The ISME Journal (2010) 4, 1398–1409
© 2010 International Society for Microbial Ecology All rights reserved 1751-7362/10
www.nature.com/ismej

ORIGINAL ARTICLE

Posttranslational modification and sequence variation of redox-active proteins correlate with biofilm life cycle in natural microbial communities

Steven W Singer[1], Brian K Erickson[2,5], Nathan C VerBerkmoes[2], Mona Hwang[1], Manesh B Shah[3], Robert L Hettich[2], Jillian F Banfield[4] and Michael P Thelen[1]

[1]*Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA;* [2]*Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA;* [3]*Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA;* [4]*Department of Earth and Planetary Sciences, University of California, Berkeley, CA, USA and* [5]*Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, TN, USA*

Characterizing proteins recovered from natural microbial communities affords the opportunity to correlate protein expression and modification with environmental factors, including species composition and successional stage. Proteogenomic and biochemical studies of pellicle biofilms from subsurface acid mine drainage streams have shown abundant cytochromes from the dominant organism, *Leptospirillum* Group II. These cytochromes are proposed to be key proteins in aerobic Fe(II) oxidation, the dominant mode of cellular energy generation by the biofilms. In this study, we determined that posttranslational modification and expression of amino-acid sequence variants change as a function of biofilm maturation. For Cytochrome$_{579}$ (Cyt$_{579}$), the most abundant cytochrome in the biofilms, late developmental-stage biofilms differed from early-stage biofilms in N-terminal truncations and decreased redox potentials. Expression of sequence variants of two monoheme *c*-type cytochromes also depended on biofilm development. For Cyt$_{572}$, an abundant membrane-bound cytochrome, the expression of multiple sequence variants was observed in both early and late developmental-stage biofilms; however, redox potentials of Cyt$_{572}$ from these different sources did not vary significantly. These cytochrome analyses show a complex response of the *Leptospirillum* Group II electron transport chain to growth within a microbial community and illustrate the power of multiple proteomics techniques to define biochemistry in natural systems.

The ISME Journal (2010) **4,** 1398–1409; doi:10.1038/ismej.2010.64; published online 20 May 2010
Subject Category: Integrated genomics and post-genomics approaches in microbial ecology
Keywords: metagenomics; proteomics; cytochrome; variation; redox

Introduction

Recent advances in genomic sequencing and mass spectrometry (MS)-based proteomics have enabled detailed, cultivation-independent analyses of protein profiles in natural microbial communities. Low-diversity communities that constitute acid mine drainage (AMD) biofilms have served as a model system for the development of these methods and for ecological investigations (VerBerkmoes *et al.*, 2009; Wilmes *et al.*, 2009b). Most of the AMD biofilms are dominated by *Leptospirillum* Group II (*Lepto*II), a Fe(II)-oxidizing, chemoautotrophic bacterium (Schrenk *et al.*, 1998; Tyson *et al.*, 2004). The biofilms exhibit distinct developmental stages that vary in microbial community composition. Early developmental stages (DS1) are dominated by *Lepto*II; however, late developmental-stage biofilms (DS2) diversify, with increasing abundance of *Leptospirillum* Group III (*Lepto*III), archaea and eukaryote populations (Wilmes *et al.*, 2009a).

Proteomic measurements of an early developmental-stage biofilm identified two atypical cytochromes expressed at high levels. These were initially identified as *Lepto*II proteins of unknown function, and later were determined to be the membrane Cytochrome$_{572}$ (Cyt$_{572}$) and periplasmic Cytochrome$_{579}$ (Cyt$_{579}$) (Ram *et al.*, 2005). Both cytochromes were purified from a mixed developmental-stage biofilm and characterized biochemically (Jeans *et al.*, 2008; Singer *et al.*, 2008). Cyt$_{572}$, localized to the outer membrane, is a 57-kDa multimeric protein that oxidizes Fe(II) at low pH, and thus is likely the Fe(II) oxidase for *Lepto*II in the biofilm (Jeans *et al.*, 2008). Inspection of metagenomic data sets showed

---

Correspondence: SW Singer, Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mail Stop 90-R1116, Berkeley, CA 94720, USA.
E-mail: SWSinger@lbl.gov

Received 11 January 2010; revised 26 March 2010; accepted 1 April 2010; published online 20 May 2010

six sequences corresponding to *Lepto*II variants of Cyt$_{572}$ in the Ultra Back A (UBA) and 5way Community Genomics (CG) databases (5way CG and UBA databases refer to environmental genomic databases obtained from DNA isolated from biofilms collected at distinct sites in the Richmond Mine. 5way CG was sampled at the 5-way convergence of streams in June 2004, and UBA was sampled in the upper A drift in June 2005. Genomic sequences were assembled and these were deposited in the databases of these names (Tyson *et al.*, 2004; Lo *et al.*, 2007)), indicating that multiple variants of the cytochrome may be expressed in the same biofilm sample. Cyt$_{579}$ was characterized as a $\sim$16-kDa monomeric protein localized to the periplasm of *Lepto*II (Singer *et al.*, 2008). Cyt$_{579}$ was isolated as a mixture of polypeptides with different N-terminal cleavage sites. Redox reactions with Fe(II) demonstrated pH-dependent Fe(II) oxidation that was inconsistent with its assignment as the Fe(II) oxidase for *Lepto*II and suggested an electron-transfer function. Our working model positions Cyt$_{572}$ as the Fe(II) oxidase on the outer membrane of *Lepto*II cells, which oxidizes Fe(II) to Fe(III) and transfers electrons to Cyt$_{579}$. This scheme is analogous to the proposed role of an outer membrane-bound *c*-type cytochrome, Cyc2, and rusticyanin, a periplasmic Cu protein, in *Acidithiobacillus ferrooxidans*, an acidophilic Fe(II)-oxidizing bacterium found in AMD environments (Castelle *et al.*, 2008).

Recent analysis of multiple AMD biofilm proteomes from different developmental stages has shown that the community switches from rapid Fe(II)-dependent autotrophic growth in early developmental stages to partitioning of fixed carbon to heterotrophs in late developmental stages (Denef *et al.*, 2010b). The effects of aging on biofilm morphology, microbial community composition and protein expression led us to speculate that electron-transfer proteins, critical for Fe(II) oxidation, could change physically over the biofilm life cycle, either from specific posttranslational modifications or from genetic variation as a result of mutation or recombination. Characterization of cytochromes from biofilms representing early and late developmental stages (DS1 and DS2, respectively) showed that both posttranslational modifications and expressed sequence variants are correlated with biofilm development.

Materials and methods

Biofilm collection

Biofilm samples were collected from the Richmond Mine (Iron Mountain, CA, USA) (40°40′38. 42″N, 122°31′19.90″W, elevation of $<$900 m); all samples were grown at the liquid–air interface. DS1 and DS2 samples were collected from $\sim$10 m into the C drift in November 2006. A DS1 sample was collected in August 2007 and a DS2 sample was collected in November 2006 from a dam at the confluence of the A and B drifts (AB-Muck dam site). DS1 samples were collected in June 2006 and November 2006 from a site $\sim$75 m into the C drift (C75 m location). Collections of the AB-End sample from November 2004 and the AB-Front sample from June 2004 have been described previously (Ram *et al.*, 2005; Goltsman *et al.*, 2009). Samples were frozen on dry ice immediately after sampling and stored at −80 °C.

Protein purification and enrichment

Both Cyt$_{579}$ and Cyt$_{572}$ were purified from the biofilms as described previously (Jeans *et al.*, 2008; Singer *et al.*, 2008) and stored at 4 °C. No change in the redox properties of samples of either protein was observed after 6 months at 4 °C, and minimal degradation was observed by SDS-PAGE. In all cases, visible spectra for both oxidized and reduced samples were identical to our previously published spectra for Cyt$_{579}$ and Cyt$_{572}$.

Enrichment of *c*-type cytochromes was achieved by fractionation of the extracellular fraction of the C75 m sample from November 2006. The 95% $(NH_4)_2SO_4$ precipitate of the acid wash fraction was dialyzed for 16 h against 20 mM H_2SO_4/100 mM $(NH_4)_2SO_4$ pH 2.2. The dialysate was loaded onto a SP-Sepharose (GE Healthcare, Piscataway, NJ, USA) fast flow column and washed with this pH 2.2 buffer. Cyt$_{579}$ was eluted by a step increase to pH 5.0 in 100 mM NaOAc, and the remaining proteins were eluted with a 0–2 M NaCl gradient in the same buffer. Characteristic visible absorption spectra for *c*-type cytochromes were observed throughout the 1.2–2.0 M fractions, with 1.4 and 1.5 M fractions containing the highest concentrations of the *c*-type cytochrome as measured by visible absorbance of the α-band at 552 nm for samples reduced with sodium ascorbate. The individual heme bands were visualized by separation of proteins on 15% SDS-PAGE by the method of Francis and Becker (1984). The stained bands were excised from the gel and digested with trypsin (Haveman *et al.*, 2006).

Redox experiments

Cyt$_{579}$ purified from C-drift GS1 (11 μg ml^{-1}) and GS2 (7 μg ml^{-1}) were reduced in the presence of 30 mM $FeSO_4$, and the extent of reduction was estimated by comparing the peak at 579 nm to a fully reduced sample, as described previously (Singer *et al.*, 2008). Redox measurements and poising were performed in 10 mM NH_4OAc pH 4.3 with 50 μM each of 1, 4-benzoquionone, N,N,N′,N′-tetramethyl-p-phenylenediamine, ferrocenecarboxylic acid and 1,1′-ferrocenedicarboxylic acid used as redox mediators (Castelle *et al.*, 2008). Concentrated solutions of sodium dithionite were used as the reductant and sodium hexachloroiridate was used as the oxidant. Concentrations used in the redox experiments were Cyt$_{579}$, 10–20 μg ml^{-1}; Cyt$_{572}$, 25 μg ml^{-1} (C-drift

DS1); and 30 μg ml^{-1} (C-drift DS2). The combination microelectrode (Microelectrodes Inc., New Bedford, MA, USA) was calibrated at pH 4 and pH 7 using quinhydrone.

PCR amplification and cloning of genes encoding Cyt$_{579}$

Genomic DNA was recovered from the biofilms by a modified phenol:chloroform extraction method and used as a template for PCR (Ram *et al.*, 2005). UBA_8062_372 (Cyt$_{579}$)was amplified using High-Fidelity Taq polymerase (Invitrogen, Carlsbad, CA, USA) with the primers 5′-ATGAGGATGTGGACA GTGGCTGTC-3′ and 5′-TTACTTTGCTGCTGTGTTC AGAAA-3′. PCR conditions were initial denaturation at 95 °C for 5 min; 30 cycles of 95 °C for 30 s, an annealing temperature of 56 °C for 30 s and 72 °C for 35 s, and an extension at 72 °C for 4 min. PCR products were cloned into TOPO TA10 (Invitrogen) using the manufacturer's conditions, and clones were sequenced at the Joint Genome Institute (Walnut Creek, CA, USA).

Cyt$_{579}$ intact mass measurement

Purified samples of Cyt$_{579}$ were further prepared for characterization of the intact proteins by high-resolution top-down MS. Cyt$_{579}$-enriched samples were desalted with Zip-Tip (C$_4$, Millipore, Billerica, MA, USA) pipette tips and eluted with 100% ACN (0.1% acetic acid, v/v). The protein fraction was then diluted into 50/50/0.1 (v/v/v) H$_2$O/ACN/acetic acid and infused into the Micromass Z-Spray source attached to a Varian (Lake Forest, CA, USA) 9.4-T (Cryomagnetics Inc., Oak Ridge, TN, USA) HiRes electrospray FT-ICR (Fourier transform ion cyclotron resonance mass spectrometer or an electrospray source coupled to the LTQ-Orbitrap-XL (Thermo Fisher Scientific, San Jose, CA, USA). MS fragmentation was achieved through collisionally activated dissociation, electron-transfer dissociation or infrared multiphoton dissociation. Parent charge states of Cyt$_{579}$ were manually selected, isolated and fragmented (collisionally activated dissociation or electron-transfer dissociation) in the ion trap before high-resolution mass measurement in the Orbitrap. For infrared multiphoton dissociation on FT-ICR, parent charge states of Cyt$_{579}$ were manually selected and isolated in the selection quadrapole before mass analysis in the FTMS analyzer cell. A 350-nM spike of ubiquitin was introduced into the C-drift DS1 and AB-Muck DS1 samples for internal mass calibration. M/z values were manually extracted from spectra, deconvoluted and plotted with Origin 8 (OriginLab, Northhampton, MA, USA).

General methods

SDS-PAGE was performed according to the method of Laemmli (1970). Protein concentration was estimated according to the method of Bradford (1976). Trypsin digestion and N-terminal sequencing of proteins were performed as described previously (Ram *et al.*, 2005). Trypsin- and pepsin-digested samples were analyzed by 6 h 2D-liquid chromatography-MS/MS as described previously (Jeans *et al.*, 2008; Singer *et al.*, 2008). Pepsin digestions of Cyt$_{572}$ preparations were performed as described previously (Jeans *et al.*, 2008).

Results

Selection of biofilms for protein extraction

Biofilm samples representing a single developmental stage and with sufficient biomass for biochemical analyses were selected for protein extraction (Table 1) (for locations, see Figure 6 in Denef *et al.* (2009)). A number of metrics were used to determine biofilm developmental stage. Visual inspection was initially used to distinguish the developmental stages, as early developmental stages are thinner (∼50 μm) and deep pink in color, whereas late developmental stages were thicker (150–200 μm) and light pink in color (Wilmes *et al.*, 2009a). Microbial community analysis was consistent with the observation that DS1 samples were dominated by *Lepto*II, and DS2 samples had an increased ratio of archaea to bacteria. These observations were complemented by detailed analysis of *Lepto*II strain populations, which indicated that the strains related to the reconstructed UBA genome dominated the morphologically identified DS1 samples and a mixture of UBA and CG types was observed in DS2 samples (Denef *et al.*, 2010a). As described in the 'Introduction' section, comparative proteomic

Table 1 Geochemical and microbial community compositions for biofilm samples

Biofilm	*Development stage*	*pH*	*Temp (°C)*	*Conductivity (ms per cm)*	*Prokaryote community composition*[a]
C-drift 10 m 11/06	1	1.01	39.5	56.3	Bacteria (89%), archaea (11%)
C-drift 10 m 11/06	2	1.01	39.5	56.3	Bacteria (59%), archaea (41%)
AB-Muck 8/07	1	0.83	37.1	75.3	Bacteria (93%), archaea (7%)
AB-Muck 11/06	2	0.84	43.0	85.1	Bacteria (67%), archaea (33%)
C75 m 11/06	1	1.00	43.0	57.9	Bacteria (65%), archaea (35%)
C75 m 6/06	1	0.70	43.0	70.5	Bacteria (42%), archaea (58%)

[a] Community bacterial and archaeal composition was determined by fluorescence *in situ* hybridization according to previously published procedure (Bond *et al.*, 2000).

analysis also distinguished biofilm developmental stage, as proteomes measured from DS1 samples were characteristic of rapid autotrophic growth, whereas DS2 proteomes were characteristic of a partitioning of fixed carbon between *Lepto*II and archaeal populations (Denef *et al.*, 2010b). The initial DS1 samples from the C-drift 10 m site were collected from the center of a flowing AMD stream. DS2 samples were collected from the stream margins in the same location, where biofilms attach to pyrite sediment (Wilmes *et al.*, 2009a). DS1 and DS2 samples were also collected from the AB-Muck site at different times; the morphology of these biofilms was very similar to the DS1 and DS2 samples collected at C-drift 10 m. Biofilms were observed on the surface of a standing pool of AMD at the C75 m. The thickness of the C75 m biofilm was consistently ~50 μm, but the pH and community composition varied over time for samples collected from the C75 m site (Table 1).

Cyt$_{579}$ N terminus is dependent on biofilm developmental stage

In our previous studies, multiple isoforms of Cyt$_{579}$ with different N-terminal truncations were purified from a mixed development-stage biofilm collected at the C-drift ~10 m site in November 2005 (Singer *et al.*, 2008). N-terminal sequencing of Cyt$_{579}$ (AKAMKPPFPV) isolated from a relatively thick, mature biofilm collected at the AB-Front site in July 2004 was identical to the N terminus of the smallest polypeptide isolated from the C-drift biofilm (Goltsman *et al.*, 2009). These preliminary results suggested that the observed N terminus of Cyt$_{579}$ correlates with the maturity of the biofilm. To test this hypothesis, Cyt$_{579}$ was isolated from the samples obtained from the DS1 and DS2 C-drift ~10 m samples, as described above. Approximately six times more Cyt$_{579}$ was isolated from the DS1 than the DS2 biofilm (Table 2). SDS-PAGE of the Cyt$_{579}$ from these samples indicated that the DS1 protein had a higher molecular weight than the DS2 cytochrome, consistent with a shorter N terminus in the late developmental stage (Figure 1). Furthermore, N-terminal sequencing by Edman degradation determined two different N termini for the DS1 (AELDILKPRV:ILKPRVPADQ; 5:1 ratio) and DS2 (AKAMKPPFPV) proteins; these corresponded to N termini determined for the Cyt$_{579}$ preparation from the mixed developmental-stage sample from the C drift (November 2005) (Singer *et al.*, 2008).

Intact protein characterization of Cyt$_{579}$

High-resolution MS measurement of the purified protein from the C-drift DS1 sample revealed two molecular species by distinct isotopic distribution packets. Identification of the most abundant isotopic masses from each distribution corresponded to molecular species of 16 131.541 and 16 119.562 kDa (Figure 2a; external calibration). The measured mass of 16 131.541 kDa corresponds to the gene UBA_8062_372_S98A (Cyt$_{579}$), with a predicted signal peptide cleaved at the N terminus (N-AELDILKP). This observation was confirmed by PCR amplification of the gene encoding for Cyt$_{579}$ from DNA recovered from the C-drift DS1 sample. A total of 29 clones were obtained after transformation of the PCR amplicon, all of which had an identical sequence to UBA_8062_372_S98A. The mature isoform of Cyt$_{579}$ corresponds to the same sequence but lacks the seven C-terminal amino acids (…FLNTAAK) ending in …GNLKPE; this was the dominant variant expressed in the previously characterized Cyt$_{579}$ preparations from C-drift biofilm samples (Singer *et al.*, 2008). The second most abundant distribution, 16 119.562 kDa, was inferred to be a modified form of Cyt$_{579}$. To obtain the most accurate mass measurement possible, an

Table 2 Yields of Cyt$_{579}$ from different biofilms

Biofilm	*Processed volume (ml)*	*Cyt$_{579}$ yield (mg)*	*Yield (μg ml^{-1} of biofilm)*
C-drift (11/06) 10 m DS1	10	1.6	160
C-drift (11/06) 10 m DS2	10	0.2	24
C 75 m (11/06)	35	2.1	60
C 75 m (6/06)	30	0.024	0.8
ABM DS2 (11/06)	20	0.27	13.5
ABM DS1 (8/07)	25	1.5	60

Abbreviations: Cyt$_{579}$, Cytochrome$_{579}$; DS, developmental stage.

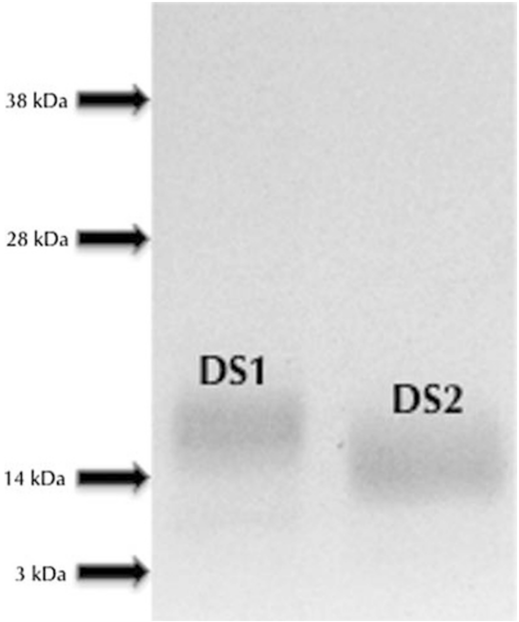

Figure 1 Cyt$_{579}$ isolated from C-drift 10 m (11/06) DS1 and DS2. Cyt$_{579}$ samples were analyzed on 4–12% Bis-Tris polyacrylamide gels (Invitrogen): lane 1, DS1 Cyt$_{579}$ (12 μg); lane 2, DS2 Cyt$_{579}$ (14 μg). The color reproduction of this figure is available on the html full text version of the manuscript.

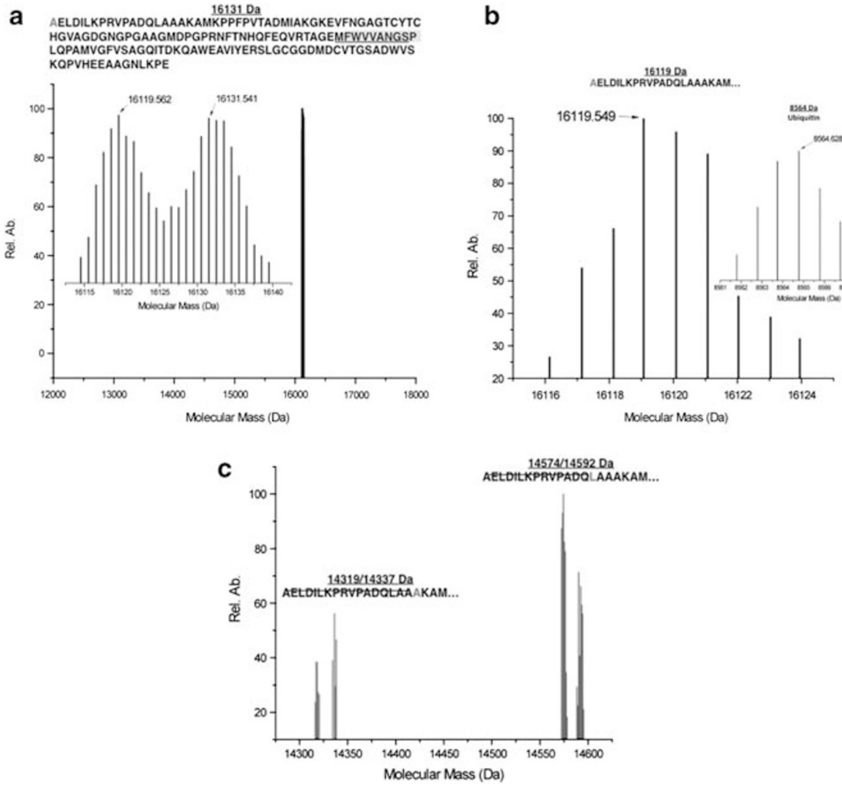

Figure 2 MS spectra of C-drift Cyt_{579} N-terminal residues are highlighted in red. (**a**) The full-length sequence of Cyt_{579} corresponding to mass 16 131 kDa. Intact mass measurement of C-drift DS1 resulted in the identification of two abundant mass distributions, including the truncated product at mass 16 119 kDa as described in text. Residues highlighted in blue indicate the unique sequence tag. (**b**) Partial sequence displaying the truncated N terminus and accurate mass measurement of C-drift DS1 (mass 16 119 kDa) following mass calibration using an internally spiked ubiquitin standard (inset). (**c**) Intact mass measurement of C-drift DS2 exhibiting two states of additional N-terminal truncation. The doublets may indicate ($-H_2O$) loss and were confirmed to be Cyt_{579} through MS fragmentation. Strikethrough highlights the cleaved sequence. The color reproduction of this figure is available on the html full text version of the manuscript.

internal calibration using ubiquitin as a standard was used and resulted in a measured mass of 16 119.549 kDa (Figure 2b). Collisional dissociation of this molecular ion resulted in 16 fragment ions, all corresponding to the sequence of Cyt_{579}, including abundant fragment ions corresponding to a sequence tag of MFWVVA, which is unique to Cyt_{579}. This exact sequence tag was also confirmed by electron-transfer dissociation measurements. In addition, infrared multiphoton dissociation fragmentation resulted in an expanded sequence tag of FWVVANGS, confirming the collisionally activated dissociation and electron-transfer dissociation results. The mass errors of the predicted versus measured fragment ions corresponding to the sequence tag were each <10 p.p.m., providing significant confidence in the identification of the 16 119 kDa species as a modified form of Cyt_{579}. Previous bottom-up peptide measurements of this modified protein verified the sequence of the expected Cyt_{579} protein, plus the presence of an oxidation of residue Met-21. Although the sum total of the fragment ions and peptide mass spectra uniquely support the assignment of the 16 119 kDa species as a Cyt_{579}, neither set provided complete sequence coverage, and thus it was impossible to unambiguously determine the modified form of this truncated version. However, on the basis of all the information, the most likely assignment of the 16 119 kDa species is modification of Cyt_{579} by oxidation (likely at Met-21), accompanied by the loss of CO from the intact protein (CO loss from intact proteins is not unusual). We searched extensively for the location of the CO loss, but were unable to pin it down in the fragmentation or peptide data. However, the calculated mass for a Cyt_{579} protein with these specific modifications is 16 119.530 kDa, in excellent agreement with the measured value (<2 p.p.m. mass error). For the C-drift DS2 sample, measured masses correspond to additional N-terminal truncations, resulting in masses of 14 574 and 14 319 kDa (N-LAAAKAMKPP and N-AKAMKPPFPV, respectively; Figure 2c). These masses were previously observed in mixed developmental-stage C-drift samples and are also derived from UBA_8062_372_S98A, with an identical cleavage at the C terminus as the Cyt_{579} isoform characterized in C-drift DS1.

Divergent redox behavior of Cyt_{579} from distinct developmental stages

Previous experiments involving reaction of Fe(II) with Cyt_{579} isolated from the Richmond Mine biofilms

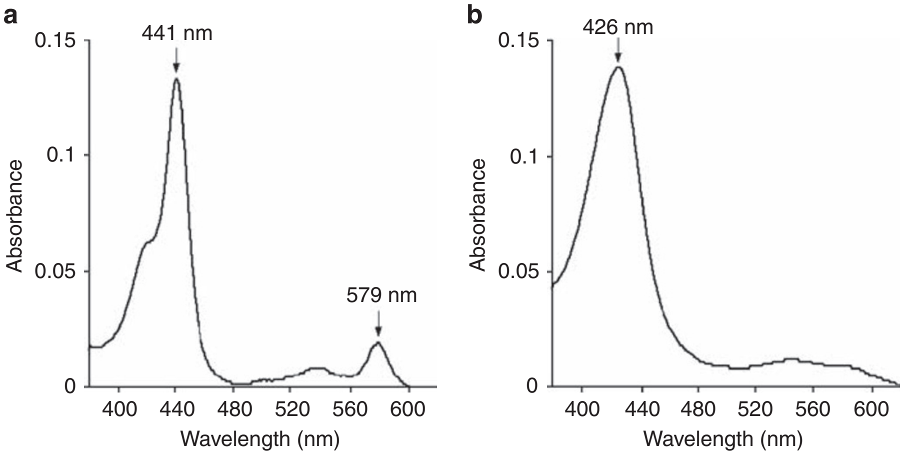

Figure 3 Fe(II) oxidation of C-drift 10 m (11/06) DS1 and DS2. Cyt_{579} preparations were exchanged into 200 mM glycine/SO_4^{2-} buffer pH 2.0 by dialysis and oxidized with a small amount of $Fe_2(SO_4)_3$. The oxidized samples were diluted into pH 2.0 buffer containing 30 mM $FeSO_4$ and spectra were acquired after 5 min: (**a**) C-drift DS1 Cyt_{579} (11 μg ml^{-1}); (**b**) C-drift DS2 (7 μg ml^{-1}).

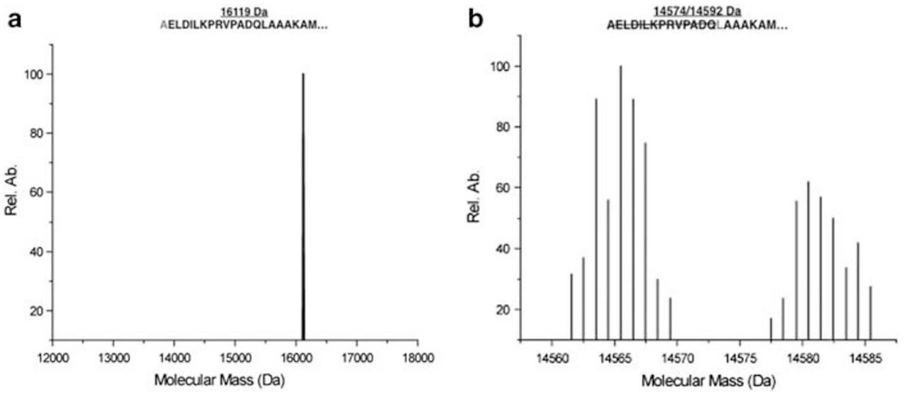

Figure 4 MS spectrum of AB-Muck. N-terminal residues are highlighted in red. (**a**) MS measurement of ABM DS1 exhibiting the same N-terminal truncation as C-drift DS1. (**b**) Intact MS measurement of ABM DS2 resulting in conformation of the LAA N-terminal truncation. The color reproduction of this figure is available on the html full text version of the manuscript.

exhibited pH-dependent redox behavior, with the oxidized form of Cyt_{579} favored at pH <3, even in the presence of a large excess of Fe(II). To test the redox behavior of the Cyt_{579} preparations from C-drift ~10 m DS1 and DS2 samples, extracted proteins were exposed to 30 mM Fe(II) at pH 2. The DS1 preparation was reduced by >95% under these conditions, whereas the DS2 preparation was reduced by <5% (Figure 3a). Consistent with these results, redox titrations at pH 4.3 with redox mediators indicated that the DS1 Cyt_{579} had a midpoint potential of 605 mV, whereas the DS2 preparation had a midpoint potential of 430 mV (Figure 3b).

Cyt_{579} isolated from AB-Muck developmental stages

To establish the generality of the developmental stage-dependent alterations of Cyt_{579}, the protein was purified from the AB-Muck DS1 samples and DS2 samples. As with C-drift samples, substantially more Cyt_{579} was extracted from the early development-stage sample (Table 2). MS analysis of intact protein from the AB-Muck DS1 samples revealed a molecular ion identical to that identified from the C-drift DS1 samples (16 119.540 kDa), indicating that the N terminus and Cyt_{579} variant were the same for both (Figure 4a). Internal mass calibration of the ABM DS1 sample resulted in a 0.022 kDa difference between the ABM DS1 and C-drift DS1 Cyt_{579}. The AB-Muck DS2 sample had molecular ions corresponding to the LAAA N terminus previously observed for the C-drift DS2 sample, but lacked the ion distribution corresponding to the AKA N terminus (Figure 2b). Redox experiments in the presence of 30 mM Fe(II) were very similar to the results obtained for the C-drift samples, and redox titrations at pH 4.3 indicated that the DS1 Cyt_{579} had a midpoint potential of 600 mV, whereas the DS2 preparation had a midpoint potential of 450 mV.

Cyt$_{579}$ from C75 m site

Although the biofilms collected from the C75 m site remained fairly constant in thickness across sampling times, the pH was lower (0.70) for the June than for the November 2006 sample, at which time the pH was 1.0. In addition, archaea were significantly more abundant than in the biofilm growing at this site in June than in November. Cyt$_{579}$ was extracted from both C75 m biofilms, which was more typical in morphology and microbial community composition to the early developmental-stage biofilms from C-drift and AB-Muck sites described above. The yield of Cyt$_{579}$ from the November C75 m sample was similar to the yields obtained from the other early growth-stage biofilms; however, the yield from the June sample was dramatically lower (75 times less Cyt$_{579}$ was extracted from the June biofilm). This result is consistent with an ultrastructural study of the same biofilm (Wilmes *et al*., 2009a), in which immunohistochemical detection of Cyt$_{579}$ showed that Cyt$_{579}$ expression was localized only at the biofilm–water interface, and that a majority of the *Lepto*II cells did not express Cyt$_{579}$. Intact protein MS characterization of Cyt$_{579}$ indicated that the proteins purified from the June and November samples had identical masses of 15 690 kDa, which is the mass of the 8062_372 S98A Cyt$_{579}$ with an ILKP N terminus, identical to the isoform observed in the previous study on Cyt$_{579}$ from the biofilms (Figure 5). Although insufficient Cyt$_{579}$ was recovered from the June C75 m sample for redox analysis, a midpoint potential of 590 mV was determined for Cyt$_{579}$ from the November C75 m sample.

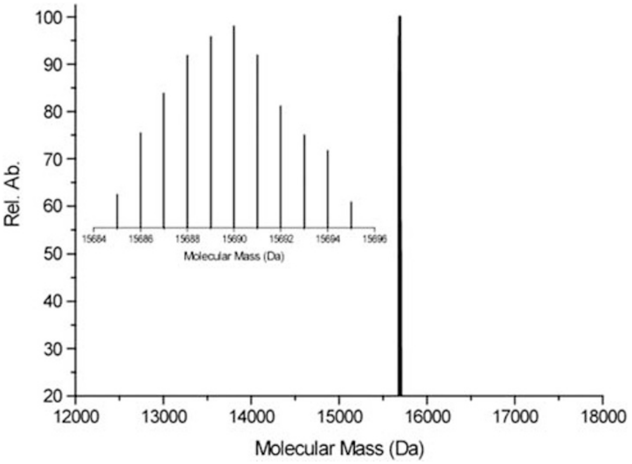

Figure 5 MS measurement of C75 m Cyt$_{579}$ resulted in an accurate mass of 15 690 kDa corresponding to an ILKP N terminus. The N-terminal residue is highlighted in red. The color reproduction of this figure is available on the html full text version of the manuscript.

Identification of c*-type cytochromes from LeptoII*

Previously, we reported that during purification of Cyt$_{579}$, a yellow protein fraction bound to the SP-Sepharose column and was eluted with a NaCl gradient (Singer *et al*., 2008). Preliminary characterization indicated that these high NaCl fractions contained *c*-type cytochromes. In concert with yields of Cyt$_{579}$, early developmental-stage samples contained more than five-fold higher levels of *c*-type cytochromes; therefore, biochemical studies focused on these biofilms. To identify these cytochromes, NaCl gradient fractions from Cyt$_{579}$ purification from the C75 m November 2006 sample were analyzed in detail. Visible spectroscopy indicated that the concentration of *c*-type cytochromes was highest in 1.4 and 1.5 M NaCl fractions (Figure 6). These spectra were indicative of *c*-type cytochromes, with an α-band for the reduced sample at 552 nm.

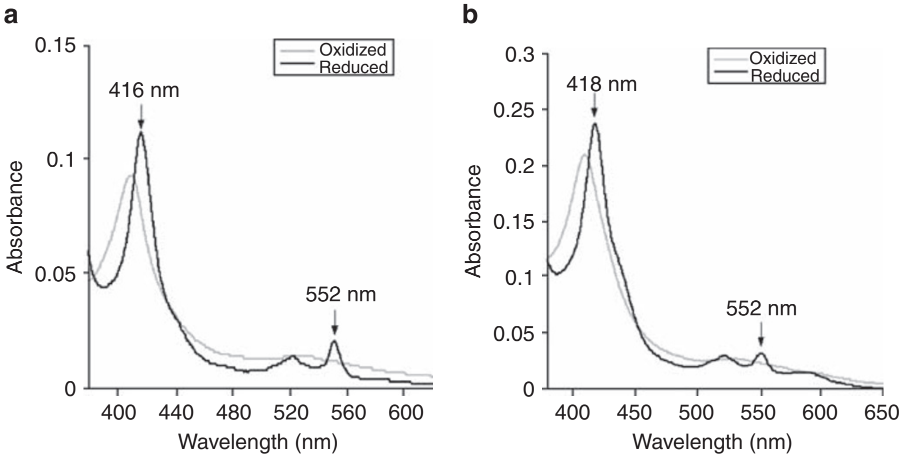

Figure 6 Visible spectra of *c*-type cytochromes. Pale yellow fractions obtained from pH 5.0 NaCl gradient elution (0–2 M) of SP-Sepharose column after elution of Cyt$_{579}$ were concentrated and measured by visible spectroscopy. Using the Soret band of oxidized samples as an indicator, fractions eluting at 1.4 M (0.1 mg ml^{-1} protein) and 1.5 M NaCl (0.2 mg ml^{-1}) contained the highest concentration of *c*-type cytochromes. Samples were reduced by addition of 1 μl of a concentrated (100 mM) solution of sodium dithionite: (**a**) 1.4 M NaCl; (**b**) 1.5 M NaCl. The color reproduction of this figure is available on the html full text version of the manuscript.

Visualization of the proteins by heme staining following SDS-PAGE revealed two prominent bands at ∼8 and 12 kDa, with a minor band at ∼10 kDa, in the 1.4 M NaCl fraction. In the 1.5 M NaCl fraction, two prominent bands were observed at ∼24 and ∼29 kDa (Figure 7). All of these bands were excised from the gel, digested with trypsin and analyzed by liquid chromatography-MS. The 8 and 12 kDa bands were identified as products of genes UBA_LII_8524_245 and 8524_197, respectively (Supplementary Information, Supplementary Figures S1 and S2). The 24 and 29 kDa proteins were identified as products of genes UBA_LII_7931_111 and 7931_112, respectively, and therefore are adjacent on the *Lepto*II genome. These protein sequences are similar to diheme cytochromes, and may be coexpressed in a predicted operon. Peptides for all of these *c*-type cytochromes were previously identified in multiple proteomic data sets obtained from extracts of AMD biofilms (Goltsman *et al.*, 2009). Peptides for an additional predicted 10 kDa monoheme *c*-type cytochrome, UBA_LII_7931_87, were detected in multiple proteomic data sets; although a corresponding band was not detected in the SP-Sepharose fractions, the faint heme-stained band at 10 kDa is a likely candidate.

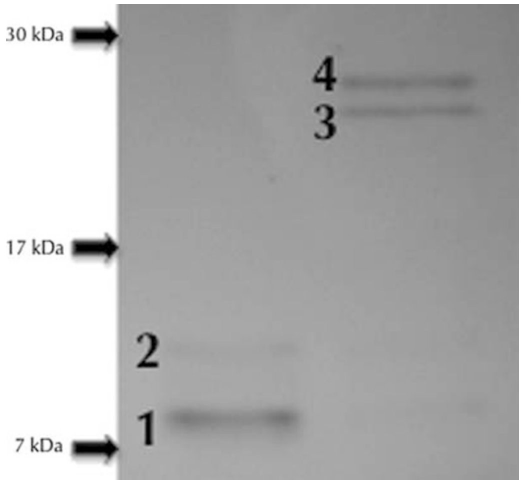

Figure 7 Heme stain of *c*-type cytochromes. 1.4 and 1.5 M NaCl fractions were treated with cold 10% TCA and the precipitated protein pelleted by centrifugation. The pellet was resuspended in 10 μl of SDS-PAGE sample buffer and protein separated by SDS-PAGE on a 15% Tris-HCl polyacrylamide gel (Bio-Rad, Hercules, CA, USA). Heme-containing proteins were then visualized by staining (Francis and Becker, 1984). The bands were excised and digested with tryspin according to a previously published procedure (Haveman *et al.*, 2006). Gene products identified are from lane 1 (1.4 M NaCl): 1. UBA_8524_235; 2. UBA_8524_197; from lane 2 (1.5 M NaCl): 3. UBA_7931_111; 4. UBA_7931_112. The color reproduction of this figure is available on the html full text version of the manuscript.

Diversification of monoheme cytochromes in late biofilm-developmental stage

Inspection of the reconstructed genomes for *Lepto*II from the 5way CG and UBA sites indicated that the sequences for the diheme *c*-type cytochromes were identical in the two genomes; however, the two monoheme *c*-type cytochromes displayed sequence variation when the two genomes were compared (Figure 8). To determine whether these variants were expressed in different developmental stages, their expression in the extracellular proteome of the AB-End, an early developmental-stage biofilm and AB-Front, a late-stage biofilm, were compared (Goltsman *et al.*, 2009). In the AB-End proteome, only peptides corresponding to the UBA variants of these two monoheme cytochromes were detected, whereas peptides for both the UBA and 5way CG variants were detected in the AB-Front proteome (Table 3), consistent with the dominance of these samples by single versus multiple genotypes.

Comparison of Cyt$_{572}$ from biofilms

To compare the effects of biofilm development stage and environmental conditions on the properties of Cyt$_{572}$, the membrane protein was purified from the November 2006 C-drift DS1 and DS2 samples and from the June and November 2006 C75 m samples. In contrast to Cyt$_{579}$, approximately equal amounts of Cyt$_{572}$ were isolated from all four biofilms; ∼0.5 mg of Cyt$_{572}$ per 5 g of material remaining after extraction of each of the biofilms with dilute acid. N-terminal sequencing of the purified proteins revealed an identical sequence (YPGFAR), corresponding to the

Cyt 1
```
CG_8_58        1 APDVKSLMNQQGCFSCHAVDQKMVGPSFKQVAARYRGKKGSLSMLAKKIISGGNGHWNDL
UBA_8524_235   1 ..............A.................D.......A...................

CG_8_58       61 TGGMMMPPHPDLKPSDAKAIAQWVLSLK
UBA_8524_235  61 ..........................V.
```

Cyt 2
```
CG_8_20        1 HGVGSRGLDAGKKIYRTNCAVCHGVDGHGEGRGGRHLNPPVPDFTQPGFLKGKSDSYLFH
UBA_8524_197   1 .........E......S..........R.D...A...........R..............

CG_8_20       61 LISNGIEDMPGWSDKLAPGQITDVLHYLRSLAGPSGDTRPPSPDRFSGE
UBA_8524_235  61 ................................A.AK....A...GL..G
```

Figure 8 Sequence comparison for monoheme *c*-type cytochromes. Sequence comparison for *Lepto*II monoheme cytochromes identified by heme stain (Simmons *et al.*, 2008; Goltsman *et al.*, 2009). For CG_8_58/UBA_8524_235, the N terminus was predicted by SignalP and observed by Edman degradation (APDVK; see Supporting Information, Figure 2). For CG_8_20/UBA_8524_197, N termini were predicted using SignalP (http://www.cbs.dtu.dk/services/SignalP/).

predicted N terminus based on the sequences for Cyt$_{572}$ in the reconstructed *Lepto*II and *Lepto*III genomes, and also identical to the previously purified preparations of Cyt$_{572}$ (Jeans *et al.*, 2008).

Analysis of the reconstructed *Lepto*II and *Lepto*III genomes has indicated a substantial number of sequence variants for the gene encoding for Cyt$_{572}$ (Jeans *et al.*, 2008). A search of the peptides obtained from each of the preparations of Cyt$_{572}$ against this genomic database showed that most of the biofilms contained multiple variants of Cyt$_{572}$ proteins that copurified (Table 4, Supplementary Figure S3). Preliminary experiments indicate that additional *Lepto*II and *Lepto*III sequence variants can be recovered from biofilm samples by shotgun genomic sequencing and PCR amplification, indicating the sequence diversity of genes encoding Cyt$_{572}$ is considerably broader than currently represented in the 5way CG and UBA genomic data sets (data not shown).

Table 3 Comparison of unique peptides[a] from monoheme cytochromes from AB-End and AB-front samples

Biofilm sample	*LII_CG_8_58*	*LII_UBA_8524_235*	*LII_CG_8_20*	*LII_UBA_8524_197*
AB-End (DS1)	0	29	0	7
AB-Front (DS2)[b]	12	10	3	4

Abbreviation: DS, developmental stage.
[a] Reported unique peptides are the average of three technical replicates.
[b] Unique peptides and associated spectra distinguishing the variants in the AB-End and AB-Front proteome are found in: compbio.ornl.gov/comparative_genomics_proteomics_of_leptospirillum/analysis.

Redox behavior of Cyt$_{572}$ from biofilms

When incubated with 1 mM Fe(II) at pH 2.0, the four preparations of Cyt$_{572}$ were fully reduced, as observed for the previously reported sample. As proteomic measurements showed that Cyt$_{572}$ preparations were mixtures of many sequence variants, formal redox titrations were not performed. However, to compare early and late developmental-stage samples, Cyt$_{572}$ preparations isolated from C-drift DS1 and DS2 samples were poised with a mixture of redox buffers at 570 mV and visible spectra recorded (Figure 9). In both cases, the preparation was reduced by 75–80%, indicating that despite the sequence variation, the overall redox behavior is similar between developmental stages.

Discussion

We isolated redox-active proteins from a natural biofilm community and showed that posttranslational

Table 4 Detected Cyt$_{572}$ variants

Variant[a]	*Molecular weight*[b]	*% Sequence coverage*	*No. of total peptides*	*No. of unique peptides*[c]
C-drift 10-12 m DS1 (11/06)				
LII_CG_variant E	24 044	84.8%	101	1
LII_CG_variant D*	61 176	67.9%	291	3
LIII_CG_variant B	23 641	53.0%	15	3
LIII_UBA_variant E	33 325	51.1%	45	13
LIII_UBA_variant G*	59 449	45.5%	49	11
LII_CG_variant A	48 428	19.8%	170	17
C-drift 10-12 m DS2 (11/06)				
LII_CG_variant B	39 445	95.7%	916	2
LII_CG_variant A	29 204	84.6%	394	2
LII_CG_variant C	27 865	83.9%	513	1
LII_CG_variant D*	61 176	74.6%	877	5
LIII_CG_variant B	23 641	63.7%	64	8
LIII_UBA_variant G*	59 449	61.6%	215	30
LIII_UBA_variant E	33 325	58.4%	93	44
C75 m (6/06)				
LII_CG_variant D*	61 176	74.7%	838	4
C75 m (11/06)				
LII_CG_variant D*	61 176	78.7%	633	2
LIII_CG_variant B	23 641	72.6%	65	24
LIII_UBA_variant G*	59 449	71.5%	181	32
LIII_UBA_variant E	33 325	67.0%	69	27

Abbreviations: Cyt$_{572}$, Cytochrome$_{572}$; DS, developmental stage.
[a] Sequences for Cyt$_{572}$ variants are listed in Supplemental Figure 2A in (Jeans *et al.*, 2008); note that in this figure sequences L3.5CG_C-I should be labeled L3.UBA_C-I as they were previously assigned to the 5CG genomic data set and later corrected.
[b] Full sequences that align with LII_CG_630_6 (composite Cyt$_{572}$ sequence) are denoted with an asterisk (*); all other sequences are partial sequences.
[c] Identity of unique peptides is listed in Supplementary Figure S3.

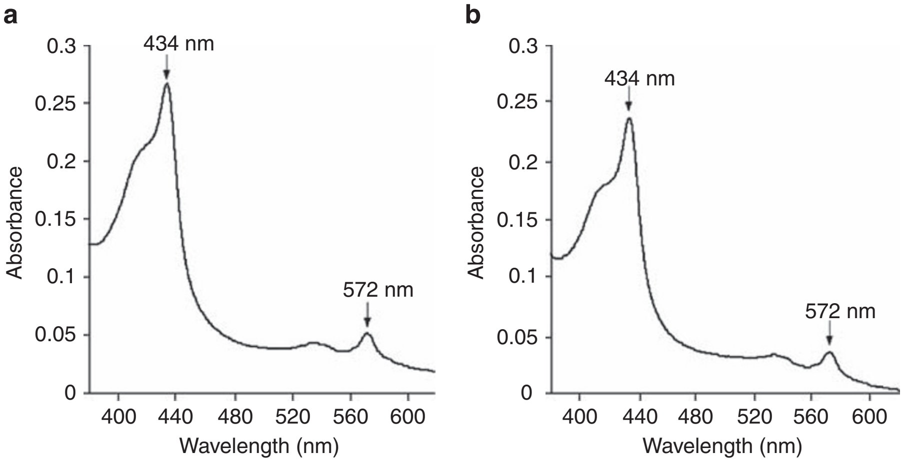

Figure 9 Poising of C-drift DS1 and DS2 Cyt$_{572}$ preparations at −570 mV. C-drift Cyt$_{572}$ preparations were diluted into 10 mM NH_4OAc containing redox mediators as described in Experimental Procedures and poised with Na_2IrCl_6 at −570 mV: (**a**) C-drift DS1 (25 μg ml^{-1}); (**b**) DS1 (30 μg ml^{-1}).

modification and expression of sequence variants are dependent on biofilm developmental stage. Isolation of Cyt$_{579}$, hypothesized to be the dominant electron-transfer protein in the periplasm of *Lepto*II (Singer *et al.*, 2008), established that protein abundance, N-terminal cleavage sites and midpoint potentials are correlated with biofilm growth stage. A C-terminal cleavage site was identified that is common to all isoforms of Cyt$_{579}$ recovered from Richmond Mine biofilms and also inferred, based on MS analysis, for intact Cyt$_{579}$ purified from the *Leptospirllum ferriphilum* P3A isolate (RC Blake II, unpublished data). Therefore, C-terminal cleavage seems to be a general posttranslational modification of Cyt$_{579}$ by *Leptospirillum* strains. The N-terminal cleavage sites do not appear to be adventitious or represent nonspecific protein degradation, as they have been observed in Cyt$_{579}$ samples prepared from multiple biofilms collected over 4 years. In addition, the N-terminal cleavage sites of late developmental-stage biofilms are unchanged when Cyt$_{579}$ purification was performed in the presence of protease inhibitors (data not shown). Therefore, the N terminus of Cyt$_{579}$ functions as a molecular 'clock' that correlates the protein isoform with the maturity of the biofilm. Geochemical conditions also affect the expression of Cyt$_{579}$, as substantially less Cyt$_{579}$ was isolated from a biofilm (C75 m in June 2006) collected under low pH conditions that had higher ratio of archaeal to bacterial community membership. Despite the significant changes in the properties of Cyt$_{579}$ in different biofilms, a striking observation enabled by high-resolution MS is that one dominant sequence variant is expressed in all of the biofilms from which the protein has been isolated.

For *c*-type cytochromes expressed by *Lepto*II, the diheme cytochromes identified by proteomic and biochemical techniques were invariant, whereas two monoheme cytochromes had distinct sequence variants in the reconstructed 5way CG and UBA genomes of *Lepto*II. The sequences identified in the *Lepto*II UBA genome are expressed in both early and late developmental-stage biofilms; however, the *Lepto*II 5way CG variants are detected by proteomics only in late developmental-stage samples. This observation is consistent with appearance of strains of *Lepto*II related to the reconstructed 5way CG strain only in late developmental-stage biofilms, as assayed by proteomic measurements and fluorescent *in situ* hybridization (Denef *et al.*, 2009). The monoheme variants from the 5way CG *Lepto*II strains expressed in the late developmental-stage biofilms may have altered redox potentials in comparison with the UBA *Lepto*II variants.

For Cyt$_{572}$, the N-terminal cleavage site and redox behavior were not altered by biofilm age and changing environmental conditions. In most of biofilm samples from which Cyt$_{572}$ was isolated, peptides that mapped to multiple genomic variants were detected from pepsin-digested samples. The significant amount of variation observed for genes coding for Cyt$_{572}$ may be a result of fine-scale environmental interactions, as distinct subpopulations of *Lepto*II colonize micro-niches in the biofilm. Localization of Cyt$_{572}$ to the outer membrane would cause the protein to interact directly with the environment and therefore be sensitive to small perturbations.

Previously, we proposed that Cyt$_{572}$ may function as the Fe(II) oxidase for *Leptospirillum* Group II, oxidizing Fe(II) on the surface of cells and transferring electrons to Cyt$_{579}$ (Figure 10a) (Jeans *et al.*, 2008; Singer *et al.*, 2008). This study suggests that *c*-type cytochromes are present at significantly lower levels and may bifurcate electrons from the dominant proteins to the cytochrome *c* oxidase and

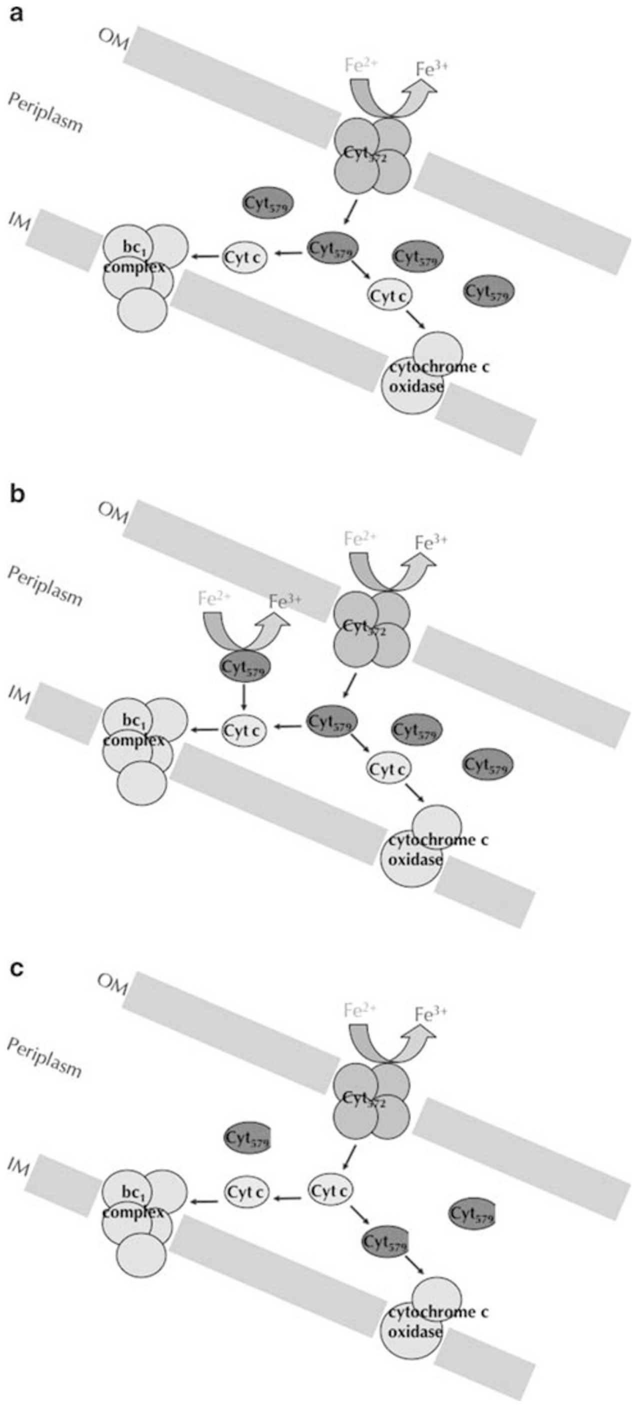

Figure 10 Working models for Fe(II) oxidation by *Lepto*II in Richmond Mine biofilms. OM represents the outer membrane of *Lepto*II cells, IM represents the cytoplasmic membrane of *Lepto*II cell: (**a**) mechanism of Fe(II) oxidation based on previously published experiments on Cyt_{579} and Cyt_{572}(Jeans *et al*., 2008; Singer *et al*., 2008); (**b**) alternative mechanism for Fe(II) oxidation in DS1 biofilms with both Cyt_{572} and Cyt_{579} as Fe(II) oxidases; (**c**) alternative mechanism for Fe(II) oxidation in DS2 biofilms with *c*-type cytochromes as initial electron acceptors from Cyt_{572}. The color reproduction of this figure is available on the html full text version of the manuscript.

cytochrome bc1 complexes. Mono-and diheme cytochromes have been shown to transfer electrons from rusticyanin to cytochrome *c* oxidase and cytochrome bc1 complex in *A. ferrooxidans* (Castelle *et al*., 2008) (Taha *et al*., 2008). However, the observation of developmental stage-dependent changes in the cytochromes complicates this model of Fe(II) oxidation. Our initial model was based on the observation that Cyt_{572} is capable of oxidizing Fe(II) at pH 1–3, whereas Cyt_{579} was only active at pH above 3. Our current study shows that Cyt_{579} isolated from carefully collected samples of early developmental-stage biofilms is readily reduced by Fe(II) at low pH and that the Cyt_{579} samples with lower redox potentials are from late developmental stages. These results suggest that Cyt_{572}, the redox potential of which is invariant as a function of developmental stage, and Cyt_{579} have similar redox potentials in early developmental-stage biofilms. Alternatively, Cyt_{579} may function as an alternative Fe(II) oxidase in early developmental-stage biofilms, as has been previously proposed (Figure 10b) (Ram *et al*., 2005). In late developmental-stage biofilms, the lowering of the redox potential of Cyt_{579} may represent an alteration of the pathway to enable the use of different electron acceptors (Figure 10c). Elucidation of these alternative pathways for electron transfer will require cytoplasmic membrane preparations from the biofilm that contain active preparations of cytochrome *c* oxidase and the cytochrome bc1 complex from *Leptospirillum* Group II.

Previous studies with single species biofilms have described changes in protein type and abundance as these biofilms mature (Southey-Pillig *et al*., 2005, Teal *et al*., 2006). Our work is unique in that it shows both posttranslational modification and sequence variation in individual proteins in the life cycle of a natural, multispecies biofilm. This study shows that combining both high-throughput proteomics measurements and targeted biochemical studies can identify highly expressed proteins in natural microbial communities that may be sensitive to changes in the environment or species composition. These observations are critical to link biochemical pathways to the functioning of natural microbial communities.

Acknowledgements

We are grateful to the Banfield laboratory members at the University of California, Berkeley, for obtaining biofilm samples, and to TW Arman, President, Iron Mountain Mines, R Sugarek, EPA, and R Carver for site access and on-site assistance. We thank Mary Ann Gawinowicz at the Columbia University Protein Core Facility for N-terminal sequence analysis, and Stephanie Malfatti at LLNL for DNA sequencing. This work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344, and was funded by the DOE Office of Science, Genomics: GTL Program grant DE-FG02-05ER64134 to JFB, RLH and MPT.

References

Bond PL, Smriga SP, Banfield JF (2000). Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl Environ Microbiol* **66**: 3842–3849.

Bradford MM. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein-utilizing principle of protein-dye binding. *Anal Biochem* **72**: 248–254.

Castelle C, Guiral M, Malarte G, Ledgham F, Leroy G, Brugna M *et al.* (2008). A new iron-oxidizing/O_2-reducing supercomplex spanning both inner and outer membranes, isolated from the extreme acidophile *Acidithiobacillus ferrooxidans*. *J Biol Chem* **283**: 25803–25811.

Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC *et al.* (2010a). Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc Natl Acad Sci USA* **107**: 2383–2390.

Denef VJ, Mueller RS, Banfield JF. (2010b). AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *ISME J* **4**: 599–610.

Denef VJ, VerBerkmoes NC, Shah MB, Abraham P, Lefsrud M, Hettich RL *et al.* (2009). Proteomics-inferred genome typing (PIGT) demonstrates inter-population recombination as a strategy for environmental adaptation. *Environ Microbiol* **11**: 313–325.

Francis RT, Becker RR. (1984). Specific indication of hemoproteins in polyacrylamide gels using a double-staining process. *Anal Biochem* **136**: 509–514.

Goltsman DSA, Denef VJ, Singer SW, VerBerkmoes NC, Lefsrud M, Mueller RS *et al.* (2009). Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing '*leptospirillum rubarum*' (group ii) and '*leptospirillum ferrodiazotrophum*' (group iii) bacteria in acid mine drainage biofilms. *Appl Environ Microbiol* **75**: 4599–4615.

Haveman SA, Holmes DE, Ding YHR, Ward JE, DiDonato RJ, Lovley DR. (2006). *c*-Type cytochromes in *Pelobacter carbinolicus*. *Appl Environ Microbiol* **72**: 6980–6985.

Jeans C, Singer SW, Chan CS, VerBerkmoes NC, Shah M, Hettich RL *et al.* (2008). Cytochrome 572 is a conspicuous membrane protein with iron oxidation activity purified directly from a natural acidophilic microbial community. *ISME J* **2**: 542–550.

Laemmli UK. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature* **227**: 680–685.

Lo I, Denef VJ, VerBerkmoes NC, Shah MB, Goltsman D, DiBartolo G *et al* (2007). Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. *Nature* **446**: 537–541.

Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC *et al.* (2005). Community proteomics of a natural microbial biofilm. *Science* **308**: 1915–1920.

Schrenk MO, Edwards KJ, Goodman RM, Hamers RJ, Banfield JF. (1998). Distribution of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans*: implications for generation of acid mine drainage. *Science* **279**: 1519–1522.

Simmons SL, DiBartolo G, Denef VJ, Goltsman DSA, Thelen MP, Banfield JF. (2008). Population genomic analysis of strain variation in *Leptospirillum* group II bacteria involved in acid mine drainage formation. *PLOS Biol* **6**: 1427–1442.

Singer SW, Chan CS, Zemla A, VerBerkmoes NC, Hwang M, Hettich RL *et al.* (2008). Characterization of cytochrome 579, an unusual cytochrome isolated from an iron-oxidizing microbial community. *Appl Environ Microbiol* **74**: 4454–4462.

Southey-Pillig CJ, Davies DG, Sauer K. (2005). Characterization of temporal protein production in *Pseudomonas aeruginosa* biofilms. *J Bacteriol* **187**: 8114–8126.

Taha TM, Kanao T, Takeuchi F, Sugio T. (2008). Reconstitution of iron oxidase from sulfur-grown *Acidithiobacillus ferrooxidans*. *Appl Environ Microbiol* **74**: 6808–6810.

Teal TK, Lies DP, Wold BJ, Newman DK. (2006). Spatiometabolic stratification of *Shewanella oneidensis* biofilms. *Appl Environ Microbiol* **72**: 7324–7330.

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM *et al.* (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**: 37–43.

VerBerkmoes NC, Denef VJ, Hettich RL, Banfield JF. (2009). SYSTEMS BIOLOGY functional analysis of natural microbial consortia using community proteomics. *Nat Rev Microbiol* **7**: 196–205.

Wilmes P, Remis JP, Hwang M, Auer M, Thelen MP, Banfield JF. (2009a). Natural acidophilic biofilm communities reflect distinct organismal and functional organization. *ISME J* **3**: 266–270.

Wilmes P, Simmons SL, Denef VJ, Banfield JF. (2009b). The dynamic genetic repertoire of microbial communities. *FEMS Microbiol Rev* **33**: 109–132.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)