^{13}C, ^{15}N and ^{1}H backbone and side chain chemical shift assignment of acid-stress bacterial chaperone HdeA at pH 6

K.A. Crowhurst

Department of Chemistry and Biochemistry
California State University Northridge
18111 Nordhoff St., Northridge CA 91330-8262

Corresponding author:

karin.crowhurst@csun.edu

Phone: 818-677-4288

Fax: 818-677-4068

Abstract

HdeA is a small chaperone found in the periplasm of several common pathogenic bacteria (*E. coli*, *S. flexneri* and *B. abortus*) which are the leading causes of dysentery worldwide, especially in developing countries. Its job is to protect other periplasmic proteins from aggregating as the bacteria pass through the low pH environment of the human stomach on their way to infect the intestines. HdeA is an inactive folded dimer at neutral pH, but becomes a disordered active monomer at pH <3. To initiate NMR characterization of HdeA at pH 6, 94% of the backbone and 86% of the side chain chemical shifts have been assigned. The loop linking helices B and C remains largely unassigned due to missing peaks in the ^{1}H-^{15}N HSQC and other spectra, most likely due to intermediate timescale chemical exchange. Many of the weakest intensity backbone peaks correspond to residues that surround this loop within the tertiary structure. Assignment experiments have therefore helped to provide preliminary clues about the region of the protein that may be most responsible for initiating unfolding as the pH drops, and constitute an important first step in improving our understanding of, and ultimately combatting, HdeA activity.

Keywords: HdeA; *E. coli*; chaperone; NMR; chemical shift assignment

Biological Context

All enteric bacteria travel through the highly acidic environment of the stomach before arriving at and infecting the intestinal walls of the host organism. In Gram-negative bacteria the periplasmic space rapidly experiences drastic changes in the pH of the surrounding environment because its outer membrane contains porins that provide easy entry to molecules <600 Da (including protons). Thus, periplasmic proteins are far more exposed to changes in environmental conditions compared to cytoplasmic proteins (Koebnik et al. 2000) and are consequently far more likely to unfold and/or aggregate after sudden pH changes.

The World Health Organization estimates that the pathogenic bacteria *Escherichia coli, Shigella flexneri* and *Brucella abortus* are responsible for at least 120 million cases of dysentery and more than 600,000 deaths each year, with a significantly higher proportion of cases in children under five and in developing countries with poor sanitation and contaminated water (Foit et al. 2013). All three of these strains have evolved the ability to survive for several hours at extremely low pH (Hong et al. 2005). This is primarily due to the presence of HdeA and HdeB, which are very small ATP-independent chaperone proteins found in the periplasm of these bacteria. There is a clear inverse correlation between the ability to survive highly acidic environments and the dose required to cause infection in the host: dysentery can be induced with very small doses of *Shigella* (which utilizes HdeA / HdeB), while infection by *Salmonella* (which does not carry the HdeA / HdeB genes) requires much larger doses (Waterman et al. 1996; Hong et al. 2012).

HdeA, at 89 residues, is one of the smallest known chaperones, second only to family member HdeB (79 residues). Its mode of action is quite intriguing, in that its inactive form at neutral pH is a folded dimer, whereas its chaperone activities are triggered when the surrounding environment transitions to pH <3 and the protein becomes partially disordered and monomeric (Hong et al. 2005; Tapley et al. 2009; Foit et al. 2013). In such an acidic environment HdeA can bind to a wide variety of other periplasmic proteins via the hydrophobic residues that previously formed its dimer interface (Gajiwala et al. 2000; Hong et al. 2005; Malki et al. 2008; Hong et al. 2012). Its job is to protect other proteins from misfolding and aggregating as the cell transitions through the harsh environment of the stomach and into the neutral environment of the intestines. Once the bacteria enter the intestinal tract, HdeA releases these proteins and refolds into its inactive dimer conformation (Tapley et al. 2009).

As the first step in our effort to characterize the biophysical properties of HdeA by NMR spectroscopy, chemical shift assignment of the folded dimer at pH 6 has been performed. The

dimer shows unexpected evidence for chemical exchange in and around the loop that connects helices B and C, evidenced by very weak and missing peaks in the ^{1}H-^{15}N HSQC and other spectra. These experiments have therefore provided not only chemical shift assignments, but also preliminary evidence that the region around this loop might be the site of unfolding initiation when the pH is lowered. Since HdeA plays a critical role in the level of pathogenicity of several species of bacteria, it has become an important target for drug design; characterizing both active and inactive states of the protein is therefore important for improving our understanding of its mechanism of action in transitioning from one state to the other.

Methods and Experiments

Protein expression and purification

Expression and purification protocols for HdeA (derived from *E. coli*) were modeled, in part, on the methods outlined in Tapley et al. (2009). Briefly, a transformation was performed in which the hdea gene in pET21a vector (obtained from James Bardwell, University of Michigan, Ann Arbor) was inserted into BL21(DE3) *E. coli* cells. One liter of minimal medium culture, which included 1 g $^{15}NH_4Cl$, 2 g ^{13}C glucose and a supplement of 5% v/v $^{13}C/^{15}N$ BioExpress (Cambridge Isotope Laboratories), was grown at 37 ^{o}C to OD_{600} ~ 1 and then induced by the addition of IPTG to a final concentration of 1 mM. The culture was grown at 37 ^{o}C for three additional hours before the cells were harvested.

The cell pellet was resuspended in periplasmic buffer (50 mM tris, pH 7.5, 50 mM NaCl, 1 mg/mL polymyxin sulfate) and stirred at 4 ^{o}C for one hour, spun at 15,000 x g for 20 minutes, then dialyzed overnight in Q Sepharose buffer A (20 mM tris, pH 8.0, 0.5 mM EDTA) (Tapley et al. 2009). The sample was filtered and injected onto a Hi-Trap Q Sepharose fast-flow anion exchange column (GE Healthcare) and eluted over a linear gradient from 0 to 0.3 M NaCl in 20 mM tris, pH 8.0, 0.5 mM EDTA. Fractions containing crude HdeA were pooled and dialyzed

overnight into SP Buffer A (20 mM sodium acetate, pH 4.0, 0.5 mM EDTA). The sample was then filtered and injected onto a Hi-Trap SP Sepharose fast-flow cation exchange column (GE Healthcare) pre-equilibrated with SP Buffer A. HdeA protein was eluted using SP Buffer B, a buffer at a higher pH (no gradient), which contained 50 mM HEPES, pH 7.0, 2 mM EDTA.

This protocol yielded ~14 mg of pure $^{13}C/^{15}N$ HdeA. The sample was concentrated and then dialyzed into NMR buffer. The final NMR sample contained 1.4 mM $^{13}C/^{15}N$ HdeA in 50 mM bis-tris, pH 6.0, 2 mM NaN_3, 0.2 mM DSS, 10% D_2O.

Assignment experiments

All experiments were recorded at 25 ^oC on a 600 MHz Agilent DD2 spectrometer with a room-temperature probe. Assignment experiments included 2D ^{15}N HSQC, ^{13}C aliphatic HSQC, ^{13}C aromatic HSQC, Hb(CbCgCd)Hd, and Hb(CbCgCdCe)He spectra, as well as 3D HNCaCb, CbCa(CO)NH, HNHA, HNCO, HCC-TOCSY-NH, CCC-TOCSY-NH, HCCH-TOCSY, aromatic HCCH-TOCSY, ^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC and aromatic NOESY experiments. Each NOESY experiment utilized a 75 ms mixing time.

Data were processed using NMRPipe v.7.2 (Delaglio et al. 1995) and analyzed with NMRViewJ v.8.2.29 (Johnson 2004).

Assignments and Data Deposition

Backbone assignments for HdeA are 94% complete (without taking into account the amide nitrogens from the four Pro residues or from the N-terminus). This average is comprised of 94.0% amide nitrogen and hydrogen, 94.6% Hα, 96.6% Cα and 89.8% carbonyl carbon atoms assigned. An annotated $^1H-^{15}N$ HSQC is provided (Fig. 1); 79 backbone amide peaks are observed in the spectrum out of a maximum of 84 expected resonances, with additional peaks coming from side chain Trp NH as well as Asn and Gln NH_2 groups.

Unassigned backbone atoms result from a stretch of residues with missing amide resonances in the ^{1}H-^{15}N HSQC: A38, N40, K42, D43 and K44, which are located at, or adjacent to, the loop connecting helices B and C (see Fig. 2a). All of these residues have unassigned backbone and (in most cases) side chain resonances. The missing backbone resonances are presumably the result of chemical exchange that is taking place on an intermediate NMR timescale and that has broadened the peaks beyond detection. Interestingly, the crystal structure (obtained at pH 4) reports missing electron density only for residues 42 – 44 in this region (Gajiwala et al. 2000); this is a smaller range than the group of resonances that are missing from the ^{1}H-^{15}N HSQC, where some of the absent (and very weak) resonances correspond to the C-terminal segment of helix B. Fig. 2b shows the wide variation in normalized peak intensities as a function of residue number: residues at the N- and C-termini have significantly more intense peaks than the rest of the protein, consistent with observations of disorder in the crystal structure (Gajiwala et al. 2000) while the peaks from the structured portion of the molecule vary over a 5-fold range of intensities.

Heating (to 35 ^{o}C) or cooling (to 10 ^{o}C) the sample does not help in resolving the missing backbone resonances (spectra not shown): the 35 ^{o}C spectrum contains the same number of peaks as are seen at 25 ^{o}C, but many resonances have clearly shifted. Notably, the spectrum recorded at 10 ^{o}C is characterized by significantly fewer peaks (70 backbone resonances seen at 10 ^{o}C compared to 79 at 25 ^{o}C). Those missing peaks, which are also among the lowest intensity at 25 ^{o}C (see Fig. 2b), correspond to residues surrounding the loop that connects helices B and C, suggesting that a larger but very specific region of HdeA is undergoing intermediate timescale chemical exchange at this lower temperature. It is improbable that these missing peaks are caused by monomer-dimer exchange, since the equilibrium dissociation constant (K_d) of HdeA at pH 6 is approximately 0.5 μM (Gajiwala et al. 2000). Taken together, these experiments indicate that even the folded HdeA dimer is exchanging between multiple

conformations and they help to provide early clues into the regions of the protein that are most likely to initiate dimer dissociation and/or unfolding when the pH is lowered.

Assignments of side chain resonances (including Cβ and Hβ) were 86% complete. The following were not included in the calculation of the total number of side chain atoms: labile protons and their associated heteroatoms (such as amine groups on lysines or hydroxyl groups on serines) with the exception of NH_2s on Asn and Gln, methyl groups at the end of Met side chains, and carbons that have no attached hydrogens. Many of the unassigned resonances are a result of our inability to assign backbone amides in the region encompassing residues 38 – 44 (as discussed above), while others are located at the *i*-1 position relative to prolines. Residues F21 and K77 are two notable positions for which side chain assignments could not be completed due to missing resonances in the NOESY and TOCSY spectra; however, the reason cannot be traced back to the loop region between helices B and C or to their proximity to prolines in the primary sequence.

These chemical shift assignments constitute the first step in the characterization of protein structure and dynamics in both the folded inactive dimer of HdeA at neutral pH and eventually the disordered active monomer at pH <3. Atomic resolution studies such as these will provide critical insight into the mechanism by which HdeA becomes activated; improved understanding of this process may aid in the development of therapeutics to inhibit HdeA activity and thereby eliminate or significantly decrease the pathogenicity of some of the more widespread and dangerous dysentery-causing bacteria.

Backbone and side chain assignments of HdeA at pH 6 (as well as time-domain data) have been deposited online at the BMRB (accession number 19165).

Acknowledgements

The author would like to thank Dr. James Bardwell (University of Michigan) for kindly providing the HdeA-expressing plasmid, as well as the Department of Chemistry & Biochemistry and the

College of Science and Mathematics at CSU Northridge for research support. The experiments were performed on a 600 MHz NMR spectrometer generously funded by NSF MRI grant CHE-1040134.

Conflict of Interest

The author declares that she has no conflict of interest in the publication of this manuscript.

References

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6: 277-293.

Foit L, George JS, Zhang BW, Brooks CL, 3rd, Bardwell JC (2013) Chaperone activation by unfolding. Proc Natl Acad Sci U S A 110: E1254-1262. doi: 10.1073/pnas.1222458110

Gajiwala KS, Burley SK (2000) HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. J Mol Biol 295: 605-612. doi: 10.1006/jmbi.1999.3347

Hong W, Jiao W, Hu J, Zhang J, Liu C, Fu X, Shen D, Xia B, Chang Z (2005) Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. J Biol Chem 280: 27029-27034. doi: 10.1074/jbc.M503934200

Hong W, Wu YE, Fu X, Chang Z (2012) Chaperone-dependent mechanisms for acid resistance in enteric bacteria. Trends Microbiol 20: 328-335. doi: 10.1016/j.tim.2012.03.001

Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278: 313-352. doi: 10.1385/1-59259-809-9:313

Koebnik R, Locher KP, Van Gelder P (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol Microbiol 37: 239-253.

Malki A, Le HT, Milles S, Kern R, Caldas T, Abdallah J, Richarme G (2008) Solubilization of protein aggregates by the acid stress chaperones HdeA and HdeB. J Biol Chem 283: 13679-13687. doi: 10.1074/jbc.M800869200

Tapley TL, Korner JL, Barge MT, Hupfeld J, Schauerte JA, Gafni A, Jakob U, Bardwell JC (2009) Structural plasticity of an acid-activated chaperone allows promiscuous substrate binding. Proc Natl Acad Sci U S A 106: 5557-5562. doi: 10.1073/pnas.0811811106

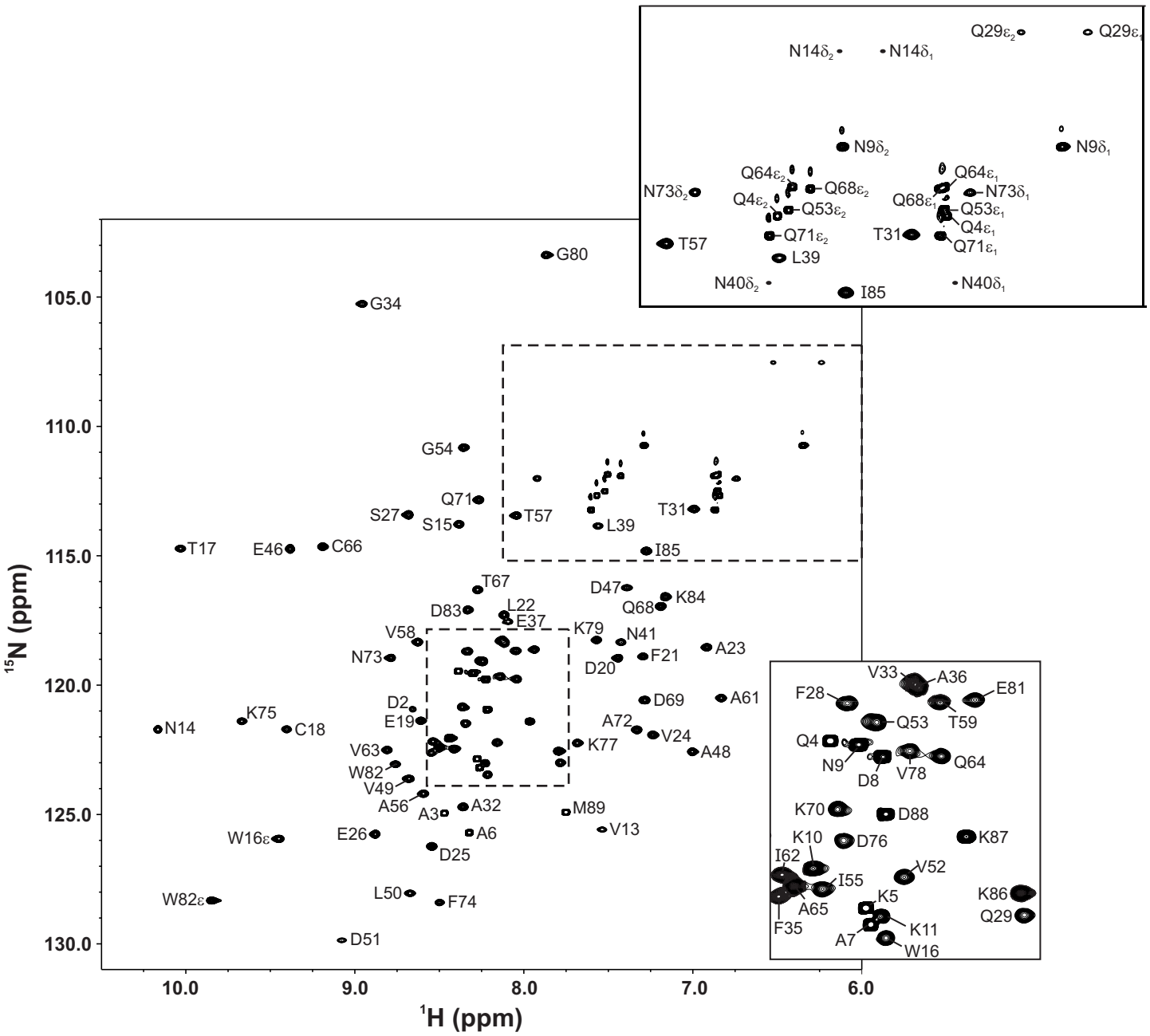
Waterman SR, Small PL (1996) Identification of sigma S-dependent genes associated with the stationary-phase acid-resistance phenotype of Shigella flexneri. Mol Microbiol 21: 925-940.

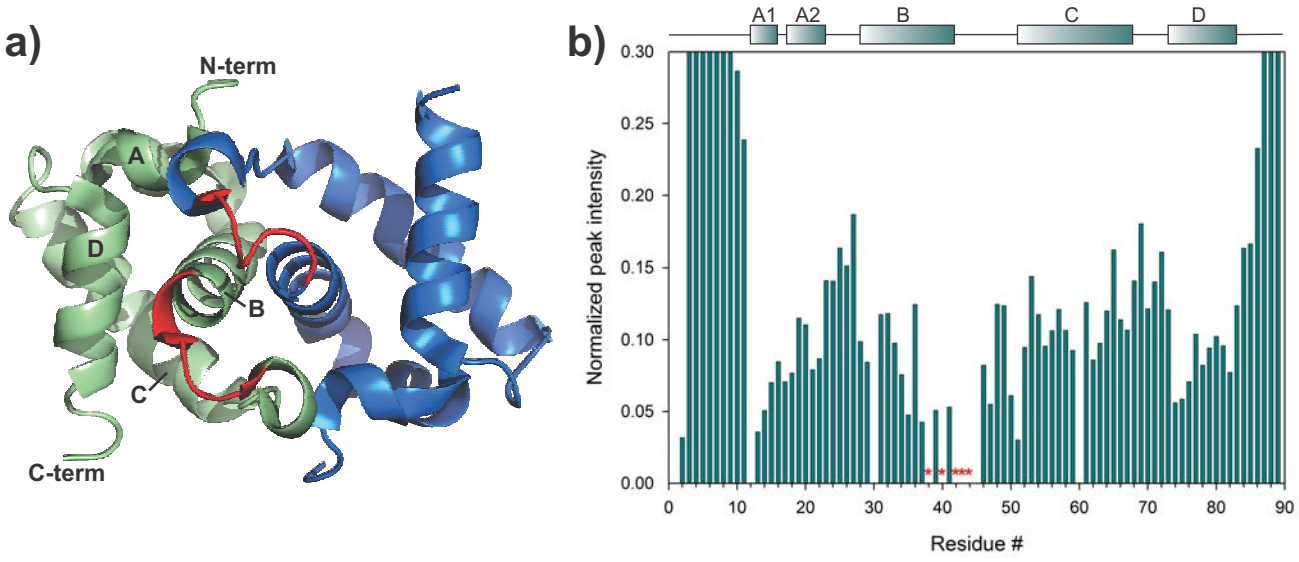
Figures

Fig. 1 ^{1}H-^{15}N HSQC spectrum of 1.4 mM $^{13}C/^{15}N$ HdeA in a 50 mM bis-tris buffer, pH 6, recorded at 25 ^{o}C on a 600 MHz NMR spectrometer. Each peak assignment is labeled with a one-letter code and residue number. Side chain NH resonances of tryptophans are additionally labeled with ε, while side chain NH_2s from Asn and Gln residues are indicated with $\delta_1/\ \delta_2$ and $\varepsilon_1/\ \varepsilon_2$, respectively. Insets at the upper and lower right side show enlargements of the side chain amide region and center of the spectrum, respectively, in order to clearly show assignment labels. Nearly complete backbone and side chain resonance assignments can be found online at the BMRB repository (BMRB accession number 19165)

Fig. 2 Mapping peak intensities with HdeA structure. **a)** Ribbon structure of the HdeA dimer (green and blue, PDB ID: 1DJ8) showing the location of amide peaks missing from the ^{1}H-^{15}N HSQC (red). Helices A – D as well as the N- and C-termini are labeled on one monomer. Note that only residues 9 – 87 are shown in this crystal structure (using crystals grown at pH 4); residues 1 – 8 and 88 – 89 are disordered and therefore not visible in the electron density map.

b) Plot of normalized amide peak intensities (from spectra recorded at pH 6) as a function of HdeA monomer residue number. The Y-axis scale is cut off at 0.3 in order to clearly show peak intensities in the structured portion of the protein sequence (residues 10 – 86). Positions marked with red stars (38, 40, 42-44) correspond to non-proline residues that have no visible amide peaks in the ^{1}H-^{15}N HSQC. The positions of helices with respect to the protein sequence are indicated with rectangles above the graph


Q29ε2
Q29ε1
N14δ2
N14δ1
N9δ2
N9δ1
Q64ε2
Q68ε2
Q64ε1
N73δ2
Q4ε2
Q53ε2
Q68ε1
N73δ1
Q53ε1
T31
Q4ε1
Q71ε2
T57
Q71ε1
L39
N40δ2
I85
N40δ1
G80
G34
G54
Q71
S27
T57
T31
S15
L39
T17
E46
C66
I85
T67
D47
K84
D83
L22
Q68
E37
K79
N41
V58
F21
A23
N73
D20
D69
A61
D2
K75
E19
A72
V24
N14
C18
K77
A48
V63
W82
V49
A56
A3
A32
M89
W16ε
E26
A6
V13
D25
W82ε
L50
F74
D51
V33
A36
E81
F28
Q53
T59
Q4
N9
Q64
D8
V78
K70
D88
K10
D76
K87
I62
V52
I55
K5
K86
A65
Q29
F35
K11
A7
W16
105.0
110.0
115.0
120.0
125.0
130.0
10.0
9.0
8.0
7.0
6.0
15N (ppm)
1H (ppm)


a)
N-term
A
D
B
C
C-term
b)
A1
A2
B
C
D
Normalized peak intensity
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
80
90
Residue #