

Northeast Structural Genomics Consortium

Determination of flexible protein regions using proton-deuterium exchange mass spectrometry (DXMS)

Disordered or unstructured regions of proteins, while often very important biologically, can pose significant technical challenges for three-dimensional structure determination of proteins by both X-ray crystallography and NMR spectroscopy. As part of the evolving target construct optimization and salvage pipeline being developed in the NESG, we employ $^1\text{H}/^2\text{H}$ exchange mass spectrometry (DXMS) (Woods Jr & Hamuro, 2001 *J Cell Biochem* S37, 89; Englander, 2006 *J Am Soc Mass Spectrom* 17, 1481) together with *in silico* prediction methods (i.e., the DisMeta Server; www-nmr.cabm.rutgers.edu/bioinformatics/disorder) for the rapid identification of disordered segments of proteins. DXMS combines the structural and dynamical insights obtained from protein $^1\text{H}/^2\text{H}$ exchange with the minute sample requirements and sequence specific information of mass spectrometry. The technique revolves around the concept that backbone amide protons in solvent-inaccessible ordered regions of a protein will exchange with solvent deuterium at a slower rate than those in flexible, disordered regions. The results are depicted in a so-called “heat map”, in which $^1\text{H}/^2\text{H}$ exchange rates are represented by colors ranging from blue (slow amide proton exchange) to red (fast amide proton exchange). DXMS has been applied in the JCSG project to the design of protein constructs with improved crystallization success as compared to the full-length proteins (Pantazatos *et al.*, 2004 *Proc Natl Acad Sci U S A* 101, 751; Spraggon *et al.*, 2004 *Protein Sci* 13, 3187). In a recent pilot study on selected targets from the NESG (Sharma *et al.*, 2009 *Proteins*, epub ahead of print Feb. 2, 2009), we demonstrated the feasibility of using DXMS to design truncated constructs yielding NMR spectra that are more amenable for NMR structure determination, while maintaining the structural integrity of the remaining ordered region of the protein. In addition, we demonstrated that the throughput of the DXMS process can be increased by analyzing mixtures of up to four proteins simultaneously without reducing the sequence coverage for each protein. The DXMS optimization of *E. coli* YiaD (NESG target ER553) serves as a representative example of how the technique can yield dramatic improvements to the quality of ^1H - ^{15}N HSQC NMR spectra, and provide structures of protein targets that could not otherwise be studied (Figure 1).

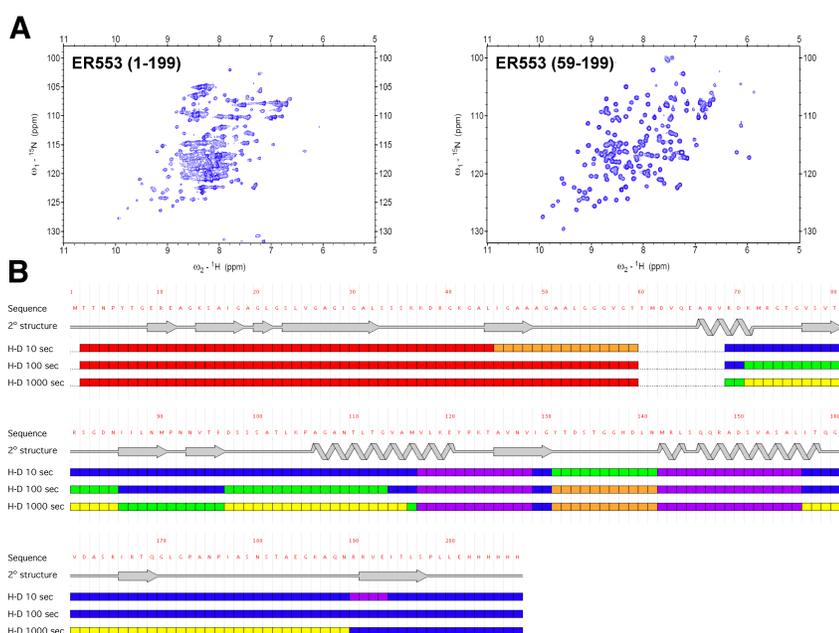


Fig. 1. DXMS-based construct optimization of *E. coli* yiaD (NESG target, ER553). **(A)** ^1H - ^{15}N HSQC spectra (20 °C) of full length (left) and construct optimized (right) *E. coli* yiaD. **(B)** DXMS results for full-length *E. coli* yiaD (10, 100, and 1000 second $^1\text{H}/^2\text{H}$ exchange durations). Note that, in contrast to the *predicted* secondary structure for the protein, the DXMS results reveal that the N-terminal ca. 60 residues of the protein are disordered; these residues may become ordered, for example, in the presence of a binding partner. The solution NMR structure of a DXMS-optimized construct (right) was ultimately solved by NESG consortium (Ramelot *et al.*, DOI 10.2210/pdb2k1s/pdb). (From Sharma *et al.*, 2009, *Proteins*, epub ahead of print Feb. 2, 2009).

We are currently investigating strategies for accelerating our DXMS protocol (i.e., robotic sampling handling; automated data analysis) with the goal of establishing DXMS as a high-throughput approach for construct optimization of partially disordered proteins selected for NMR structure determination in structural biology and structural genomics projects.