and J. C. Forest, Standardization in Laboratory Medicine: New Challenges, *Clinica Chimica Acta*, **355**(1–2), 1-12 (2005).
13. *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM), 2008 Version with Minor Corrections* (Date Accessed: 1 NOV 2013) Available from: http://www.bipm.org/.
14. G. Koumantakis, Traceability of Measurement Results, *Clinical Biochemist Reviews*, **29 (Suppl 1)**(AUG), S61–S66 (2008).
15. BIPM: Bureau International des Poids et Mesures, *Evaluation of Measurement Data — Guide to the Expression of Uncertainty in Measurement (JCGM 100:2008 GUM 1995 with Minor Corrections)*. Sèvres, France. (Date Accessed: 9 SEP 2013) Available from: http://www.bipm.org.
16. *ISO 5725 (1994) Accuracy (Trueness and Precision) of Measurement Methods and Results, Five Parts*. Geneva, Switzerland. (Date Accessed: 3 NOV 2013) Available from: http://www.iso.org/.
17. *ISO/IEC 17025 (2005): General Requirements for the Competence of Testing and Calibration Laboratories*. Geneva, Switzerland. (Date Accessed: 1 NOV 2013) Available from: http://www.iso.org/.
18. *EURACHEM/ CITAC Guide No. 4: Quantifying Uncertainty in Analytical Measurement*. (Date Accessed: 4 NOV 2013) Available from: http://www.eurachem.org.
19. Rigorously, the measurand of Fourier transform instruments is the frequency and amplitude of an image current and in quadrupole and time-of-flight instruments it is the ion

multiplier output current. In practice, the instrument manufacturer provides the customer the traceable calibration hierarchy information relevant to accounting for the I^{obs} vs. *m/z* spectrum. The investigator must verify the claimed traceability of the instrument. The final traceability to the instrument performance (*i.e*., mass calibration) remains the responsibility of the laboratory operator.
20. J. Zhang, P. Ramachandran, R. Kumar and M. L. Gross, H/D Exchange Centroid Monitoring Is Insufficient to Show Differences in the Behavior of Protein States, *Journal of the American Society for Mass Spectrometry*, **24**(3), 450-453 (2013).
21. Z. Zhang and D. L. Smith, Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation, *Protein Science*, **2**, 522-531 (1993).
22. J. J. Englander, C. Del Mar, W. Li, S. W. Englander, J. S. Kim, D. D. Stranz, Y. Hamuro and V. L. Woods, Jr., Protein Structure Change Studied by Hydrogen-Deuterium Exchange, Functional Labeling, and Mass Spectrometry, *Proceedings of the National Academy of Sciences USA*, **100**(12), 7057-7062 (2003).
23. E. Theodorsson, Validation and Verification of Measurement Methods in Clinical Chemistry, *Bioanalysis*, **4**(3), 305-320 (2012).
24. A. Menditto, M. Patriarca and B. Magnusson, Understanding the Meaning of Accuracy, Trueness and Precision, *Accreditation and Quality Assurance*, **12**(1), 45-47 (2007).
25. K. M. Burns, M. Rey, C. A. H. Baker and D. C. Schriemer, Platform Dependencies in Bottom-up Hydrogen/Deuterium Exchange Mass Spectrometry, *Molecular & Cellular Proteomics*, **12**(2), 539-548 (2013).
26. I. Apostol, I. Krull and D. Kelner. Analytical Method Validation for Biopharmaceuticals. (2012. In: Analytical Chemistry [Internet]. Rijeka, Croatia.: InTech, [cited 31 OCT 2913]; [pp. 115-134]. Available from: http://www.intechopen.com/books/analytical-chemistry/analytical-method-validation-for-biopharmaceuticals.
27. L. Mayne, Z.-Y. Kan, P. S. Chetty, A. Ricciuti, B. T. Walters and S. W. Englander, Many Overlapping Peptides for Protein Hydrogen Exchange Experiments by the Fragment Separation-Mass Spectrometry Method, *Journal of the American Society for Mass Spectrometry*, **22**, 1898-1905 (2011).
28. Protein sample was prepared by diluting bovine ubiquitin in PBS buffer (10 mmol/L sodium phosphate, 138 mmol/L NaCl, pH 7.4) to prepare a 3 μmol/L final analytical concentration, and the sample was allowed to equilibrate at 4 °C for 2 h. Sample manipulations were conducted by an HDX PAL robot (LEAP Technologies, Carrboro, NC). By operation of an automated syringe mounted on a robot arm, protein solution (5 μL) was diluted into 25 μL D_2O buffer (20 mmol/L sodium phosphate, 500 mmol/L NaCl, pD 7.6) at 3 °C. This action initiated hydrogen-deuterium exchange reactions. After 30 seconds the HDX reaction process was quenched by mixing each sample with 35 μL of 3 mmol/L urea, 1% TFA, pH 2.5 at 1 °C. Within seconds a syringe on the second robot arm removed the quenched solution and injected it into the sample loop of a refrigerated (1 °C) chromatographic apparatus. Subsequently, the quenched solution was digested on-line with an immobilized pepsin column, and the digested protein solution was trapped on a C18 guard column (1.0 mm dia., 5 μm, Grace Discovery Sciences). The peptide mixture was separated with a C18 analytical column (1.0 mm dia. x 5 cm length, 1.9 μm, Hypersil GOLD, Thermo Scientific) via a Dionex Ultimate 3000 UPLC with a 9.5 min gradient and analyzed in a Thermo LTQ Orbitrap Elite (Thermo Fisher, San Jose, CA). To minimize mass bias in the mass analyzer, the resolution was set to 60,000. Deuterium uptake of

each peptide was calculated using HDX Workbench (See Refs. 29, 30). The data are not corrected for back exchange.
29. Certain commercial materials and equipment are identified in this paper in order to adequately specify the experimental procedure. Such identification implies neither recommendation or endorsement by the National Institute of Standards and Technology nor that the material or equipment identified is the best available for the purpose.
30. B. Pascal, S. Willis, J. Lauer, R. Landgraf, G. West, D. Marciano, S. Novick, D. Goswami, M. Chalmers and P. Griffin, HDX Workbench: Software for the Analysis of H/D Exchange MS Data, *Journal of the American Society for Mass Spectrometry*, **23**, 1512−1521 (2012).
31. M. J. Chalmers, S. A. Busby, B. D. Pascal, Y. He, C. L. Hendrickson, A. G. Marshall and P. R. Griffin, Probing Protein Ligand Interactions by Automated Hydrogen/Deuterium Exchange Mass Spectrometry, *Analytical Chemistry*, **78**(4), 1005-1014 (2006).
32. W. Burkitt and G. O'Connor, Assessment of the Repeatability and Reproducibility of Hydrogen/Deuterium Exchange Mass Spectrometry Measurements, *Rapid Communications in Mass Spectrometry*, **22**(23), 3893-3901 (2008).
33. J. R. Clarkson, Z. F. Cui, R. C. Darton and J. R. Clarkson, Protein Denaturation in Foam - I. Mechanism Study, *Journal of Colloid and Interface Science*, **215**(2), 323-332 (1999).
34. J. B. Hedges, S. Vahidi, X. F. Yue and L. Konermann, Effects of Ammonium Bicarbonate on the Electrospray Mass Spectra of Proteins: Evidence for Bubble-Induced Unfolding, *Analytical Chemistry*, **85**(13), 6469-6476 (2013).
35. G. B. Strambini and E. Gabellieri, Proteins in Frozen Solutions: Evidence of Ice-Induced Partial Unfolding, *Biophysical Journal*, **70**(2), 971-976 (1996).
36. J. Fang, K. D. Rand, P. J. Beuning and J. R. Engen, False EX1 Signatures Caused by Sample Carryover During HX MS Analyses, *International Journal of Mass Spectrometry*, **302**(1-3), 19-25 (2011).
37. Y. Wu, S. Kaveti and J. R. Engen, Extensive Deuterium Back-Exchange in Certain Immobilized Pepsin Columns Used for H/D Exchange Mass Spectrometry, *Analytical Chemistry*, **78**(5), 1719-1723 (2006).
38. J. G. Sheff, M. Rey and D. C. Schriemer, Peptide-Column Interactions and Their Influence on Back Exchange Rates in Hydrogen/Deuterium Exchange-MS, *Journal of the American Society for Mass Spectrometry*, **24**(7), 1006-1015 (2013).
39. P. Cioni and G. B. Strambini, Effect of Heavy Water on Protein Flexibility, *Biophysical Journal*, **82**(6), 3246-3253 (2002).
40. G. W. Slysz, A. J. Percy and D. C. Schriemer, Restraining Expansion of the Peak Envelope in H/D Exchange-MS and Its Application in Detecting Perturbations of Protein Structure/Dynamics, *Analytical Chemistry*, **80**(18), 7004-7011 (2008).
41. J. Pan, J. Han, C. H. Borchers and L. Konermann, Conformer-Specific Hydrogen Exchange Analysis of Aβ(1–42) Oligomers by Top-Down Electron Capture Dissociation Mass Spectrometry, *Analytical Chemistry*, **83**(13), 5386-5393 (2011).
42. M. Guttman, M. Scian and K. K. Lee, Tracking Hydrogen/Deuterium Exchange at Glycan Sites in Glycoproteins by Mass Spectrometry, *Analytical Chemistry*, **83**(19), 7492-7499 (2011).
43. R. Y. C. Huang and J. W. Hudgens, Effects of Desialylation on Human A1-Acid Glycoprotein–Ligand Interactions, *Biochemistry*, **52**(40), 7127-7136 (2013).

44. D. Goswami, S. Devarakonda, M. Chalmers, B. Pascal, B. Spiegelman and P. Griffin, Time Window Expansion for HDX Analysis of an Intrinsically Disordered Protein, *Journal of the American Society for Mass Spectrometry*, **24**(10), 1584-1592 (2013).
45. B. T. Walters, A. Ricciuti, L. Mayne and S. W. Englander, Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment, *Journal of the Americaon Society for Mass Spectrometry*, **23**(12), 2132-2139 (2012).
46. R. P. Buck, S. Rondinini, A. K. Covington, F. G. K. Baucke, C. M. A. Brett, M. F. Camões, M. J. T. Milton, T. Mussini, R. Naumann, K. W. Pratt, P. Spitzer and G. S. Wilson, Measurement of pH. Definition, Standards, and Procedures (IUPAC Recommendations 2002), *Pure and Applied Chemistry*, **74**(11), 2169–2200 (2002).
47. J. D. Venable, L. Okach, S. Agarwalla and A. Brock, Subzero Temperature Chromatography for Reduced Back-Exchange and Improved Dynamic Range in Amide Hydrogen/Deuterium Exchange Mass Spectrometry, *Analytical Chemistry*, **84**(21), 9601-9608 (2012).
48. J. D. Venable, W. Scuba and A. Brock, Feature Based Retention Time Alignment for Improved HDX MS Analysis, *Journal of the American Society for Mass Spectrometry*, **24**(4), 642-645 (2013).
49. HDExaminer Sierra Analytics, Inc., (Modesto, CA); 2009.
50. M. Guttman, D. D. Weis, J. R. Engen and K. K. Lee, Analysis of Overlapped and Noisy Hydrogen/Deuterium Exchange Mass Spectra, *Journal of the American Society for Mass Spectrometry*, **24**(12), 1906-1912 (2013).
51. Z. Y. Kan, B. T. Walters, L. Mayne and S. W. Englander, Protein Hydrogen Exchange at Residue Resolution by Proteolytic Fragmentation Mass Spectrometry Analysis, *Proceedings of the National Academy of Sciences USA*, **110**(41), 16438-16443 (2013).
52. Z. Q. Zhang, A. Zhang and G. Xiao, Improved Protein Hydrogen/Deuterium Exchange Mass Spectrometry Platform with Fully Automated Data Processing, *Analytical Chemistry*, **84**(11), 4942-4949 (2012).
53. D. M. Truckses, K. E. Prehoda, S. C. Miller, J. L. Markley and J. R. Somoza, Coupling between Trans/Cis Proline Isomerization and Protein Stability in Staphylococcal Nuclease, *Protein Science*, **5**(9), 1907-1916 (1996).
54. J. J. Skinner, W. K. Lim, S. Bédard, B. E. Black and S. W. Englander, Protein Hydrogen Exchange: Testing Current Models, *Protein Science*, **21**(7), 987-995 (2012).
55. J. J. Skinner, W. K. Lim, S. Bédard, B. E. Black and S. W. Englander, Protein Dynamics Viewed by Hydrogen Exchange, *Protein Science*, **21**(7), 996-1005 (2012).
56. *ICH Harmonised Tripartite Guideline, Topic Q6B. Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products. Step 4, Consensus Guideline*. (Date Accessed: 15 OCT 2013) Available from: http://www.ich.org/.
57. M. Panteghini, Traceability, Reference Systems and Result Comparability, *Clinical Biochemist Reviews*, **28**(3), 97-104 (2007).
58. *ISO 17511:2003 Metrological Traceability for in Vitro Diagnostic Medical Devices*. (Date Accessed: 3 NOV 2013) Available from: http://www.iso.org/.
59. D. W. Tholen, ISO/IEC 17043: The New International Standard for Proficiency Testing, *Accreditation and Quality Assurance*, **13**(12), 727-730 (2008).