Published in final edited form as:
Proteins. 2009 September ; 76(4): 882–894. doi:10.1002/prot.22394.

Construct Optimization for Protein NMR Structure Analysis Using Amide Hydrogen / Deuterium Exchange Mass Spectrometry

Seema Sharma[1,2], **Haiyan Zheng**[1], **Yuanpeng J. Huang**[1,2], **Asli Ertekin**[1,2], **Yoshitomo Hamuro**[3], **Paolo Rossi**[1,2], **Roberto Tejero**[1,2], **Thomas B. Acton**[1,2], **Rong Xiao**[1,2], **Mei Jiang**[1,2], **Li Zhao**[1,2], **Li-Chung Ma**[1,2], **G. V. T. Swapna**[1,2], **James M. Aramini**[1,2,*], and **Gaetano T. Montelione**[1,2,4,*]

[1]Center for Advanced Biotechnology and Medicine, Rutgers University, 679 Hoes Lane, Piscataway, New Jersey 08854

[2]Northeast Structural Genomics Consortium, Rutgers University, Piscataway, New Jersey 08854

[3]ExSAR Corporation, Monmouth Junction, New Jersey 08852

[4]Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854

Abstract

Disordered or unstructured regions of proteins, while often very important biologically, can pose significant challenges for resonance assignment and three-dimensional structure determination of the ordered regions of proteins by NMR methods. In this paper, we demonstrate the application of $^{1}H/^{2}H$ exchange mass spectrometry (DXMS) for the rapid identification of disordered segments of proteins and design of protein constructs that are more suitable for structural analysis by NMR. In this benchmark study, DXMS is applied to five NMR protein targets chosen from the Northeast Structural Genomics project. These data were then used to design optimized constructs for three partially disordered proteins. Truncated proteins obtained by deletion of disordered N- and C-terminal tails were evaluated using ^{1}H-^{15}N HSQC and ^{1}H-^{15}N heteronuclear NOE NMR experiments to assess their structural integrity. These constructs provide significantly improved NMR spectra, with minimal structural perturbations to the ordered regions of the protein structure. As a representative example, we compare the solution structures of the full length and DXMS-based truncated construct for a 77-residue partially disordered DUF896 family protein YnzC from *Bacillus subtilis*, where deletion of the disordered residues (ca. 40% of the protein) does not affect the native structure. In addition, we demonstrate that throughput of the DXMS process can be increased by analyzing mixtures of up to four proteins without reducing the sequence coverage for each protein. Our results demonstrate that DXMS can serve as a central component of a process for optimizing protein constructs for NMR structure determination.

Keywords

DXMS; hydrogen-deuterium exchange; mass spectrometry; NMR; partially disordered proteins; protein construct optimization; structural genomics

[*]To whom correspondence should be addressed: CABM-Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, Tel: (732) 235-5321; FAX: (732) 235-5633, e-mail: jma@cabm.rutgers.edu, guy@cabm.rutgers.edu.

Introduction

NMR spectroscopy routinely provides high-accuracy solution-state structures of proteins, and is a powerful tool for probing dynamics and interactions with other biological molecules, including small molecule ligands and within multidomain complexes.[1-3] With the emergence of worldwide structural genomics efforts, NMR has served as a complement to X-ray crystallography as a means of rapidly obtaining high-quality three-dimensional (3D) structures of biologically interesting proteins, as well as increasing the coverage of protein sequence space by using these structures as templates for large-scale homology modeling.[4]

Natively disordered or unstructured regions in proteins are both common and biologically important, particularly in modulating intermolecular recognition processes.[5-7] From a practical point of view, however, such disordered regions often pose significant challenges for structure determination by either X-ray or NMR methods. Although, NMR is especially advantageous for analyzing proteins with intrinsically disordered regions that sometimes prevent crystallization,[8] obtaining NMR resonance assignments and structural data for partially unfolded proteins is complicated by a number of factors, including (*i*) NMR resonances for disordered residues tend to be highly overlapped, making their assignment difficult, (*ii*) at one extreme of the relaxation range, disordered residues are characterized by very intense NMR peaks, which can mask underlying resonances from structured regions of the protein, (*iii*) these intense signals from disordered residues can also lead to troublesome spectral artifacts that can complicate the analysis of resonances from ordered residues, (*iv*) disordered termini of proteins may also cause aggregation, precipitation, and sample instability, precluding NMR structure analysis. A potential remedy is to remove disordered N- or C-terminal residues from the protein sequence, provided this procedure does not compromise the structural integrity of the native protein. This strategy was used, for example, in a recent study of *Escherichia coli* ribosome-binding factor A (RbfA), where a 25-residue deletion from the C-terminus of the protein resulted in dramatic improvements in NMR spectral quality and sample stability, and ultimately lead to a solution structure that was not possible for the full length protein.[9] The success of such an approach clearly requires *a priori* residue-specific knowledge of the unfolded region(s) in the protein of interest.

The combination of NMR and hydrogen/deuterium ($^{1}H/^{2}H$) exchange is a well-established technique for monitoring protein dynamics and folding at a residue-specific level.[10-13] In recent years, mass spectrometric measurements of backbone amide proton exchange rates have been successfully implemented to acquire complementary information on smaller sample quantities for the identification of protein-protein or protein-ligand interaction surfaces, to determine the structural stability of proteins and protein complexes, and to characterize flexibility in localized regions of proteins.[10-12,14-22] Analyzing proteins using $^{1}H/^{2}H$ exchange mass spectrometry (DXMS),[19,23-25] where the backbone amide proton exchange rates are used to detect the solvent accessibility of backbone amide groups, has allowed the design of protein constructs with improved crystallization success as compared to the full-length protein.[24,25] Mass spectrometry can also be combined with limited proteolysis (LPMS) to elucidate domain boundaries, ultimately leading to the design of constructs providing diffraction quality crystals.[26]

Here we describe a process for addressing certain classes of challenging proteins using mass spectrometry based construct optimization of partially disordered proteins selected for NMR structure determination by the Northeast Structural Genomics (NESG) Consortium (www.nesg.org). Our general strategy of construct optimization for structure determination in the NESG in shown in Figure 1A. Initial ^{1}H-^{15}N HSQC and ^{1}H-^{15}N hetNOE NMR

screening experiments are used to identify candidate proteins for construct optimization; typically these exhibit ^{1}H-^{15}N peak dispersion indicating structured residues, together with overlapping cross-peaks with ^{1}H-^{15}N chemical shifts characteristic of disordered residues (suggesting some structural disorder). These data reveal that there are disordered segments of the protein, but in the absence of resonance assignments, do not provide information on their location(s) in the sequence. Efforts are next made to identify the polypeptide sequence(s) corresponding to these putative disordered regions using a consensus set of disorder prediction methods (see, for example, Supplementary Fig. S1). If this consensus prediction indicates, with high reliability, a disordered N- or C-terminal segment, several constructs lacking these terminal disordered "tail" residues are generated. However, when no clear consensus is obtained from the various disorder prediction programs, or multiple disordered regions are predicted, DXMS experiments are performed to determine approximate boundaries between ordered and disordered regions. Constructs designed and produced on the basis of either DXMS or disorder predictions are then expressed and purified, and reassessed using ^{1}H-^{15}N HSQC experiments. Ideally, for optimal constructs, deletion of flexible regions does not affect the tertiary structure of the protein but significantly improves the quality of data that can be obtained. This can be validated using an HSQC NMR comparison metric. The truncated protein constructs designed by deletion of the disordered residues are then used for NMR assignment and structure determination.

Here, we describe DXMS studies of five NESG target proteins: brain specific protein C32E8.3 from *Caenorhabditis elegans* (NESG target WR33); DUF896 family protein YnzC from *Bacillus subtilis* (NESG target SR384); protein YjcQ from *Bacillus subtilis* (NESG target SR346); cytoplasmic protein Q8ZRJ2 from *Salmonella typhimurium* (NESG target StR65), and *Escherichia coli* lipoprotein YaiD (NESG target ER553). Using the first four of these proteins, we compared the DXMS-based protein disorder results with site-specific flexibility data obtained from ^{1}H-^{15}N heteronuclear Nuclear Overhauser Effect (hetNOE) experiments. Using ^{1}H-^{15}N HSQC NMR spectra and complete 3D solution NMR structure determination, we demonstrate that removal of disordered tail regions in C32E8.3 and YnzC does not disturb the NMR resonances in the folded regions, while at the same time providing samples that are more suitable for rapid NMR assignment and 3D structure determination. The DXMS optimization of YaiD serves as a striking example of how the technique can yield dramatic improvements to the quality of ^{1}H-^{15}N HSQC NMR spectra, ultimately leading to structures of protein targets which could not otherwise be studied. We further demonstrate that the DXMS results for YjcQ and Q8ZRJ2 are consistent with the predominantly ordered solution structures obtained for these controls. Finally, amide proton exchange rates were analyzed for four proteins individually and as a mixture, to illustrate the potential for higher-throughput DXMS-based protein disorder determination for sets of non-interacting proteins.

Materials and methods

Protein expression, cloning, and purification

All proteins used for the $^{1}H/^{2}H$ exchange mass spectrometry experiments and NMR analysis were expressed, cloned and purified based on methodologies previously published by our laboratory.27 Briefly, the full length gene and truncated construct of C32E8.3 from *C. elegans* were cloned into modified pET15 expression vectors containing a short N-terminal purification tag (MGHHHHHHSH).28 Full length genes for *B. subtilis* YnzC and YjcQ, *S. typhimurium* Q8ZRJ2, and *E. coli* YaiD, as well as the various truncated constructs of *B. subtilis* YnzC and *E. coli* YaiD were cloned into pET21 expression vectors containing a short C-terminal affinity tag (LEHHHHH). All vectors were transformed into codon enhanced BL21 (DE3) pMGK *E. coli* cells, which were cultured at 37 °C in MJ minimal medium.29 Samples for DXMS studies were prepared without isotopic

enrichment. ^{13}C,^{15}N-double labeled samples required for NMR structure determination were fermented using $(^{15}NH_4)_2SO_4$ and *U*-^{13}C-glucose as the sole sources of nitrogen and carbon, respectively. Protein expression was induced at reduced temperature (17 °C) by IPTG (isopropyl-β-D-thiogalactopyranoside). Expressed proteins were purified using an AKTAexpress (GE Healthcare) two-step protocol consisting of HisTrap HP affinity and HiLoad 26/60 Superdex 75 gel filtration chromatography. Sample purity (> 95%) was confirmed using SDS-PAGE and MALDI-TOF mass spectrometry.

$^1H/^2H$ exchange mass spectrometry

Protein $^1H/^2H$ exchange experiments were conducted following the methods described by Spraggon et al.25 Our general protocol is shown in Figure 1B. A 5 μl aliquot of protein sample (~ 25 - 50 μg unlabeled protein in 10 mM Tris HCl, 150 mM NaCl, pH 7.5, unless otherwise indicated) was mixed with 15 μl of deuterium oxide (^2H_2O) containing 10 mM Tris HCl, 150 mM NaCl, pH 7.5 and incubated on ice for set time points before being quenched by the addition of 30 μl of quench solution containing 1.0 M Gu-HCl and 0.5% formic acid (FA). This quench solution reduces the sample pH to ~ 2.5, quenching the rate of amide proton exchange; the GuHCl partially unfolds the protein ensuring more efficient cleavage by pepsin in the subsequent step of the process. The resulting mixture was frozen immediately on dry ice (-80 °C). For the zero-time-point experiment, 5 μl of protein sample was mixed with 15 μl of 10 mM Tris HCl, 150 mM NaCl, pH 7.5 in H_2O, and then 'quenched' and analyzed in the same way as the $^1H/^2H$ exchanged samples. Samples containing multiple proteins were studied by mixing equal volumes of the protein solutions and analyzing 5 μl aliquots of the resulting mixture in the same manner as for the individual protein samples. Therefore, the final protein concentrations in the mixed samples were at 25% of the concentrations used in the individual protein analyses.

HPLC solvent bottles, connection lines including the sample loop, the pepsin column and the analytical column were all kept on ice. Frozen samples were thawed on ice and immediately manually injected into a pre-column (66 μl bed-volume, Upchurch) packed in-house with immobilized pepsin (PIERCE) at a flow rate of 100 μl/min at 0 °C, followed by injection of solution A at 100 μl/min into a 200 μl sample loop (A: 0.1% formic acid in water). Proteins are thus subjected to pepsin cleavage at 0 °C for < 1 min (66 μl pepsin column bed volume, 100 μl/min flow rate). The protocol ensures extensive pepsin cleavage, which produces large numbers of overlapping peptides and provides extensive coverage of the protein sequence and high resolution in the exchange heat map.

After pepsin digestion, the sample loop was next brought with a C18 HPLC column (Discovery, BioWide Pore C18-3, 5 cm × 2.1 mm 3 um, Supelco) online by valve switching. Following a three minute wash step with 2% solution B (solution B: 0.1% formic acid in acetonitrile) at a flow rate of 200 μl/min, the digested peptides were separated by a linear acetonitrile gradient of 2-50% solution B over 17 min at 200 μl/min. The eluate was then analyzed by an electrospray-linear ion-trap mass spectrometer (LTQ, ThermoFisher). For measurement of the mass shift in $^1H/^2H$ exchange experiments, MS was set to perform full-scan in the m/z range of 300-2000 in profile mode for the entire LC-MS run. Peaks corresponding to the deuterated peptides were manually extracted based on approximate retention time and m/z. The average m/z was calculated as the centroid of the isotopic mass distribution averaged over a retention-time window defined at 30% peak height. The amount of deuteration of each peptide was quantified by the difference of the average m/z at each time point from that of the zero-time-point sample (fully protonated state). For the correction of back exchange during the pepsin cleavage and chromatographic seperation process, a completely exchanged sample was produced as described by Hamuro et al.;30 5 μl of the protein sample was mixed with 15 μl of 0.5% formic acid in ^2H_2O and incubated at room temperature for 24 h. The sample was then quenched and analyzed using the same

conditions as the $^{1}H/^{2}H$ exchange experiments. The formula of Zhang & Smith15 (Eqn. 1) was used for calculation of normalized deuterium incorporation levels for each peptide.

$$D_t = \frac{m(labeled, t) - m(unlabeled)}{m(control) - m(unlabeled)} \times 100 \quad \text{(Eqn 1)}$$

Peptide Identification

For peptide identifications, a sample was processed the same way as the zero-time-point sample described above for amide $^{1}H/^{2}H$ exchange measurements. The mass spectrometer was set to perform one full-scan MS in the m/z range 300-2000, followed by zoom scans of the top 5 most intense ions and MS/MS of multiply charged ions. Dynamic exclusion conditions were set to exclude parent ions that were selected for MS/MS twice within 30 sec, and the exclusion duration was 60 sec. Acquired data was then searched using Sequest software against a homemade sequence database composed of 83,095 entries of NESG target proteins, plus the *E. coli* sequence database and sequences of common contaminants, such as human keratins. The search parameters were set to use no enzyme and parent tolerance of +/- 2 amu and fragment ion tolerance of +/-1 amu. The search results were confirmed manually.

Solution structure determination of full length and truncated construct of *B. subtilis* YnzC

A complete description of the methods used in the solution NMR structure determinations of full length and truncated *B. subtilis* YnzC are presented elsewhere.31 Briefly, samples of uniformly ^{13}C,^{15}N-enriched full length YnzC and truncated YnzC(1-46) for NMR structure determination were prepared at protein concentrations of 1.1 to 1.4 mM in 20 mM MES, 100 mM NaCl, 5 mM $CaCl_2$, 10 mM DTT, 5% $^{2}H_2O$ / 95% H_2O, pH 6.5. All NMR data were collected at 20 °C on Varian INOVA 500 and 600 MHz and Bruker AVANCE 600 and 800 NMR spectrometers. Complete ^{1}H, ^{13}C, and ^{15}N resonance assignments for full length *B. subtilis* YnzC and YnzC(1-46), were determined using GFT NMR data collection methods32,33 and conventional triple resonance NMR methods, respectively,34 and deposited in the BioMagResDB (BMRB accession numbers 7225 and 15476). ^{1}H-^{15}N heteronuclear NOEs were measured with gradient sensitivity-enhanced 2D heteronuclear NOE approaches.35,36 The full length YnzC structure was determined using the program AutoStructure 2.1.137 interfaced with XPLOR-NIH 2.11.2.38 The folded N-terminal residues (1-42) of the 20 lowest energy structures out of 100 calculated were deposited in the Protein Data Bank (PDB ID, 2HEP). The structure of YnzC(1-46) was calculated using CYANA 2.139,40 followed by refinement by restrained molecular dynamics in explicit water using CNS 1.2.41,42 The final refined ensemble of structures (excluding the C-terminal His_6) were deposited in the Protein Data Bank (PDB ID, 2JVD). Structural statistics and global quality scores43,44 for the full length and truncated YnzC solution NMR structures are presented elsewhere.31

Results

DXMS construct optimization of TPPP family protein *C. elegans* C32E8.3

The partially unfolded 180-residue *C. elegans* protein CE32E8.3 belongs to the family of Tubulin Polymerization Promoting Proteins (TPPP). The prototype of this family, TPPP/p25 alpha, promotes aberrant tubulin polymerization, is known to block mitotic spindle formation in *Drosophila* embryo, and has been identified as a marker for alpha-synucleinopathies.45-48 While there is some controversy in the literature as to whether TPPP/p25 alpha is natively unfolded or flexible but natively folded,49,50 the solution

structure of the *C. elegans* homologue CE32E8.3 (NESG target WR33), with 37.5% sequence similarity to human TPPP/p25 alpha, has been solved by the NESG consortium (DOI 10.2210/pdb1pul/pdb). It consists of five helices with an intrinsically-disordered region in the C-terminal one-third of the protein sequence. The protein construct that was used for the NMR structure determination was designed utilizing backbone NMR spectral assignments for the full-length protein,28 since the backbone chemical shift and ^{1}H-^{15}N hetNOE data indicated high flexibility in this region of the protein. Figure 2 compares the NMR ^{1}H-^{15}N HSQC spectra for the full-length CE32E8.3 protein (1-180) and the truncated protein construct (1-115) that was used for the solution structure determination. This comparison shows the presence of many overlapping peaks in the full-length protein with chemical shift values typical of disordered residues. These peaks are absent in the truncated protein construct. The amide ^{15}N and ^{1}H resonance frequencies for the remaining ordered residues are identical in both spectra, confirming that deletion of the disordered sequence does not disturb the structure of the remaining protein.

The process of deducing the NMR backbone resonance assignments for the full-length 180-residue CE32E8.3 protein in order to identify disordered residues was slow and laborious, requiring milligram quantities of the protein. It was particularly challenging to complete resonance assignments for disordered residues that exhibit overlap with resonances from the ordered helical residues. Mass spectrometry, which uses only microgram amounts of protein, facilitates rapid construct optimization for subsequent NMR studies. The relative rates of $^{1}H/^{2}H$ exchange for backbone amide protons were utilized to identify the solvent accessible/flexible regions in this protein. The normalized deuterium incorporation levels averaged for overlapping peptides are presented in Figure 3A. The first three rows in this figure provide the protein sequence, NMR secondary structure, and ^{1}H-^{15}N hetNOE values determined using nearly complete backbone resonance assignments. The next three rows (labeled ‘I’ for ‘individual protein’) denote the degree of $^{1}H/^{2}H$ exchange at ~ 0 °C, color-coded to illustrate the deuterium uptake levels when the exchange reaction is quenched after 10 sec, 100 sec, and 1000 sec, respectively. The results reveal faster exchanging amide protons in the ~ 60 C-terminal residues of the protein (residue 122 onwards), consistent with hetNOE data obtained for this more flexible region of the protein. The deuterium exchange levels at 100 sec and 1000 sec time points for residues 32-41 and 58-68 also reflect high solvent accessibility of amide proton sites in these regions of the structure, consistent with the locations of interhelical loops in these segments identified by the NMR chemical shift data and by the solved 3D structure. Thus, the DXMS data, recorded on smaller quantities of protein sample and generated much more rapidly than resonance assignments and hetNOE data for this 180-residue protein, provide sufficiently accurate determination of disordered regions of the protein to allow construct designs similar to those provided by the extensive NMR studies.

DXMS construct optimization of the partially disordered protein *B. subtilis* YnzC

Further validation of mass spectrometry based disorder identification was achieved by comparing the DXMS results for the 77-residue, 8.8 kDa partially unstructured putative cytoplasmic protein YnzC from *B. subtilis* (NESG target SR384) with ^{1}H-^{15}N hetNOE values obtained using NMR spectroscopy. The solution structure for this protein, a monomer based on static light scattering data,31 is comprised of two antiparallel alpha helices followed by an extended high flexibility region in the C-terminal region of the protein.31 Results for the time series $^{1}H/^{2}H$ exchange experiments conducted for protein YnzC are shown in Figure 3B. Once again we observe excellent agreement between the flexible regions of the protein as evidenced by low (or negative) ^{1}H-^{15}N hetNOE values (residue 42 and above) and the mass spectrometry results indicating high (greater than 70%) deuterium incorporation levels for residues 42 onward. The deuterium levels for the peptides that were

selected for $^{1}H/^{2}H$ exchange analysis of this protein are listed in Supplementary Table S1. Note that the peptide comprising residues 26-35 shows ~ 20% deuterium uptake after 10 sec of exchange, while the peptide comprising residues 41-51 shows 80% deuterium uptake, implying that the disorder boundary lies somewhere within the boundaries for the latter. We therefore prepared five truncated protein constructs covering the length of this peptide (constructs 1-40, 1-43, 1-46, 1-49 and 1-52). To confirm that DXMS based truncations did not disturb the protein structure in the ordered region of this protein, we also conducted NMR ^{1}H-^{15}N HSQC experiments for all five of these protein constructs (Figure 4A). It is interesting to note that the removal of up to 37 amino acids from the disordered C-terminal region of the protein (close to 50% of the entire protein sequence) does not significantly affect the amide ^{15}N and ^{1}H resonance frequencies for the structured region of this protein. Figure 4B shows the solution structures for the full-length protein (1-77) and the truncated protein construct (1-46).31 The backbone root-mean-square deviation (RMSD) of 0.84 Å between the folded regions (residues 5-19 and 22-38) of the average structures for the two ensembles confirm that deletion of the disordered C-terminal half of the protein does not perturb the native state structure of the N-terminal half of this protein. This result further demonstrates the value of the DXMS technique in construct optimization for preparation of NMR samples.

DXMS construct optimization of the partially disordered protein *E. coli* YaiD

In our structural genomics effort, the DXMS approach routinely results in the design of constructs with dramatically enhanced NMR spectral properties. This is illustrated by the construct optimization of *E. coli* YaiD. The *yaiD* gene from *E. coli* encodes for a 219-residue, 22.2 kDa bacterial lipoprotein precursor, that is cleaved at a cysteine near its N-terminus (C21), which in turn is covalently linked to the periplasmic face of the inner membrane. In initial NMR screening experiments, the mature 199-residue YaiD protein exhibited marginal quality ^{1}H-^{15}N HSQC spectra (Figure 5A, left). DXMS analysis of the full-length protein revealed a ~ 60-residue disordered region in the N-terminal region of the protein, which was not anticipated from disorder prediction and predicted secondary structure in this region of the sequence (Figure 5B).51 Constructs were designed with varying N-terminal truncations based on these DXMS data. The YaiD[59-199] construct yielded the best ^{1}H-^{15}N HSQC spectra in subsequent screening experiments (Figure 5A, right), a striking improvement in spectral quality compared to the full-length protein, and pattern of ^{1}H-^{15}N amide resonance frequencies essentially identical to corresponding peaks in the spectrum of the full-length protein. The solution NMR structure of this DXMS-optimized construct was ultimately solved in our NESG consortium (DOI 10.2210/pdb2k1s/pdb).

Simultaneous DXMS on multiple samples

$^{1}H/^{2}H$ exchange mass spectrometry of purified recombinant proteins serves as a rapid means of identifying protein disorder using much smaller quantities of sample compared to amide exchange NMR experiments, which also requires resonance assignment information. However, each analysis requires 2 - 3 days for sample preparation, MS data collection and data analysis. Consequently, we became interested in further increasing the throughput by analyzing mixtures of purified proteins in a single set of $^{1}H/^{2}H$ exchange experiments. Such an experimental set up should give results similar to the individual protein experiments as long as there are no protein-protein interactions, no sequence similarities for the different proteins, and the sequence coverage is not significantly affected due to overlapping mass spectra from co-eluting peptides. To this end, we analyzed a mixture of four proteins, where two proteins (protein C32E8.3, DUF896 protein YnzC) are partially disordered proteins while the other two proteins (NESG target SR346 - YjcQ protein from *B. subtilis*, and NESG target StR65 - cytoplasmic protein Q8ZRJ2 from *S. typhimurium*) adopt highly

ordered folds based on their solution NMR structures (DOI 10.2210/pdb2jn8/pdb and DOI 10.2210/pdb2hgc/pdb, respectively).

The $^{1}H/^{2}H$ exchange data for four proteins, C32E8.3, YnzC, YjcQ, Q8ZRJ2, analyzed individually and as a four component mixture, are shown in Figures 3A, 3B, 6A, and 6B, respectively. There is generally good agreement between $^{1}H/^{2}H$ exchange data for all four proteins when analyzed individually and in the protein mixture. In the case of the fully structured protein Q8ZRJ2 (Figure 6A), slowly-exchanging protons are observed for the entire protein sequence in both sets of experiments, although sequence coverage of the data obtained on the protein mixture is slightly lower (88%).

Initial results for YjcQ protein (Figure 6B), both in the single protein $^{1}H/^{2}H$ exchange experiment (I) and in the protein mixture (M), indicate disorder in the C-terminal region of the protein (residue 72 onwards). At first glance, this result appears to contradict the NMR structure showing an additional alpha helix in this region of the protein, and the high ^{1}H-^{15}N hetNOE values (residues 72-80) consistent with low flexibility for this region of the protein. We traced this discrepancy in the different methods of assessing intrinsic disorder to the differences in pH of the buffers employed in the NMR structural study versus the amide hydrogen exchange measurements. While the NMR experiments were conducted at pH 5.5, the first DXMS experiment was conducted at pH 7.5. The DXMS protocol was then repeated using the same pH 5.5 buffer employed in the NMR studies of YjcQ (Figure 6C). Under these conditions, the 10 sec deuterium incorporation level for peptide 72-82 goes down from 88% at pH 7.5 to less than 50% at pH 5.5. This result highlights the importance of making DXMS measurements under conditions similar to those to be used for NMR structural studies, in order to generate data for proper construct optimization and subsequent protein structure analysis.

Discussion

Intrinsically unfolded regions in proteins often have significant biological relevance. 5 For example, the C-terminal third of TPPP CE32E8.3 protein studied here is the most strongly conserved region of the sequence, and is probably involved in specific protein-protein interactions. However, when the goal is the structural analysis of the *ordered* region of the protein structure, removing such disordered regions is a practical approach for obtaining crystals and/or improved NMR data, and the spectroscopic and structural studies afforded by the resulting optimized constructs may serve to bootstrap future studies of the full length protein and/or order/disorder transitions which accompany complex formation.

In practice we find that the DXMS method is well suited for identifying disordered regions of proteins in the pH 5.5 to 7.5 range, a window typically employed in protein NMR. In principle, this does not preclude examining $^{1}H/^{2}H$ exchange in proteins at more extremes of pH, but this has not been necessary in our studies of NESG proteins to date.

Pepsin is an excellent choice for the DXMS method, which requires efficient proteolysis in the low pH range (pH 2 - 3). Pepsin shows broad substrate specificity, typically cutting the unfolded polypeptide chain at high frequency on the C-terminal side of bulky hydrophobic amino acid residues. In our experience with many proteins, pepsin provides extensive sequence coverage with overlapping peptide fragments. While other acidic proteases are available, they generally have more limited sequence specificity, are less robust, and are therefore not particularly useful for our applications. In our hands, the immobilized pepsin column can be used repeatedly for many DXMS runs.

DXMS is an alternative approach to limited proteolysis (LP)52،53 and LPMS26 for protein construct design and optimization, combining the advantages of mass spectrometry,

including low sample consumption (high sensitivity) and high-throughput potential, with the solvent-accessibility information afforded by $^{1}H/^{2}H$ exchange studies. While LP and LPMS also allow rapid identification of the flexible regions in a protein, these studies may require a customized experimental protocol for each case, since different combinations of enzymes (depending on the properties of individual proteins) may be required to achieve sufficient resolution. For the purpose of protein construct optimization, amide exchange mass spectrometry is highly suitable because it provides comprehensive information regarding exchange rates for the flexible and structured regions of the protein, thus allowing the design of protein constructs with clearly defined boundaries. DXMS also allows unambiguous distinction between flexible tail regions and flexible internal loops. Overall, compared to LP and LPMS, DXMS is a higher resolution technique providing more complete information on the locations of disordered regions.

In several cases described in the Results section, the DXMS data reveal the locations of internal disordered loops in the protein structure. In all of these cases, these loops are relatively small, and no attempts have been made to surgically remove disordered loops from the middle of ordered regions of proteins. However, this strategy could in principle be used to design constructs lacking large internal disordered loops. Moreover, DXMS data are also useful for identifying disordered linker regions between ordered domains of multidomain proteins.

Knowledge of unstructured regions in proteins gleaned from DXMS can also be used for other applications, such as identifying potential binding partners for these proteins including other proteins, polynucleic acids, ligands, or stabilizing metal ions. This is particularly relevant to eukaryotic genomes where disordered proteins/domains are more prevalent. Furthermore, building a database of experimental protein disorder data will aid bioinformatics applications by providing a training set for the development and benchmarking of new or existing disorder prediction and structure modeling programs. We are developing such a database of DXMS data on NESG proteins for these purposes.

Our analysis of four-protein mixtures demonstrates that we can accelerate DXMS data accumulation by analyzing mixtures of proteins in a single set of experiments, without significantly compromising the sequence coverage for each protein. Wales et al55 have also recently demonstrated DXMS measurements on a mixture of four proteins. In the mixed-sample DXMS experiments presented here, we note a subtle yet systematic difference between the results for each protein measured in a mixture compared to individually. Namely, there appears to be slightly more $^{1}H/^{2}H$ exchange in the mixed samples compared to studies using individual proteins. DXMS data are measured over many fragments. When the experiment is done as a mixture of proteins data is obtained for fewer number of fragments per residue and the accuracy of the measurement is reduced. Hence, mixed samples provide lower statistical averaging of exchange data. Regardless, the qualitative exchange pattern is the same, meaning that DXMS analysis of mixtures is applicable to deducing the approximate boundaries between ordered and disordered regions in each protein. However, for more quantitative amide proton exchange rate measurements, it is preferable to analyze proteins individually.

Finally, while the mixed DXMS experiments provide reliable results and they are a first step toward speeding up this approach, we are actively exploring other means of increasing the throughput of the DXMS approach. For instance, several steps in the protocol, including sample handling and injection, can potentially be automated using robotics, which are already implemented in the protein production pipeline within our structural genomics consortium27. Also, mass spectral data analysis is currently a manual and time-consuming process. As a result we and others56 are designing automation software to speed up the data

analysis for these DXMS experiments. Taken together, these strategies will help establish the DXMS technique as a high-throughput approach for construct optimization of partially disordered proteins selected for NMR structure determination in both structural biology and structural genomics projects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Peter Lobel, Patricia Weber (ExSAR Corp.), John Everett, and Binchen Mao for helpful discussions. This work was supported by a grant from the National Institute of General Medical Sciences Protein Structure Initiative (U54-GM074958).

References

1. Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., III; Skelton, NJ.; Rance, M. Protein NMR Spectroscopy: Principles and Practice. 2. New York: Elsevier Academic Press; 2007. p. 912
2. Kay LE. NMR studies of protein structure and dynamics. J Magn Reson. 2005; 173:193–207. [PubMed: 15780912]
3. Bax A, Grishaev A. Weak alignment NMR: A hawk-eyed view of biomolecular structure and dynamics. Curr Opin Struct Biol. 2005; 15:563–570. [PubMed: 16140525]
4. Liu J, Montelione GT, Rost B. Novel leverage of structural genomics. Nature Biotechnol. 2007; 25:849–851. [PubMed: 17687356]
5. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nature Rev. 2005; 6:197–208.
6. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J. 2005; 272:5129–5148. [PubMed: 16218947]
7. Huang YJ, Montelione GT. Proteins flex to function. Nature. 2005; 438:36–37. [PubMed: 16267540]
8. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. Chem Rev. 2004; 104:3607–3622. [PubMed: 15303830]
9. Huang YJ, Swapna GVT, Rajan PK, Ke H, Xia B, Shukla K, Inouye M, Montelione GT. Solution NMR structure of ribosome-binding factor A (RbfA), a cold-shock adaptation protein from *Escherichia coli*. J Mol Biol. 2003; 327:521–536. [PubMed: 12628255]
10. Englander SW, Downer NW, Teitelbaum H. Hydrogen exchange. Annu Rev Biochem. 1972; 41:903–924. [PubMed: 4563445]
11. Englander SW, Mayne L. Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. Annu Rev Biophys Biomol Struct. 1992; 21:243–265. [PubMed: 1525469]
12. Englander SW, S T, Englander JJ, Mayne L. Mechanisms and uses of hydrogen exchange. Curr Opin Struct Biol. 1996; 6:18–23. [PubMed: 8696968]
13. Redfield C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. Methods. 2004; 34:121–132. [PubMed: 15283921]
14. Thevenon-Emeric G, Kozlowski J, Zhang Z, Smith DL. Determination of amide hydrogen exchange rates in peptides by mass spectrometry. Anal Chem. 1992; 64:2456–2458. [PubMed: 1466454]
15. Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Protein Sci. 1993; 2:522–531. [PubMed: 8390883]
16. Ghaemmaghami S, Fitzgerald MC, Oas TG. A quantitative, high-throughput screen for protein stability. Proc Natl Acad Sci U S A. 2000; 97:8296–8301. [PubMed: 10890887]
17. Kaltashov IA, Eyles SJ. Studies of biomolecular conformations and conformational dynamics by mass spectrometry. Mass Spectrom Rev. 2002; 21:37–71. [PubMed: 12210613]

18. Powell KD, Ghaemmaghami S, Wang MZ, Ma L, Oas TG, Fitzgerald MC. A general mass spectrometry-based assay for the quantitation of protein-ligand binding interactions in solution. J Am Chem Soc. 2002; 124:10256–10257. [PubMed: 12197709]
19. Englander JJ, Del Mar C, Li W, Englander SW, Kim JS, Stranz DD, Hamuro Y, Woods VL Jr. Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. Proc Natl Acad Sci U S A. 2003; 100:7057–7062. [PubMed: 12773622]
20. Horn JR, Kraybill B, Petro EJ, Coales SJ, Morrow JA, Hamuro Y, Kossiakoff AA. The role of protein dynamics in increasing binding affinity for an engineered protein-protein interaction established by H/D exchange mass spectrometry. Biochemistry. 2006; 45:8488–8498. [PubMed: 16834322]
21. Englander SW. Hydrogen exchange and mass spectrometry: A historical perspective. J Am Soc Mass Spectrom. 2006; 17:1481–1489. [PubMed: 16876429]
22. Wales TE, Engen JR. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev. 2006; 25:158–170. [PubMed: 16208684]
23. Woods VL Jr, Hamuro Y. High resolution, high-throughput amide deuterium exchange-mass spectrometry (DXMS) determination of protein binding site structure and dynamics: utility in pharmaceutical design. J Cell Biochem Suppl. 2001; S37:89–98. [PubMed: 11842433]
24. Pantazatos D, Kim JS, Klock HE, Stevens RC, Wilson IA, Lesley SA, Woods VL Jr. Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. Proc Natl Acad Sci U S A. 2004; 101:751–756. [PubMed: 14715906]
25. Spraggon G, Pantazatos D, Klock HE, Wilson IA, Woods VL Jr, Lesley SA. On the use of DXMS to produce more crystallizable proteins: structures of the *T. maritima* proteins TM0160 and TM1171. Protein Sci. 2004; 12:3187–3199. [PubMed: 15557262]
26. Gao X, Bain KB, Bonanno JB, Buchanan M, Henderson D, Lorimer D, Marsh C, Reynes JA, Sauder JM, Schwinn K, Thai C, Burley SK. High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. J Struct Funct Genomics. 2005; 6:129–134. [PubMed: 16211509]
27. Acton TB, Gunsalus KC, Xiao R, Ma LC, Aramini J, Baran MC, Chiang YW, Climent T, Cooper B, Denissova NG, Douglas SM, Everett JK, Ho CK, Macapagal D, Rajan PK, Shastry R, Shih LY, Swapna GV, Wilson M, Wu M, Gerstein M, Inouye M, Hunt JF, Montelione GT. Robotic cloning and protein production platform of the Northeast Structural Genomics Consortium. Methods Enzymol. 2005; 394:210–243. [PubMed: 15808222]
28. Monleon D, Chiang Y, Aramini JM, Swapna GVT, Macapagal D, Gunsalus KC, Kim S, Szyperski T, Montelione GT. Backbone ^{1}H, ^{15}N, and ^{13}C assignments for the 21 kDa *Caenorhabditis elegans* homologue of 'brain-specific' protein. J Biomol NMR. 2004; 28:91–92. [PubMed: 14739645]
29. Jansson M, Li Y-C, Jendeberg L, Anderson S, Montelione GT, Nilsson B. High level production of uniformly ^{15}N- and ^{13}C-enriched fusion proteins in *Escherichia coli*. J Biomol NMR. 1996; 7:131–141. [PubMed: 8616269]
30. Hamuro Y, Wong L, Shaffer J, Kim JS, Stranz DD, Jennings PA, Woods VL Jr, Adams JA. Phosphorylation driven motions in the COOH-terminal Src kinase, CSK, revealed through enhanced hydrogen-deuterium exchange and mass spectrometry (DXMS). J Mol Biol. 2002; 323:871–881. [PubMed: 12417200]
31. Aramini JM, Sharma S, Swapna GVT, Huang YJ, Ho CK, Shetty K, Cunningham K, Ma LC, Zhao L, Owens LA, Jiang M, Xiao R, Liu J, Baran MC, Acton TB, Rost B, Montelione GT. Solution structure of the SOS response protein YnzC from *Bacillus subtilis*. Proteins. 2008; 72:526–530. [PubMed: 18431750]
32. Atreya H, Szyperski T. G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. Proc Natl Acad Sci USA. 2004; 101:9642–9647. [PubMed: 15210958]
33. Liu G, Shen Y, Atreya H, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TB, Arrowsmith CH, Montelione GT, Szyperski T. NMR data collection and analysis protocol for high-throughput protein structure determination. Proc Natl Acad Sci. 2005; 102:10487–10492. [PubMed: 16027363]

34. Aramini JM, Huang YJ, Swapna GVT, Cort JR, Rajan PK, Xiao R, Shastry R, Acton TB, Liu J, Rost B, Kennedy MA, Montelione GT. Solution NMR structure of *Escherichia coli* ytfP expands the structural coverage of the UPF0131 protein domain family. Proteins. 2007; 68:789–795. [PubMed: 17523190]
35. Li Y-C, Montelione GT. Solvent saturation-transfer effects in pulsed-field-gradient heteronuclear single-quantum-coherence (PFG-HSQC) spectra of polypeptides and proteins. J Magn Reson Ser B. 1993; 101:315–319.
36. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. Biochemistry. 1994; 33:5984–6003. [PubMed: 7514039]
37. Huang YJ, Tejero R, Powers R, Montelione GT. A topology-constrained distance network algorithm for protein structure determination from NOESY data. Proteins. 2006; 62:587–603. [PubMed: 16374783]
38. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. J Magn Res. 2003; 160:66–73.
39. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol. 1997; 273:283–298. [PubMed: 9367762]
40. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol. 2002; 319:209–227. [PubMed: 12051947]
41. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. 1998; D54:905–921.
42. Linge JP, Williams MA, Spronk CAEM, Bonvin AMJJ, Nilges M. Refinement of protein structures in explicit solvent. Proteins. 2003; 50:496–506. [PubMed: 12557191]
43. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins. 2007; 66:490–493.
44. Huang YJ, Powers R, Montelione GT. Protein NMR Recall, Precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J Am Chem Soc. 2005; 127:1665–1674. [PubMed: 15701001]
45. Hlavanda E, Kovacs J, Olah J, Orosz F, Medzihradszky KF, Ovadi J. Brain-specific p25 protein binds to tubulin and microtubules and induces aberrant microtubule assemblies at substoichiometric concentrations. Biochemistry. 2002; 41:8657–8664. [PubMed: 12093283]
46. Tirian L, Hlavanda E, Olah J, Horvath I, Orosz F, Szabo B, Kovacs J, Szabad J, Ovadi J. TPPP/p25 promotes tubulin assemblies and blocks mitotic spindle formation. Proc Natl Acad Sci U S A. 2003; 100:13976–13981. [PubMed: 14623963]
47. Kovacs GG, Laszlo L, Kovacs J, Jensen PH, Lindersson E, Botond G, Molnar T, Perczel A, Hudecz F, Mezo G, Erdei A, Tirian L, Lehotzky A, Gelpi E, Budka H, Ovadi J. Natively unfolded tubulin polymerization promoting protein TPPP/p25 is a common marker of alpha-synucleinopathies. Neurobiol Dis. 2004; 17:155–162. [PubMed: 15474353]
48. Kovacs GG, Gelpi E, Lehotzky A, Hoftberger R, Erdei A, Budka H, Ovadi J. The brain-specific protein TPPP/p25 in pathological protein deposits of neurodegenerative diseases. Acta Neuropathol (Berl). 2007; 113:153–161. [PubMed: 17123092]
49. Orosz F, Kovacs GG, Lehotzky A, Olah J, Vincze O, Ovadi J. TPPP/p25: from unfolded protein to misfolding disease: prediction and experiments. Biol Cell. 2004; 96:701–711. [PubMed: 15567525]
50. Otzen DE, Lundvig DM, Wimmer R, Nielsen LH, Pedersen JR, Jensen PH. p25alpha is flexible but natively folded and binds tubulin with oligomeric stoichiometry. Protein Sci. 2005; 14:1396–1409. [PubMed: 15883183]
51. Rost B, Yachdav G, Liu J. The PredictProtein server. Nucleic Acids Res. 2004; 32:W321–W326. [PubMed: 15215403]

52. Fontana A, Fassina G, Vita C, Dalzoppo D, Zamai M, Zambonin M. Correlation between sites of limited proteolysis and segmental mobility. Biochemistry. 1986; 25:1847–1851. [PubMed: 3707915]

53. Fontana A, de Lauretto PP, Spolaore B, Frare E, Picotti P, Zambonin M. Probing protein structure by limited proteolysis. Acta Biochim Pol. 2004; 51:299–321. [PubMed: 15218531]

54. Huang, YJ.; Montelione, GT. DisMeta disorder prediction Metaserver. Rutgers University; http://www-nmr.cabm.rutgers.edu/bioinformatics/disorder/

55. Wales TE, Fadgen KE, Gerhardt GC, Engen JR. High-speed and high-resolution UPLC separation at zero degrees Celsius. Anal Chem. 2008; 80:6815–6820. [PubMed: 18672890]

56. Pascal BD, Chalmers MJ, Busby SA, Mader CC, Southern MR, Tsinomeras NF, Griffin PR. The Deuterator: software for the determination of backbone amide deuterium levels from H/D exchange MS data. BMC Bioinformatics. 2007; 8:156. [PubMed: 17506883]

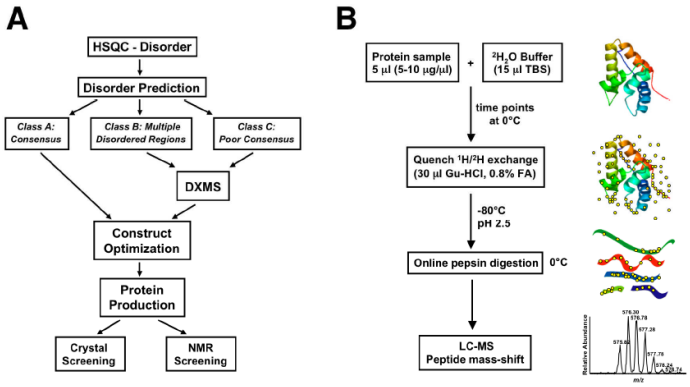

Figure 1.

(A) General strategy for construct optimization of targets for structure determination in the NESG consortium. After initial NMR screening, disorder prediction results for targets exhibiting evidence of partial disorder are classified into three groups: A) consensus, B) multiple disordered regions, and C) poor consensus. Construct optimization for Class A targets is based exclusively on the bioinformatics predictions. Class B and C targets, however, are subsequently analyzed by DXMS so as to accurately determine the ordered/disordered boundaries, and optimized constructs are designed on the basis of these data. Finally, construct-optimized targets are re-evaluated by ^{1}H ^{-15}N HSQC NMR and sent for crystallization screening. **(B)** General protocol for DXMS analysis of protein targets in the NESG consortium. The major steps in the protocol are as follows: mixing the protein sample(s) with $^{2}H_{2}O$ (depicted on the right with yellow circles), quenching the $^{1}H/^{2}H$ exchange at specific time points by lowering the pH, pepsin digestion, and separation of peptide fragments by LC-MS. See Material and Methods section for a complete description of the DXMS protocol used in this work.

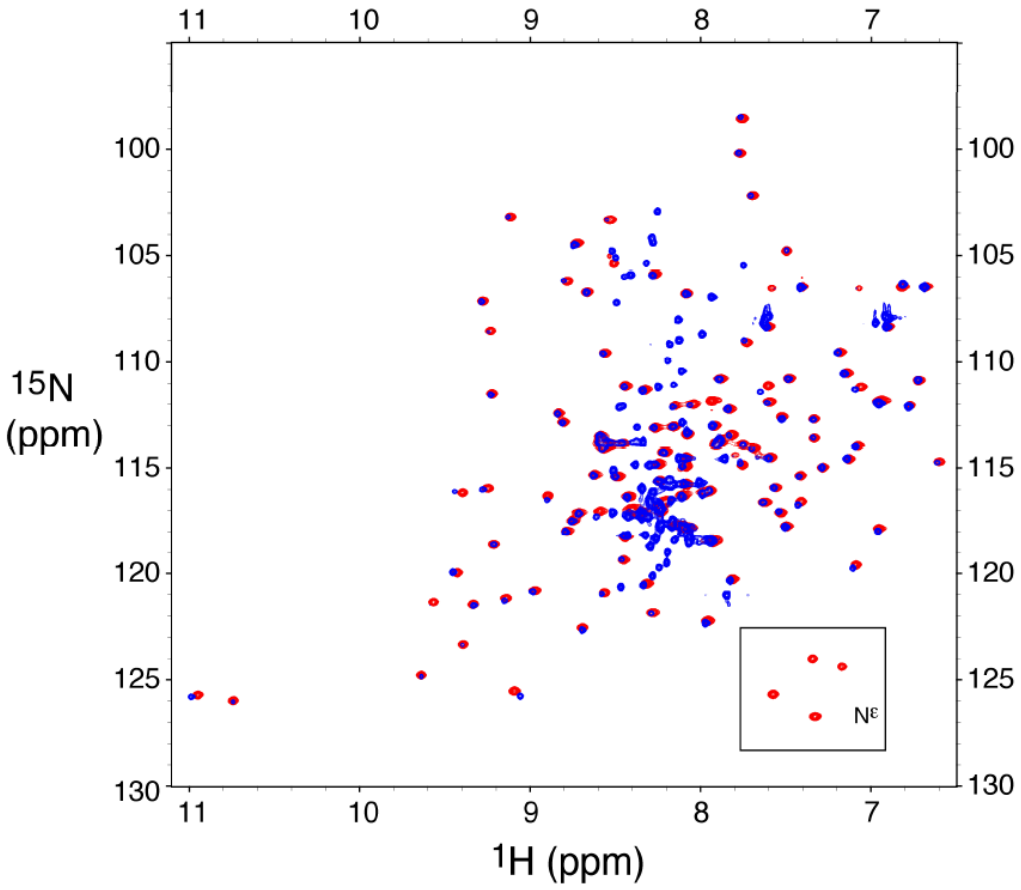

Figure 2.
Overlay of ^{1}H ^{-15}N HSQC NMR spectra (25 °C) of TPPP family protein C32E8.3 from *C. elegans* (NESG target WR33) for the full length (1 - 180) protein (blue), and truncated (1 - 115) protein construct (red). Side chain Arg peaks aliased in the ^{15}N-dimension of the spectrum of truncated C32E8.3 are boxed.

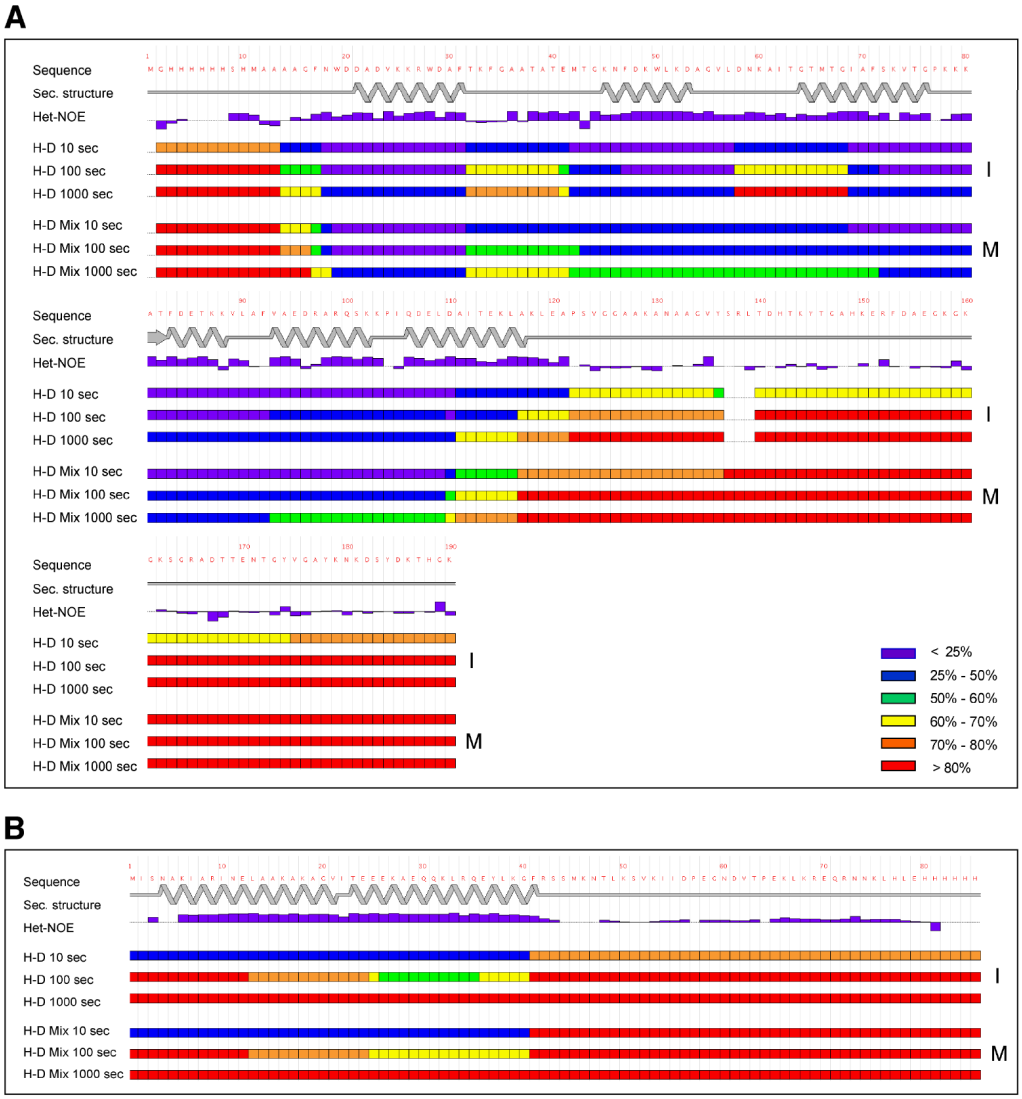

Figure 3.

Protein sequence, NMR secondary structure, ^{1}H-^{15}N hetNOE, and DXMS results (10, 100, and 1000 second exchange durations; pH 7.5 and temperature ~ 0 °C) analyzed individually (I) and in a four protein mixture (M) for **(A)** protein C32E8.3 from *C. elegans* (NESG target WR33) and **(B)** protein YnzC from *B. subtilis* (NESG target SR384).

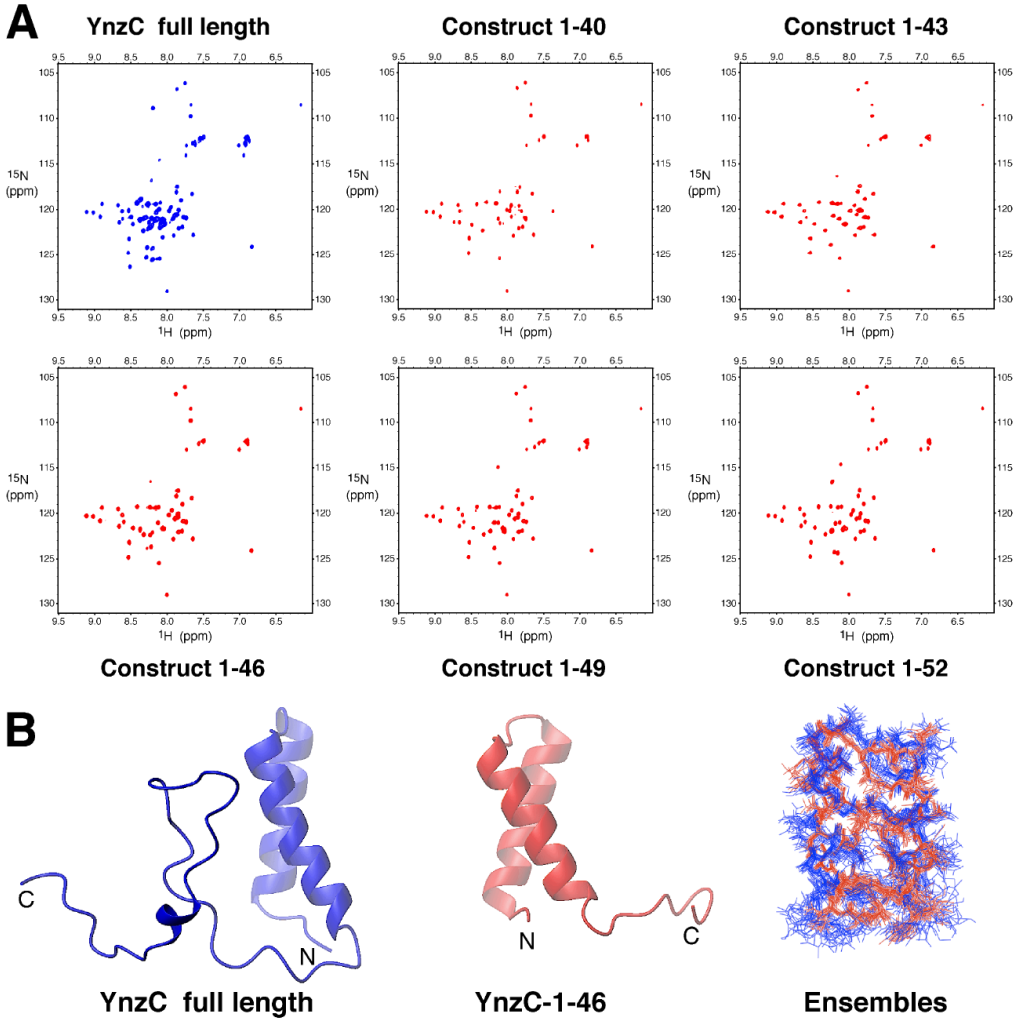

Figure 4.

(A) ^{1}H-^{15}N HSQC NMR spectra (20 °C) for the full length (1 - 77) protein YnzC from *B. subtilis* (NESG target SR384) and five truncated protein constructs that were designed based on the DXMS results. **(B)** NMR solution structures of full length (blue; PDB ID, 2HEP) and truncated (red; PDB ID, 2JVD) protein YnzC from *B. subtilis* (NESG targets SR384 and SR384-1-46, respectively). The first two images show ribbon diagrams of representative (lowest energy) conformers of full length and truncated YnzC. The superimposed final ensembles of structures (rotated 180°) are presented on the right (20 models each; heavy atoms for residues 2 to 40 are shown). The backbone RMSD between the mean coordinates of the ordered residues encompassing the helices (5-19 and 22-38) of each ensemble is 0.84 Å.

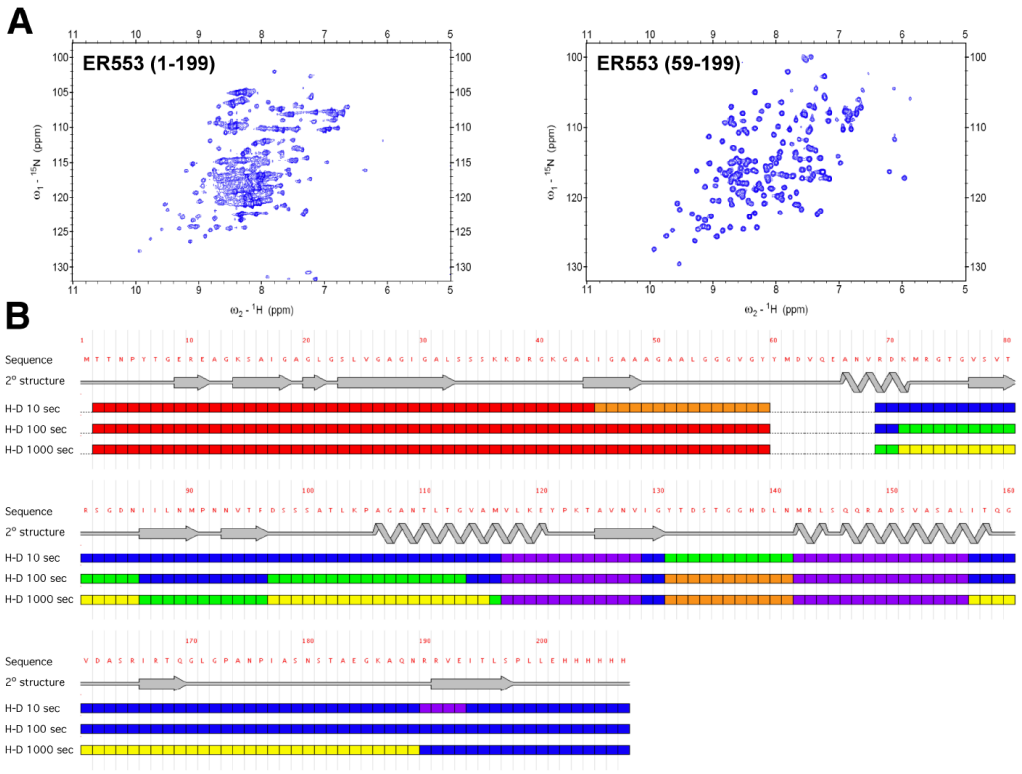

Figure 5.

DXMS-based construct optimization of *E. coli* yiaD (NESG target, ER553). **(A)** ^{1}H-^{15}N HSQC spectra (20 °C) of full length (left) and construct optimized (right) *E. coli* yiaD (59-199). **(B)** DXMS results for full length *E. coli* yiaD (10, 100, and 1000 second exchange durations at pH 7.5 and ~ 0 °C). The PROF51 secondary structure prediction results are shown above the DXMS data.

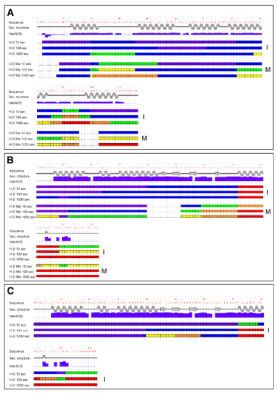

Figure 6.

(A) Protein sequence, experimentally-determined secondary structure, ^{1}H-^{15}N hetNOE, and DXMS results (10, 100, and 1000 second exchange durations at pH 7.5 and ~ 0 °C) for cytoplasmic protein Q8ZRJ2 (NESG target StR65) analyzed individually (I) and in a four protein mixture (M). **(B** and **C)** Protein sequence, experimentally-determined secondary structure, ^{1}H-^{15}N hetNOE, and DXMS results (10, 100, and 1000 second exchange durations) for YjcQ protein from *B. subtilis* (NESG target SR346). **(B)** Results at pH 7.5 and ~ 0 °C for the protein analyzed individually (I) and in a four protein mixture (M). **(C)** Results at pH 5.5 and ~ 0 °C (in 20 mM ammonium acetate, 0.1 M NaCl, 5 mm $CaCl_2$) for the protein analyzed individually (I).