Protein Science (1997), *6*:2203–2217. Cambridge University Press. Printed in the USA.

Human recombinant [C22A] FK506-binding protein amide hydrogen exchange rates from mass spectrometry match and extend those from NMR

ZHONGQI ZHANG,[1] WEIQUN LI,[2] TIMOTHY M. LOGAN,[2] MING LI,[2] AND ALAN G. MARSHALL[1,2]

[1]Center for Interdisciplinary Magnetic Resonance, National High Magnetic Field Laboratory, Florida State University, Tallahassee, Florida 32310
[2]Department of Chemistry, Florida State University, Tallahassee, Florida 32310

(RECEIVED April 7, 1997; ACCEPTED May 29, 1997)

Abstract

Hydrogen/deuterium exchange behavior of human recombinant [C22A] FK506 binding protein (C22A FKBP) has been determined by protein fragmentation, combined with electrospray Fourier transform ion cyclotron resonance mass spectrometry (MS). After a specified period of H/D exchange in solution, C22A FKBP was digested by pepsin under slow exchange conditions (pH 2.4, 0 °C), and then subjected to on-line HPLC/MS for deuterium analysis of each proteolytic peptide. The hydrogen exchange rate of each individual amide hydrogen was then determined independently by heteronuclear two-dimensional NMR on ^{15}N-enriched C22A FKBP. A maximum entropy method (MEM) algorithm makes it possible to derive the distributions of hydrogen exchange rate constants from the MS-determined deuterium exchange-in curves in either the holoprotein or its proteolytic segments. The MEM-derived rate constant distributions of C22A FKBP and different segments of C22A FKBP are compared to the rate constants determined by NMR for individual amide protons. The rate constant distributions determined by both methods are consistent and complementary, thereby validating protein fragmentation/mass spectrometry as a reliable measure of hydrogen exchange in proteins.

Keywords: conformation; FK506 binding protein; FT-ICR; hydrogen exchange; mass spectrometry; maximum entropy method; NMR; solvent accessibility.

Protein hydrogen exchange has become a powerful tool for probing the higher-order structure and dynamics of proteins (Woodward et al., 1982; Englander & Kallenbach, 1984; Gregory & Rosenberg, 1986; Englander et al., 1996). Among the many different techniques for determining backbone amide hydrogen exchange rates, high-resolution multidimensional NMR spectroscopy has been the most successful and definitive, because it can determine hydrogen exchange rates of individual amide protons. However, such experiments require large quantities of protein (typically tens of mg), and the upper mass limit is typically ≤30,000 Da.

A method based on tritium exchange and proteolytic digest has been described by Rosa and Richards (1979) as well as by Englander et al. (1985) for investigating large proteins. In that fragment separation method, after the protein was labeled with tritium, the protein was fragmented with an acid protease, followed by HPLC separation of the proteolytic peptides. The levels of exchange into different segments of the protein were quantified from the radioactivity and amount of each peptide. That fragment separation approach has been applied to large proteins. However, the fragment separation method requires that each proteolytic peptide be isolated and quantified within a time period significantly shorter than the deuterium back-exchange half-lives, thereby limiting the number of segments available as structural probes (Englander et al., 1992). Zhang and Smith (1993) extended the method by replacing tritium with deuterium, and determining the extent of deuteration for each proteolytic peptide by means of directly coupled HPLC and mass spectrometry (MS). Because the amide deuterium level was determined directly from the molecular weight of the peptide, it was not necessary to isolate the peptide before the analysis. Therefore, the number of peptides available as structural probes was greatly increased.

Mass spectrometry has become increasingly popular for examining high-order structural features of proteins (Smith & Zhang, 1994), and protein fragmentation combined with mass analysis has been applied to many protein structural problems (Zhang & Smith, 1993, 1996; Johnson & Walsh, 1994; Liu & Smith, 1994; Dharmasiri & Smith, 1996; Johnson, 1996; Zhang et al., 1996,

Reprint requests to: Alan G. Marshall, Center for Interdisciplinary Magnetic Resonance, National High Magnetic Field Laboratory, 1800 East Paul Dirac Drive, Florida State University, Tallahassee, Florida 32310; e-mail: marshall@magnet.fsu.edu.

1997; Smith et al., 1997). Compared to H/D exchange observed by NMR, mass spectrometry is vastly more sensitive (picomole quantity) and extends to proteins and complexes of much higher mass. However, mass spectrometry usually yields overall hydrogen exchange for the whole protein or its segments, and back-exchange (see below) can vitiate the results. Although it is obviously desirable to validate the mass spectrometric method by direct comparison to NMR, no systematic comparison of protein hydrogen exchange rates from MS and NMR has yet been performed, mainly because the rates of one or more amide hydrogens are typically missing from the NMR data (due to peak overlap and/or unassigned amide resonances), and the mass spectrometric data typically do not resolve individual amide proton exchange rates.

For a particular amide hydrogen, the exchange reaction follows first-order kinetics, i.e., the number of unexchanged hydrogens decreases exponentially with time during the exchange period. Because different amide hydrogens in a protein segment usually have difference exchange rates, the deuterium exchange-in time course for such a segment is described by a distribution of rate constants. As a first approximation, Zhang and Smith (1993) have used a three-component model, in which three rate constants of the amide hydrogens in each protein segment are used to fit the deuterium exchange-in time courses. However, dividing all amide hydrogens into three groups is clearly arbitrary. As a next step, Zhang et al. (1996) recently described the deuterium exchange of all N backbone amide hydrogens in a segment as a sum of N exponentially decreasing terms, from which the number of distinguishable rate constants is derived iteratively. However, because that method does not consider experimental errors explicitly in the fitting process, the experimental data can be overinterpreted. Both analytical and numerical Laplace inversions have also been used to derive the rate constant distribution in proteins (Knox & Rosenberg, 1980; Gregory, 1983). In this paper, we employ a maximum entropy method (MEM) to deconvolve the deuterium exchange-in time course into a distribution of rate constants. Because MEM inherently incorporates experimental errors, it provides the most unbiased result, thereby avoiding over- or underinterpretation of the data. Moreover, random noise is suppressed in the entropy maximization process, to yield a smoother and less noisy result without sacrificing accuracy significantly. From the MEM distribution of rate constants, the missing NMR data are not so serious a problem, and the MS and NMR data may be compared directly.

Fourier-transform ion cyclotron resonance mass spectrometry [FT-ICR MS (Comisarow & Marshall, 1974; Marshall & Schweikhard, 1992; Speir et al., 1992; Buchanan & Hettich, 1993; Wilkins, 1994)] is especially well-suited for protein hydrogen exchange experiments because its ultrahigh mass-resolving power resolves isotopic distributions in all but very large proteins (McLafferty, 1994; Wu et al., 1995; Senko et al., 1996). [Double-depletion of ^{13}C and ^{15}N further extends the upper mass limit for FT-ICR MS by almost an order of magnitude (Marshall et al., 1997).] The isotopic distributions for a deuterated protein and its proteolytic fragments reveal not only their deuterium contents, but also how the deuterons are distributed—a unique advantage of mass spectrometry, making it potentially complementary to NMR for understanding protein structure and dynamics (Miranker et al., 1993). The high mass accuracy of FT-ICR MS also makes possible the identification of the proteolytic fragments of a protein based on accurate mass measurement alone, thereby greatly simplifying interpretation of the protein fragmentation method used to determine hydrogen exchange rates of different segments of a protein (Zhang & Smith, 1993).

In this paper, hydrogen exchange behavior of a mutant of the FK506 binding protein (C22A FKBP) has been measured by both electrospray ionization (ESI) FT-ICR MS and heteronuclear two-dimensional NMR. FKBP is a small (107 amino acids) protein that exhibits peptidyl-prolyl *cis–trans* isomerase activity in vitro, but whose natural ligands have not been identified. FKBP complexed to ascomycin or rapamycin is involved in immunosuppression through the formation of ternary complexes with calcineurin or FKBP-rapamycin-associated protein (FRAP), respectively (Fruman et al., 1994). For each deuterium exchange-in time course measured by protein fragmentation/mass spectrometry, a distribution of hydrogen exchange rate constants can be derived by MEM. The distributions of rate constants for segments (or differences between two overlapped segments) are then compared to the available NMR rate constants for individual amide hydrogens.

Results and discussion

Analysis of protein structure and dynamics may be complicated by the formation of non-native disulfide bonds. Although FKBP contains no disulfides, there is a single cysteine at position 22. To minimize complications due to intermolecular disulfide bond formation during folding and unfolding, we have replaced this residue with alanine (C22A FKBP). NMR studies reveal no significant difference in the structure of C22A versus wild-type FKBP (M. Li & T.M. Logan, unpubl.), and previous enzymatic assays reveal no change in peptidyl-prolyl *cis–trans* isomerase activity (Park et al., 1992).

NMR determination of H/D exchange rates of individual amide hydrogens of C22A FKBP

The ^{1}H-^{15}N HSQC spectrum of ^{15}N-enriched C22A FKBP collected at four different deuteration stages are presented in Figure 1. The NMR resonances are assigned from a combination of standard triple-resonance experiments. Specifically, the HNCO, HNCACB, and CBCA(CO)NH experiments (Muhandiram & Kay, 1994) were used to identify the individual resonances and provide sequential connectivities via the carbon chemical shifts. These assignments were confirmed by analysis of 3D ^{15}N-separated TOCSY-HSQC and 3D ^{15}N-separated NOESY-HSQC experiments. Analysis of the ^{13}C chemical shifts and NOESY connectivities indicate little change in the structure of C22A FKBP compared to wild-type. The residue labels on each spectrum of Figure 1 denote the amide position on each residue. The labels are in reverse order of the exchange rates of corresponding amide hydrogens. Thus, the peaks remaining in the spectrum after the longest exchange period (48 h) are labeled first ($k < 0.04\ h^{-1}$), then the additional peaks seen in the spectrum after 5 h of exchange ($0.04\ h^{-1} < k < 0.6\ h^{-1}$), followed by additional peaks in the spectrum after 10 min of exchange ($0.6\ h^{-1} < k < 6\ h^{-1}$). The remaining hydrogens in the nonexchanged spectrum are those that exchange very fast ($k > 6\ h^{-1}$).

The hydrogen exchange rate constant of each resolved amide hydrogen was calculated by plotting peak volume versus exchange period. These volumes were then fitted to a single exponential decay (first-order kinetics). The p^2H value of the exchange solution read directly from a pH meter (pD_{read}) was 7.07, or ~0.03 pH unit higher than the pD_{read} value of the exchange solution for MS measurement ($pD_{read} = 7.04$). Because hydrogen exchange is base-

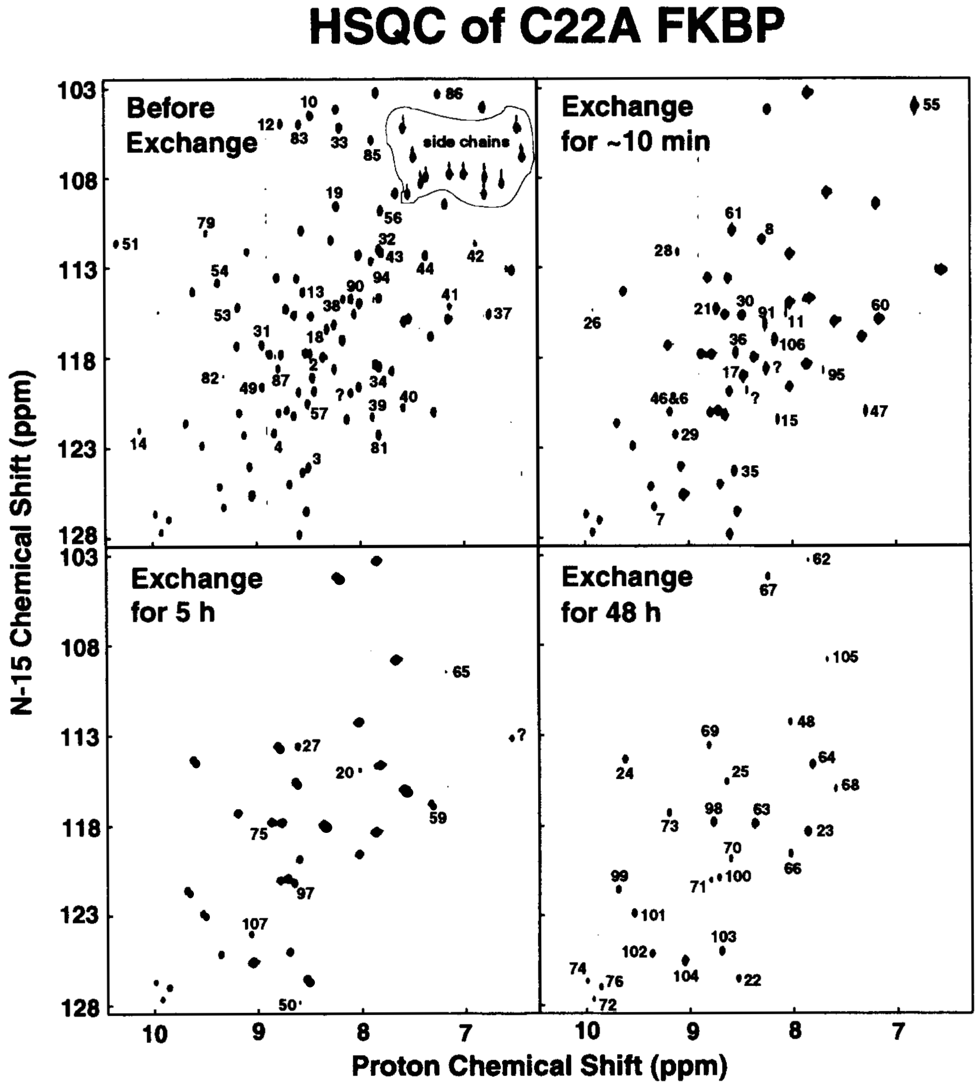

Fig. 1. HSQC spectra of ^{15}N-labeled C22A FKBP after each of four different H/D exchange periods, in which each amide resonance is labeled as its assigned amino acid residue. Assigned amide hydrogens in the spectrum before exchange (upper left) have rate constant $>6\ h^{-1}$; those after ~10 min of exchange (upper right) have rate constants of 0.6–6 h^{-1}; those after 5 h of exchange (lower left) have rate constants, 0.04–0.6 h^{-1}; and those after 48 h of exchange (lower right) have rate constants $<0.04\ h^{-1}$.

catalyzed at pH ~ 7, the exchange rate should be proportional to the hydroxide concentration. The first-order exchange rate constants determined by NMR were therefore corrected by dividing by a factor of $10^{0.03}$ ~ 1.07 for comparison to MS data. C22A FKBP has 107 residues, in which 7 proline residues and the N-terminal residue do not have an amide hydrogen. Therefore, there are 99 total backbone amide hydrogens in C22A FKBP. Due to resonance overlap, only 94 rate constants were obtained from the NMR data. Figure 2 (right) shows a bar graph of the exchange rate constants for amide hydrogens on 90 different residues (exchange rates for amide hydrogens on some residues were not available due to ambiguities in assignment caused by differences in pH between this experiment and the pH for the experiments on which the assignments were based). The minimum measured exchange rate constant for C22A FKBP is ~0.002 h^{-1} and the maximum resolvable rate constant is ~10 h^{-1}. For resonances exhibiting no detectable peak volume at the shortest exchange period (~10 min), the exchange rate constants are $>10\ h^{-1}$. The three-dimensional X-ray crystal structure of FKBP (van Duyne et al., 1991) is shown in Figure 2 (left), and the α-helix and β-sheets are identified on the bar graph of exchange rate constants (Fig. 2, right). It is clear that the exchange rate is reduced significantly in both α-helix and β-sheet secondary structural segments, and is much higher in residues forming loops and turns.

Although the E5 and L104 amide hydrogen resonances overlap, they include a proton with one of the slowest exchange rates in the protein. From the three-dimensional structure of FKBP (Fig. 2, left) and NMR-determined hydrogen exchange rate constant map (Fig. 2, right), it can be seen that residues L97–K105 form the stable interior β-sheet and have the slowest hydrogen exchange rate in the protein (0.002–0.05 h^{-1}), whereas E5 is in a relatively unstable β-strand and is bracketed spatially by several faster exchanging amide hydrogens ($>2\ h^{-1}$). Thus, inclusion of only the later time points (>5 h) to calculate the rate

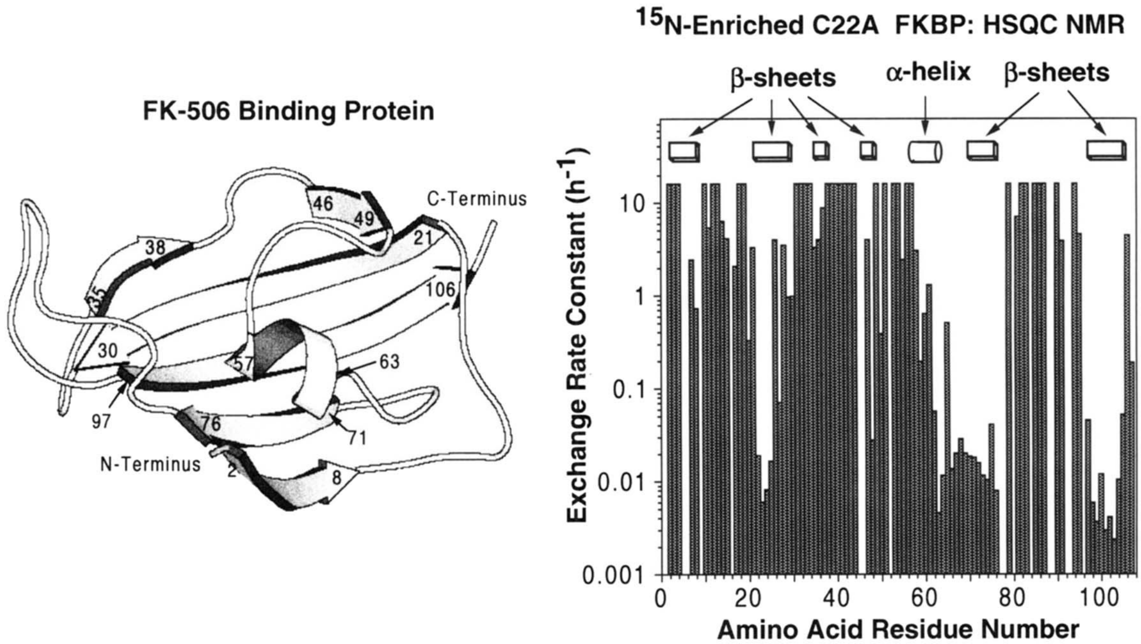

Fig. 2. Left: Three-dimensional structure of FKBP (van Duyne et al., 1991), drawn with MOLSCRIPT (Kraulis, 1991). Right: Hydrogen exchange rate constants determined by NMR, for each of the NMR-assignable backbone amide hydrogens of C22A FKBP. Segments containing α-helix and β-sheets are represented by the symbols ▭ and ▭, respectively.

contant for this resonance should yield a good approximation of the L104 amide hydrogen. That assumption is supported by the good fit between MS and NMR data shown in Figures 5 and 8 (discussed below).

Mass spectrometric identification of proteolytic fragments of C22A FKBP

C22A FKBP was digested by pepsin at pH 2.4 and 0 °C for 3 min at a substrate:enzyme ratio (w/w) of 1:1. The protein digest was then subjected to ESI FT-ICR MS analysis with online desalting. The proteolytic peptides were assigned by combining the following methods: (1) MS/MS spectra of peptides generated by use of infrared multiphoton dissociation (IRMPD) (Little et al., 1994); (2) known proteolytic specificity of pepsin (Powers et al., 1977; Smith et al., 1997); and (3) accurate mass measurements based on identified peptides as internal mass calibration (<10 ppm mass accuracy). Some 34 fragments, spanning the full C22A FKBP primary amino acid sequence were identified. After excluding fragments whose MS signals were too weak to yield accurate deuterium content, we were able to determine the deuterium incorporation into 22 segments, of which 13 (spanning 100% of the C22A FKBP primary sequence) are presented here.

Mass spectrometric determination of H/D exchange for C22A FKBP and its segments

C22A FKBP was exposed to D_2O for different lengths of time and the exchange was quenched in each case by decreasing the pH of the solution to 2.4 and lowering the temperature to 0 °C. Part of this solution was analyzed by online desalting and ESI FT-ICR MS. Figure 3 (left) shows mass spectra of intact C22A FKBP following exposure to D_2O for differing periods of time.

To establish the site(s) of deuterium incorporation, we digested the balance of the ^2H-exchanged C22A FKBP under quenching conditions (pH 2.4, 0 °C) for 3 min, at a substrate:enzyme ratio of 1:1. The protein digest was then subjected to online LC/MS analysis using a reverse-phase perfusion column and ESI FTICR MS. Figure 3 (right) shows mass spectra of the proteolytic peptide, G1–M29 at several different ^2H exchange times. Because the online desalting (of intact C22A FKBP) and HPLC separation (of its pepsin-cleaved fragments) were performed in H_2O, all fast-exchangeable hydrogens on the side chains of the protein or peptides were back-exchanged from deuterium to hydrogen, leaving deuteriums only for the peptide backbone amides.

The top and bottom spectra in Figure 3 (both left and right) were obtained from the nondeuterated protein and fully deuterated protein, respectively. Fully deuterated protein was generated by incubating C22A FKBP in 4 M urea at 40 °C for 50 h in D_2O (the equilibrium denaturation midpoint for C22A at 25 °C is 4.0 M urea; M. Li & T.M. Logan, unpubl.). Both the nondeuterated protein and fully deuterated protein went through exactly the same analysis procedure as the partially deuterated protein (the second and third spectra). The deuterium content, D, for each partially deuterated species was calculated from (Zhang & Smith, 1993):

$$D = \frac{m - m_{0\%}}{m_{100\%} - m_{0\%}} N, \tag{1}$$

in which N is the total number of backbone amide hydrogens in the protein or its segment (namely, $N = 99$ for C22A FKBP and $N = 26$ for G1–M29, after taking into account that the N-terminal residue and prolines do not have amide hydrogens), and $m_{0\%}$, m, and $m_{100\%}$ are the average molecular weights calculated from the centroid of the isotope envelopes of the nondeuterated species (top-

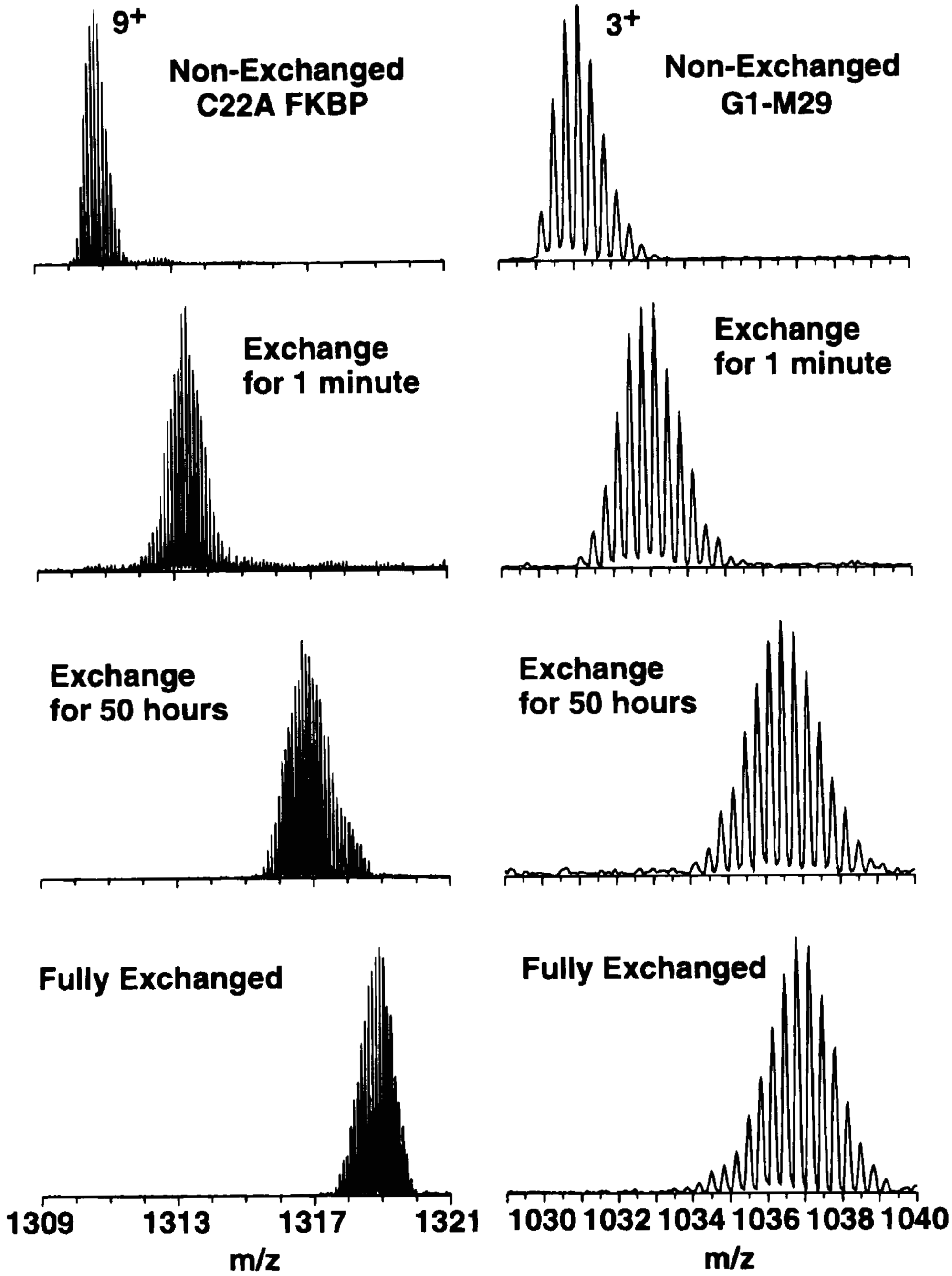

Fig. 3. ESI FT-ICR mass spectra of intact C22A FKBP ($[M+9H]^{9+}$) and segment G1–M29 ($[M+3H]^{3+}$) following various deuteration stages. Top and bottom spectra for both C22A FKBP and G1–M29 represent zero-deuteration and full-deuteration controls, respectively.

most spectra of Fig. 3), partially deuterated species (next two rows of spectra in Fig. 3), and fully deuterated species (bottom spectra of Fig. 3). Deuterium uptake is described by Equation 1 because some deuterium gain (during digestion in D_2O) or loss (during digestion and desalting or HPLC separation in H_2O) always occurs during analysis; thus, the nondeuterated and fully deuterated samples serve as controls to correct for such deuterium gain or loss. All deuterium contents described in this paper are calculated from Equation 1.

Mass spectrometric deuterium exchange-in time course and its resultant hydrogen exchange rate constant distribution; NMR hydrogen exchange rate constant distribution

Figure 4 (left) shows the deuterium exchange-in time courses of intact C22A FKBP (top) and segment G1–M29 of C22A FKBP (bottom). (The deuterium exchange-in time course of C22A FKBP was derived from two sets of exchange data performed and measured on different days.) Unfortunately, we cannot compare di-

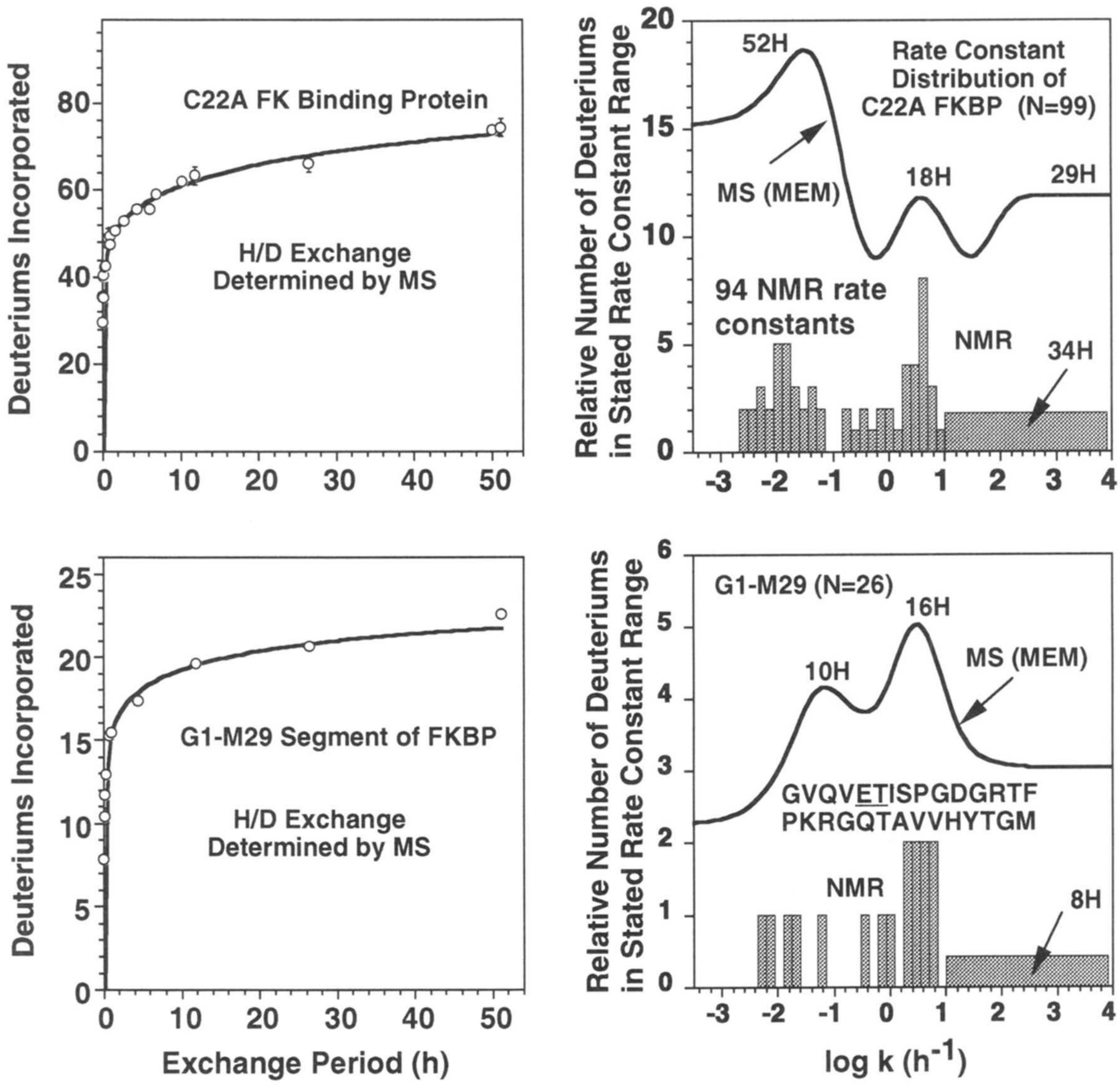

Fig. 4. H/D exchange time courses of C22A FKBP (upper left) and segment G1–M29 (lower left) and the distributions (right, smooth curves) of their rate constants derived from their MS-determined H/D exchange time courses. The number of amide hydrogens for each resolved peak in the MEM-derived rate constant distribution is shown above that peak. The NMR-determined rate constants of amide hydrogens are shown as bars (or as a total number denoted by an arrow when the exchange rate is too fast to be measured by the NMR experiment). Note that NMR lacks the rate constants for five hydrogens in C22A FKBP, and amides E5 and T6 (underlined in the sequence) for G1–M29.

rectly such MS exchange-in time courses with those constructed from individual amide hydrogen exchange rates determined by NMR, because the unavailable amide exchange data from the NMR experiments (due to overlapped resonances and ambiguities in resonance assignment) would distort the comparison. However, we can convert the mass spectrometric results to a form that is comparable directly to the NMR results, by using MEM to extract a distribution of hydrogen exchange rate constants from the MS-determined hydrogen exchange time course data for both intact protein and its proteolytic fragments (see Materials and methods).

Figure 4 (right) shows the rate constant *continuous* distribution (smooth curve) of C22A FKBP and G1–M29 derived by MEM (see Materials and methods) from the MS-determined time course data, along with the *discrete* distribution of rate constants determined by NMR (bars). The total number, N, of amide hydrogens for intact C22A FKBP and its G1–M29 segment is given in the figure. The continuous curves in Figure 4 (left) give rise to the MEM-based rate constant distribution curves shown in Figure 4 (right). In Figure 4 (right), each vertical bar *height* for the NMR-derived data represents the number of amide hydrogens with exchange rate constants in the stated range. For rate constant distributions derived from the MS data, the ordinate in Figure 4 (right) has been scaled such that the total *area* under the curve corresponds to the total number of amide hydrogens in the protein or segment over that range of rate constants. The number of hydrogens (calculated from the area under each resolved peak in the MEM-derived distribution curve) appears at the top of each peak, as a measure of the number of amide hydrogens of each resolvably different exchange rate constant. For both MS and NMR data, the horizontal axis is the logarithm of exchange rate constant—a log scale was chosen to cover a wide range of exchange rates.

A final aspect of the format of Figure 4 requires explanation, before we can interpret the results. Under the conditions for this experiment ($pD_{read} \approx 7$, 25 °C), the hydrogen exchange rate constant for a totally exposed peptide amide hydrogen (and hence the maximum rate constant for any amide hydrogens in the protein) is $\sim 10^4$ h^{-1} (Bai et al., 1993). Thus, the number of amide hydrogens with $\log k > 4$ should be zero. However, because the maximum rate constant that can be determined from the NMR data is ~ 10 h^{-1}, all amide hydrogens with faster exchange rates are considered to have equal probability throughout the range, $1 \leq \log k \leq 4$. In that case, bars with equal abundance are used to represent amide hydrogens with $k > 10$ h^{-1}; the number of hydrogens in this exchange regime is designated by an arrow. For amide hydrogens of slower exchange rate ($\log k < 1$), the number of amide hydrogens in each "bin" of NMR-determined exchange rate constants can be determined directly from the graph. In the rate constant distribution of C22A FKBP determined by NMR (bar graph at upper right of Fig. 4), rate constants determined from all 94 observed HSQC amide proton peaks were used. Similarly, for the MEM distribution based on MS data, MEM assigns equal probability to all rate constants that are too fast or too slow to measure—hence, the MEM curves flatten at each end of the rate constant distribution range. Finally, in cases for which the MEM-determined exchange rate constant falls at the limit of the detectable range (e.g., bottom curves in Fig. 7, see below), the MEM distribution exhibits a sigmoid shape, and the most probable rate constant corresponds to the midpoint of the shoulder.

Comparison of MS- and NMR-derived amide hydrogen exchange rate constant distributions for intact C22A FKBP and its G1–M29 segment

For intact C22A FKBP, the MS- and NMR-derived amide hydrogen exchange rate constant distributions agree quite well: each shows three partially resolved rate constant groupings. The differences between the two distributions result in part from different deuterium back-exchange rates at different amide positions that occur during the MS experiment. Specifically, the number of deuteriums at slow-exchanging sites (due to residual structure present during online desalting) will be overestimated by the MS method (see below for further discussion of this point). Also note that the NMR data for intact C22A FKBP lacks rate constants for five amide hydrogens due to overlapped NMR resonances. Although two amide hydrogens (E5 and T6, which overlap with L104 and F46, respectively) are missing from the NMR data for the G1–M29 segment (Fig. 4, lower right), there is still reasonable agreement with the MS MEM-based rate constant distribution for that segment.

For the rate constant distributions determined by MS for both C22A FKBP and its G1–M29 segment (Fig. 4, right), there is non-zero probability of amide hydrogens with very small exchange rate constants ($k < 10^{-3}$ h^{-1}), whereas NMR shows no amide hydrogens with exchange constants in that range. The reason is that when the exchange rates of several amide hydrogens in a segment are all slow, the MS-determined deuterium exchange-in rate constants are additive; although the sum of those rate constants can be calculated, individual rate constants are not distinguishable and MEM assigns equal probability to all rate constants below the measurable threshold [i.e., the lower $\log k$ limit of each MEM-derived distribution curve in Fig. 4 (right) effectively extends indefinitely to the left]. NMR, on the other hand, yields a definite rate constant because it can resolve individual rate constants. The behavior at the high $\log k$ limit is different from the limit at low $\log k$ because the highest $\log k$ is limited by the exchange rate of an amide hydrogen totally exposed to solvent, which is $\log k \approx 4$ under our experimental conditions. Thus, the MEM-based distribution of rate constants drops effectively to zero for $\log k > 4$.

Note that because (1) the shortest exchange period is ~ 1 min for MS measurements and ~ 10 min for NMR measurements, and (2) MS is more sensitive in detecting small deuterium content, the maximum detectable rate constant from MS is higher than from NMR. In Figure 4 (right), the MEM-derived rate constant distribution curve from MS levels off at $\log k > 2.3$; thus, any rate constants greater than ~ 200 h^{-1} are not distinguishable. In other words, the maximum detectable rate constant is ~ 200 h^{-1} for MS, compared to ~ 10 h^{-1} for NMR. That difference leads to different apparent rate constant distributions for the two methods in the range $1 \leq \log k \leq 2.3$, as shown in Figure 4 (upper right) and Figure 5, because the NMR data do not provide an accurate rate constant for $\log k > 1$.

Comparison of MS- and NMR-derived amide hydrogen exchange rate constant distributions for 12 selected segments of C22A FKBP

Figure 5 shows distributions of amide hydrogen exchange rate constants for 12 different segments of the C22A FK506-binding protein determined from MS and NMR. The amide hydrogens with rate constants unavailable from the NMR data are underlined in each sequence. The two sets of data have reasonable fits for most of the segments. For example, segment S67–L97 exhibits two distinct rate constant groups resolved by the MS data. From the areas under the MEM-derived rate constant distribution curve, 11 of 26 amide hydrogens have rate constants < 0.2 h^{-1}, and 15 of 26 have rate constants > 2 h^{-1}. By comparison, rate constants determined from NMR reveal that 10 amide hydrogens have rate constants in the range of 0.007–0.05 h^{-1} and 11 amide hydrogens have rate constant > 3 h^{-1}. Rate constants for the remaining five amide hydrogens are not available from NMR. Similar comparisons are evident for segments T75–L97, Y82–L97, and L103–E107, for which two distinct rate constant groups are well resolved in each case.

For other segments that do not have well-resolved groups of exchange rate constants, the MS- and NMR-derived data are in reasonable agreement. For some segments (e.g., S67–L97, T75–L97 and T75–F99), the MS- and NMR-derived rate constant distributions differ in the range $1 \leq \log k \leq 2.3$, due to the higher MS-measurable exchange rate constant, as noted above.

Finally, for at least one very short segment (L103–E107), the agreement between MS- and NMR-derived amide hydrogen exchange rate constants is near-quantitative (see Fig. 8 below), suggesting that comparison between MS and NMR-derived rates would be improved by use of short peptide segments. We are thus led to seek ways to shorten the effective length of a segment for MS analysis, as described in the next section.

Use of overlapped segments to improve the resolution of MS-derived amide hydrogen exchange rate constant distributions for comparison to NMR results

The key to improving the MS resolution lies in the availability of peptides whose primary amino acid sequences overlap by just a few

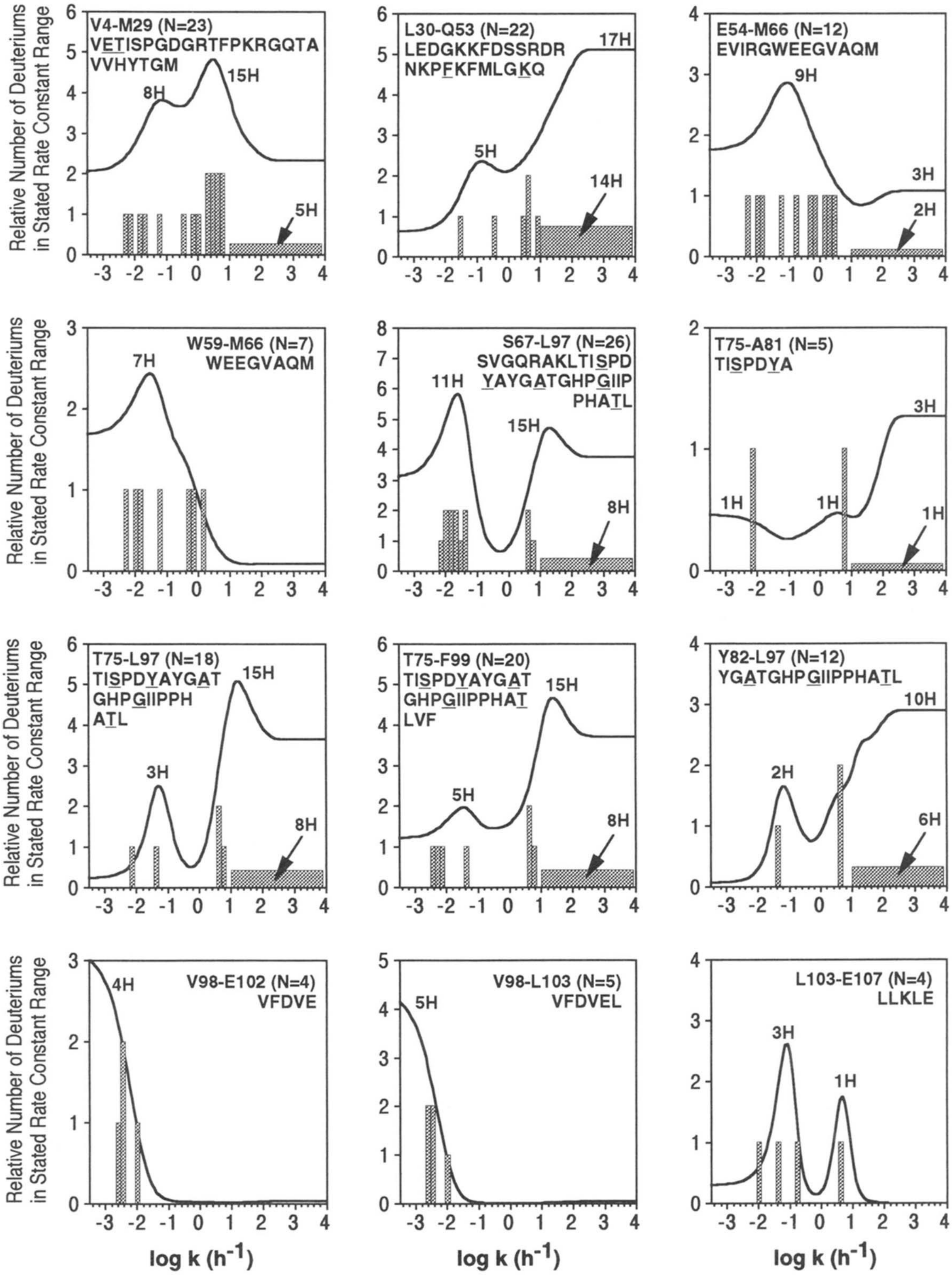

Fig. 5. Distribution of rate constants for many different segments of C22A FKBP derived from MS-determined H/D exchange time courses (smooth curves) and from HSQC NMR (bars). N is the total number of backbone amide hydrogens in that segment. The number of amide hydrogens for each resolved peak in the MEM-derived rate constant distribution is shown above that peak. The NMR-determined rate constants of amide hydrogens are shown as bars (or as a total number denoted by an arrow when the exchange rate is too fast to be measured by NMR). Note the generally excellent agreement between the MS and NMR results. The underlined residues in the sequence denote amide hydrogens for which no NMR-determined rate constant is available due to resonance overlap and/or ambiguity in assignment.

residues. For example, deuterium incorporation into the segment V2–V4 may be determined from the difference in deuterium contents of peptides G1–M29 and V4–M29 (note that the first residue of a peptide does not have an amide hydrogen; and the deuterium on the N-terminal primary amine is washed away in the HPLC separation step). Figure 6 (left) shows the deuterium exchange-in time course for the segment V2–V4, calculated from the difference in the time courses for the above-mentioned two overlapping peptides. MEM may then be applied as before to derive the distribution of rate constants for this three-residue segment (Fig. 6, right). NMR data tell us that amide hydrogens of all three residues exchange rapidly ($>10\ h^{-1}$). However, the MS data show that one hydrogen exchanges slowly ($<0.4\ h^{-1}$). The discrepancy is caused by errors in the original MS data for G1–M29 and V4–M29, because the deuterium content of V2–V4 is determined from a small difference (3 amide protons) between two large segments (26 and 23 amide protons), so that a 4% error in the MS data from the longer segment due to back exchange could lead to an apparent change of one hydrogen. Thus, although spatial resolution may be improved greatly by taking differences between overlapped segments, care must be taken in interpreting the result, especially for a small difference between two large segments. Figure 7 shows the MS-based rate constant distributions for four more short segments derived from data for other overlapping peptides. The MS data for these shorter peptides provide an excellent match to the NMR results in each case.

Comparison of directly observed MS-based and back-calculated NMR-based exchange time courses for selected short segments of C22A FKBP

Among the 12 protein segments presented in Figure 5, 5 have complete corresponding NMR data (E54–M66, W59–M66, V98–E102, V98–L103, and L103–E107). For those segments, rate constants of all amide hydrogens in a segment are known by NMR, and the MS-based deuterium exchange-in time course can be compared directly to the time course generated from the NMR-determined exchange rate constants. Figure 8 shows such comparisons for W54–E66, V98–E102, and L103–E107. The MS-determined deuterium contents after various exchange periods are shown as scattered points and the back-calculated NMR time courses are shown as smooth curves. The MS data match the NMR data perfectly for the L103–E107 segment. The discrepancy between the MS- and NMR-based time courses of W54–E66 and V98–E102 may result from inaccuracy in the deuterium back-exchange correction for the MS data. Typically, the deuterium loss from deuterium back-exchange during digestion and HPLC fractionation is 20–40% of the deuterium (calculated from the observed deuterium content in the fully deuterated control), and the deuterium gains are 1–5% of the nondeuterated amide positions (calculated from the observed deuterium content in the nondeuterated control), varying from one peptide to another. If the deuterium content of a peptide is calculated from Equation 1, we may correct for the deuterium gains or losses by use of the two control samples. In Equation 1, it is assumed that all amide hydrogens in a segment exchange at the same rate during analysis (Zhang & Smith, 1993); however, under quenched conditions (pH 2.4, 0 °C), the hydrogen exchange rates for amide hydrogens in V98–E102 can differ by as much as a factor of 12, with half-lives ranging from 11 to 127 min; and by a factor of 50, with half-lives ranging from 9 to 440 min for W54–E66 [calculated with a program written by the author according to a previously published algorithm (Bai et al., 1993)]. Thus, deuterium back-exchange during analysis is one of the major sources of discrepancies between the MS and NMR results; therefore, efforts were made at every step of the analysis to reduce deuterium back-exchange. These efforts include careful control of pH and temperature during peptic digestion and HPLC separation, as well as reducing analysis time by use of a capillary perfusion column (Afeyan et al., 1990b, 1990a, 1991) for peptide separation.

For some segments, deuterium may be retained on some side chains during LC/MS analysis, leading to another source of devi-

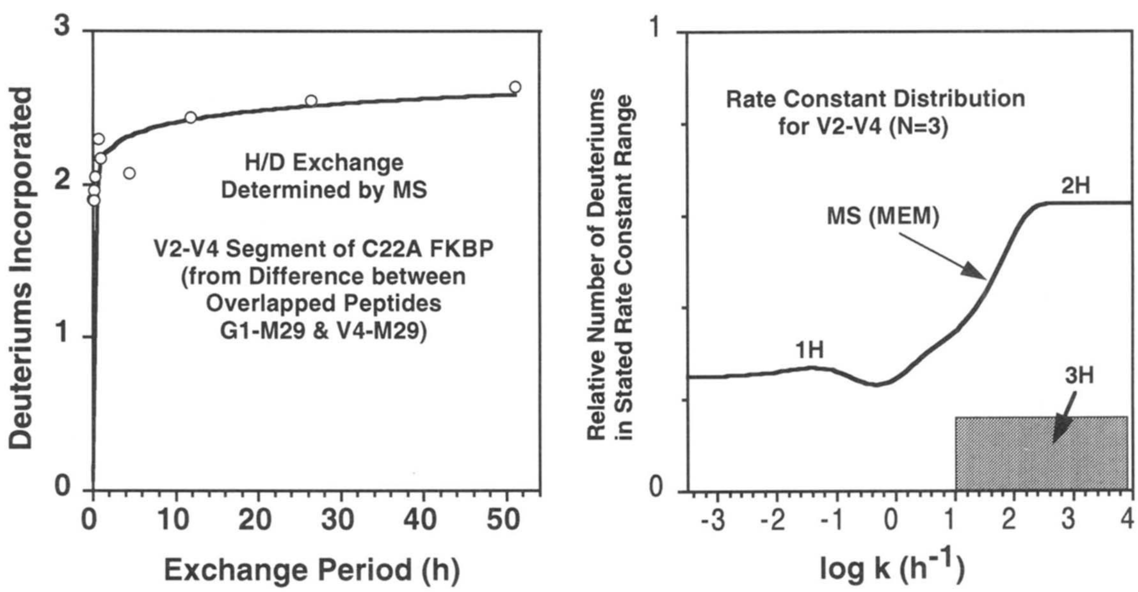

Fig. 6. H/D exchange time course (left) and MEM-derived distribution of H/D exchange rate constants (right) for the short V2–V4 segment (left). The left-hand data represent the difference between similar curves for two overlapped peptides, G1–M29 and V4–M29. The NMR-determined rate constants are shown as bars (right).

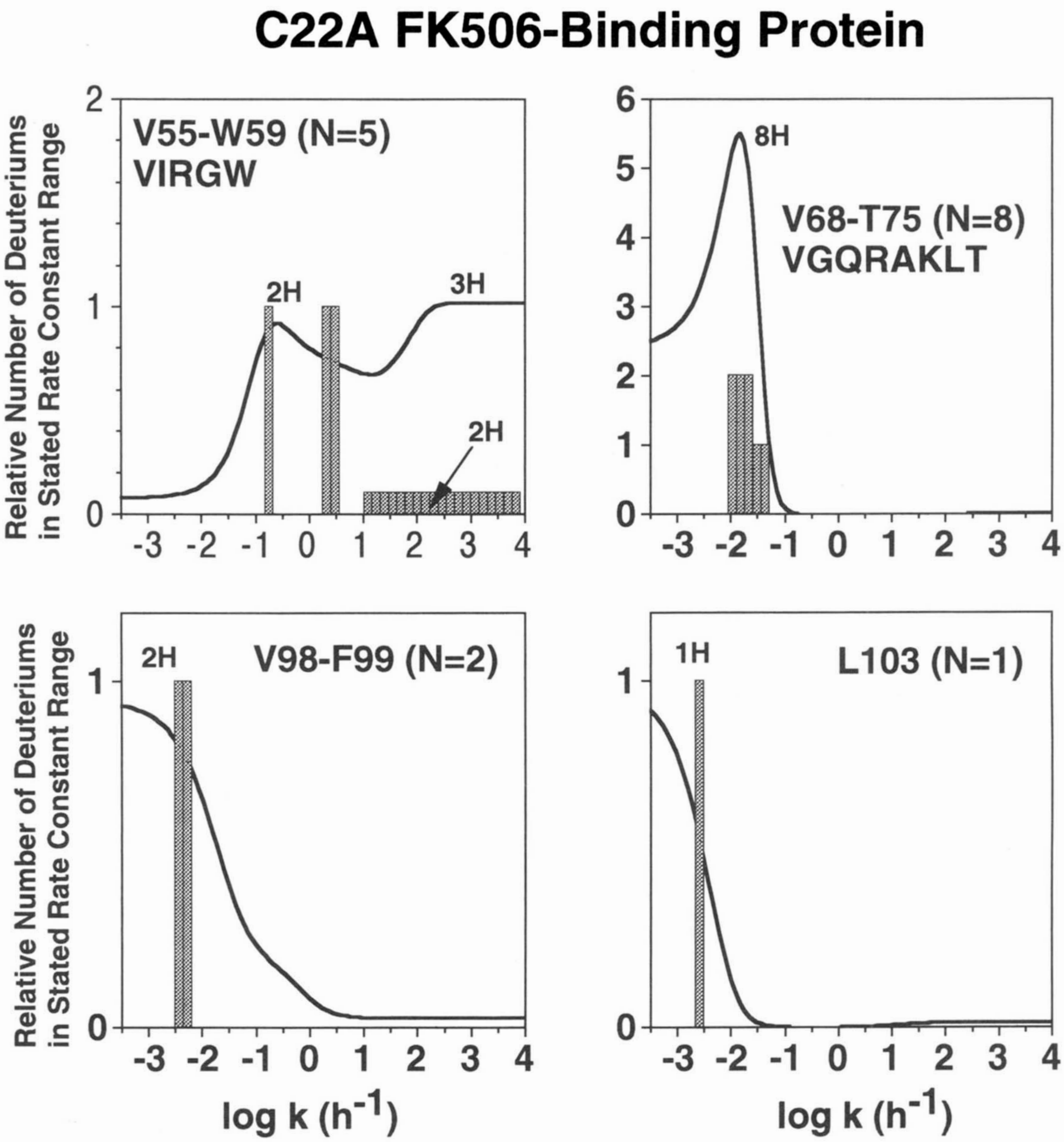

Fig. 7. Distribution of rate constants for four small segments of C22A FKBP, derived from H/D exchange time courses from MS of overlapped peptides (smooth curves), and NMR (bars). Deuterium content in segment V55–W59 was calculated from the difference in deuterium content between E54–M66 and W59–M66; segment V68–T75 from the difference between S67–L97 and T75–A97; segment V98–F99 from the difference between T75–F99 and T75–L97; and L103 from the difference between V98–L103 and V98–E102.

ation between MS and NMR data (Johnson, 1996). In Equation 1, we assume that all side-chain hydrogens are completely washed out in the desalting or HPLC separation step: i.e., that the total number of exchangeable hydrogens is the total number of backbone amide hydrogens in the peptide. However, the NδH of Arg have exchange half-lives of a few minutes under quenched conditions (Bai et al., 1993) and may not wash out completely, especially considering that the whole analysis takes only about 7 min. Moreover, the C-2 hydrogen of His may be exchanged after a prolonged exchange period and that hydrogen is quite stable at low pH and temperature (Gregory & Rosenberg, 1986).

How then are the discrepancies caused by uncertainties in the back-exchange correction reflected in the rate constant distribution? Because the back-exchange correction is applied equally to all of the time course data, the *number* of protons having a calculated rate constant distribution derived from MS versus NMR may differ, but the *rate constants* themselves should not. For example, consider the MEM-derived rate constant distribution of intact C22A FKBP. Although the quenched condition used here is acidic (pH 2.4, 0 °C), some residual structure may protect some amide hydrogens from exchange, and that residual structure may be related to the native structure. As a result, those amide hydrogens that exchange slowly in the

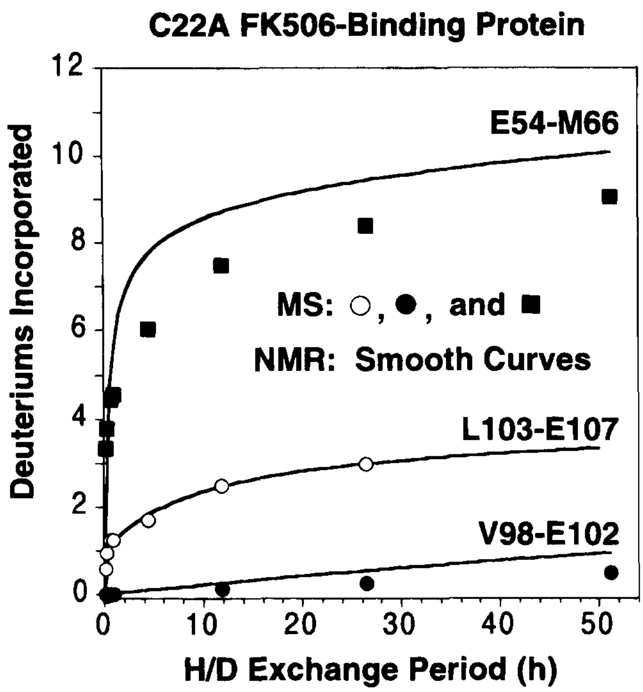

Fig. 8. Comparison of deuterium incorporation as a function of H/D exchange period, determined by MS (data points) and from time courses constructed from NMR-determined exchange rate constants (smooth curves) for C22A FKBP segments E54–M66 (sequence EVIRGWEEGVAQM, solid squares), V98–E102 (sequence VFDVE, solid circles), and L103–E107 (sequence LLKLE, open circles).

native structure may also exchange slowly under quenched conditions. If the same back-exchange factor is used to correct for the deuterium back-exchange for all amide hydrogens, the abundance of slow-exchanging amide hydrogens may be overestimated, leading to a less accurate representation of the rate constant distribution for the intact protein (Fig. 4, upper right). Nevertheless, the positions of peaks in the MS-determined distribution matches quite well to the NMR-determined rate constant distribution.

Conclusions

The primary application of protein hydrogen-deuterium exchange is to characterize protein conformational and dynamical changes. NMR is widely used for studying the exchange process because it provides information on (essentially) all backbone amides, but suffers from a lack of sensitivity and an upper limit of approximately 20 kDa. MS complements the NMR method very well because it has very high inherent sensitivity (which provides access to different exchange rate regimes), and can easily distinguish different conformational states. The problem with MS determination of H/D exchange rates is that exchange rates for individual amides are more difficult to determine.

In this paper, we compare the H/D exchange rates in C22A FKBP determined by NMR and MS, and introduce two methods that allow a semi-quantitative comparison of the H/D exchange rates determined by these two methods. The first approach is to extract the exchange rate distribution in peptide fragments by a maximum entropy method. Laplace inversion has been used to derive hydrogen exchange rate constant distributions for proteins; however, Laplace inversion is a notoriously ill-conditioned problem. In general, even for data with arbitrarily low noise levels, there exist a large number of solutions that fit the experimental data comparably well (Gregory, 1983). In contrast, the present MEM algorithm completely avoids inversion in deriving the distribution. Compared to the previous methods to derive the rate constant distribution (Gregory & Rosenberg, 1986), the MEM algorithm gives a more reliable and less biased solution. In addition, noise is minimized in the entropy maximization process. In the numerical Laplace inversion algorithm described by Gregory (1983), the smoothest solution that is consistent with the data is chosen as the final result, which is similar to the MEM algorithm described in this paper. However, smoothness was judged by visual inspection, introducing subjectivity into the process, whereas MEM finds the smoothest curve based on entropy maximization. Note that back-exchange during the MS method leads to an overestimate of the total number of protons exchanging in a given peptide fragment, but no difference in the exchange rate distributions.

A second improvement is to compare the exchange rate distributions in overlapping peptides to assign differences in the exchange rates of these peptides to individual residues (or to very short primary segments). Pepsin usually generates many overlapping peptides because of its low specificity. Taking difference of overlapping peptides to identify deuterium incorporation of shorter segments is an important approach toward improving the spatial resolution of the MS method (Zhang & Smith, 1993; Zhang et al., 1996). Here we have demonstrated that this method allows a quantitative comparison of exchange rates between NMR and MS for relatively short peptides. In some cases, the difference in deuterium contents of overlapped peptides may even have better accuracy (not precision) than the deuterium content determined from the original peptides. For example, an amide hydrogen on the second residue from the N-terminus of a peptide exchanges relatively fast under quenched conditions (half-life usually < 10 min) due to the inductive effect of the positive charge on the N-terminus (Molday et al., 1972; Bai et al., 1993). As a result, because the deuterium on this position is lost more easily than others during analysis, a large error will result if that position is deuterated, However, the *difference* in deuterium content of two N-terminal overlapping peptides will not be affected. Using this difference method, we achieve a spatial resolution of $\sim$1–4 residues (*average* resolution for the whole protein is $\sim$11 residues); average spatial resolution to as few as 5 residues has been achieved for other proteins (Zhang et al., 1996).

Finally, FT-ICR offers several unique advantages over other forms of mass spectrometry for hydrogen exchange analysis. First, due to its ultrahigh mass *resolution*, isotopic peaks of an intact protein can be resolved. Thus, in a proteolytic digest experiment, FT-ICR MS isotopic peaks of all protein fragments are well resolved, facilitating charge state assignment. With other mass spectrometers, isotopic peaks of all fragments are usually *not* resolved, and a better HPLC separation is needed in order to assign charge states, with concomitant increase in HPLC separation time and thus increased deuterium back-exchange. Second, because of FT-ICR's ultrahigh mass measurement *accuracy*, most of the proteolytic fragments of a protein can be identified by their accurate masses, thereby reducing the need for MS/MS or other partial sequencing of each fragment. Because peptide identification is the limiting factor (in both time and sample consumption) (Smith et al., 1997), FT-ICR MS greatly reduces the experimental time and effort, as well as sample consumption, compared to other mass analyzers.

Materials and methods

Materials

Sephadex G-15 from Pharmacia Biotech; D_2O and pepsin from Sigma Chemical Co. were used without further purification. $^{15}NH_4Cl$ was from Cambridge Isotope Labs. All other chemicals were of reagent grade or higher.

Protein preparation

FKBP was mutated to remove the single cysteine (C22A) by use of a mega-primer PCR protocol (Barettino et al., 1994; M. Li & T.M. Logan, unpubl.). Mutant C22A FKBP was overexpressed in *Escherichia coli* strain JM109 by use of a di-cistronic vector described previously (Pilot-Mathias et al., 1993). The bacteria were grown in LB medium containing 50 mg/L ampicillin and induced with 1 mM isopropyl thiogalactose (IPTG) at 0.6 OD. Protein expression was continued for 8–10 h after induction before harvesting by centrifugation. The cell paste was suspended in lysis buffer (50 mM Na phosphate, pH 7.4, containing 10 mM EDTA and 1 mM phenylmethyl sulfonyl fluoride), and broken by passage through a French press. Cellular debris was removed by centrifugation (10,000 rpm for 20 min in a Beckman JA20 rotor) and loaded onto a DE52 column (2.5 × 10 cm), pre-equilibrated in elution buffer (10 mM HEPES, pH 7.5). At this pH, FKBP is not retained on the column and elutes as the first peak (Holzman et al., 1991). Fractions containing FKBP were pooled, and brought to 90% saturation with ammonium sulfate. The resulting precipitate was pelleted by centrifugation, re-dissolved in 2 mL elution buffer, and applied to a gel filtration column (G75, 2.5 × 120 cm). FKBP-containing fractions were pooled, concentrated using a Centriprep 10 (Amicon), and further purified by reverse-phase HPLC (C4, Vydac) on a Shimadzu HPLC system. The 4.6 mm × 250 mm column was developed at a flow rate of 1.0 mL/min with a linear gradient (10–80%) of solvent B (acetonitrile containing 0.1% TFA) in solvent A (water containing 0.1%TFA) in 30 min, with two LC-10AD solvent delivery modules. FKBP eluted at 37% acetonitrile. C22A uniformly enriched in ^{15}N was isolated as described above from bacteria grown in M9 minimal medium containing 4 gm/L glucose and 1 g/L $^{15}NH_4Cl$.

Hydrogen exchange determined by NMR

H/D exchange of ^{15}N-labeled C22A-FKBP was initiated by rapidly replacing H_2O by D_2O (10 mM KH_2PO_4/K_2HPO_4) by use of a centrifuged Sephadex G-15 gel filtration column. The p^2H value read directly from the pH meter was pD = 7.07. Two-dimensional HSQC spectra (Stonehouse et al., 1994) were collected at 16.9 T on a Varian Unityplus console with a minimal two-step phase cycle (Marion et al., 1989). The data were collected as 64 and 1,024 complex points with sweep widths of 1,650 and 10,000 Hz for ^{15}N and ^{1}H, respectively), for a total experiment time of approximately 8.5 min (for two scans). HSQC spectra were collected continuously for the first 72 min after initiating exchange, and then the experiment time was doubled (by increasing the number of scans) and several additional HSQC spectra collected. This process was repeated over a period of ~50 h. Changes in the signal-to-noise ratio due to increased number of scans was taken into account when comparing the integrated resonances. Spectra were processed on Silicon Graphics computers with nmrPipe (Delaglio et al., 1995), and analyzed with NMRView (Johnson & Blevins, 1994).

Hydrogen exchange determined by mass spectrometry

H/D exchange of C22A-FKBP was initiated by rapidly replacing H_2O by D_2O (0.01 M KH_2PO_4/K_2HPO_4) by use of a centrifuging Sephadex G-15 gel filtration column (1 × 3 cm). The p^2H reading was 7.04. The protein/D_2O solution (~100 μM) was incubated at 25 °C for each of several exchange periods and then combined with a fivefold excess of 0.1 M H_3PO_4/KH_2PO_4 buffer, pH 2.4, to decrease the pH to 2.4, thereby quenching the exchange reaction. The samples were stored at −70 °C until analysis.

To adjust for deuterium gain or loss under quenched conditions, we prepared two control samples (Zhang & Smith, 1993). A zero-deuteration control was prepared by diluting the protiated protein solution directly into a solution so as to make the final solution composition the same as the quenched sample. A full-deuteration control was prepared by incubating the protein/D_2O solution at 40 °C for 50 h in 4 M urea. Both control samples were analyzed as for the other samples.

For determination of the global hydrogen exchange behavior of C22A-FKBP, each sample was thawed and analyzed by electrospray FT-ICR mass spectrometry with on-line desalting. For determination of the extent of deuterium incorporation into different segments of the protein primary sequence, deuterated protein was thawed and digested for 3 min at 0 °C with pepsin (1:1 substrate:enzyme ratio). Each peptide mixture was then analyzed by on-line LC/MS, and the extent and distribution of deuterium incorporation was determined from their ESI FT-ICR mass spectra.

The LC/MS system was constructed from two Shimadzu LC-10AD pumps. For on-line desalting of protein and peptides, a reverse-phase C-8 microbore guard column (15 × 1 mm, Microtech) was used, with a total flow rate of 40 μL/min. After sample injection, the column was first desalted by washing with aqueous solvent for 1.5 min, then washed with 60% acetonitrile to elute the proteins or peptides. The total analysis period for each sample was ~2.5 min. For on-line LC/MS separation and analysis of the proteolytic peptides, a capillary perfusion column (50 × 0.3 mm, LC Packings) was used, at a flow rate of 40 μL/min. After sample injection, the column was first washed with aqueous solvent for 1 min and the peptides were then eluted with a gradient (0–60% acetonitrile over 5 min) in which both solvents contained 0.9% formic acid. The perfusion column decreases the separation time and thus reduces deuterium loss during separation. The complete separation period was typically less than 4 min. To reduce deuterium back-exchange, both columns were immersed in ice water during analysis. Because protiated solvents were used for both on-line desalting and HPLC separation, all fast-exchanging hydrogens on the side chains of the protein are washed away. Thus, the deuterium contents measured from the mass spectra constitute a direct measure of deuterium incorporated into the peptide backbone amide linkages.

Mass spectrometric analyses were performed with a homebuilt FT-ICR mass spectrometer, with a shielded 9.4 Tesla superconductive magnet, and equipped with a homebuilt external electrospray interface described elsewhere (Senko et al., 1996). An Odyssey™ data system (Finnigan FTMS, Madison, Wisconsin) was used to acquire data and process the data into a peak list. The peak list was further analyzed with software written by the author to obtain the extent of deuterium in the protein or peptides.

Identification of proteolytic fragments

C22A-FKBP was digested with pepsin (~1:1 substrate:enzyme weight ratio) at pH 2.4 and 0 °C for 3 min, then subjected to

FT-ICR MS analysis with on-line desalting. Mass resolving power, $m/\Delta m_{50\%}$, in which $\Delta m_{50\%}$ is the mass-domain peak full width at half-maximum peak height, was between 50,000 and 120,000, depending on the peptide mass-to-charge ratio, m/z (Marshall et al., 1979). Identification of some peptides could be achieved from their measured molecular weights (the accuracy was usually within 100 ppm based on external calibration) and the known specificity of pepsin (Powers et al., 1977; Smith et al., 1997), as well as their MS/MS spectra generated by infrared multiphoton dissociation (IRMPD) with a 40 W continuous-wave CO_2 laser (Little et al., 1994). Once these peptides were identified, very accurate (usually to within less than 10 ppm) molecular weights of other peptides could be determined by internal mass calibration based on the masses of the identified peptides. Thus, most of the proteolytic peptides could be identified by their accurate masses. Some of them were further confirmed by IRMPD. A given parent ion was isolated selectively by stored-waveform inverse Fourier transform (SWIFT) broad-band ejection of ions of all other m/z values (Marshall et al., 1985; Guan & Marshall, 1996) and was fragmented by infrared irradiation at 30–40 W for 400–1,000 ms.

Hydrogen exchange rate constant distribution by maximum entropy method

Different exchangeable hydrogens in the protein or segment of the protein generally have different exchange-in rates. Hydrogen exchange behavior for a protein or any one of its segments may thus be characterized by a distribution in exchange rate constants. This section describes an algorithm for extraction of the rate constant distribution by use of MEM. Although the algorithm is used here to derive a distribution of hydrogen exchange rate constants, it can be used to derive the rate constant distribution for any reaction system consisting of several parallel first-order reactions.

MEM is a statistical method that gives the most-possible solution based on the available experimental data and noise. Entropy is a measure of amount of uncertainty in a probability distribution. If there are no "constraints" for the system (i.e., no information is known about the system), then the statistically most probable solution is that each event has equal probability: namely, a solution with maximized entropy (in our case, a "flat" distribution).

MEM finds the highest-entropy solution subject to the error range of the original experimental data. By maximizing entropy, MEM gives the most impartial estimate of the probabilities of the events, based on the available experimental data, without any other assumptions. Thus, MEM avoids overinterpreting the data and gives the most unbiased result. Another characteristic of the MEM-based result is that it contains the least noise, because random noise is minimized in the entropy maximization process. The maximum entropy method described here is similar to that from a previous paper that used MEM to deconvolve the natural isotope abundance from a mass spectrum (Zhang et al., 1997).

For a system with N exchangeable hydrogens, hydrogen exchange kinetics may be described as

$$D = N - \sum_{i=1}^{N} \exp(-k_i t), \tag{2}$$

in which D is the number of deuteriums incorporated into the protein or one of its segments, k_i is the rate constant for deuterium incorporation at the ith (of N possible) sites, and t is the reaction period.

Equation 2 may also be written as

$$D = N - \int_0^{\infty} \exp(-kt) p'_k \, dk, \tag{3}$$

in which k is the rate constant, $p'_k dk$ is the relative abundance of hydrogens whose rate constants fall between k and $k + dk$. $p'_k dk$ may also be thought of as the probability that a given hydrogen has an exchange rate constant between k and $k + dk$. Because k usually covers a wide range (e.g., many orders of magnitude for protein amide hydrogen exchange rate constants), it is convenient to represent k on a logarithmic scale. Equation 3 then becomes

$$D = N - \int_{-\infty}^{\infty} \exp(-kt) p_k \, d(\log k), \tag{4}$$

or, more practically,

$$D = N - \sum_{k=k_1}^{k_2} \exp(-kt) p_k \Delta \log k, \tag{5}$$

in which $p_k \Delta \log k$ is the probability that the logarithm of the rate constant lies between $\log k$ and $\log k + \Delta \log k$, or (equivalently) the relative abundance of hydrogens whose (logarithmically scaled) rate constants range from $\log k$ to $\log k + \Delta \log k$. k_1 and k_2 denote the lowest and highest rate constants to be included (usually limited by the experimentally accessible range). Let $f_k = p_k \Delta \log k$. Equation 5 then becomes

$$D = N - \sum_{k=k_1}^{k_2} f_k \exp(-kt), \tag{6}$$

in which f_k is the probability that reaction system exhibits (logarithmically scaled) rate constants between $\log k$ and $\log k + \Delta \log k$.

Here, we follow Skilling's definition of entropy (Skilling & Bryan, 1984)

$$S = -\sum_{k=k_1}^{k_2} f_k \left[\ln\left(\frac{f_k}{A}\right) - 1 \right], \tag{7}$$

in which f_k is the probability for event k to occur (in this context, event k denotes that the logarithm of the rate constant lies between $\log k$ and $\log k + \Delta \log k$) and A is a user-defined parameter (see below). The first derivative of the entropy is

$$\frac{\partial S}{\partial f_k} = \ln(A) - \ln(f_k). \tag{8}$$

Thus, if all f_k are equal (namely, a flat distribution) with (common) magnitude, A, the entropy is at its maximum. Therefore, when S is maximized, A is the most probable value of f_k when no other constraints are available. Here, A is defined as the average abundance, i.e., $A = N/N_k$, in which N is the number of amide hydrogens and N_k is the number of increments (typically 50) in rate constant that define the resolution of the rate constant distribution. In (many) cases for which enough constraints are available, value of A is not critical. (Note that the probabilities are automatically positive because the entropy does not exist if any f_k value is negative.)

Next, we seek to maximize S. As described above, without constraints, S is maximized when all f_k are equal. That is one particular distribution with maximal entropy. However, we also need to fit the distribution to the experimental data to within experimental error. That is to say, we want to maximize S subject to a constraint, C. The most often used constraint is χ^2, according to

$$C = \chi^2 = \sum_t \frac{(D_t^{calc} - D_t^{exp})^2}{\sigma_t^2}, \tag{9}$$

in which t is the set of times at which the extent of deuterium incorporation is measured. D_t^{calc} is the reaction product deuterium content at time, t, calculated from Equation 6; D_t^{exp} is the reaction product deuterium content observed experimentally after time, t; and σ_t is the standard deviation of reaction product deuterium content at time, t.

The goal of MEM is to find a solution that exhibits the maximum possible entropy *and* has a C value within a specified range. If the unconstrained maximum entropy solution is the same as the constrained solution, then all of the rate constants have the same probability, which usually means the data are too noisy to interpret. Otherwise, the MEM solution will always lie on the boundary of the constraint C (Skilling & Bryan, 1984; Zhang et al., 1997). In all cases described in this paper, a 99% confidence level in the χ^2 distribution served as the constraint.

The method most often used to find the maximum, S, subject to a constraint starts from a so-called Lagrange multiplier, λ ($\lambda > 0$), used to form another function, Q,

$$Q = S - \lambda C. \tag{10}$$

One then maximizes Q by varying λ until a user-defined C-value is found. MEM thus becomes a multidimensional function maximization problem. By maximizing Q, we are actually maximizing S and minimizing C at the same time. λ can be thought of as a "weight factor" between the two. A set of f_k values is obtained by maximizing Q, and p_k can be calculated from $p_k = f_k/\Delta \log k$. A plot of p_k versus $\log k$ yields the rate constant distribution presented in Results. Note that because D approaches zero as t approaches zero, Equation 6 requires that

$$\sum_{k=k_1}^{k_2} p_k \Delta \log k = \sum_{k=k_1}^{k_2} f_k = N. \tag{11}$$

That is, for a plot of p_k versus $\log k$, the area under the curve should equal the total number of first-order rate constants N (total number of amide hydrogens).

The method of choice for multidimensional maximization depends on the size and complexity of the problem (Press et al., 1992). The second derivative of Q can be written as

$$\frac{\partial^2 Q}{\partial f_k^2} = \frac{\partial^2 S}{\partial f_k^2} - \lambda \frac{\partial^2 C}{\partial f_k^2}. \tag{12}$$

From Equation 8, we have

$$\frac{\partial^2 S}{\partial f_k^2} = -\frac{1}{f_k}. \tag{13}$$

From Equation 9,

$$\frac{\partial C}{\partial f_k} = \sum_t \frac{2(D_t^{calc} - D_t^{exp})}{\sigma_t^2} \frac{\partial D_t^{calc}}{\partial f_k}. \tag{14}$$

D_t^{calc} is calculated from Equation 6:

$$\frac{\partial D_t^{calc}}{\partial f_k} = -\exp(-kt). \tag{15}$$

Substituting Equation 15 into 14 yields

$$\frac{\partial C}{\partial f_k} = -\sum_t \frac{2(D_t^{calc} - D_t^{exp})}{\sigma_t^2} \exp(-kt). \tag{16}$$

Taking the derivative of Equation 16, and using Equation 15 again, we have

$$\frac{\partial^2 C}{\partial f_k^2} = \sum_t \frac{2[\exp(-kt)]^2}{\sigma_t^2}. \tag{17}$$

Substituting Equation 13 and 17 into Equation 12, we obtain the second derivative of Q as

$$\frac{\partial^2 Q}{\partial f_k^2} = -\frac{1}{f_k} - \lambda \sum_t \frac{2[\exp(-kt)]^2}{\sigma_t^2}. \tag{18}$$

One can see that the second derivative of Q is always negative, which means that the function has only one maximum. For the present application, there are no local maxima, and the size of the data set is relatively small; ergo, simple methods such as Powell's method, conjugate gradient method, and the variable metric method are suitable (Press et al., 1992). Most of the results discussed below were performed by the variable metric method.

A program based on the above described methodology was written in ANSI C under LabWindows/CVI (National Instruments, Austin, Texas) for Windows. The function maximization routines, including Powell's method, conjugate gradient method, and variable metric method, are from *Numerical Recipes in C, the Art of Scientific Computing* (Press et al., 1992). Most MEM processes could be completed within 2 s, for the variable metric method running on a 200 MHz Pentium Pro computer. The program is available from the author.

In MEM, standard deviations for different time points are needed. For deuterium exchange of intact C22A FKBP, the standard deviations at different time points were obtained from multiple experiments. For deuterium exchange of different segments of the protein, multiple experiments were done for one time point, and it was found that the standard deviations for different peptides are typically about 1% of the total number of amide hydrogens (N) in the peptide if the MS peaks have reasonable signal-to-noise ratio (i.e., mass measurement is not the limiting factor). Thus, 1% of N was used as the standard deviation for each of the time points for MEM analyses of protein segments.

Acknowledgments

We thank C.L. Hendrickson and J.P. Quinn for helpful discussions and for their continuing efforts in modifying and developing the instrumentation used in the present experiments. This work was supported by NSF (CHE-94-13008), NIH (GM-54035), NIH (GM-31683), Florida State University, and the National High Magnetic Field Laboratory in Tallahassee, Florida.

References

Afeyan NB, Fulton SP, Gordon NF, Mazsaroff I, Varady L, Regnier FE. 1990a. Perfusion chromatography: An approach to purifying biomolecules. *Bio/Technology* 8:203–206.

Afeyan NB, Fulton SP, Regnier FE. 1991. Perfusion chromatography packing materials for proteins and peptides. *J Chromatogr 544*:267–279.

Afeyan NB, Gordon NF, Mazsaroff I, Varady L, Fulton SP, Yang YB, Regnier FE. 1990b. Flow-through particles for the high-performance liquid chromatographic separation of biomolecules: Perfusion chromatography. *J Chromatogr 519*:1–29.

Bai Y, Milne JS, Mayne L, Englander SW. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct Funct Genet 17*:75–86.

Barettino D, Feigenbutz M, Valcarcel R, Stunnenberg HG. 1994. Improved method for PCR-mediated site-directed mutagenesis. *Nucleic Acids Res 22*:541–542.

Buchanan MV, Hettich RL. 1993. Characterization of large biomolecules by Fourier transform mass spectrometry. *Anal Chem 65*:245A–259A.

Comisarow MB, Marshall AG. 1974. Fourier transform ion cyclotron resonance spectroscopy. *Chem Phys Lett 25*:282–283.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: A multidimensional spectra processing system based on UNIX pipes. *J Biomol NMR 6*:277–293.

Dharmasiri K, Smith DL. 1996. Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins. *Anal Chem 68*:2340–2344.

Englander JJ, Rogero JR, Englander SW. 1985. Protein hydrogen exchange studied by the fragment separation method. *Anal Biochem 147*:234–244.

Englander SW, Englander JJ, McKinnie RE, Ackers GK, Turner GJ, Westrick JA, Gill SJ. 1992. Hydrogen exchange measurement of the free energy of structural and allosteric change in hemoglobin. *Science 256*:1684–1687.

Englander SW, Kallenbach NR. 1984. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quart Rev Biophys 16*:521–655.

Englander SW, Sosnick TR, Englander JJ, Mayne L. 1996. Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol 6*:18–23.

Fruman DA, Burakoff SJ, Beirer BE. 1994. Immunophilins in protein folding and immunosuppression. *FASEB J 8*:391–400.

Gregory RB. 1983. Comparison of analytically and numerically derived hydrogen exchange-rate distribution functions. *Biopolymers 22*:895–909.

Gregory RB, Rosenberg A. 1986. Protein conformational dynamics measured by hydrogen isotopic exchange techniques. *Methods Enzymol 31*:448–508.

Guan S, Marshall AG. 1996. Stored waveform inverse Fourier transform (SWIFT) ion excitation in trapped-ion mass spectrometry: Theory and applications. *Int J Mass Spectrom Ion Processes 157/158*:5–37.

Holzman TL, Egan DA, Edalji R, Simmer RL, Helfrich R, Taylor A, Burres NS. 1991. Preliminary characterization of a cloned neutral isoelectric form of the human peptidyl prolyl isomerase cyclophilin. *J Biol Chem 226*:2474–2479.

Johnson BA, Blevins RA. 1994. NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR 4*:603–614.

Johnson RS. 1996. Mass spectrometric measurement of changes in protein hydrogen exchange rates that result from point mutations. *J Am Soc Mass Spectrom 7*:515–521.

Johnson RS, Walsh KA. 1994. Mass spectrometric measurement of protein amide hydrogen exchange rates of apo- and holo-myoglobin. *Protein Sci 3*:2411–2418.

Knox DG, Rosenberg M. 1980. Fluctuations of protein structure as expressed in the distribution of hydrogen exchange rate constants. *Biopolymers 19*:1049–1068.

Little DP, Speir JP, Senko MW, O'Cornnor PB, McLafferty FW. 1994. Infrared multiphoton dissociation of large multiply-charged ions for biomolecule sequencing. *Anal Chem 66*:2809–2815.

Liu Y, Smith DL. 1994. Probing high order structure of proteins by fast-atom bombardment mass spectrometry. *J Am Soc Mass Spectrom 5*:19–28.

Marion D, Ikura M, Tschudin R, Bax A. 1989. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J Magn Reson 85*:393–399.

Marshall AG, Comisarow MB, Parisod G. 1979. Relaxation and spectral line shape in Fourier transform ion cyclotron resonance spectroscopy. *J Chem Phys 71*:4434–4444.

Marshall AG, Schweikhard L. 1992. FT/ICR/MS: Technique developments. *Int J Mass Spectrom Ion Processes 118/119*:37–70.

Marshall AG, Senko MW, Li W, Li M, Dillon S, Guan S, Logan TM. 1997. Protein molecular weight to 1 Da by ^{13}C, ^{15}N double-depletion and FT-ICR mass spectrometry. *J Am Chem Soc 119*:433–434.

Marshall AG, Wang T-CL, Ricca TL. 1985. Tailored excitation for Fourier transform ion cyclotron resonance mass spectrometry. *J Am Chem Soc 107*:7893–7897.

McLafferty FW. 1994. High-resolution tandem FT mass spectrometry above 10 kDa. *Acc Chem Res 27*:379–386.

Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. 1993. Detection of transient protein folding populations by mass spectrometry. *Science 262*:896–900.

Muhandiram R, Kay LE. 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J Magn Reson 103B*:203–216.

Park ST, Aldape RA, Futer O, DeCenzo MT, Livingston DJ. 1992. PPIase catalysis by human FT506-binding protein proceeds through a conformational twist mechanism. *J Biol Chem 267*:3316–3324.

Pilot-Mathias T, Pratt SD, Lane BC. 1993. High-level synthesis of the 12-kDa Human FK506-binding protein in *Escherichia coli* using translational coupling. *Gene 128*:219–225.

Powers JC, Harley AD, Myers DV. 1977. Substrate specificity of porcine pepsin. *Adv Exp Med Biol 95*:141–57.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992. *Numerical recipes in C. The art of scientific computing, 2nd ed.* Cambridge, England: Cambridge University Press.

Rosa JJ, Richards FM. 1979. An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on proteins: Application to ribonuclease S peptide. *J Mol Biol 133*:399–416.

Senko MW, Hendrickson CL, Pasa-Tolic L, Marto JA, White FM, Guan S, Marshall AG. 1996. Electrospray ionization FT-ICR mass spectrometry at 9.4 Tesla. *Rapid Commun Mass Spectrom 10*:1824–1828.

Skilling J, Bryan RK. 1984. Maximum entropy image reconstruction general algorithm. *Mon Not R Astr Soc 211*:111–124.

Smith DL, Deng Y, Zhang Z. 1997. Probing the noncovalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom 32*:135–146.

Smith DL, Zhang Z. 1994. Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrom Rev 13*:411–429.

Speir JP, Gorman GS, Amster IJ. 1992. Laser desorption, chemical ionization, and laser desorption/chemical ionization applications with FTMS. In: Gross ML. ed. *Mass spectrometry in the biological sciences: A tutorial.* Dordrecht, The Netherlands: Kluwer Academic Publishers. pp 199–212.

Stonehouse J, Shaw GL, Keeler J, Laue ED. 1994. Minimizing sensitivity losses in gradient-selected ^{15}N-^{1}H HSQC spectra of proteins. *J Magn Reson 107A*:178–184.

van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. 1991. Atomic Structure of FKBP-FK506, an Immunophilin-Immunosuppressant Complex. *Science 252*:839–842.

Wilkins CL. 1994. Fourier transform mass spectrometry. *Trends Anal Chem 13* (Special issue):223–251.

Woodward C, Simon I, Tuchsen E. 1982. Hydrogen exchange and the dynamic structure of proteins. *Mol Cell Biochem 48*:135–60.

Wu Q, Van Orden S, Cheng X, Bakhtiar R, Smith RD. 1995. Characterization of cytochrome *c* variants with high-resolution FTICR mass spectrometry: Correlation of fragmentation and structure. *Anal Chem 67*:2498–2509.

Zhang Z, Guan S, Marshall AG. 1997. Enhancement of the effective resolution of mass spectra of high-mass biomolecules by maximum-entropy based deconvolution to eliminate the isotopic natural abundance distribution. *J Am Soc Mass Spectrom 8*:659–670.

Zhang Z, Post CB, Smith DL. 1996. Amide hydrogen exchange determined by mass spectrometry: Application to rabbit muscle aldolase. *Biochemistry 35*:779–791.

Zhang Z, Smith DL. 1993. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci 2*:522–531.

Zhang Z, Smith DL. 1996. Detection of thermal-induced unfolding domains in aldolase by amide hydrogen exchange and mass spectrometry. *Protein Sci 5*:1282–1289.