Supplemental Material

Molecular structure and peptidoglycan recognition of *Mycobacterium tuberculosis* ArfA (Rv0899).

Yong Yao[1], Neha Barghava[1], Johnny Kim[1], Michael Niederweis[2], Francesca M. Marassi[1 ‖]

[1]Sanford Burnham Medical Research Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
[2]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA.

Table S1. NMR restraints and Structure statistics.[a]

	ArfA-c(D236A)	ArfA-c
Experiment Restraints		
NOE distances	1967	1897
dihedral angles	217	167
HN RDC orientations	114	98
Root Mean Square Deviation (RMSD)		
RMSD from experimental restraints		
NOE (Å)	0.029	0.039
dihedral angle (deg)	0.456	0.478
HN RDC (Hz)	0.523	0.364
RMSD from idealized covalent geometry		
bonds (Å)	0.006	0.006
angles (deg)	0.579	0.630
impropers (deg)	0.899	0.916
Coordinate Precision[b]		
Average RMSD from mean (Å)		
backbone atoms (A)	0.83±0.16	0.77±0.24
all heavy atoms (Å)	1.31±0.17	1.38±0.26
Ramachandran Plot ϕ/ψ Angle Statistics[c]		
most favored regions	90.5%	86.60%
additional allowed regions	8.6%	12.20%
generously allowed regions	0.3%	0.70%
disallowed regions	0.6%	0.50%

[a]Calculated for the 20 lowest-energy structures of a total 200 calculated structures. [b]Calculated for the structured core of the protein from residue 211-326 for ArfA-c(D236A) or residues 211-225 and 237-326 for ArfA-c. [c]Calculated with PROCHECK.

Table S2. Distance restraints used to model the ArfA-c(D236)/MDP complex. [a]

atom 1	atom 2	Å	(+ /-)
`assign (resid 1 and name H8*)`	`(resid 302 and name   HN)`	4.0	2.2 1.5
`assign (resid 4 and name   NZ)`	`(resid 262 and name  OD*)`	4.0	2.2 0.5
`assign (resid 4 and name   NZ)`	`(resid 262 and name    O)`	4.0	2.2 0.5
`assign (resid 4 and name OH*)`	`(resid 319 and name  NH*)`	3.5	1.7 0.5
`assign (resid 4 and name OH*)`	`(resid 277 and name  NH*)`	3.5	1.7 0.5
`assign (resid 4 and name   HA)`	`(resid 269 and name  HD*)`	4.0	2.2 2.0
`assign (resid 4 and name   HA)`	`(resid 269 and name HG2*)`	4.0	2.2 2.0

[a] Distances were derived from observations of the chemical shift mapping and STD experiments. They do not represent experimental NOEs.
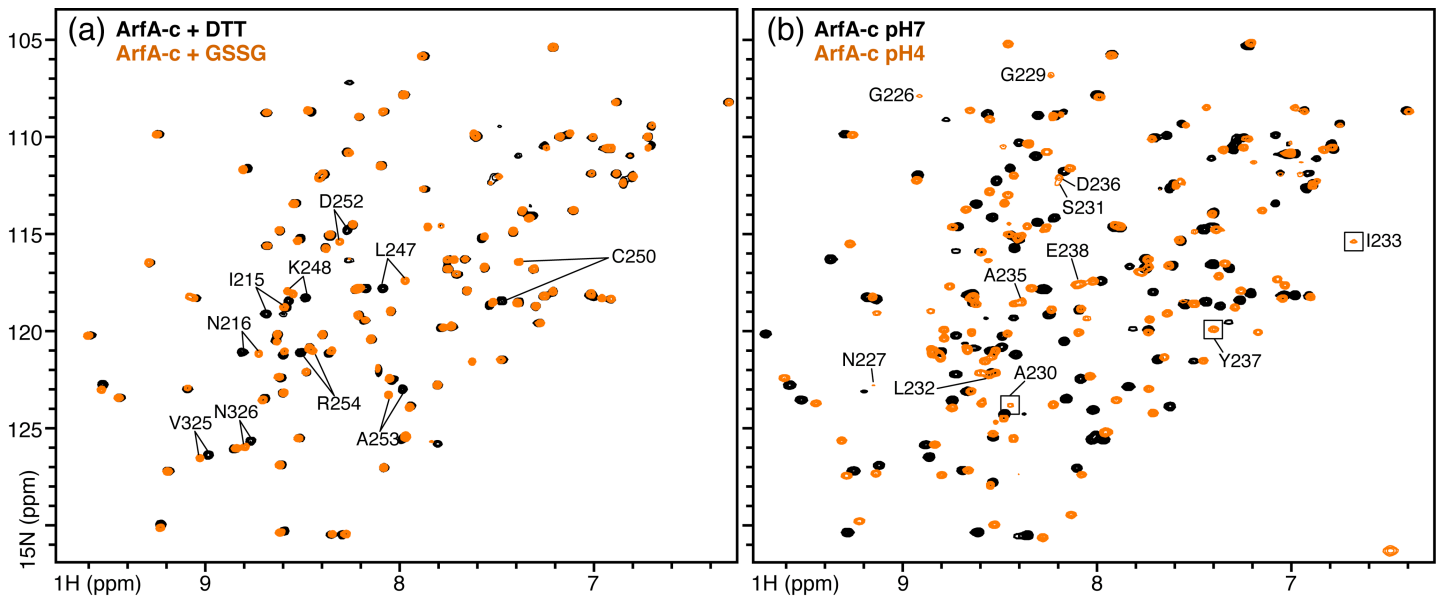
Fig. S1. Effect of C208-C250 disulfide bond and pH on the NMR spectrum of the C domain of ArfA.
All ^{1}H/^{15}N HSQC spectra were obtained at 40°C and pH7. (a) NMR spectra of ArfA-c obtained after forming the disulfide bond with GSSG (orange) or after reducing the disulfide bond with DTT (black). (b) NMR spectra of ArfA-c obtained at pH7 (black) or pH4 (orange). Peaks in boxes are observed at a lower contour level.
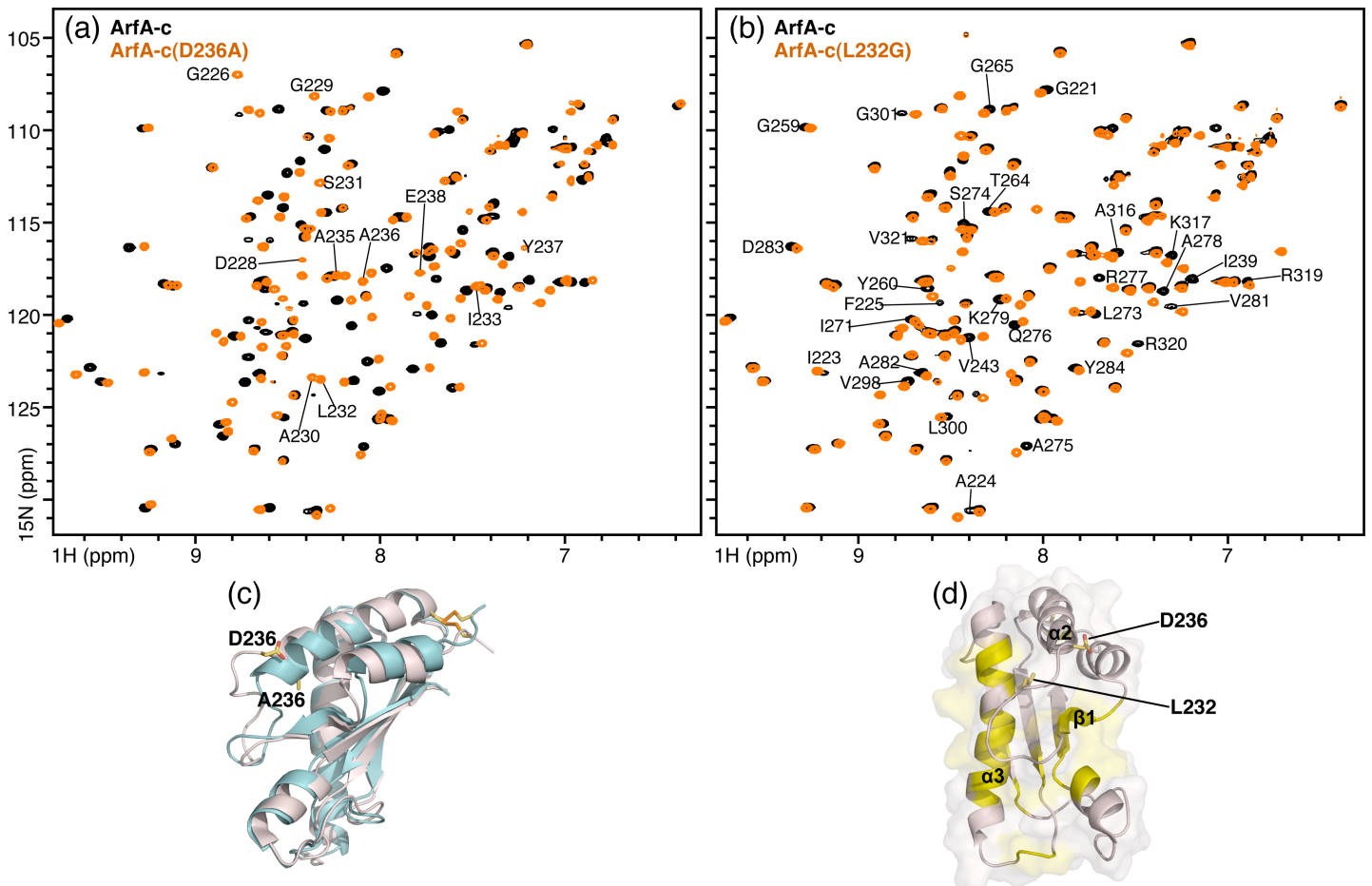

Fig. S2. Effect of D236A and L232G mutations on the NMR spectrum of the C domain of ArfA.
All ^{1}H/^{15}N HSQC spectra were obtained at 25°C and pH7, after treatment with GSSG to form the disulfide bond. (a) Spectra of wild-type ArfA-c (black) and ArfA-c(D236A) (orange). (b) Spectra of wild-type ArfA-c (black) and ArfA-c(L232G) (orange). (c) Superimposed lowest energy structures of ArfA-c (pink) and ArfA-c(D236A) (cyan). (d) Significant chemical shift changes (≥0.03 ppm) due to the L232G mutation map to key residues (yellow) in helix α3 and in the β-sheet of ArfA-c.
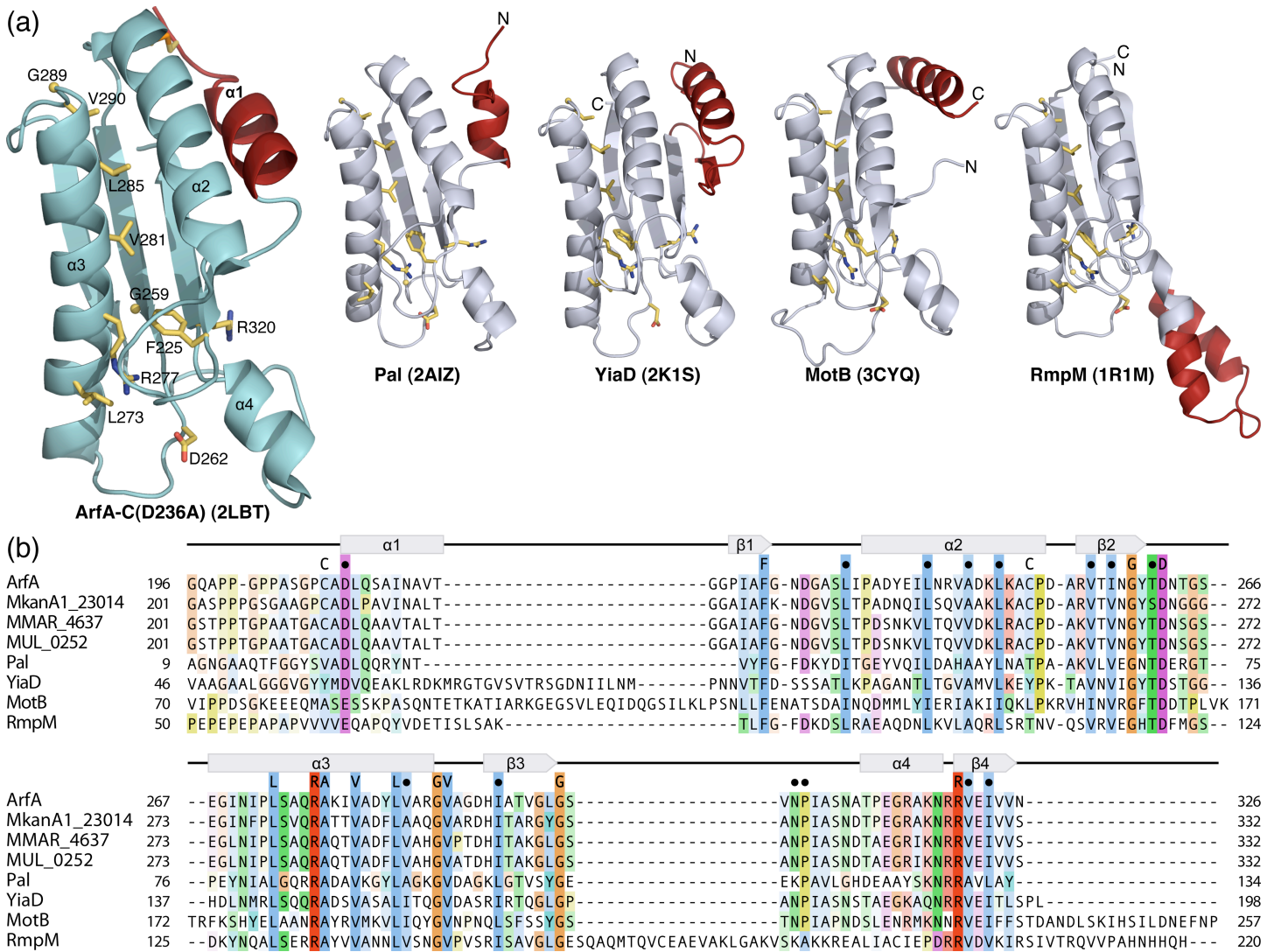
Fig. S3. Structure-based sequence alignment of the C domain of *M. tuberculosis* ArfA with representative OmpA-like domains (pfam00691) of known structure.

(a) Structures of *M. tuberculosis* ArfA-c(D236A); *H. influenzae* Pal [21]; *E. coli* YiaD; *H. Pylori* MotB [22]; and *N. meningitis* Rmpm [23]. PDB accession codes are provided in parentheses for each structure. Structural elements colored red are unique to each protein and not part of the core βαβαβαβ OmpA-like fold. (b) Sequences are for: *M. tuberculosis* ArfA (NP_215414); *M. kansasii* MkanA1_23014 (ZP_04750863); *M. marinum* MMAR_4637 (YP_001852895); *M. ulcerans* MUL_0252 (YP_904461); *H. influenzae* Pal (NP_438542); *E. coli* YiaD (NCBI: YP_003034446); *H. pylori* MotB (NP_207609); and *N. meningitis* Rmpm (NP_273431). Perfectly conserved residues and Cys (single letter), highly conserved residues (●), and secondary structure of *M. tuberculosis* ArfA, are marked above the sequences. Residues with at least 20% conservation are colored with the ClustalX scheme. Citation numbers are from the main text.
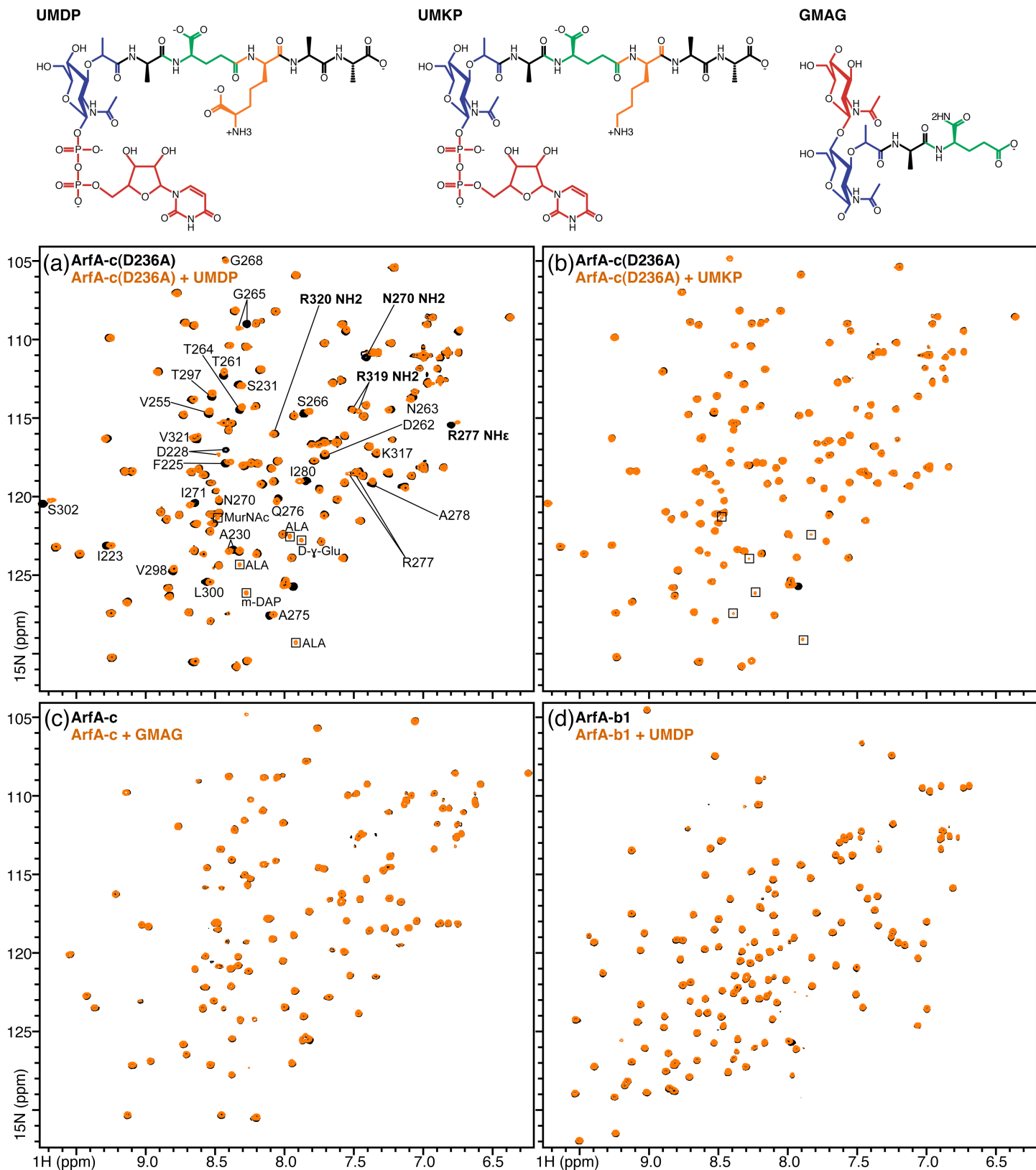
Fig. S4. Effect of peptidoglycan peptide analogs on the NMR spectra of the C domain of ArfA.
^{1}H/^{15}N HSQC spectra of ArfA were obtained at pH7, with (orange) or without (black) ~20 molar equivalents of peptidoglycan peptide analogs. (a) ArfA-c(D236A) ± UMDP at 25°C. Peaks from the natural abundance peptide are enclosed in boxes. (b) ArfA-c(D236A) ± UMKP at 25°C. (c) ArfA-c ± GMAG dipeptide at 40°C. (d) ArfA-b1 ± UMDP at 25°C. The peptide structures are shown at top.
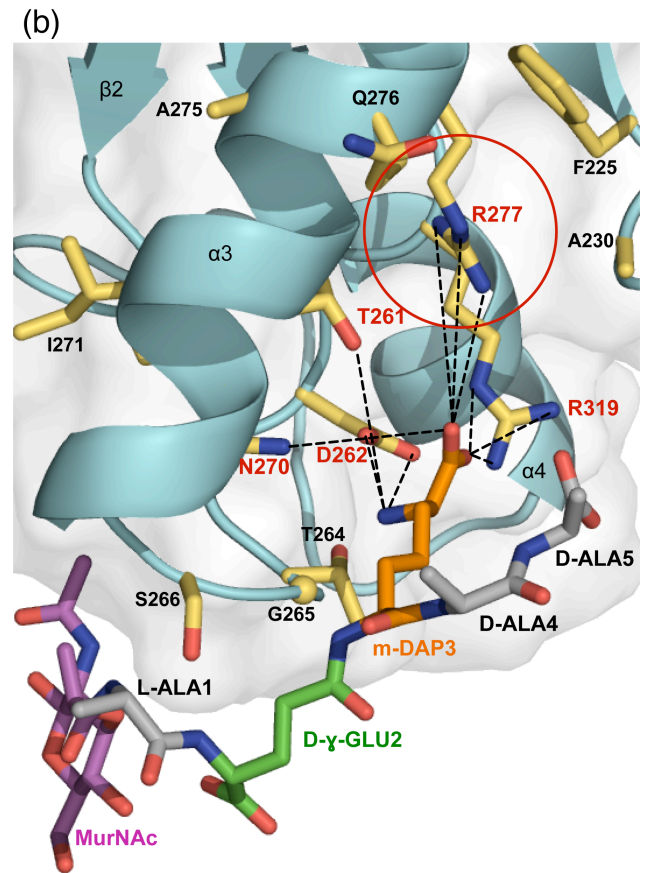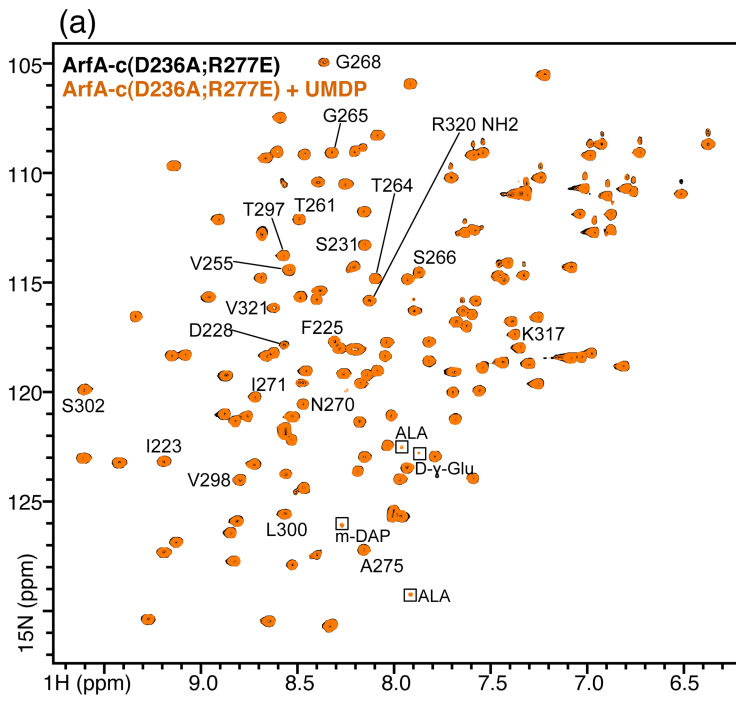
Fig. S5. Effect of R277E mutation on the ability of ArfA-c to bind the peptidoglycan peptide precursor UMDP.
(a) ^{1}H/^{15}N HSQC spectra of ArfA-c(D236A;R277E) double mutant were obtained at pH7 and 25°C, with (orange) or without (black) ~20 molar equivalents of the peptidoglycan peptide precursor UMDP. Peaks from the natural abundance peptide are enclosed in boxes. Labeled peaks serve as comparison to the data in Fig. S4. (b) Close up view of the peptidoglycan binding pocket with structural model of bound MDP (MurNAc–L-Ala–D-γ-Glu–m-DAP–D-Ala–D-Ala) derived from the data. Key residues important for peptidoglycan binding are labeled in red. R277 is in the red circle.
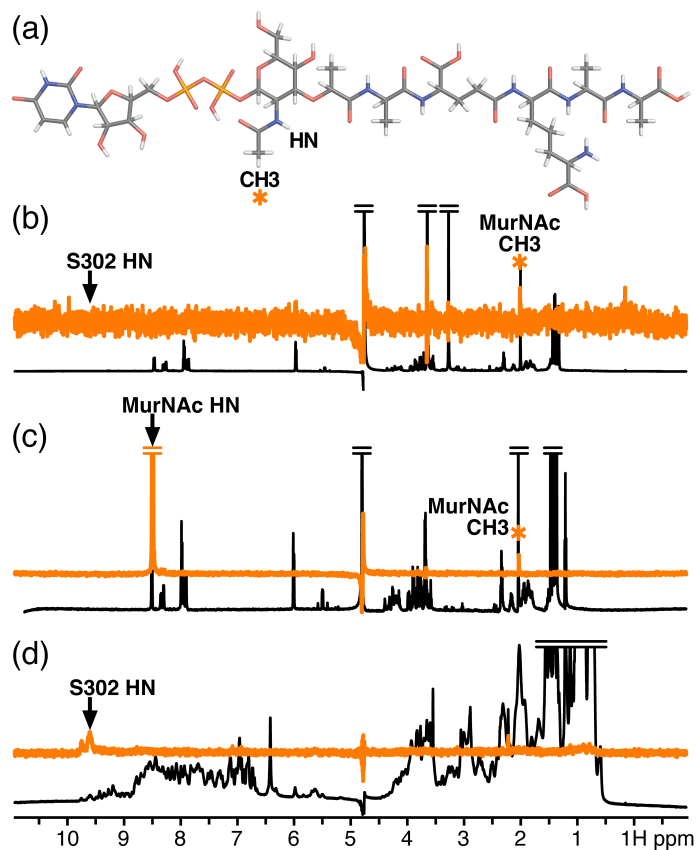
Figure S6. Saturation transfer difference NMR spectra showing the proximity of ArfA-c S302 to the MurNAc N-acetyl methyl group of UMDP.
One-dimensional ^{1}H NMR spectra were obtained without (black) or with (orange) saturation at the frequency designated by the arrows. (a) Chemical structure of UMDP. The MurNAc methyl and amide groups are labeled. (b) Spectra for a mixture of ArfA-c(D236A) and UMDP (1/20 molar ratio). Saturation at the backbone amide HN frequency of S302 transfers to the MurNAc N-acetyl CH3 group (orange asterisk). (c) Spectra for free UMDP. Saturation at the MurNAc HN frequency transfers to the neighboring N-acetyl CH3 group (orange asterisk). (d) Spectra for free ArfA-c(D236A). Saturation at the backbone amide HN frequency of S302 transfers to neighboring protein sites.