

Chemical shift assignments for Rv0577, a putative glyoxylase associated with virulence from *Mycobacterium tuberculosis*

Garry W. Buchko · Heungbok Kim ·
Peter J. Myler · Thomas C. Terwilliger ·
Chang-Yub Kim

Received: 23 April 2011 / Accepted: 27 June 2011 / Published online: 15 July 2011
© Springer Science+Business Media B.V. (outside the USA) 2011

Abstract Approximately one-third of mankind has been exposed to *Mycobacterium tuberculosis*, the etiological agent responsible for tuberculosis (TB). As part of an effort to develop a new generation of anti-TB agents, the chemical shifts for the 261-residue, virulence-associated protein Rv0577 from *M. tuberculosis* has been extensively assigned.

Keywords Tuberculosis · Infectious diseases · Virulence factors · Host-pathogen interactions · SSGCID

Biological content

Mycobacterium tuberculosis is the etiological agent responsible for tuberculosis (TB), a chronic infectious disease that killed approximately 1.6 million people and infected another 8–9 million in 2008 (World Health Organization 2009). Indeed, it is estimated that one-third of the human population has been infected with this gram-positive tubercle bacillus (Enarson 2003) with most of these cases in 22 “high-burden” nations where the disease is endemic (Russell et al. 2010). While effective public health care systems keeps TB under control in the Western world, the emergence of multi-drug and extremely drug-resistant *M. tuberculosis* strains could result in a sudden loss of this control. Consequently, there is an urgency to develop a new generation of intervention strategies to treat and control TB (Myler et al. 2009). One current tactic is to focus on better understanding the molecular biology of the *M. tuberculosis* gene products, especially with regards to the interaction of virulence factors in the microenvironment in the host (Russell et al. 2010). One potential virulence factor is the protein Rv0577. A biological function for this 261-residue protein has not yet been identified, however, it has been associated with the methylglyoxal detoxification pathway (Pethe et al. 2010). More importantly, Rv0577 has been shown to be the responsible component for neutral red staining of virulent strains of *M. tuberculosis* (Andreu et al. 2004). Such a correlation, coupled with the isolation of Rv0577 in *M. tuberculosis* culture filtrates, suggests that Rv0577 may be a virulence factor. To assist function identification and structure-based drug design that targets Rv0577, a structure for this protein has recently been determined using X-ray diffraction methods and deposited in the RBSC PDB (3OXH). Once the coordinates for the structure of Rv0577 are released the

G. W. Buchko (✉)
Biological Sciences Division and Seattle Structural Genomics
Center for Infectious Disease, Pacific Northwest National
Laboratory, Richland, WA 99352, USA
e-mail: garry.buchko@pnl.gov

H. Kim · T. C. Terwilliger · C.-Y. Kim (✉)
Bioscience Division, Los Alamos National Laboratory,
Los Alamos, NM 87545, USA
e-mail: cykim@lanl.gov

P. J. Myler
Seattle Biomedical Research Institute and Seattle Structural
Genomics Center for Infectious Disease, Seattle,
WA 98109-5219, USA

P. J. Myler
Department of Medical Education and Biomedical Informatics,
University of Washington, Seattle, WA 98195, USA

P. J. Myler
Department of Global Health, University of Washington, Seattle,
WA 98195, USA

amide chemical shift assignments presented here will facilitate chemical shift perturbation studies and assist in the identification of ligand binding surfaces on the protein (Zuiderweg 2002).

Materials and methods

Cloning, expression, and purification

The *Rv0577* gene was amplified by PCR using the genomic DNA of *M. tuberculosis* strain H37Rv and the oligonucleotide primers 5'-AGATATACATATGCCCAAGAGAAGCGAATACAGG-3' (forward) and 5'-AATTCGGA TCCTTGCTGCGGTGCGGGCTTCA-3' (reverse) (Invitrogen, Carlsbad, CA) containing NdeI and BamHI restriction enzyme digestion sites, respectively (underlined). Following NdeI and BamHI digestion, the amplified DNA was inserted into a modified pET28b expression vector (Novagen, Madison, WI) such that the expressed gene product contained an eight amino acid extension, -RSHHHHHH, at the C-terminus of the native protein to assist protein purification by metal chelation chromatography. The recombinant plasmid was then transformed into *Escherichia coli* BL21(DE3) cells (Novagen, Madison, WI) by a heat shock method. Uniformly ^{15}N -, ^{13}C -labeled Rv0577 was obtained by growing the transformed cells (310 K) in minimal medium (Miller) containing $^{15}\text{NH}_4\text{Cl}$ (1 mg/mL) and $\text{D-}[^{13}\text{C}_6]\text{glucose}$ (2.0 mg/mL) supplemented with NaCl (50 $\mu\text{g}/\text{mL}$), MgSO_4 (120 $\mu\text{g}/\text{mL}$), CaCl_2 (11 $\mu\text{g}/\text{mL}$), Fe_2Cl_3 (10 ng/mL) and the antibiotic kanamycin (34 $\mu\text{g}/\text{mL}$). After an OD_{600} reading of ~ 0.8 was reached, the cells were cooled to 298 K and protein expression induced with isopropyl β -D-1-thiogalactopyranoside (0.026 $\mu\text{g}/\text{mL}$). Approximately 5 h later the cells were harvested by mild centrifugation and frozen at 193 K. Following the thawing of the frozen pellet, Rv0577 was purified with a conventional two-step protocol involving Ni-NTA affinity purification (Qiagen, Valencia, CA) followed by gel-filtration chromatography on a Superdex75 HiLoad 16/60 column (GE Healthcare, Piscataway, NJ) (Buchko et al. 2006). The latter step simultaneously exchanged Rv0577 into the buffer used for the NMR studies (100 mM NaCl, 20 mM Tris, 1.0 mM dithiothreitol, pH 7.1).

Nuclear magnetic resonance spectroscopy

All the NMR data was collected at 303 K on a double-labeled (^{13}C , ^{15}N) sample (~ 1.5 mM) using Varian Inova-600, and -750 spectrometers equipped with triple resonance probes and pulse field gradients. Assignment of the ^1H , ^{13}C , and ^{15}N chemical shifts of the backbone and side chain resonances

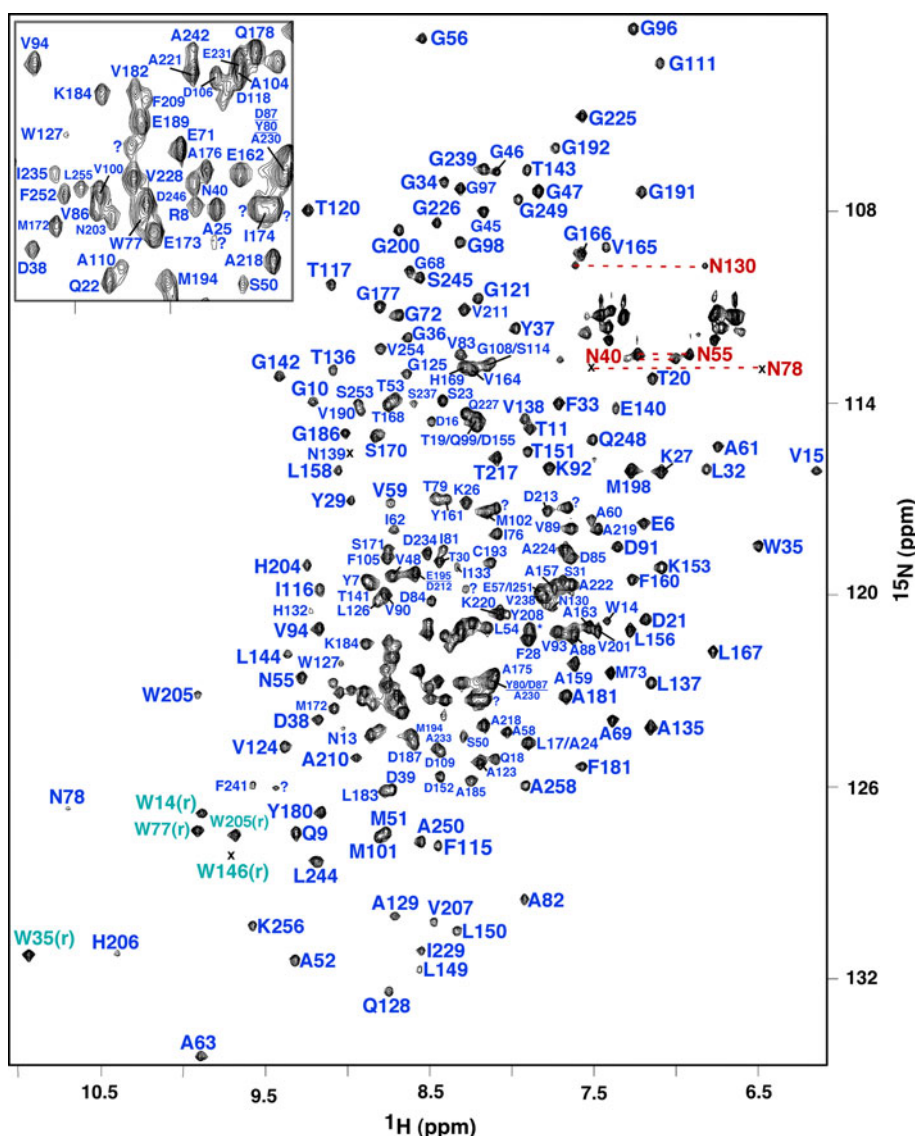
were made from standard two-dimensional ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, HBCBCGCDHD, and HBCBCGDCHE experiments and three-dimensional HNCACB, CBCA (CO)NH, HCC-TOCSY-NNH, CC-TOCSY-NNH, and HNCO experiments using the Varian Biopack suite of pulse programs. Three-dimensional ^{15}N -edited NOESY-HSQC and ^{13}C -edited NOESY-aromatic-HSQC experiments collected with a mixing time of 85 ms were analyzed to assist the ^1H assignments. An overall rotational correlation time (τ_c) for Rv0577 was estimated from backbone amide ^{15}N $T_{1\rho}$ / $T_{1\rho}$ ratios (Szyperski et al. 2002). All NMR data was processed using Felix2007 (MSI, San Diego, CA) software and analyzed with the program Sparky (v3.115). Indirect methods (DSS = 0 ppm) were used to reference the ^1H , ^{13}C , and ^{15}N chemical shifts that were deposited into the BioMagResBank database (<http://www.bmrb.wisc.edu>) under the accession number BMRB-17597.

Results and discussion

Rv0577 has a molecular weight of 28.4 kDa, (27.3 kDa plus a 1.1 kDa, 8-residue C-terminal tag). Behavior of the protein in size exclusion chromatography and NMR spectroscopy experiments both indicated that Rv0577 was a monomer in solution. First, using a flow rate of 1.0 mL/min, Rv0577 eluted off a 16/60 Superdex75 size exclusion column with a retention time of 72 min, a value characteristic of an ~ 28 kDa protein on such a column (data not shown). Second, the estimated isotropic overall rotational correlation time (τ_c) for Rv0577, as inferred from ^{15}N spin relaxation times (Szyperski et al. 2002), was 14.2 ± 0.5 ns at 293 K, a value consistent with an ~ 30 kDa protein. Note that while these two experiments indicated that Rv0577 was a monomer in solution, at 28.4 kDa, the ^1H - ^{15}N HSQC spectrum was not especially well resolved and featured broad cross peak line shapes at 293 K (data not shown). The spectral properties of the ^1H - ^{15}N HSQC spectrum were significantly improved by collecting the backbone assignment data at 303 K, a temperature at which circular dichroism spectroscopy showed no changes in global structure (data not shown) and the protein proved to be stable over long periods of time.

Figure 1 illustrates the ^1H - ^{15}N HSQC spectrum for Rv0577 with the unambiguous amide (blue) and side chain $-\text{NH}_2$ (red) assignments labeled. The recombinant protein contains an eight residue, C-terminal tag and no amide cross peaks were observed for any of these residues. Excluding this "invisible" C-terminal tag, amide cross peaks were identified for 218 out of the 237 (261 - (23P + M1)) expected residues of Rv0577. The unassigned residues are highlighted (blue) in the amino acid sequence shown for the protein in Fig. 2. Seven of the 19 "missing" amide cross

Fig. 1 The ^1H - ^{15}N HSQC spectrum for Rv0577 (~1.5 mM) collected at a proton resonance frequency of 750 MHz, 303 K, in 100 mM NaCl, 20 mM Tris, 1 mM DTT, pH 7.1. Size exclusion chromatography and the estimated correlation time indicate that the 269-residue protein (8-residue tag) is a monomer in solution. The assigned amide cross peaks are labeled in *blue*, the assigned side chain resonances are identified with a *red horizontal line*, cross peaks below the contour threshold marked with an “x”, and the unassigned amide cross peaks labeled with a *question mark*. The *inset* is an expansion of the congested region near the center of the spectrum



peaks are at termini and likely reflect greater dynamic motion at these regions. While some of these “missing” amide cross peaks may exist underneath assigned residues (overlapped) or be responsible for cross peaks labeled with a question mark in Fig. 1, another possibility is that some of these “missing” residues, especially those in small clusters, may truly be “missing” because the chemical shifts have broadened out beyond detection due to motion in the intermediate timescale (ms to μs) and/or exchange between many different conformational states (local disorder). Such regions have previously been associated with binding and catalysis in many proteins. As shown in Fig. 2, three “missing” amide cross peak clusters of two, three, and four consecutive residues are present in Rv0577, and consequently, may be biologically significant in Rv0577.

Chemical shifts for $^1\text{H}^{\text{N}}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, and ^{13}CO were used to predict regions of helical and β -strand secondary

structure in Rv0577 using the chemical shift index program (Wishart and Sykes 1994). The regions of secondary structure predicted by this program are shown in the sequence of Rv0577 in Fig. 2 and illustrate that the protein contains 12 β -strands (cyan) and eight α -helices (red). The tertiary arrangements of these elements of secondary structure will be known after the coordinates for the crystal structure (PDB ID 3OXH) are released. Once these coordinates are released, the assignment of the ^1H - ^{15}N HSQC spectrum for Rv0577 will facilitate chemical shift perturbation studies and assist in the identification of ligand binding surfaces on the protein (Zuiderweg 2002). Such experiments will also help verify the biological function of Rv0577 and mechanisms of drug resistance, potentially accelerating efforts towards the development of a new generation of intervention strategies to treat and control TB (Myler et al. 2009). For example, Rv0577 has been

MPKRSEYRQG TPNWVDLQTT DQSAAKFYT SLFGWGYDDN PVPGGGGVYS 50
 MATLNGEAVA AIAPMPPGAP EGMPPIWNTY IAVDDVDAVV DKVVPGGGQV 100
 MMPAFDIGDA GRMSFITDPT GAAVGLWQAN RHIGATLVNE TGTLIWNELL 150
 TDKPDLALAF YEAVVGLTHS SMEIAGQNY RVLKAGDAEV GGCMEPPMPG 200
 VPNHWHVYFA VDDADATAAK AAAAGGQVIA EPADIPSVGR FAVLSDPQGA 250
 IFSVLKPAPQ QRSHHHHH 269

Fig. 2 The amino acid sequence for Rv0577 highlighting, in *blue*, residues for which amide cross peaks could not be unambiguously assigned in the ^1H - ^{15}N HSQC spectrum. Regions of helical and β -strand secondary structure predicted by chemical shift index analysis of the $^1\text{H}^{\text{N}}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, and ^{13}CO chemical shifts are highlighted in *red* and *cyan*, respectively. The eight-residue C-terminal tag is shown in *magenta*

associated with *M. tuberculosis* resistance to novel pyrimidine-imidazoles, a group of compounds that are potent to cells cultured in vitro with media containing glycerol but not potent to cells grown in vivo where glycerol is absent (Pethe et al. 2010).

Acknowledgments Funding for this research was provided by the National Institute of Allergy and Infectious Diseases, National Institute of Health, Department of Health and Human Services, under Federal Contract number HHSN272200700057C. The SSGCID internal ID for Rv0577 is MytuD.17269.a. Much of this research was performed at the W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility located at Pacific Northwest National Laboratory (PNNL) and sponsored by US Department of Energy's Office of Biological and Environmental Research (BER) program. Battelle operates PNNL for the US Department of Energy.

References

Andreu N, Soto CY, Roca I, Martin C, Gibert I (2004) *Mycobacterium smegmatis* displays the *Mycobacterium tuberculosis*

- virulence-related neutral red character when expressing the Rv0577 gene. FEMS Microbiol 231:283–289
- Buchko GW, Kim C-Y, Terwilliger TC, Kennedy MA (2006) Solution structure of the conserved hypothetical protein Rv2302 from *Mycobacterium tuberculosis*. J Bacteriol 188:5993–6001
- Enarson DA (2003) Tuberculosis as a global public health problem. In: Kaufmann SHE, Hahn H (eds) Mycobacteria and TB. Karger, Basel, pp 1–16
- Myler PJ, Stacy R, Stewart LJ, Staker BL, Van Voorhis WC, Varani G, Buchko GW (2009) The Seattle Structural Genomics Center for Infectious Disease (SSGCID). Infect Diseases Drug Targets 9:493–506
- Pethe K, Sequeira PC, Agarwalla S, Rhee K, Kuhen K, Phong WY, Patel V, Beer D, Walker JR, Duraiswamy J, Jiricek J, Keller TH, Chatterjee A, Tan MP, Ujjini M, Rao SPS, Camacho L, Bifani P, Mak PA, Ma I, Barnes SW, Chen Z, Plouffe D, Thayalan P, Ng SH, Au M, Lee BH, Tan BH, Ravindran S, Nanjundappa M, Lin X, Goh A, Lakshminarayana SB, Shoen C, Cynamon M, Kreiswirth B, Dartois V, Peters EC, Glynn R, Brenner S, Dick T (2010) A chemical genetic screen in *Mycobacterium tuberculosis* identifies carbon-source-dependent growth inhibitors devoid of in vivo efficiency. Nature Comm 1:57
- Russell DG, Barry CE III, Flynn JL (2010) Tuberculosis: what we don't know can, and does, hurt us. Science 328:852–856
- Szyperski T, Yeh DC, Sukumaran DK, Moseley HNB, Montelione GT (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. Proc Natl Acad Sci USA 99:8009–8014
- Wishart DS, Sykes BD (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. J Biomol NMR 4:171–180
- World Health Organization (2009) Globular tuberculosis control: a short update to the 2009 report. Geneva
- Zuiderweg ERP (2002) Mapping protein–protein interactions in solution using NMR spectroscopy. Biochemistry 41:1–7