Biomol NMR Assign (2011) 5:1–5
DOI 10.1007/s12104-010-9252-7

ARTICLE

Complete ^{1}H, ^{13}C and ^{15}N NMR assignments for donor-strand complemented AafA, the major pilin of aggregative adherence fimbriae (AAF/II) from enteroaggregative *E. coli*

Yi Yang · Andrea A. Berry · Wei-Chao Lee · James A. Garnett · Jan Marchant · Jonathan A. Levine · Peter J. Simpson · Sarah A. Fogel · Kristen M. Varney · Steven J. Matthews · James P. Nataro · Keith G. Inman

Received: 1 April 2010 / Accepted: 2 August 2010 / Published online: 17 August 2010

Abstract Aggregative adherence fimbriae (AAF) are the primary adhesive factors of enteroaggregative *Escherichia coli* (EAEC) and are required for intestinal colonization. They mediate binding to extracellular matrix proteins of the enteric mucosa and display proinflammatory effects on epithelial cells in vitro. Among the simplest of bacterial fimbriae, these passive hairlike appendages are composed primarily of a single 16-kDa structural and adhesive subunit, AafA. Oligomerization occurs by incorporating the N-terminal strand of each AafA subunit into an otherwise incomplete β-sheet of an adjacent AafA subunit. We have engineered a highly soluble AafA monomer by positioning the N-terminal "donor strand" at the C-terminus, following a turn and short linker that were introduced to allow access of the donor strand to the recipient cleft of the same subunit. The resulting "donor-strand complemented" AafA subunit, or AafA-*dsc* folds autonomously, is monodisperse in solution, and yields high quality NMR spectral data. Here, we report the ^{1}H, ^{13}C, and ^{15}N chemical shift assignments for AafA-*dsc*.

Keywords AafA · Aggregative adherence fimbriae · Adhesin · Donor strand · Enteroaggregative · NMR resonance assignment

Biological context

Enteroaggregative *Escherichia coli* (EAEC) are a clinically important cause of diarrhea in both developing and industrialized countries. Infection can lead to watery diarrhea that is often persistent and inflammatory (Okeke and Nataro 2001). AAF-mediated attachment to host epithelium is a required first step in the infection process. The four known AAF alleles, AAF/I through AAF/IV share little homology and are antigenically distinct, yet all four have been implicated in mediating the same "stacked brick" pattern of HEp-2 cell adherence that defines the enteroaggregative pathotype. Studies of AAF/II mutants suggest that they have important roles in pathogenesis including induction of IL-8 release, disruption of tight junctions, and binding to extracellular matrix proteins (Farfan et al. 2008; Harrington et al. 2005). The major structural subunit (or *pilin*) of AAF/II is the protein AafA (GenBank: AAB82330.1), which is also the principal adhesin of EAEC strains expressing this allele. AafA has been shown to bind fibronectin, laminin, and type IV collagen in vitro but not BSA or type I collagen. Evidence

Y. Yang · W.-C. Lee · J. A. Garnett · J. Marchant · P. J. Simpson · S. J. Matthews
Division of Molecular Biosciences, Department of Life Sciences, Imperial College London, South Kensington, London SW7 2AZ, UK

A. A. Berry · J. A. Levine · S. A. Fogel · J. P. Nataro
Division of Infectious Diseases and Tropical Pediatrics, University of Maryland School of Medicine, Center for Vaccine Development, 685 W. Baltimore St. Room 480, Baltimore, MD 21201, USA

K. M. Varney
Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, 685 W. Baltimore St. Room 480, Baltimore, MD 21201, USA

K. G. Inman (✉)
Department of Clinical Investigation, Walter Reed Army Medical Center, Building 7 Room 203, 6900 Georgia Avenue, NW, Washington, DC 20307, USA
e-mail: inman.kg@gmail.com

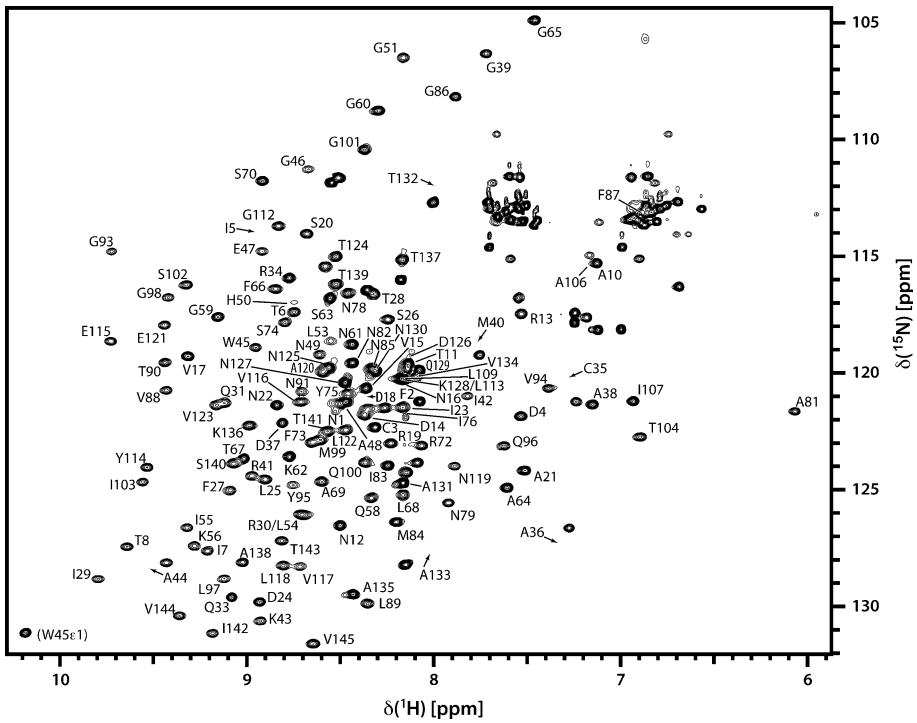

Fig. 1 Assigned 2D ^{1}H-^{15}N HSQC spectrum of AafA-*dsc* recorded with 4 transients at 500 MHz using ^{15}N, ^{13}C-labeled AafA-*dsc* at 308 K. Assigned residues that are too weak to be visible at this level are indicated with *arrowheads*

suggests fibronectin is the primary receptor for EAEC adherence to cultured colonic epithelial cells (Farfan et al. 2008).

AAF are distantly related to the Afa/Dr adhesins, which are expressed by a subset of uropathogenic *E. coli* and diffusely adherent *E. coli* (Boisen et al. 2008). AafA exhibits 44.0% similarity and 14.8% identity to AfaE, the major pilin from AFA-III diffusely adherent *E. coli* that assembles into a flexible fimbrial structure. Despite the low identity between AafA and AfaE-III, significant sequence conservation is found among the chaperone and usher accessory proteins, as well as in the overall structure of the gene cluster. Due to theses similarities, the two systems are generally expected to possess similar macrostructure (Anderson et al. 2004a). Under this model, AAF are comprised of three subunit types: an usher, AafC that anchors the fibril to the outer membrane; a major subunit, AafA that polymerizes to form the fibril; and a minor subunit or tip antigen, AafB, found at its distal end. As with chaperone-usher systems generally, each AafA subunit should form an immunoglobulin fold that lacks the C-terminal or G-strand. Instead, that strand is found at the N-terminus of the mature subunit where it cannot fold into the G-strand cleft but instead occupies that cleft on an adjacent subunit, forming the basis for polymerization and fiber formation.

To make monomeric AafA for solution study, we have chosen an approach that has been successful for AfaE-III and other chaperone-usher systems (Anderson et al. 2004a; Choudhury et al. 1999). Specifically, we moved the N-terminal donor strand of AafA to the C-terminus along with and following an artificial turn and positioning linker that allowed it to complete the fold of the same subunit. The resulting monomeric "donor-strand-complemented" AafA, or AafA-*dsc* yields monodisperse multinuclear NMR spectra with broad chemical shift dispersion as expected for a well-folded protein (Fig. 1). Moreover, AafA-*dsc* retains the principal adhesive properties of the intact fimbriae, binding fibronectin, laminin, and type IV collagen but not BSA or type I collagen, and blocks bacterial binding either to plates coated with those proteins or to cultured colonic epithelial cells, demonstrating its relevance to in vitro studies of EAEC adherence (Farfan et al. 2008). Structural studies of AafA-*dsc* are underway, in support of our overall goal to develop anti-adhesive interventions that prevent or alleviate EAEC infection. We report here the ^{1}H, ^{15}N, and ^{13}C resonance assignments for AafA-*dsc* as a first major step toward determining its structure.

Methods and experiments

Protein expression and purification

The design of AafA-*dsc* was based on a successful approach used for the related AfaE-III (Anderson et al. 2004b). The first 10 N-terminal amino acids of the native AafA (KTATSTITVV) were removed and inserted at the C-terminus, downstream of a turn (DNKQ) and six additional

residues from the uncleaved AafA N-terminus (NATAVA), added to position the donor strand and allow autonomous folding of the subunit. A synthetic *aafA-dsc* gene was inserted into plasmid pQE-30 as an N-terminal hexahistidine-tagged fusion which was then expressed in *E. coli* strain M15[pREP4] (Qiagen, Gaithersburg, MD). Cultures were grown in either lysogeny broth or M9 minimal medium containing either $^{15}NH_4Cl$ or both that and ^{13}C-glucose (Cambridge Isotope Laboratories, Andover, MA). Media were supplemented with 50 $\mu g \cdot mL^{-1}$ carbenicillin or 50–100 $\mu g \cdot mL^{-1}$ ampicillin. Cultures were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37°C for 4–6 h, after which cells were collected by centrifugation at 6,000×*g* for 20 min, 4°C. Decanted pellets were frozen at −20°C. Cell pellets from 1 L media were resuspended in 20 mL denaturing buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 8 M urea) and lysed by French Press (SLM Instruments, Urbana, IL). The fusion protein was purified by histidine affinity chromatography using NTA-agarose resin loaded with Ni^{+2} (Qiagen) under denaturing conditions. The purified protein was dialysed into phosphate buffer containing 1 M urea (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 M urea) and then into 50 mM acetate buffer with 50 mM NaCl at pH 5.0. Refolded protein was concentrated to 0.5 mM for the NMR experiments.

NMR spectroscopy

NMR samples were prepared in 90% H_2O/10% D_2O at pH 5.0. Spectra were recorded at 308 K on Bruker DRX500, DRX600, Avance II 800 MHz, and Avance II 950 MHz

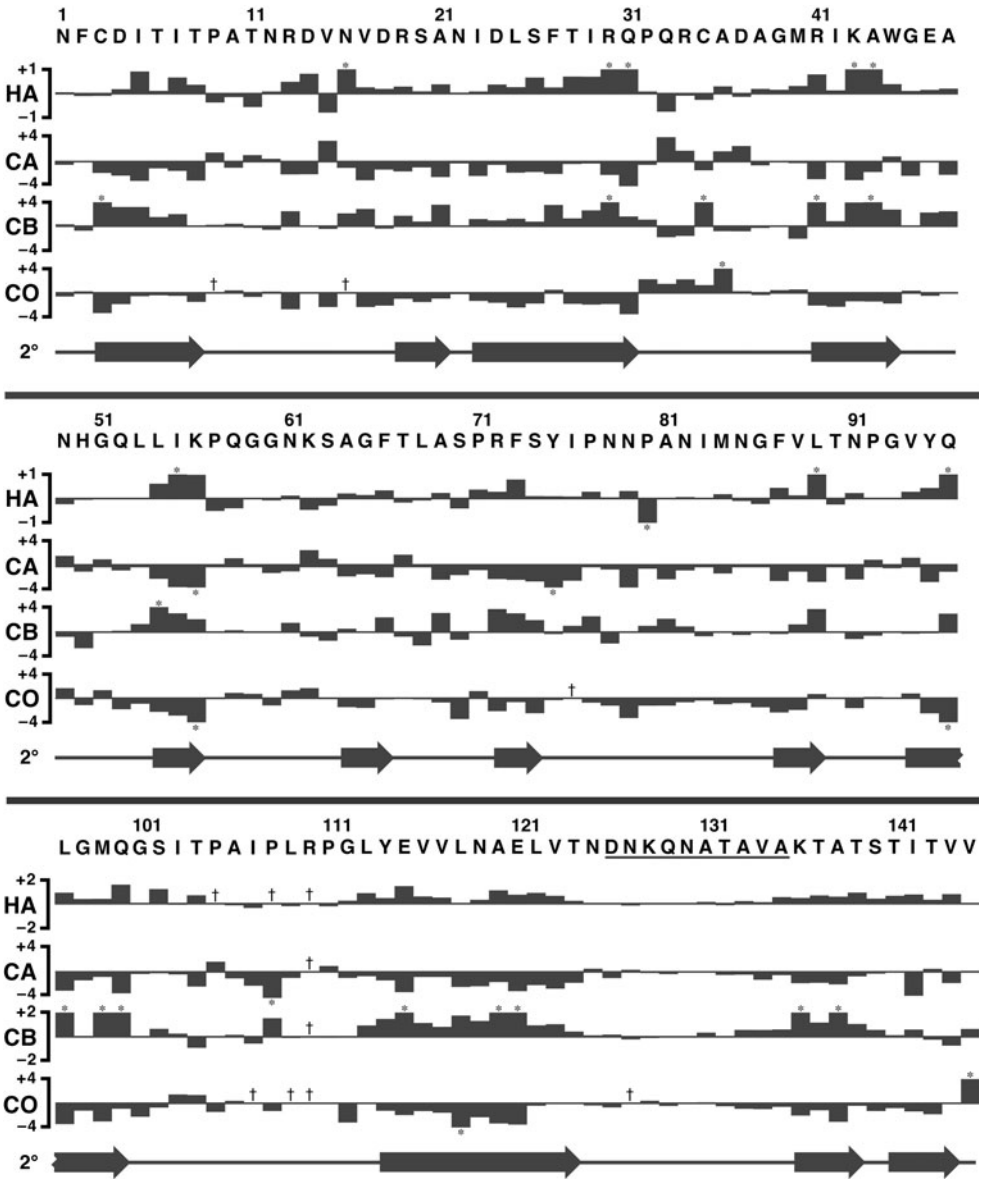

Fig. 2 Secondary structure predictions for AafA-dsc. Chemical shift deviations from random coil values are shown for $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, and backbone carbonyl ^{13}C. *Arrows* indicate probable β-strands based on TALOS predictions and NOE correlations. Residues that were introduced between the native C-terminus and the displaced donor strand are *underlined*. Deviations larger than ranges shown are indicated with an *asterisk* (*). Unassigned nuclei are marked with a *dagger* (†)

spectrometers equipped with triple resonance, z-axis pulsed-field gradient cryoprobes. The following experiments were performed (reviewed in Sattler et al. (1999) unless otherwise noted): CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, (H)CCH-TOCSY, H(C)CH-TOCSY, HBHA(CBCAC-O)NH (Grzesiek and Bax 1993), (H)CCCONH (Celda and Montelione 1993), ^{15}N-edited NOESY-HSQC and ^{13}C-edited HMQC-NOESY. Assignments were extended by resolving overlap using 4D ^{13}C/^{15}N-edited HMQC-NOESY-HSQC data collected at 600 MHz, and 4D ^{13}C/^{13}C-edited HMQC-NOESY-HMQC data collected at 800 MHz on a sample prepared in 100% D_2O. Pulse programs were written in-house or downloaded from the pulse program library at the National Magnetic Resonance Facility at Madison (NMRFAM) or provided by the spectrometer manufacturer. The data were processed with NMRPipe (Delaglio et al. 1995) and analyzed with NMRView (Johnson and Blevins 1994) or Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, Univ. of California, San Francisco, unpublished source).

Extent of assignment and data deposition

Connectivities between residues were established using MARS in an iterative manner, adding or adjusting assignments manually (Jung and Zweckstetter 2004). Figure 1 shows the ^{1}H-^{15}N HSQC spectrum of AafA-*dsc* with assignments labeled. The backbone chemical shifts of two non-proline residues excluding the N terminal affinity tag were not observable in triple resonance spectra, namely Q52 and R110. Furthermore, eight residues appear very weakly on ^{15}N-HSQC spectra, namely I5, D18, C35, A36, M40, A44, T132, and A133, indicated by arrowheads on the figure. Chemical exchange is suspected in regions of missing assignments or weak signal; experiments are currently underway to explore this possibility and the time-scales on which it may be occurring. Assignments were determined for 98% of each of backbone ^{13}C, non-proline backbone ^{15}N, and backbone amide ^{1}H atoms. Sidechain assignments were determined for 95% of ^{1}H atoms attached to carbon, and 83% of ^{13}C atoms (96% excluding those not typically assigned, namely Arg C_ζ, Asp C_γ, Asn C_γ, Glu C_δ, Gln C_δ, His C_γ, Phe C_γ, Trp C_γ, $C_{\delta 2}$, $C_{\varepsilon 2}$, and Tyr C_γ and C_ζ). Chemical shift values have been deposited in the Biological Magnetic Resonance Bank database (http://www.bmrb.wisc.edu/) under accession number 16748.

Secondary structure

Secondary structure was predicted on the basis of chemical shift index (CSI) using TALOS and from sequential NOE correlations observed in 3D and 4D NOE data. The CSI shows that AafA-*dsc* folds to establish 12 β-strands, consistent with the solution structure of AfaE (Anderson et al. 2004b; Wishart and Sykes 1994). Although less than 20% homology is shared between AafA-*dsc* and AfaE-*dsc*, their secondary structure is similar and comprised of β-sheets with similar registries and NOE correlations. Inter-strand NOEs between the N-terminal donor strand and the native C-terminus provide good evidence that the structure is stably folded via donor strand complementation. Differences between the two secondary structures include the absence in AafA-*dsc* of a short 3–10 helix that is seen in AfaE-*dsc*, which may yet become evident in the final structure (Fig. 2).

Acknowledgments KGI acknowledges support from the University of Maryland School of Medicine Department of Pediatrics. The current employer of KGI requires the following statement: "The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army, Department of Defense, or U.S. Government."

References

Anderson KL, Billington J, Pettigrew D, Cota E, Simpson P, Roversi P, Chen HA, Urvil P, du Merle L, Barlow PN, Medof ME, Smith RAG, Nowicki B, Le Bouguenec C, Lea SM, Matthews S (2004a) An atomic resolution model for assembly, architecture, and function of the Dr adhesins. Mol Cell 15:647–657

Anderson KL, Cota E, Simpson P, Chen HA, du Merle L, Le Bouguenec C, Matthews S (2004b) Letter to the editor: complete resonance assignments of a 'donor-strand complemented' AfaE: the afimbrial adhesin from diffusely adherent *E. coli*. J Biomol NMR 29:409–410

Boisen N, Struve C, Scheutz F, Krogfelt KA, Nataro JP (2008) New adhesin of enteroaggregative Escherichia coli related to the Afa/Dr/AAF family. Infect Immun 76:3281–3292

Celda B, Montelione GT (1993) Total correlation spectroscopy (TOCSY) of proteins using coaddition of spectra recorded with several mixing times. J Magn Reson B 101:189–193

Choudhury D, Thompson A, Stojanoff V, Langermann S, Pinkner J, Hultgren SJ, Knight SD (1999) X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. Science 285:1061–1066

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

Farfan MJ, Inman KG, Nataro JP (2008) The major pilin subunit of the AAF/II fimbriae from enteroaggregative Escherichia coli mediates binding to extracellular matrix proteins. Infect Immun 76:4378–4384

Grzesiek S, Bax A (1993) Amino-acid type determination in the sequential assignment procedure of uniformly C-13/N-15-enriched proteins. J Biomol NMR 3:185–204

Harrington SM, Strauman MC, Abe CM, Nataro JP (2005) Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative Escherichia coli. Cell Microbiol 7:1565–1578

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

Jung YS, Zweckstetter M (2004) Mars—robust automatic backbone assignment of proteins. J Biomol NMR 30:11–23

Okeke IN, Nataro JP (2001) Enteroaggregative Escherichia coli. Lancet Infect Dis 1:304–313

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 34:93–158

Wishart DS, Sykes BD (1994) The C-13 chemical-shift index—a simple method for the identification of protein secondary structure using C-13 chemical-shift data. J Biomol NMR 4:171–180