

Screening Combinatorial Libraries of de Novo Proteins by Hydrogen–Deuterium Exchange and Electrospray Mass Spectrometry

Daniel M. Rosenbaum,[†] Sushmita Roy,[‡] and Michael H. Hecht*

Contribution from the Department of Chemistry, Princeton University, Princeton, New Jersey 08544-1009

Received June 2, 1999

Abstract: Combinatorial methods have emerged as valuable tools for the discovery of proteins, nucleic acids, and small molecules with novel structures and properties. While combinatorial methods can generate de novo proteins with native-like properties, finding such proteins in libraries containing an abundance of non-native structures has proved difficult and tedious. To overcome these difficulties, we developed a rapid screen for native-like properties. The screen uses electrospray mass spectrometry (ESMS) to monitor the hydrogen–deuterium (H–D) exchange kinetics in semicrude samples of de novo proteins expressed in *Escherichia coli*. To demonstrate the utility of the approach, we screened two libraries of de novo sequences and identified proteins whose amide protons were protected from exchange with solvent. The results of the screen correlate well with orthogonal methods for detecting native-like structures. As protection of amide protons from exchange is a hallmark of well-folded proteins, this screen can be used to identify native-like proteins from combinatorial libraries containing both native-like and molten globule-like structures. Moreover, since the screen can be applied to semicrude samples and does not require extensive protein purification, it can be used for medium throughput screening of large combinatorial libraries.

Introduction

De novo protein design has progressed to the point where achieving abundant secondary structure is no longer challenging. The major hurdle today is to devise novel proteins with structural and thermodynamic properties that are native-like rather than molten globule-like. Native protein structures typically contain not only secondary structure and a hydrophobic core, but also close packing of nonpolar side chains. Such packing confers a degree of structural rigidity to the macromolecule.¹ In contrast, the molten globule state is characterized as a dynamic ensemble with secondary structure and a loose hydrophobic core, but without unique packing.² Molten globules have been observed as protein folding intermediates, and some natural proteins can be induced to form equilibrium molten globules under conditions of extreme pH, mild denaturation, or absence of cofactors.³

Early efforts to design novel proteins typically yielded molten globules.⁴ More recently, however, several approaches have produced de novo proteins that appear to mimic the native state of natural proteins.⁵ Typically, these approaches have used “rational design” or computational methods to design a small number of novel sequences that are synthesized and characterized one at a time.

In contrast, we developed a combinatorial method for the design of large libraries of novel amino acid sequences.⁶ These

libraries are constructed using a “binary code” strategy in which the sequence locations of polar and nonpolar amino acids are specified explicitly, but the identities of these side chains are allowed to vary. Combinatorial diversity is made possible by the binary organization of the genetic code: Polar residues are encoded by the degenerate DNA codon NAN, while nonpolar residues are encoded by the degenerate DNA codon NTN (where N represents a specified mixture of DNA nucleotides A, G, T, and C). Initial work using the binary code strategy focused on the design of 4-helix bundles.⁶ More recently, we have extended this approach to design novel β -sheet structures.⁷

We showed previously that the majority of our designed α -helical sequences fold into collapsed structures dominated by α -helical secondary structure.^{6,8} Biophysical characterization of

* To whom correspondence should be addressed.

[†] Present address: Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge MA 02138.

[‡] Present address: The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037.

(1) Richards, F. M.; Lim, W. A. *Q. Rev. Biophys.* **1994**, *26*, 423–498.
(2) (a) Alexandrescu, A. T.; Evans, P. A.; Pitkeathly, M.; Baum, J.; Dobson, C. M. *Biochemistry* **1993**, *32*, 1707–1718. (b) Kay, M. S.; Baldwin, R. L. *Nat. Struct. Biol.* **1996**, *3*(5), 439–445.

(3) Kuwajima, K. *Proteins* **1989**, *6*, 87–103.

(4) DeGrado, W. F.; Raleigh, D. P.; Handel, T. *Curr. Opin. Struct. Biol.* **1991**, *1*, 984–993.

(5) (a) Munson, M.; Balasubramanian, S.; Fleming, K. G.; Nagi, A. D.; O'Brien, R.; Sturtevant, J. M.; Regan, L. *Protein Sci.* **1996**, *5*, 1584–1593. (b) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342–345. (c) Betz, S. F.; Liebman, P. A.; DeGrado, W. F. *Biochemistry* **1997**, *36*, 2450–2458. (d) Lazar, G. A.; Desjarlais, J. R.; Handel, T. M. *Protein Sci.* **1997**, *6*, 1167–1178. (e) Fezoui, Y.; Connolly, P. J.; Osterhout, J. J. *Protein Sci.* **1997**, *6*, 1869–1877. (f) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82–86. (g) Schafmeister, C. E.; LaPorte, S. L.; Miercke, L. J.; Stroud, R. M. *Nat. Struct. Biol.* **1997**, *4*(12), 1039–1045. (h) Roy, S.; Ratnaswamy, G.; Boice, J. A.; Fairman, F.; McLendon, G.; Hecht, M. H. *J. Am. Chem. Soc.* **1997**, *119*, 5302–5306. (i) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. *Protein Sci.* **1998**, *7*(6), 1404–1414. (j) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. *Science* **1998**, *282*, 1462–1467. (k) Walsh, S. T.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5486–5491. (l) Schenck, H. L.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4869–4870. (m) Skalicky, J. J.; Gibney, B. R.; Rabanal, F.; Bieber Urbauer, R. J.; Dutton, P. L.; Wand, A. J.; *J. Am. Chem. Soc.* **1999**, *121*, 4941–4955. (n) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1999**, *121*, 4952–4960.

(6) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. *Science* **1993**, *262*, 1680–1685.

(7) West, M. W.; Wang, W.; Patterson, J.; Mancias, J. D.; Beasley, J. R.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, in press.

(8) Roy, S. Ph.D. Thesis, Princeton University, 1998.

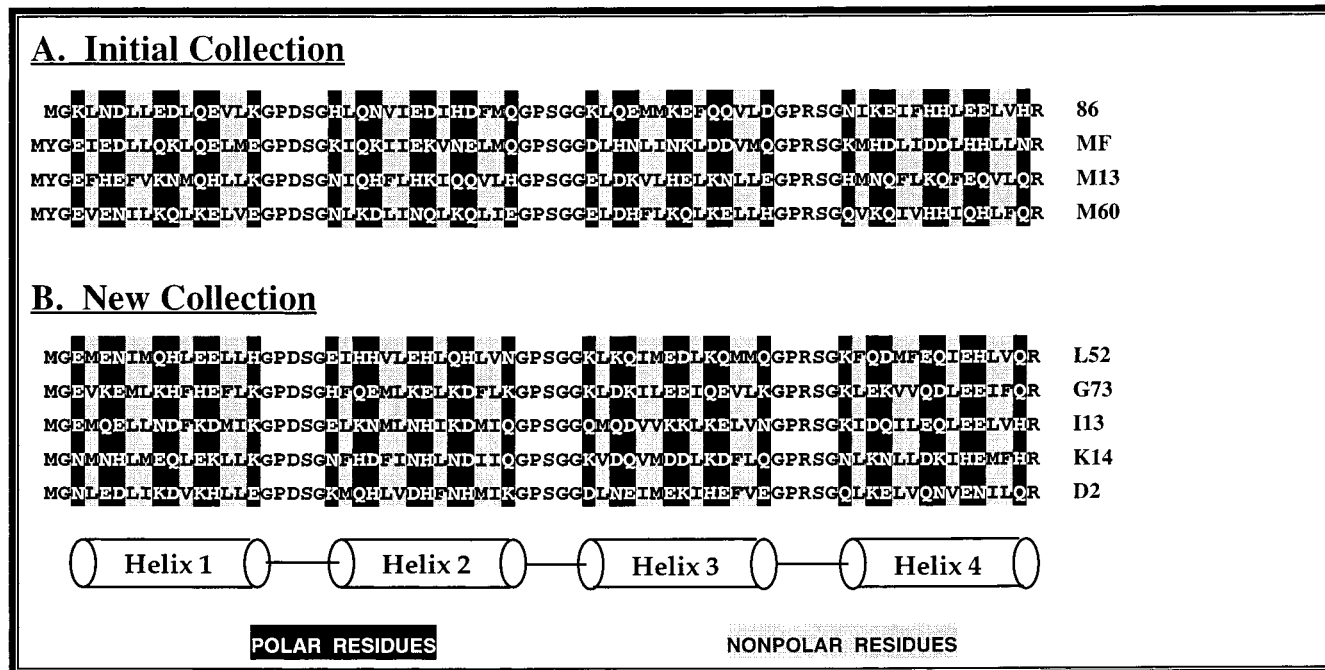


Figure 1. Sequences of de novo proteins designed by binary patterning of polar and nonpolar amino acids. Amino acid sequences are listed using the single letter code. Combinatorially varied α -helices are shaded with polar residues shown as white letters on black background, and nonpolar residues as black letters on gray background. Interhelical turn sequences were held constant. Sequences were confirmed by electrospray mass spectrometry. “Initial Collection” refers to the library of sequences reported by Kamtekar et al.⁶ “New Collection” refers to an expanded library of ~500 sequences¹¹ based upon the same binary code patterning.

an individual protein from this collection showed that it recapitulated many properties of native proteins.^{5h}

How frequently do such native-like structures occur in binary code libraries? While it seems reasonable that binary patterning of polar and nonpolar amino acids can generate amphiphilic secondary structures that collapse into tertiary structures with hydrophobic cores, it is not clear how frequently native-like molecules will occur in libraries of proteins for which interior packing is not designed a priori. Assessing the frequency of native-like proteins in combinatorial libraries requires a screen for some property that can serve as a “signature” of native-like structures.

The traditional signatures of native-like structures are cooperative chemical and thermal denaturation, lack of nonspecific aggregation, NMR chemical shift dispersion, and protection from amide proton exchange.⁹ For individual sequences derived from rational or computational design, one can readily isolate each protein and analyze it for this full range of properties. However, for the large number of sequences generated by combinatorial methods, extensive purification and full characterization is not practical. Instead, it would be more productive to prepare samples that do not require arduous purification procedures, and then screen many such samples for a single property that correlates well with native-like structures. Here, we describe such a screen. The screen measures protection from amide proton exchange by using electrospray mass spectrometry to monitor the time course of bulk hydrogen–deuterium exchange in crude (unpurified) samples of proteins derived from combinatorial libraries of de novo sequences. The observed exchange profiles correlate well with the results of orthogonal methods for detecting native-like structures, thereby demonstrating that the screen can distinguish native-like proteins from less stable structures.

Results

Protein Samples. The proteins used in this study were obtained from two combinatorial libraries of binary code sequences. Proteins M60, MF, M13, and 86 are derived from the initial collection of 29 sequences reported by Kamtekar et al.⁶ [The “M” prefix indicates the original sequence was modified by insertion of a tyrosine following the initiator methionine¹⁰]. Proteins G73, L52, I13, D2, and K14 were taken from a larger collection of several hundred sequences, which has not been published and was characterized only minimally.¹¹ Amino acid sequences of the proteins are shown in Figure 1. All sequences in both collections are constrained by the binary pattern of polar and nonpolar residues designed to favor the formation of amphiphilic helices. Yet, as specified by the combinatorial nature of the design, there is substantial diversity of amino acid identities at most positions in the α -helices. No pair of these proteins has more than 30% sequence identity.

The protein samples used in this study were not purified. Instead, following expression in *E. coli*, a rapid freeze–thaw procedure¹² was used to enrich the desired protein relative to bacterial contaminants. The resulting semicrude samples contain ~50–60% de novo protein in a background of cellular contaminants. The gel in Figure 2 shows several of the concentrated freeze–thaw samples prior to their dilution into D₂O (the binary code proteins run at the bottom of the gel). Judging from the intensities of the Coomassie stained bands, the concentrations of recombinant protein in each lane are fairly similar. On the basis of gels of protein samples of known concentration, each recombinant protein is judged to be ~50–100 μ M. Aside from oligomerization effects, H–D exchange

(10) Xiong, H. Ph.D. Thesis, Princeton University, 1995.

(11) (a) Kamtekar, S. Ph.D. Thesis, Princeton University, 1995. (b) Mclean, J. Senior Thesis, Princeton University, 1997. (c) Helmer, K. Senior Thesis, Princeton University, 1997.

(12) Johnson, B. H.; Hecht M. H. *BioTechnology* **1994**, *12*, 1357–1360.

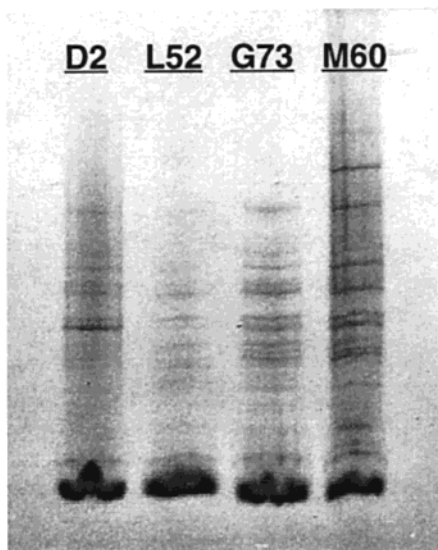


Figure 2. Polyacrylamide gel of crude samples (see Experimental Section) of four proteins characterized in this study. De novo protein (~8500 Da) is the dark band at the bottom of the gel. Equal volumes (10 μ L) of each sample were loaded.

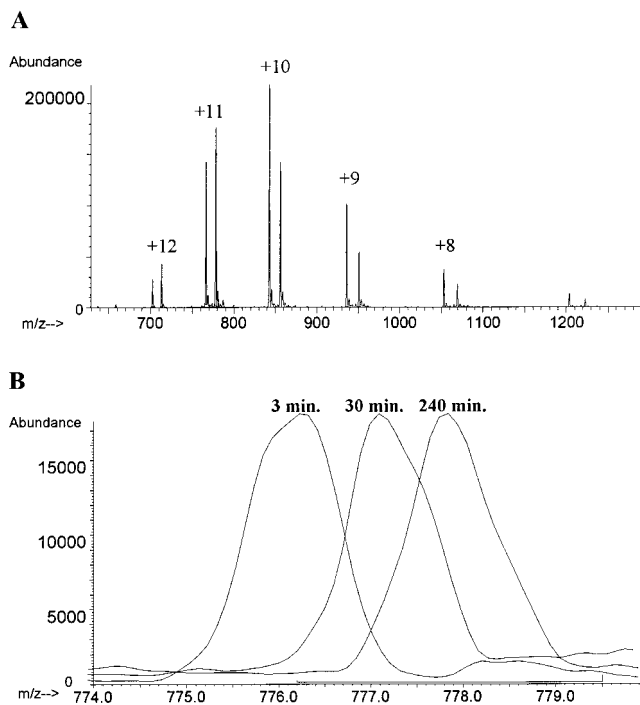


Figure 3. (A) Mass spectrum of protein G73 before initiation of exchange. Each charge state has two peaks: lower m/z is the $-Met$ species and higher m/z is the $+Met$ species. (B) Overlay of the $+11$ charge state peak ($-Met$) at different exchange times as indicated.

is not concentration dependent, so we did not attempt to equalize concentration. The concentration of impurities clearly varies from sample to sample.

H–D Exchange of de Novo Proteins Monitored by Electrospray Mass Spectrometry. Figure 3A shows the mass spectrum of protein G73 before exchange. Note the presence of doublet charge state peaks. There are clearly two species of similar mass in the crude sample of this protein. Deconvolution of the masses of the two sets of peaks showed that the two species are the intact sequence of G73 shown in Figure 1 (8550 amu) and a protein from which the initiator methionine has been cleaved (8419 amu). The presence of two populations is

expected because the protein is expressed in *E. coli*, and methionine aminopeptidase is known to cleave N-terminal Met residues in vivo. Exchange profiles of the $+Met$ and $-Met$ forms of G73 derived from the same set of mass spectra were almost identical, indicating that cleavage of the initiator Met has little effect on H/D exchange (not shown).

Figure 3B shows the $+11$ charge state peak of protein G73 migrating during the time course of the exchange. Indicated above each peak is the time elapsed after initiation of H–D exchange. At each time point, exchange was terminated by quenching the reaction at pH 2.5 and freezing at -80 $^{\circ}$ C. The peaks shown in Figure 3B are only ~ 2 m/z units wide. However, since the charge state is $+11$, the masses of the species giving rise to these signals actually span a range of 22 amu. This reflects the range of deuteration states and natural ^{13}C abundance. In the analyses of exchange rates described below, the average m/z for each peak was approximated as the m/z at maximum signal intensity.

Time Course of H–D Exchange for de Novo Proteins.

Figure 4A shows the exchange profiles for proteins from the initial library of Kamtekar et al.⁶ The number of protons protected is calculated by subtracting the observed mass from the mass expected from a fully exchanged sample. For all samples, fewer than 50 protons remain protected by the first time point (1 min). This is expected since labile side chain protons and fully exposed amide protons exchange very rapidly. Slowly exchanged protons are presumed to indicate protected backbone amides. There is a clear distinction between the exchange profile of protein M60 and those of the other three proteins. Whereas M13, MF, and 86 lose nearly all their protons within the first 20 min, M60 retains over 20 protected protons after 4 h. (The limit in all samples of ~ 7 –10 protons remaining unexchanged at long time points presumably results from back exchange with residual H_2O introduced by the original samples; see the Experimental Section.)

Figure 4A demonstrates clearly that protein M60 is more resistant to exchange than the other proteins in the initial collection. This result agrees with earlier work, which showed that M60 possesses numerous properties of native-like structures, including cooperative thermal denaturation and chemical shift dispersion in its NMR spectrum.^{5h} The correlation between the current results (Figure 4A) and those more detailed studies demonstrates that a rapid screen using mass spectrometry to monitor H/D exchange can distinguish native-like de novo proteins from those that resemble molten globules.

On the basis of the success of this screen in distinguishing M60 from less native-like structures, we next applied the method to screen a more recent library of proteins whose sequences have not been published, and which have been characterized only minimally.¹¹ Figure 4B shows the exchange profiles for five proteins from this new collection. As with the earlier library, one protein clearly stands out from the rest. Protein G73 has over 20 protons protected after 4 h. After 1 day of exchange, G73 retains 15 protons. After 1 week G73 reaches the baseline level of 7 protons. Purification of protein G73 and analysis of its 1-dimensional NMR spectrum showed considerably more chemical shift dispersion than the spectra of other proteins in this collection,⁸ thereby confirming the ability of the ESMS/H–D exchange screen to identify native-like structures.

Discussion

Several factors modulate H–D exchange rates in proteins. Most amide groups are involved in hydrogen-bonding interactions within elements of secondary structure. Such interactions

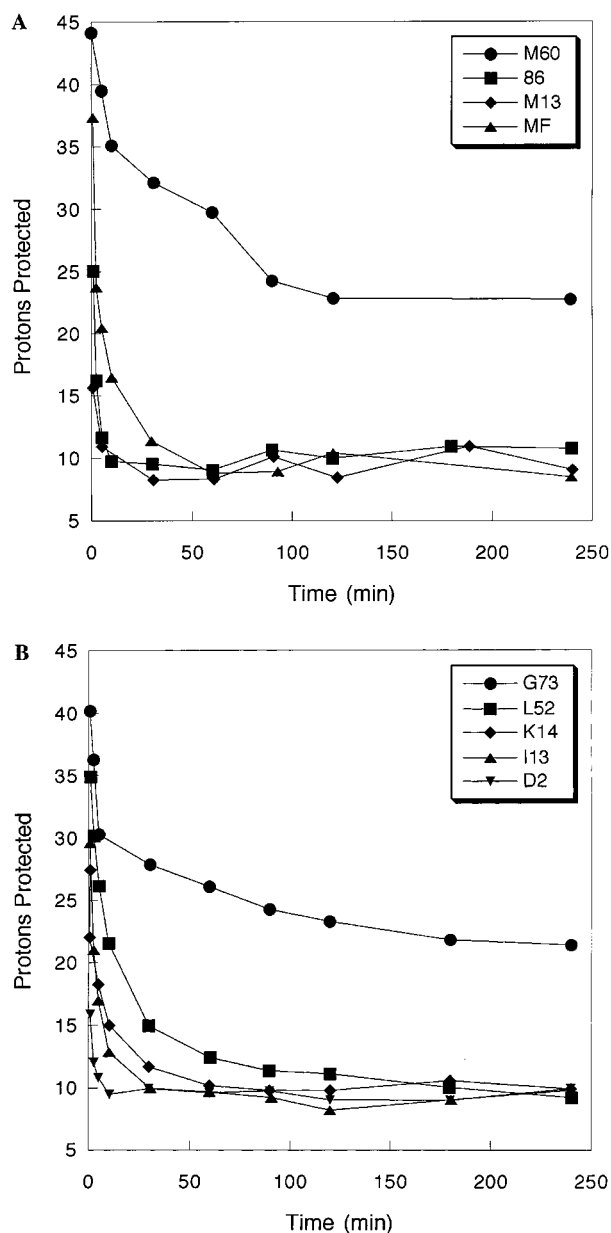


Figure 4. Hydrogen–deuterium exchange profiles of the proteins from the α -helical binary code libraries: (A) initial library⁶ and (B) expanded library.¹¹ The Y-axis shows the average number of protons protected as calculated from the fully exchanged mass minus the mass observed at a given time point during the course of the exchange. The X-axis refers to the time after initiation of exchange (at pH 5.7) when aliquots were quenched and frozen. After long time points, an average of ~ 7 –10 protons remain on the proteins. This equilibrium value is consistent with the fraction of protonated solvent in the exchange solutions (crude protein samples are initially in H₂O, then diluted into D₂O).

impede formation of H-bonds to solvent catalysts and thereby slow exchange. Amide protons can also be protected from exchange by burial in the hydrophobic interior of a protein. Such protons must become exposed through local or global unfolding events before exchange can occur.¹³

In the so-called EX2 limit, which is observed in most experiments, the unfolding rate is much faster than the inherent rate of amide proton transfer, so the observed rate depends on both the unfolding equilibrium and the rate of chemical exchange.¹³ In H–D exchange experiments at the EX2 limit

measured by NMR, it is possible to determine the exchange rate of a specific amide proton. The protection factor and the resulting Gibbs free energy of the unfolding equilibrium easily follow.¹⁴

Electrospray mass spectrometry can also be used to follow H–D exchange in proteins, as demonstrated initially by Katta and Chait.¹⁵ Since then ESMS/H–D exchange has been used for a number of applications in protein chemistry.¹⁶ These include quality control analysis of batch-to-batch variations in the production of recombinant protein drugs,¹⁷ and the detection of equilibrium molten globules,¹⁸ protein folding intermediates,¹⁹ chaperonin-assisted protein folding,²⁰ two-state unfolding transitions,²¹ and ligand-binding interactions.²²

ESMS/H–D exchange experiments typically follow the exchange of all labile protons simultaneously, and individual protection factors are usually not determined. Recent studies indicate that residue-specific protection factors can be determined by collision-induced mass spectrometry.²³ However, resolution of individual protons is not necessary for screening combinatorial libraries. Instead, we can interpret our ESMS data in terms of bulk exchange profiles. Proteins that have more protected protons in their bulk exchange profiles contain more amides with slower exchange rates, and therefore more peptide groups with unfolding equilibria favoring a folded state. Thus more bulk protection suggests structural rigidity common to the native state of natural proteins.

The ESMS method described here has four distinct benefits that facilitate moderate throughput screening: (i) The method is amenable to rapid processing; (ii) it can be applied to small amounts of protein; (iii) dilute samples can be used to minimize aggregation; and (iv) as shown in this study, even significant impurities are tolerated.

To establish the validity of the screen, it is important to demonstrate that the results correlate with other methods for distinguishing native-like structures. We previously developed a 1D proton NMR screen for the binary code library.^{11c,24} After collecting spectra for each protein, the de novo proteins were ranked according to dispersion of peaks in the amide region and the presence of upfield-shifted methyl peaks. Of all the proteins studied in the current report, only M60 and G73 showed both dispersion in the amide region and upfield-shifted peaks. As shown in Figure 4, these same two proteins are set apart from the others in the H–D exchange profiles observed in the ESMS screen. NMR dispersion and H–D exchange/ESMS agree that G73 and M60 are the most native-like proteins.

Twenty eight proteins from the initial collection of Kamtekar et al.⁶ have been purified by HPLC and subjected to thermal

(14) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins* **1994**, *20*, 4–14.

(15) (a) Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 214–217. (b) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1993**, *115*, 5(14), 6317–6321.

(16) Woodward, C. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 672.

(17) Ramanathan, R.; Gross, M. L.; Zielinski, W. L.; Layloff, T. P. *Anal. Chem.* **1997**, *69*, 5142–5145.

(18) Chung, E. W.; Nettleton, E. J.; Morgan, C. J.; Grob, M.; Miranker, A.; Radford, S. E.; Dobson, C. M.; Robinson, C. V. *Protein Sci.* **1997**, *6*, 1316–1324.

(19) Miranker, A.; Robinson, C. V.; Radford, S. E.; Aplin, R. T.; Dobson, C. M. *Science* **1993**, *262*, 896–900.

(20) Robinson, C. V.; Gross, M.; Eyles, S. J.; Ewbank, J. J.; Mayhew, M.; Hartl, F. U.; Dobson, C. M.; Radford, S. E. *Nature* **1994**, *372*, 646–651.

(21) Yi, Q.; Baker, D. *Protein Sci.* **1996**, *5*, 1060–1066.

(22) Wagner, D. S.; Anderegg, R. J. *Anal. Chem.* **1994**, *66*, 706–711.

(23) Deng, Y.; Pan, H.; Smith, D. L. *J. Am. Chem. Soc.* **1999**, *121*, 1966–1967.

(24) Roy, S.; Helmer, K. J.; Hecht, M. H. *Fold. Des.* **1997**, *2*(2), 89–92.

(13) Englander, S. W.; Kallenbach, N. R. *Q. Rev. Biophys.* **1984**, *16*(4), 521–655.

denaturation studies.²⁵ Of the four proteins from that collection characterized in this paper, M60 had the most cooperative melting profile, with $T_m = 55$ °C and $\Delta H(\text{van't Hoff}) = 24.5$ kcal/mol. Protein 86 underwent thermal denaturation with only slight cooperativity ($T_m = 42$ °C and $\Delta H = 13.5$ kcal/mol), while MF and M13 had noncooperative melts. Protein M60 is the only candidate in this group to display significant protection from exchange in Figure 4A.

The H–D exchange/ESMS screen has some limitations. One property of de novo proteins that can potentially bias H–D exchange data is oligomerization. If a protein is present primarily in dimeric or higher aggregation states, it is possible that amide protons will be protected from exchange by the oligomerization interface. All candidates studied here were subjected to gel filtration chromatography (not shown) to determine approximate molecular sizes, albeit at higher concentrations and with a different buffer (50 mM phosphate + 200 mM NaCl, pH 8) than that used for exchange. Protein M60 was observed in monomer and dimer states; G73 was primarily dimeric; MF and 86 were primarily monomeric; M13 populated both monomer and dimer states; I13 was monomeric; K14 was trimeric; L52 was present as dimer, trimer, and tetramer; and D2 was present as monomer, dimer and tetramer. Overall, it is clear that the presence of higher aggregation states does not distinguish the protected proteins (M60, G73) from those that are unprotected. Along with the correlation with other screens, this suggests that aggregation effects are not solely responsible for the observed differences in protection. It should be noted that biases introduced by aggregation would complicate any screen for native-like properties. Indeed, H–D exchange/ESMS is less likely to suffer from these biases because the experiments are performed with more dilute protein.

In conclusion, measurement of H–D exchange by ESMS has successfully distinguished native-like proteins from less stable structures. Proteins with the greatest degree of protection from exchange are the same as those previously found to exhibit the most native-like NMR spectra. The ESMS/H–D screen can be performed on small aliquots of relatively crude samples, and thus has the potential for large-scale screening of novel libraries.

Materials and Methods

Preparation of Crude Protein Samples. The de novo proteins were expressed under control of the T7 promoter as described previously.⁶ Samples were prepared according to a previously published freeze–thaw protocol.¹² This method relies on the release of recombinant protein from the bacterial cytoplasm upon disruption of the cell envelope by repeated cycles of freezing and thawing. Cultures of 250 mL of *E. coli*, each containing a recombinant plasmid, were grown to $OD_{600} = 0.8$. IPTG was added to 100 $\mu\text{g}/\text{mL}$ to induce expression and cultures were grown for an additional 3 h. Following centrifugation, the cell pellet was subjected to 3 cycles of freezing on dry ice/ethanol (10 min) and thawing at room temperature (10 min). After resuspension in 100 mM MgCl_2 , the cell debris was pelleted. The supernatant was acidified by addition of Acetate buffer (pH 4.1) to 50 mM, and impurities were precipitated. Finally, protein samples were exchanged into Milli-Q H_2O

(Millipore) and concentrated using a Centriprep 3 (Amicon). Aliquots of these “crude preps” were analyzed by ESMS. SDS-PAGE was performed with the Pharmacia Phastgel system. Gels were stained with Coomassie brilliant blue dye.

Hydrogen–Deuterium Exchange. Ultrapure (99.96%) D_2O was obtained from Sigma; CD_3COOD was from Cambridge Isotope Laboratories; NH_4OAc was from EM Science.

All exchange solutions were buffered in 5 mM Ammonium acetate/*d*₄-acetic acid. Sodium in the buffer was avoided because Na^+ ions interfere with mass-spectrometry analysis. A dilute buffer was used to limit back-exchange due to NH_4OAc . Before the initiation of exchange, 500 μL of ultrapure D_2O was mixed with 25 μL of buffer [100mM $\text{NH}_4\text{OAc}/\text{CD}_3\text{COOD}$ in D_2O , pH 5.7]. H–D exchange was initiated by diluting 25 μL of a crude protein sample into the buffered D_2O solution. During the initial dilution and subsequent extraction of aliquots, exposure to air was minimized. Ambient moisture can enter aqueous solutions, which increases the amount of protonated solvent and thus promotes back exchange. At different time points, a small volume of the exchanging solution was removed and added to 200 μL of quenching solution [1% CD_3COOD in D_2O , pH 2.5]. The size of each aliquot removed varied from protein to protein. For each protein tested, a test mass spectrum of a crude protein sample showed what dilution factor would provide a reasonable signal intensity. Typically 10–20 μL of exchanging solution was removed at each time point. The quenched aliquots were immediately flash-frozen in a dry ice/ethanol slurry and stored at -80 °C for later analysis by ESMS.

Electrospray Mass Spectrometry. After thawing, acidified protein samples were injected into a HP 5989 Electrospray MS Engine and the desolvated molecular ions were analyzed with a quadrupole mass filter. Electrospray instruments are normally operated with a carrier solvent including water and methanol. Methanol was eliminated from the carrier to avoid protein denaturation and D_2O was used instead of water to avoid back exchange. Each injection comprised of 50–100 quadrupole scans averaged over 1–2 min.

Analysis of Data. The sequences shown in Figure 1 were used to calculate expected values for the fully exchanged mass of each protein. Each data point in Figure 4 is derived from a mass spectrum averaged over at least three charge state peaks and over the injection time. To average the masses from different charge states, the mass of adduct ions must first be subtracted from the mass of each peak. Initial deprotection of over half of the exchangeable protons in our proteins was assumed to result from exchange of labile side chain protons. The adduct in our calculations was the deuteron, with mass 2.014 amu. The +Met and –Met forms of the proteins showed almost identical exchange profiles (not shown). In cases where the \pm Met species of a protein were both observed exchanging, only one of the species is shown in the exchange profiles of Figure 4.

After long time points, an average of ~ 7 –10 protons remain on the proteins. For example, after 1 week, G73 retains a baseline of 7 protons. This represents 5% of the 136 labile protons (including backbone and side chains) expected from the sequence of this protein. This equilibrium value of 5% is consistent with that expected from the initial mixing of a 25 μL sample with 525 μL of deuterated buffer (yielding a final H_2O content of 4.5%).

Acknowledgment. We thank Dorothy Little and John Eng for assistance with electrospray mass spectrometry. We also thank Michael Gross and Ragulan Ramanathan for helpful discussions, and for help with pilot experiments.

(25) Roy, S.; Hecht, M. H. Manuscript submitted for publication.