Biomolecular NMR Assignments (2022) 16:247–251
https://doi.org/10.1007/s12104-022-10087-2

Mike P. Williamson
m.williamson@sheffield.ac.uk
1 School of Biosciences, University of Sheffield, Firth Court, Western Bank, S10 2TN Sheffield, UK

\textbf{ARTICLE}

\section*{Abstract}

\textit{Enterococcus faecalis} is a major causative agent of hospital acquired infections. The ability of \textit{E. faecalis} to evade the host immune system is essential during pathogenesis, which has been shown to be dependent on the complete separation of daughter cells by peptidoglycan hydrolases. AtlE is a peptidoglycan hydrolase which is predicted to bind to the cell wall of \textit{E. faecalis}, via six C-terminal repeat sequences. Here, we report the near complete assignment of one of these six repeats, as well as the predicted backbone structure and dynamics. This data will provide a platform for future NMR studies to explore the ligand recognition motif of AtlE and help to uncover its potential role in \textit{E. faecalis} virulence.

\textbf{Keywords} Peptidoglycan \cdot Hydrolase \cdot \textit{E. faecalis} \cdot AtlE \cdot R6

\section*{Biological context}

\textit{E. faecalis} is a leading cause of nosocomial infection, causing life-threatening infections in immunocompromised patients and patients with antibiotic-induced dysbiosis (Arias and Murray 2012; Diekema et al. 2019). The virulence of \textit{E. faecalis} and its resistance to antimicrobials is associated with the presence of a surface rhamnopolysaccharide called the Enterococcal Polysaccharide Antigen (EPA; Xu et al. 2000). EPA is made of a rhamnan backbone substituted by “decorations” that underpin its biological activity (Smith et al. 2019). Although EPA decorations vary between strains, one gene encoding a putative peptidoglycan hydrolase (AtlE) is ubiquitous, suggesting an important contribution to EPA activity (Fig. 1A). AtlE contains a signal peptide, a glycosyl hydrolase family 25 (GH25) catalytic domain, and a C-terminal domain with six conserved repeats (R1-R6) of unknown function (Fig. 1B). The major peptidoglycan hydrolase expressed in \textit{E. faecalis}, AtlA (Qin et al. 1998; Mesnage et al. 2008), also has six C-terminal repeats that are essential for substrate binding and overall catalytic activity (Eckert et al. 2006). The R1-R6 domains of AtlE are therefore likely to also be involved in substrate binding and may ultimately contribute towards the biological activity of AtlE.

Here, we report the complete assignment of one of the conserved repeats from the C-terminal domain of OG1RF AtlE, and the corresponding structure prediction. These NMR assignments represent a step forward towards the identification of the cell wall motif recognised by the R1-R6 domain in AtlE, and of how this domain contributes to \textit{E. faecalis} virulence.

\section*{Methods and experiments}

\textbf{Protein expression and purification}

Due to the high sequence similarity of the OG1RF R1-R5 sequences (Fig. 1C), the most divergent repeat (R6, residues 742–819 of AtlE) was selected for cloning, expression, and purification. The sequence encoding the R6 domain was PCR amplified using oligonucleotides R6_Fw (ATACCATGGCGACAGATCACGAAATACCAGTATGACAACAAC) and R6_Rev (ATGGGATCCATCGTTTTTGGTATCATACAGATCACGAAATACCAGTATGACAACAAC) and cloned into pET2818 using NcoI and BamH1 sites to generate pET2818_R6. The plasmid encodes an 86-residue polypeptide with a hexahistidine tag at the C-terminus.
UK). Gel filtration was performed on the concentrated protein using a Superdex 75 26/200 column pre-equilibrated in buffer C (40 mM phosphate buffer, pH 6.0). Fractions containing R6 were collected and concentrated as described above to a final concentration of 1.2 mM.

**NMR experiments**

All NMR experiments were recorded at 298 K using a Bruker Neo 600 MHz NMR spectrometer with a 5 mm TCI cryo-probe running TopSpin version 4.0.5. NMR experiments were performed in 5-mm NMR tubes containing 1 mM R6, 1 mM trimethylsilylpropanoic acid (TSP), 2 mM sodium azide, and 10% v/v $^2$H$_2$O, in 40 mM phosphate buffer at pH 6.0, with a total volume of 550 µL. Two-dimensional $^{15}$N-$^1$H and $^{13}$C-$^1$H HSQC, and an assortment of three-dimensional NMR experiments, HNCO, HNCA, HNCOCA, HNCA, HNOCACB, HNOCAB, HCCH-TOCSY and
CCH-TOCSY, were performed for the assignment of R6. The assignment of arginine Nε-H side-chain resonances required an additional three-dimensional TOCSY-HSQC experiment, with a mixing time of 120 ms. Using information obtained from the assigned $^{15}$N-$^1$H HSQC spectrum, HA and HB resonances were assigned from a three-dimensional HBHA(CO)NH experiment, which were then used with the HCCH-TOCSY and CCH-TOCSY for sidechain assignments. In all cases, standard Bruker pulse sequences were used. The $^1$H chemical shifts were referenced according to the internal $^1$H signal of TSP resonating at 0.00 ppm. $^{13}$C and $^{15}$N chemical shifts were then referenced indirectly according to nuclei-specific gyromagnetic ratios.

**Extent of assignment and data deposition**

Chemical shifts corresponding to the $^1$H, $^{15}$N, $^{13}$Cα, $^{13}$Cβ, $^{13}$C′ of the R6 backbone were assigned using the standard triple resonance approach (Gardner and Kay 1998). Spectra were processed and analysed using TopSpin version 4.0.2 and FELIX (FELIX NMR, Inc.). The “asstools” assignment program (Reed et al. 2003) was employed to align and match spin systems to the R6 sequence for the assignment of the R6 backbone. $^1$Hα, $^1$Hβ and arginine side chain resonances (Nε-Hε) were assigned manually following the method of Ohlenschläger et al. (1996), using the R6 backbone assignments as reference. Sidechain resonances were assigned using HCCH-TOCSY and CCH-TOCSY experiments.

Figure 2 shows the assigned $^{15}$N HSQC spectrum for the recombinant R6 protein. The spectrum is of high resolution with clear, well-defined peaks. Excluding the “difficult” signals (N-terminal residue and His-tag, non-protonated aliphatic and aromatic C and N, Argη, Lysζ), 98.7% of all backbone $^{15}$N and amide protons were assigned (missing only A2), 100% of all Cα, Cβ, C′, Hα and Hβ backbone signals were obtained, as well as 100% of all asparagine sidechain Nδ, Hδ1 and Hδ2 signals, 100% of all glutamine Nε, Hε1 and Hε2, and 100% of arginine Nε-Hε signals. 89.5% of sidechain signals were assigned, with most of the missing signals being from aromatic rings. Arginine side chain signals (Fig. 2, green) are folded in the nitrogen dimension. The full list of assigned shifts can be found within the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 51184.

The TALOS-N webserver (Shen and Bax 2013) was then used to predict the dynamics and secondary structure of R6 from the reported backbone chemical shifts (Fig. 3 A). The random coil index values (RCI-S2) shown in Fig. 3Ai indicate R6 may have three dynamic regions, specifically at the N-terminus, C-terminus and between residues 54 to 58. Secondary structure prediction (Fig. 3Aii) suggests R6 to contain seven β-sheets (β1: K17-V22, β2: R25-Y27, β3: P38-K40, β4: T44-Y52, β5: R59-T62, β6: G65-T68, β7: V74-K76), with a flexible region predicted to fall between β4 and β5.

Alphafold-2 (AF2) is another protein structure predictor which allows the accurate prediction of protein structures using only the primary sequence (Jumper et al. 2021). In theory this platform could therefore be used in tandem with TALOS-N as a validation method. AF2 predicted a similar overall protein secondary structure for R6, as compared to the TALOS-N output, as shown in Fig. 3B. Whilst eight...
In summary, our results suggest that we have produced a reliable prediction of the dynamics and secondary structure of R6, and that AF2 can be used as a tool to complement chemical shift-based protein structure predictions. The assignments and structural details reported here will be used to explore the binding of this domain to the cell wall, to begin to understand the biological activity of AtlE, and ultimately, its potential contribution to *E. faecalis* virulence.

**Funding** JLD is funded by a BBSRC DTP studentship (grant BB/M011151/1, studentship N°2283067).

**Data Availability** The NMR chemical shift assignments have been deposited at the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu) under the BMRB accession number 51184.
Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The experiments described here comply with current UK laws.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Arias CA, Murray BE (2012) The rise of the Enterococcus: beyond vancomycin resistance. Nat Rev Microbiol 10:266–278. https://doi.org/10.1038/nrmicro2761

Diekema DJ, Hsueh PR, Mendes RE, Pfüller MA, Rolston KV, Sader HS, Jones RN (2019) The microbiology of bloodstream infection: 20-Year trends from the SENTRY antimicrobial surveillance program. Antimicrob Agents Chemother 63:e00355–e00319. https://doi.org/10.1128/AAC.00355-19

Eckert C, Lecerf M, Dubost L, Arthur M, Mesnage S (2006) Functional analysis of AtlA, the major N-acetylglucosaminidase of Enterococcus faecalis. J Bacteriol 188:8513–8519. https://doi.org/10.1128/JB.01145-06

Gardner KH, Kay LE (1998) The use of 1H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct 27:357–406. https://doi.org/10.1146/annurev.biophys.27.1.357

Jumper J, Evans R, Pritzel A et al (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2

La Rosa SL, Snipen LG, Murray BE, Willems RJ, Gilmore MS, Diep DB, Nes IF, Brede DA (2015) A genomic virulence reference map of Enterococcus faecalis reveals an important contribution of phage03-like elements in nosocomial genetic lineages to pathogenicity in a Caenorhabditis elegans infection model. Infect Immun 83:2156–2167. https://doi.org/10.1128/IAI.02801-14

Mesnage S, Chau F, Dubost L, Arthur M (2008) Role of N-acetylglucosaminidase and N-acetylmuramidase activities in Enterococcus faecalis peptidoglycan metabolism. J Biol Chem 283:19845–19853. https://doi.org/10.1074/jbc.M802323200

Ohlenschläger O, Ramachandran R, Flemming J, Gührs KH, Schlott B, Brown LR (1996) NMR secondary structure of the plasminogen activator protein staphylokinase. J Biomol NMR 9:273–286. https://doi.org/10.1023/a:1018678925512

Reed MA, Hounslow AM, Sze KH, Barsukov IG, Hosszu LL, Clarke AR, Craven CJ, Waltho JP (2003) Effects of domain dissection on the folding and stability of the 43 kDa protein PGK probed by NMR. J Mol Biol 330:1189–1201. https://doi.org/10.1016/s0022-2836(03)00625-9

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR, 56, 227–241(2013). https://doi.org/10.1007/s10858-013-9741-y

Smith RE, Salamaga B, Szkuta P, Hajdamowicz N, Prajsnar TK, Bulmer GS, Fontaine T, Kołodziejczyk J, Herry JM, Hounslow AM, Williamson MP, Serror P, Mesnage S (2019) Decoration of the enterococcal polysaccharide antigen EPA is essential for virulence, cell surface charge and interaction with effectors of the innate immune system. PLOS Pathog 15:1007730. https://doi.org/10.1371/journal.ppat.1007730

Qin X, Singh KV, Xu Y, Weinstock GM, Murray BE (1998) Effect of disruption of a gene encoding an autolysin of Enterococcus faecalis OG1RF. Antimicrob Agents Chemother 42:2883–2888. https://doi.org/10.1128/AAC.42.11.2883

Xu Y, Singh KV, Qin X, Murray BE, Weinstock GM (2000) Analysis of a gene cluster of Enterococcus faecalis involved in polysaccharide biosynthesis. Infect Immun 68:815–823. https://doi.org/10.1128/iai.68.2.815-823.2000

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.