Sequence analysis and confirmation of type IV pili-associated proteins PilY1, PilW and PilV in *Acidithiobacillus thiooxidans*

Type IV pili-associated proteins PilY1, PilW and PilV in *Acidithiobacillus thiooxidans*

Elvia Alfaro-Saldaña[1,2], Araceli Hernández-Sánchez[2], O. Araceli Soberano-Patrón[3], Marizel Astello-García[1], J. Alfredo Méndez-Cabañas[2], J. Viridiana García-Meza[1*]

[1] Geomicrobiología, Metalurgia, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

[2] Biofísica Molecular, Instituto de Física, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

[3] División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica, San Luis Potosí, México

* Corresponding author

E-mail: jvgm@uaslp.mx (JVGM)

JVGM is senior author

Abstract

Acidithiobacillus thiooxidans is an acidophilic chemolithoautotrophic bacterium widely used in the mining industry due to its metabolic sulfur-oxidizing capability. The biooxidation of sulfide minerals is enhanced through the attachment of *A. thiooxidans* cells to the mineral surface. The Type IV pili (TfP) of *At. thiooxidans* may play an important role in the bacteria attachment, since among other functions, TfP play a key adhesive role in the attachment to and colonization of different surfaces. In this work, we reported for the first time the confirmed mRNA sequences of three TfP proteins from *At. thiooxidans*, the protein PilY1 and the TfP pilins PilW and PilV. The nucleotide sequences of these TfP proteins show changes of some nucleotide positions with respect to the corresponding annotated sequences. The bioinformatic analyses and 3D-modeling of protein structures sustain their classification as TfP proteins, as structural homologs of the corresponding proteins of *P. aeruginosa*, results that sustain the role of PilY1, PilW and PilV in pili assembly. Also, that PilY1 comprises the conserved *Neisseria*-PilC (superfamily) domain of the tip-associated adhesin, while PilW of the superfamily of putative TfP assembly proteins and PilV belongs to the superfamily of TfP assembly protein. Also, the analyses suggested the presence of specific functional domains involved in adhesion, energy transduction and signaling functions. The phylogenetic analysis indicated that the PilY1 of *Acidithiobacillus* genus forms a cohesive group linked with iron- and/or sulfur-oxidizing microorganisms from acid mine drainage or mine tailings. This work

44 enriches knowledge regarding colonization, adhesion and biooxidation of inorganic

45 sulfurs by *A. thiooxidans*.

46

Introduction

Acidithiobacillus thiooxidans is an acidophilic chemolithoautotroph that uses reduced sulfurs as a source of electrons and reducing power, including elemental sulfur (S^0), polysulfides (S_n^{2-}) and sulfide minerals, such as pyrite (FeS_2), chalcopyrite ($CuFeS_2$) or sphalerite (ZnS).

Bacterial attachment to mineral surfaces influences the rate of dissolution of the mineral because of surficial phenomena: Mixed potential decreases, changes in kinetics and mass-transport processes [1]. Accordingly, bacterial attachment is due to self-organization by a bioelectrochemical evolution on the interface. Interfacial studies on charge and mass transfer demonstrate that S^0 biooxidation by *At. thiooxidans* begins in the early stages of interaction (1 to 24 h) when the biofilm is not constituted, and it is primarily controlled by surficial characteristics that pivoted the bacterial attachment to the hydrophobic S^0; such attachment is an energy-dependent process in which *At. thiooxidans* essentially activates or modifies the reactive properties of S^0 [2, 3]. The hydrophobic character of the interface "determines the free energy of the adhesion process" [4]. The Type IV pili (TfP) of *At. thiooxidans* may play an important role in the bacterial attachment and bioelectrochemical evolution on the bacteria-mineral interface, *e.g.,* surpassing hydrophilic interactions.

Valdés *et al*. [5] and Li *et al*. [6] suggested that the efficiency of *At. ferrooxidans* to attach and colonize mineral surfaces (*e.g.,* FeS_2 or $CuFeS_2$) and solid reduced sulfur depends on TfP as well as its multiple copies of genes for pili biosynthesis. Other

69 *Acidithiobacillus* species such as *At. caldus* and *At. thiooxidans* contain genes
70 coding for TfP assembly proteins [7-9] that are related to biofilm formation, acting as
71 c-di-GMP effector proteins [10].

72 The TfP are semiflexible polymeric filaments of pilins anchored to the cellular
73 membrane, of 5-7 nm diameter and from 4-5 μm up to several micrometers in length
74 [11,12]. The TfP have been grouped based on the aminoacidic homology of the pilin
75 subunits, which are relatively conserved in prokaryotes. Pilins share an N-terminal
76 cleavage/methylation (N-methylphenylalanine) domain (NTD) of approximately 25
77 amino acids (aa) followed by a stretch of hydrophobic residues forming an extended
78 α-helix and a disulfide bond at the C-terminal domain (CTD) [11, 13]. Pilins interact
79 via their conserved NTD α-helix, which forms a hydrophobic core that provides
80 extreme mechanical strength [13]. Further, it has been reported that this conserved
81 NTD in prepilins is also in the type II secretion systems (T2SS), known as
82 pseudopilins [14-17].

83 Pelicic [18] suggested that the minimal machinery needed for TfP assembly includes
84 pilins and other TfP proteins: (i) Major pilin with NTD (pilin-like motif), (ii) specific
85 peptidase that processes the precursors of pilins or prepilins (*e.g.,* the PilD of
86 *Neisseria* spp.), (iii) traffic ATPase that powers TfP assembly (PilT), (iv) internal inner
87 membrane protein (PilG), (v) integral outer membrane or secretin necessary for the
88 emergence of TfP on the cell surface (PilQ), and (vi) proteins also found in T2SS,
89 the pilin-like proteins.

Interesting, “the TfPs are universal in prokaryotes and have shown extreme **functional** versatility”. Among other **functions** (motility, cell signaling, pathogenic functions, protein secretion, DNA uptake, electrical conductance, and so on), TfP are “sticky organelles” Berry and Pelicic [17] that play a key role as adhesive to stick bacteria to one another and to attach to and to colonize to a wide variety of surfaces, leading to the formation of colonies and even biofilms with cells embedded within an EPS matrix, including “2-D” (thin liquid) and 3-D biofilms [12, 17].

The adhesive ability of TfP is due to the presence of a non-pilin protein, an adhesin (integrin) located on the distal tip of the TfP [17, 19-21]. This adhesin, designated PilC or PilY1, is expressed in multiple species. For instance, PilY1 of *Pseudomonas aeruginosa* is the orthologue of the meningococcal PilC of *Neisseria gonorrhoeae*, and both proteins Although PilY1 and PilC share partial sequence homology, they have a high degree of structural similarity [22, 23]. PilC is a phase-variable protein that belongs to a Conserved Protein Domain Family of several PilC protein sequences [23]. Further, PilC/PilY1 is dispensable for TfP assembly at early stages of TfP biogenesis, for instance, in the absence of pilus retraction [17, 22-24]. Our previous analysis on,showed that the tip-associated adhesin PilY1 of *At. thiooxidans* Licanantay (WP_031573362] exhibited 55% and 86% identity with the type IV pilin biogenesis protein of *At. thiooxidans* ATCC 19377 (WP_010638975.1) and of *At. albertensis* (WP_075322776.1), respectively.

Other proteins involved in adhesion are the minor but highly conserved pilin proteins PilW and PilV of *Neisseria* spp. [17, 24, 25]. The outer-membrane lipoprotein PilW

participates in pilus biogenesis for the stabilization of the pilus but not for their assembly, as well as to allow bacterial adherence [26]. Another possible function for PilW is the transfer of electrons, as has been proposed for *At. ferrooxidans* [27] and other species with TfP [28].

Thus, the putative proteins PilY1, PilW and PilV of *At. thiooxidans* were examined in this work. We chose these proteins because of PilY1's possible function as an adhesin and the role of PilW and PilV as pilus assembly proteins (structural pilins) of the TfP of *At. thiooxidans*. Most genes of *At. thiooxidans* are annotated in GenBank. In this study, the sequences of the TfP proteins PilY1, PilW and PilV were first confirmed and uploaded to GenBank. In addition, bioinformatic analyses and 3-D modeling of each TfP protein were performed.

Materials and methods

Acidithiobacillus thiooxidans culture and maintenance

At. thiooxidans strain ATCC 19377 was cultured in ATCC 125 medium (in g/L: 0.4 $(NH_4)_2SO_4$, 0.5 $MgSO_4 \cdot 7H_2O$, 0.25 $CaCl_2$, 3 KH_2PO_4, 0.005 $FeSO_4 \cdot 7H_2O$, and 10 S^0, pH 2.0 adjusted with concentrated H_2SO_4). Cultures were aerobically incubated at 29 ± 1 °C under orbital agitation (110 to 120 rpm) for 5, 10, 15 and 21 days.

Multiplex PCR and cDNA sequencing of *pilY1*, *pilW*, and *pilV*

We used the *At. thiooxidans* ATCC 19377 T draft genome AFOH01000001 originally described by Valdés et al. [7]. The annotated protein sequence of PilY1, PilW and PilV as well as the 16S ribosomal subunit with the accession numbers WP_010638975.1, WP_010638981.1, WP_010638979.1 and AJ459803, respectively, were. Primers were designed against conserved regions of the putative sequences of *pilY1* (ATHIO_RS0106065), *pilW* (ATHIO_RS0106075) and *pilV* (ATHIO_RS0106080) using the software Primer3 [29] and Vector NTI (InforMax-Invitrogen, USA). As a positive control for each PCR reaction, we designed a pair of primers that amplified the 16S rRNA.

Samples for RNA extraction were extracted after 5, 10, 15 and 21 days. At these timepoints, cells were collected by centrifugation at 0.08 x g to separate the cells from S^0; cells were concentrated and washed by centrifugation at 21.1 x g for 1 min using saline phosphate buffer (PBS; in g/L: 8 NaCl, 1.44 Na_2HPO_4, 0.24 KH_2PO_4 (J.T. Baker, USA), and 0.2 KCl (Sigma-Aldrich, USA); pH 7.4).

RNA extraction was performed using TRIzol (Invitrogen, USA) reagent according the manufacturer's protocol. After quantification using a NanoDrop 2000c (Thermo Scientific, USA), cDNA synthesis was performed according to the manufacturer's recommendation, using 200U M-MLV (Invitrogen, USA) in a 20 µL total volume reaction. The amplification products were purified using the QIAquick kit (Qiagen, USA) following the recommended procedure, quantified and sequenced. The

152 products were sequenced by the Sanger method at LANBAMA (National Laboratory
153 of Biotechnology, San Luis Potosí, Mexico) and LANGEBIO (National Laboratory of
154 Genomic for the Biodiversity, Guanajuato, México) facilities. Each sequence was
155 confirmed at least five times by analyzing amplification products obtained from
156 different culture replicates. The generated sequences were compared with the
157 annotated sequences of each corresponding gene using the Multalin software [30].

158 In silico (bioinformatic) analyses of proteins

159 The confirmed sequences of PilY1, PilW and PilV were analyzed to search for
160 domains, families and functional sites with bioinformatic tools, including Simple
161 Modular Architecture Search, SMART [31], PROSITE [32], and EMBL-EBI [33] and
162 the CDD/SPARCLE domain analysis function [34]. The protein structure homology-
163 modeling was achieved by using the server SWISS-MODEL [35]. The 3-D model of
164 each protein was generated in I-TASSER [36, 37]]. The analyses of possible ligands
165 were performed in the Ligand-Protein Binding Database, BioLip [38].

166 Phylogenetic relationships

167 Phylogenetic relationships were analyzed for the confirmed nucleotide and protein
168 sequences of PilY1, PilW and PilV from *At. thiooxidans* ATCC19377 against those
169 of the NCBI database using BLASTN 2.8.0+ [39, 40], and UniProt [41] bioinformatic
170 resources.

171 The sequences obtained from NCBI were aligned with those obtained in this work
172 using MEGA7 software [42] by the Neighbor-Joining method. Thus, the evolutionary

history was inferred by using the Maximum Likelihood method (ML) based on the Jones-Taylor Thornton (JTT) model [43]; for the bootstrap consensus tree, 500 replicates were performed [44]. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model; the tree with the highest log likelihood is shown. All positions with less than 95% site coverage were eliminated.

Transmission electron microscopy (TEM) analyses

After 4 days of incubation, *A. thiooxidans* cells were negative stained with 1% phosphotungstic acid (PTA) or 2% uranyl acetate. One aliquot was directly used, and the other was fixed previously to reduce the insoluble S^0 present in the sample and stabilize pili. Both samples were washed three times at 1500 rpm for 5 min and stained with PTA. The pili were observed in TEM (JEOL 200 CX, Japan) at 100 kV.

Results and Discussion

TEM analyses

After the microscopic analyses, we identified pili in all tested conditions (Fig 1a): i) Negative staining with 2% uranyl made the flagellum, pili, and S^0 visible; ii) Upon negative staining with 1% PTA, the pili and flagellum were surrendered by extracellular polymeric substance; iii) *A. thiooxidans* were previously washed to remove S^0 from the medium, then stained with 1% PTA, although S^0 was eliminated,

194 and many flagella and pili were lost as well. We fixed the bacteria before washing,
195 and in this condition, we preserved the pili, arranged in a pili network (iv, insert).

196

197 **Figure 1. Representative micrographs and multiplex PCR amplification**
198 **products of *pilY1*, *pilW*, and *pilV*.**

199 The TEM microphotographs of *At. thiooxidans* showed the pili (triangles) and flagella
200 (arrows) (**a**). The multiplex PCR was obtained after 5 days of culture; the positive
201 control corresponds to a region of 598 bp of the rRNA 16S (**b**).

202

203 Sequence analyses

204 For the sequencing of *pilY1*, *pilW* and *pilV*, primers were designed against conserved
205 regions of the putative sequences of 16S from *At. thiooxidans* (GenBank
206 NZ_AFOH01000047.1; [7]. Due to its size (3499 bp), we identified conserved regions
207 of the annotated sequence of *pilY1* comparing it with its homologue in *At.*
208 *ferrooxidans* strain ATCC 53993 (WP_064218310.1), according to the BLAST
209 analysis. The eight pairs of primers designed (*pilY-1* to *pilY-8*; S2 Table) were used
210 to amplify, purify and sequence (per triplicate) each PCR product of such regions;
211 each obtained amplicon was aligned against the annotated sequence of the *pilY1*
212 gene (ATHIO_RS0106065) to highlight changes between them. After analyzing each
213 sequenced amplicon, the complete sequence was rejoined, and new pairs of primers
214 were designed (*pilY-9* to *pilY-11*, S2 Table) to assess if the highlighted changes were

215 *bonafide*; each pair of these three primers initiates just in the regions with the most
216 significant insertions and deletions previously found. The new obtained amplicons
217 were individually aligned against their corresponding fragment of the reassembled
218 sequence and then with the annotated sequence of *At. thiooxidans*.

219 A similar strategy was performed for the sequencing of *pilW* and *pilV*, using 5 and 4
220 pairs of designed primers, respectively (S2 Table); the obtained amplicons were
221 compared with the corresponding annotated sequence of *pilW* (ATHIO_RS0106075)
222 and *pilV* (ATHIO_RS0106080)

223 Once the complete sequences were obtained, the expression of *pilY1*
224 (MH021598.1), *pilW* (MH021599.1) and *pilV* (MH021600.1) was evaluated at
225 different culture times (Fig 1b).

226 Thus, the nucleotide sequences of *pilY1*, *pilW* and *pilV* show changes at some
227 nucleotide positions with respect to the corresponding annotated sequences (S1
228 Table). The changes are more significant for *pilY1* mainly within the N-terminal 1560
229 bp region. In this region, the difference between the annotated and the confirmed
230 sequences is approximately 25% due to changes of some bases as well as
231 insertions and deletions of 25 and 7 nucleotides, respectively. In the second half of
232 the sequence (1991 bp), the differences between both annotated and confirmed
233 sequences is just 3% due to two insertions of six nucleotides in the confirmed
234 sequence.

235 However, annotated and confirmed protein sequences showed a pairwise distance
236 of 0.13 (Fig 2a). The PilY1 (AWP39905.1) sustains 85% identity (1007 identical

237 position) with the annotated WP_010638975.1 reported in GenBank. According to
238 BLAST analysis, PilY1 (AWP39905.1) sustains 99% and 77% identity with the
239 sequence of the type IV pilin biogenesis protein of *At. thiooxidans*
240 (WP_065968128.1) and the hypothetical protein of *At. ferrooxidans*
241 (WP_064218310.1); these sequences comprise the *Neisseria*-PilC superfamily
242 domain [34].

243

244 **Figure 2**. **Comparation between the confirmed and annotated sequences**.

245 (**a**) PilY1 AWP39905.1 and WP_010638975.1, (**b**) PilW AWP39906.1 and
246 WP_010638979.1 and (**c**) PilV AWP39907.1 and WP_010638981.1, by multiple
247 sequence alignment accuracy and high throughput (MUSCLE) and toggling
248 conserved sites at 100% level (in colors); computed pairwise distance (Poisson
249 model) is also shown. Analyses were done in MEGA7 [42].

250

251 Although *pilW* (MH021599.1) shows 20 synonymous changes, the protein sequence
252 (AWP39906.1) shows 98% identity against the annotated and translated sequence
253 of the prepilin-type N-terminal cleavage/methylation domain-containing protein,
254 WP_010638979.1 (ATHIO_RS0106075; Fig 2b). Both sequences comprise a region
255 of 132 aa, the PilW superfamily of putative TfP assembly proteins, as well as *At.*
256 *ferrooxidans* ATCC 23270 (WP_064218312.1; 83% identity with *At. thiooxidans*
257 AWP39906.1).

Finally, the *pilV* sequence of (MH021600.1) exhibited three synonymous changes that resulted in an identical protein to the annotated and translated WP_010638981.1 (Fig 2c). PilV (AWP39907.1) belongs to the PilV super family of TfP assembly protein, an extracellular structure involved in cell motility [34]. Blast analysis for PilV (AWP39907.1) indicated 92% identity with the pilin (putative) of *At. ferrooxidans* ATCC 23270 (ACK80286.1).

Proteins bioinformatic analyses and phylogenetic trees

PilY1

The bioinformatic analyses of PilY1 (AWP39905.1) confirm that it is a non-pilin protein, essentially hydrophilic (gravy -0.081), and its last 462 aa (48.19 kDa region 715-1176 aa) shares sequence and structural homologies with the C-terminal domain (CTD) of PilY1 of *Ps. aeruginosa* (3hx6.1) [23] that belongs to the conserved domain *Neisseria*-PilC superfamily [34]. In contrast, the NTD of PilC of *Neisseria* spp., PilY1 of *P. aerugionosa* and the PilY1 reported in this work are non-conserved, divergent sequences. Both NTD and CTD sequences of the *At. thiooxidans* PilY1 share similarities with the TfP biogenesis protein of *Acidithiobacillus* spp. (Fig 3).

Figure 3. Multiple sequence alignment of PilY1 (AWP39905.1) of *At. thiooxidans*.

278 Showing two regions: (**a**) part of the vWA domain within the NTD of some
279 *Acidithiobacillus* spp. showing the MIDAS (DxSxS) motif, and (**b**) the PQQ domain
280 with the functional motif LYxxxxxG within the CTD of some *Acidithiobacillus* spp. and
281 other bacteria with PilY1/PilC.

282

283 The previous is evidenced in the phylogenetic tree (Fig 4), wherein the
284 *Acidithiobacillus* genus forms a cohesive group, deeply related with iron- and/or
285 sulfur-oxidizing microorganisms from acid mine drainage (AMD) or mine tailings
286 such as *Th. bhubaneswarensis*, *Ac. thiooxydans*, *Ga. acididurans*, *Sulfuriferula* sp.,
287 and those of the AMD metagenome [45-49]. This phylogenetic tree also reveals that
288 PilY1 and PilC are homologues; all the analyzed sequences comprise the *Neisseria*-
289 PilC beta-propeller domain. Thus, the phylogenetic analysis of PilY1/PilC suggests
290 homologating the nomenclature of this TfP protein (Fig 4), perhaps as the *Neisseria*
291 spp. pilus assembly/adherence protein PilC.

292

293 **Figure 4**. **Molecular phylogenetic analysis of PilY1 (AWP39905.1)**.
294 The evolutionary history obtained by ML (log likelihood: -8108.55) of PilY1/PilC,
295 using 21 aa sequences and the PilY1 (AWP39905.1). There were 212 positions in
296 the final dataset. Evolutionary analyses were conducted in MEGA7 [42]. The aa
297 sequences used for alignment and phylogenetic tree were derived from: acid mine
298 drainage (AMD) metagenome (CBI05240.1), *A. warmingii* DSM 173 (SDX64909.1),
299 *Acidifferobacter thiooxydans* ZJ (OCX45123.1), *At. albertensis* DSM 14366
300 (WP_075322776), *At. caldus* ATCC 51756 (AIA55059.1 and WP_064306242.1), *At.*

ferrivorans YL15 (WP_085537817.1), *At. ferrooxidans* ATCC 23270 (ACK78377.1) *At. ferrooxidans* ATCC 53993 (ACH83324.1), *At. thiooxidans* (WP_010638975.1) *At. thiooxidans* Licanantay (WP_031573362.1), *Ga. acididurans* (isolate ShG14-8; KXS31914.1), *Ge. sulfurreducens* KN400 (ADI84872.2), *Methyloglobulus morosus* KoM1 (ESS74050.1); *Ns. meningitidis* (WP_101069668.1), *Nitrosomoas communis* Nm110 (SDW55994.1), *Ps. aeruginosa* PAO1 (AAA93502.1), *Rhodoferax* sp. DCY110 (WP_076201095.1), PilC of *Sulfuriferula* sp. AH1 (WP_087447088.1), *Te. thermophilus* JCM 19170 (CUB07558.1), *Th. bhubaneswarensis* DSM 18181 (CUA94317.1), *Thiorhodococcus drewsii* AZ1 (EGV31806.1). Tfp: type IV pilin.

Our results are consistent with the predicted 3-D model of *At. thiooxidans* PilY1 (AWP39905.1; Fig 5). This model was generated by I-TASSER based on structures from the PDB (5j44A, 3hx6, 6emkA, 3ja4A and 3iylW), highlighting the PilY1 CTD structure of *P. aeruginosa* (3hx6A and 3hx6) [23]. According to the homology-modeling, PilY1 of *At. thiooxidans* (AWP39905.1) showed similarity (22.73%) with such PilY1of *P. aeruginosa*, mainly among the CTD that corresponds to the *Neisseria*-PilC beta-propeller domain of the tip-associated adhesin PilY1.

Figure 5. **3-D models of the non-pilin protein PilY1 (AWP39905.1).**

The model shown has a TM-score of 0.61±0.14 with *Ps. aeruginosa* (TM-score >0.5 indicates a correct topology model, while TM-score <0.17 indicates a random similarity) **(a)**. A region of vWA showing the binding ligands L156, Y156, Y-319 and

T-320 of MIDAS, for Ca^{2+} binding (**b** and **b'**). a region CTD of PilY1/PilC (TM-score of 0.62±0.14) showing the binding residues for carbohydrates (*e.g.,* α-D-glucose, β-glucose, α-dextrose) by S-754, L-774, G-776, M-778, D-834, L-835, Q-836, L-837 **(c** and **c')**; 3-D modeling conducted in I-TASSER [36, 37] and BioLip [38]. Model of the PQQ region predicted by homology [35] and the anchor motif LYxxxxTG (**d** and **d'**).

At. thiooxidans PilY1 (AWP39905.1) bioinformatic analysis confirmed the presence of five regions or motifs (Figs 3 and 5) also found in TfP proteins from *At. ferrivorans* SS3 (WP_085537817.1), *At. ferrooxidans* ATCC 53993 (ACH83324.1), *At. albertis* (WP_075322776), *At. caldus* MTH-04 (WP_064306242.1), and *At. thiooxidans* Licanantay (WP_031573362):

(i) A von Willebrand factor type A (vWA) domain (540 aa) was found in the region from amino acid 42 to 581, folded into a α/β Rossmann fold (alternating β-strand with α-helix). Cellular functions such as cell migration, adhesion and signaling have been associated with the vWA domain [34, 50]. E-value (BLAST):$1x10^{-28}$.

(ii) Within the vWA, a metal ion-dependent adhesion sites (MIDAS motif, for Mn^{2+}, Mg^{2+} or Ca^{2+}) of 35 aa residues (region 153-187) folded as a short coil followed by an α-helix. MIDAS are commonly present in cell surface-adhesion receptors or molecules (CAMs), *e.g.,* integrins [51]; thus, such components are involved in cell-cell and cell-matrix interactions through its adhesion function. MIDAS comprised the conserved components DxSxS, T, and D (Fig 3a) [34, 52]. Modeling by homology [35], the

MIDAS motif of *A. thiooxidans* PilY1 (AWP39905.1) exhibited 34.38% similarity (covering 91%, from 2 to 33 aa) with the model of a cell wall surface anchor family protein (PDB: 3tw0.1.A), specifically with the adhesive tip pilin of *Streptococcus agalactiae* [50]. The constructed 3-D model specifies that the MIDAS motif within vWA includes two different loops (α-helix) exposed in the protein surface (Fig 5b and b'), a location that allows it to interact with divalent metal ions such as Ca^{2+} via the binding ligands I-154, N-156 and T-320 of MIDAS.

(iii) The aforementioned conserved CTD or *Neisseria*-PilC domain is 462 aa (715-1176 aa) mainly composed by aliphatic aa (G, A, V, L and S, T; 60.2%) and slightly hydrophilic (gravy: -0.094). The 3-D model of this CTD (Fig 5c) was based only on structures of the CTD of *P. aeruginosa* PilY1, 3hx6 [23], showing a normalized Z-score of the threading alignments, up to 7.52. Other structural analogs predicted using I-TASSER, are oxidoreductases -substrate oxidation and electron transfer (PDB: 1h4jE, 1lrwA, 2d0vI, 1yiqA1, 1FLG), with Ca^{2+} or pyrroloquinoline quinone (PQQ) as ligands (see below).

(iv) Within the *Neisseria*-PilC domain, the 3-D model also predicts binding sites for carbohydrates such as glucose (Fig 5c), previously described for tail-spike proteins for recognition and adhesion of *Salmonella* and *E. coli* bacteriophage HK620 [53].

(v) Also within the *Neisseria*-PilC domain is a motif that occurs in propellers of PQQ cofactor binding domains (SMART and InterPro accession numbers SM00564 and IPR018391, respectively) of 27 aa (region 926-952) that catalyzes redox reactions; *e.g.*, a β-propeller repeat occurring in enzymes with PQQ as cofactor in prokaryotic

367 quinoprotein dehydrogenases that are involved in electron transfer processes, and
368 thus energy transduction [54]. *Ps. aeruginosa* PilY1 (AAA93502) and PilC of *N.*
369 *meningitidis* 22404 (WP_101069668) also include the PQQ cofactor binding
370 domain, while most of the *Acidithiobacillus* spp. and other sulfur-oxidizing
371 microorganisms comprise the anchor motif LYxxxxTG (Figs 5b and d).

372 The iron- and/or sulfur-oxidizing microorganisms from AMD (*e.g.*, *Th.*
373 *bhubaneswarensis*, *Ac. thiooxydans*, *Ga. acididurans*, *Sulfuriferula* sp.; Fig 4) also
374 contain the PQQ domain (Fig 3b), while the PilY1-related protein (fragment) from the
375 AMD metagenome (CBI05240.1) mostly corresponds to the vWA domain within the
376 NTD that also comprises the MIDAS motif, DxSxSxxxxxxxxxT (Fig 3a). Further, the
377 other sulfur-oxidizing microorganisms presented in the phylogenetic tree (Fig 4), *Te.*
378 *thermophilus* and *T. drewsii* AZ1, included the PQQ domain in their tip-associated
379 adhesin PilY1/PilC, according to SMART analyses [31]. The presence of PQQ in
380 such PilY1/PilC proteins suggests that this adhesin initiates the biooxidation of
381 reduced compounds (*e.g.,* S^0 or metal sulfides as FeS_2, and $CuFeS_2$) and may be
382 involved in electron transfer between the substrate and other components. *Sensu* Li
383 and Li [27], the biooxidation of Fe(III) by *At. ferrooxidans* depends on functional pili
384 which transfer electrons from the reduced Fe to the cell surface, while *At.*
385 *ferrooxidans* attaches strongly to solid surfaces such as FeS_2 [6].

386

Pilins PilW and PilV

The 3D-modeling of PilW and PilV of *At. thiooxidans* (Figs 6a and b) revealed high similarities to the pilin of *Ps*. *aeruginosa* strain K, PAK (PDB key: 1oqw). For PilW, the model was also based on structures of a protein with structural similarity to flagellin of *Burkholderia pseudomallei* (4ut1A), a putative peptide-binding domain (adhesin) of *Marinomonas primoryensis* (5k8g), and the pilin FimA for adhesion from *Dichelobacter nodosus* (3sok).

Figure 6. **Bioinformatic analyses of the confirmed sequences of *At. thiooxidans* PilW (AWP39906.1) and PilV (AWP39907.1)**.

(**a**) 3-D model of PilW (TM-score: 0.42±0.14) and (**b**) PilV (TM-score: 0.49±0.15) conducted in I-TASSER [36, 37] and BioLip [38]; the models were obtained based on structures from the PDB data base; for PilW: 5k8g, 1oqw, 4ut1A, 3sok, 4m00, 5gaoE; for PilV: 1oqw, 3sok, 5bw0, 1ay2 and 3ci0. Modeling by homology of the NTD of (**c**) PilW (1-81 aa) and (**d**) PilV (11-62 aa) using as template the 3nje.1 minor pseudopilin from the *Ps. aeruginosa* [15]. for PilW, and the 5vxy.1.E pilin of *Ps. aeruginosa* and *N. gonorrhoeae* [55, 56] for PilV. (**e**) Multiple sequence alignment conducted in MEGA7 [42], of the region NTD (25-28 aa) of the confirmed PilW (AWP39906.1) and PilV (AWP39907.1) sequences of *At. thiooxidans*, and of other sequences from the gene bank of prepilin containing proteins of *At. thiooxidans* ATCC 19377 (WP_010638979.1), *At. thiooxidans* Licanantay (WP_051690664.1). *At. ferrooxidans* ATCC 593993 (ACH83326.1), *At. ferrivorans* (WP_035191506.1),

409 *At. albertensis* (WP_075323115.1) and *At. caldus* (WP_070113768.1), as well as
410 TfP pilins or prepilins of *Ps. auriginosa* PAO1 (NP_253215.1) *N. gonorrehaea*
411 (SBO76855.1), *E. coli* (PIM50236.1), *Microcystis aeruginosa* (WP_012267210.1),
412 *Desulfovibrio magneticus* (EKO40131.1), *Shewanella baltica* (WP_012588380.1)
413 and *Mariprofundus micogutta* (WP_083530569.1).

414

415 As the sequences 1oqw and 3sok, PilW and PilV of *At. thiooxidans* and other TfP
416 pilins have leader peptides of approximately 20-25 aa in the NTD with a highly
417 conserved G and the GFXXXXE domain (Fig 6e). According to Dalrymple and
418 Mattick [57] and Mattick [58], all the necessary information for the processing of
419 subunits and assembly of pili is in these first aa of the α1-N; thus, mutations of NTD
420 residues can markedly affect pilus assembly [13]. Specifically, the conserved
421 glutamic acid E5 is the only charged residue in NTD (Fig 6b'), which is essential for
422 pilus assembly and seems to be required for the methylation step [34, 55, 59]. In *At.*
423 *thiooxidans*, E5 interacts with divalent cations such as Ca^{2+} (Fig 6b).

424 The CTD of PilW includes a region from aa residue 239 to 371 that is identified as
425 the "TfP pilus assembly protein PilW" (pfam16074). Its predicted secondary structure
426 is mainly strands and coil, that define a globular head (Fig 6a), a head also observed
427 in the 3-D model of PilV (Fig 6b). This globular head domain is a structurally variable
428 region among the different pili, variability that imposes different functions, *e.g.*,
429 specific subunit interactions and packing arrangements within the filament, pilus-
430 pilus interaction, and interactions between the pili and their environment [55].

431

432 The molecular phylogenetic analysis confirms that PilW and PilV of *At. thiooxidans*
433 are core subunits of TfP, similar to PilW and PilV of *Ps. aeruginosa* PAO1 (Fig 7).

434

435 **Figure 7**. **Phylogenetic relationships based on the pilins PilW (AWP39906.1)**
436 **and PilV (AWP39907.1) of *At. thiooxidans*.**

437 The evolutionary history obtained by ML (log likelihood -1678.78) using 12 amino
438 acids sequences. There was a total of 97 positions in the final dataset. Evolutionary
439 analyses were conducted in [42]. The aa sequences used were derived from: *At.*
440 *caldus* ATCC 51756 (WP_014002880.1); *At. ferrivorans* SS3 for PilW
441 (WP_035191506.1), *At. ferrivorans* CF27 (CDQ09112.1), *At. ferrooxidans*
442 (ACH83326.1), *At. thiooxidans* ATCC 19377 for PilW (WP_010638979.1) and PilV
443 (WP_010638981.1), *Ps. aeruginosa* PilW, (NP_253242.1) and PilV (NP_253241.1)
444 and *T. thermophilus* (AAM55484.1).

445

446 Resuming, PilY1 (AWP39905.1) comprises the conserved *Neisseria*-PilC
447 (superfamily) beta-propeller domain of the tip-associated adhesin while PilW
448 (AWP39906.1) of the superfamily of putative TfP assembly proteins, and PilV
449 (AWP39907.1) belongs to the super family of TfP assembly protein. Further analysis
450 will be required to elucidate the specific function of PilY1, PilW and PilV as well as
451 the molecular mechanisms of pili assembly in *At. thiooxidans*.

452

453 Acknowledgments

454 We thank LINAN-IPICyT, for providing laboratory support.

455

456 References

457 01. Crundwell FK. The dissolution and leaching of minerals: Mechanisms, myths
458 and misunderstandings. Hydrometallurgy. 2013;139: 132-148. doi:
459 10.1016/j.hydromet.2013.08.003.
460 02. Bevilaqua D, Acciari HA, Arena FA, Benedetti AV, Fugivara CS, Tremiliosi-
461 Filho G, et al. Utilization of electrochemical impedance spectroscopy for
462 monitoring bornite (Cu_5FeS_4) oxidation by *Acidithiobacillus ferrooxidans*.
463 Miner Eng. 2009;22: 254-262. doi: 10.1016/j.mineng.2008.07.010.
464 03. López-Cázares MI, Patrón-Soberano A, García-Meza JV. Bioelectrochemical
465 changes during the early stages of chalcopyrite interaction with
466 *Acidithiobacillus thiooxidans* and *Leptospirillum* sp. Minerals. 2017;7: 156.
467 doi:10.3390/min7090156.
468 04. Echeverría-Vega A, Demergasso C. Copper resistance, motility and the
469 mineral dissolution behavior were assessed as novel factors involved in
470 bacterial adhesion in bioleaching. Hydrometallurgy. 2015;157: 107-115. doi:
471 10.1016/j.hydromet.2015.07.018
472 05. Valdés J, Pedroso I, Quatrini R, Dodson RJ, Tettelin H, Blake R, et al.
473 *Acidithiobacillus ferrooxidans* metabolism: From genome sequence to

industrial applications. BMC Genomics. 2008;9: 597. doi:10.1186/1471-2164-9-597.

06. Li YQ, Wan DS, Huang SS, Leng FF, Yan L, Ni YQ, Li HY. Type IV pili of *Acidithiobacillus ferrooxidans* are necessary for sliding, twitching motility, and adherence. Current Microbiol. 2010;60: 17-24. doi: 10.1007/s00284-009-9494-8.

07. Valdés J. Ossandon, F, Quatrini R, Dopson M, Holmes DS. Draft genome sequence of the extremely acidophilic biomining bacterium *Acidithiobacillus thiooxidans* ATCC 19377 provides insights into the evolution of the *Acidithiobacillus* genus. J Bacteriol. 2011;193: 7003-7004. doi: 10.1128/JB.06281-11.

08. Acuña LG, Cárdenas JP, Covarrubias PC, Haristoy JJ, Flores R, Nuñez H, et al. Architecture and gene repertoire of the flexible genome of the extreme acidophile *Acidithiobacillus caldus*. PLoS One. 2013;8: e78237. doi: 10.1371/journal.pone.0078237.t003.

09. Travisany D, Cortés MP, Latorre M, Genova A, Budinich M, Bobadilla-Fazzani R, Parada P, et al. A new genoma of *Acidithiobacillus thiooxidans* provides insights into adaptation to a bioleaching environment. Res Microbiol. 2014;165: 743-752. doi: 10.1016/j.resmic.2014.08.004.

10. Díaz M, Castro M, Copaja S, Guiliani N. Biofilm formation by the acidophile bacterium *Acidithiobacillus thiooxidans* involves c-di-GMP pathway and Pel exopolysaccharide. Genes. 2018;9: 113. doi: 10.3390/genes9020113.

11. Gerven N, Waksman G, Remaut H. Pili and Flagella: Biology, structure, and biotechnological applications. Prog Mol Biol Transl Sci. 2011;103: 21-72. doi: 10.1016/B978-0-12-415906-8.00005-4.

12. Conrad JC. Physics of bacterial near-surface motility using flagella and type IV pili: implications for biofilm formation. Res Microbiol. 2012;163: 619-629. doi: 10.1016/j.resmic.2012.10.016.

13. Craig L, Pique ME, Tainer JA. Type IV pilus structure and bacterial pathogenicity. Nat Rev Microbiol. 2004;2: 363-378. doi: 10.1038/nrmicro885.

14. Pugsley AP. The complete general secretory pathway in gram-negative bacteria. Microbiol Rev. 1993;57: 50-108.

15. Franz LP, Douzi, B, Durand, E, Dyer, DH, Voulhoux, R, Forest, KT. Structure of the minor pseudopilin XcpW from the *Pseudomonas aeruginosa* type II secretion system. Acta Cryst Sect D. 2011; 67: 124-130. doi: 10.1107/s0907444910051954.

16. Hartung S, Arvai AS, Wood T, Kolappan S, Shin DS, Craig L, et al. Ultrahigh resolution and full-length pilin structures with insights for filament assembly, pathogenic functions, and vaccine potential. J Biol Chem. 2011;286: 44254-44265. doi: 10.1074/jbc.M111.297242.

17. Berry J, Pelicic V. Exceptionally widespread nanomachines composed of type IV pilins: The prokaryotic Swiss Army knives. FEMS Microbiol Rev. 2015;39: 134-154. doi: 10.1093/femsre/fuu001.

18. Pelicic V. Type IV pili: e pluribus unum? Mol Microbiol. 2008;68: 827-837. doi: 10.1111/j.1365-2958.2008.06197.x.

519 19. Rudel T, Scheuerpflug I, Meyer TF. *Neisseria* PilC protein identified as type-
520 4 pilus tip-located adhesin. Nature. 1995;373: 357-359. doi:
521 10.1038/373357a0.
522 20. Nassif X, Marceau M, Pujol C, Pron B, Beretti JL, Taha MK. Type-4 pili and
523 meningococcal adhesiveness. Gene. 1997;192: 149-153. doi:
524 10.1016/S0378-1119(96)00802-5.
525 21. Craig L, Volkmann N, Arvai AS, Pique ME, Yeager M, Egelman EH, et al.
526 Type IV pilus structure by cryo-electron microscopy and crystallography:
527 Implications for pilus assembly and functions. Mol Cell. 2006;23: 651-662.
528 doi: 10.1016/j.molcel.2006.07.004.
529 22. Heiniger RW, Winther-Larsen HC, Pickles RJ, Koomey M, Wolfgang MC.
530 Infection of human mucosal tissue by *Pseudomonas aeruginosa* requires
531 sequential and mutually dependent virulence factors and a novel
532 pilus-associated adhesin. Cell Microbiol. 2010;12: 1158-1173. doi:
533 10.1111/j.1462-5822.2010.01461.x.
534 23. Orans J, Johnson MD, Coggan KA, Sperlazza JR, Heiniger RW, Wolfgang
535 MC, et al. Crystal structure analysis reveals *Pseudomonas* PilY1 as an
536 essential calcium-dependent regulator of bacterial surface motility. Proc Natl
537 Acad Sci USA. 2010;107: 1065-1070. doi:10.1073/pnas.0911616107.
538 24. Goosens VJ, Busch A, Georgiadou M, Castagnini M, Forest KT, Waksman
539 G, Pelicic V. Reconstitution of a minimal machinery capable of assembling
540 type IV pili. Proc Natl Acad Sci USA. 2017;114: E4978-E4986. doi:
541 10.1073/pnas.1618539114.
542 25. Winther-Larsen HC, Wolfgang M, Dunham S, Van Putten JP, Dorward D,
543 Løvold C, et al. A conserved set of pilin-like molecules controls type IV pilus

544 dynamics and organelle-associated functions in *Neisseria gonorrhoeae*. Mol
545 Microbiol. 2005;56: 903-917. doi: 10.1111/j.1365-2958.2005.04591.x.
546 26. Carbonnelle E, Hélaine S, Prouvensier L, Nassif X, Pelicic V. Type IV pilus
547 biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after
548 pilus assembly, essential for fibre stability and function. Mol Microbiol.
549 2005;55: 54-64. doi: 10.1111/j.1365-2958.2004.04364.x.
550 27. Li Y, Li H. Type IV pili of *Acidithiobacillus ferrooxidans* can transfer electrons
551 from extracellular electron donors. J Basic Microbiol. 2013;54: 226-231.
552 doi:10.1002/jobm.201200300.
553 28. Reardon PN, Mueller KT. Structure of the type IVa major pilin from the
554 electrically conductive bacterial nanowires of *Geobacter sulfurreducens*. J
555 Biological Chem. 2013;288: 29260-29266. doi: 10.1074/jbc.M113.498527
556 29. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et
557 al. Primer3- New capabilities and interfaces. Nucleic Acids Res. 2012:40:
558 e115. doi: 10.1093/nar/gks596
559 30. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucl
560 Acids Res. 1988;16: 10881-10890. doi: 10.1093/nar/16.22.10881. Available
561 from: http://multalin.toulouse.inra.fr/multalin/
562 31. Letunic I. Doerks T, Bork P. SMART: recent updates, new developments and
563 status in 2015. Nucleic Acids Res. 2015;D257-D260. doi:
564 10.1093/nar/gku949. Available from: http://smart.embl-heidelberg.de/
565 32. PROSITE. Available from: https://prosite.expasy.org
566 33. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The
567 Pfam protein families database: towards a more sustainable future. Nucleic

568 Acids Res. 2016;44: D279-D285. doi: 10.1093/nar/gkv1344. Available from:
569 https://pfam.xfam.org
570 34. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al.
571 CDD/SPARCLE: Functional classification of proteins via subfamily domain
572 architectures. Nucleic Acids Res. 2017;45: D200-D203. doi:
573 10.1093/nar/gkw1129. Available from:
574 https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
575 35. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al.
576 SWISS-MODEL: Modelling protein tertiary and quaternary structure using
577 evolutionary information. Nucleic Acids Res. 2014;42: W252- W258. doi:
578 10.1093/nar/gku340. Available from: https://www.swissmodel.expasy.org/
579 36. Roy A, Yang J, Zhang Y. COFACTOR: An accurate comparative algorithm
580 for structure-based protein function annotation. Nucleic Acids Res. 2012;40:
581 W471-W477. 10.1093/nar/gks372 Available from:
582 https://zhanglab.ccmb.med.umich.edu
583 37. Yang J, Zhang Y. I-TASSER server: New development for protein structure
584 and function predictions, Nucleic Acids Res. 2015;43: W174-W181. doi:
585 10.1093/nar/gkv342. Available from: https://zhanglab.ccmb.med.umich.edu
586 38. Yang J, Roy A, Zhang Y. BioLiP: A semi-manually curated database for
587 biologically relevant ligand-protein interactions. Nucleic Acids Res. 2013;41:
588 D1096-D1103. doi: 10.1093/nar/gks966.
589 39. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al.
590 Gapped BLAST and PSI-BLAST: A new generation of protein database

591 search programs. Nucleic Acids Res. 1997;25: doi: 3389-3402.
592 10.1093/nar/25.17.3389. Available from: https://blast.ncbi.nlm.nih.gov
593 40. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schäffer AA, et
594 al. Protein database searches using compositionally adjusted substitution
595 matrices. FEBS J. 2005;272: 5101-5109. Doi: 10.1111/j.1742-
596 4658.2005.04945.x.
597 41. UniProt. Available from: http://www.uniprot.org
598 42. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics
599 Analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33: 1870-1874.
600 doi: 10.1093/molbev/msw054
601 43. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data
602 matrices from protein sequences. Bioinformatics. 1992;8: 275-282. doi:
603 10.1093/bioinformatics/8.3.275.
604 44. Felsenstein J. Confidence limits on phylogenies: An approach using the
605 bootstrap. Evolution. 1985;39: 783-791. doi:
606 10.1016/j.hydromet.2015.07.018.
607 45. Panda SK, Jyoti V, Bhadra B, Nayak KC, Shivaji S, Rainey FA, et al.
608 *Thiomonas bhubaneswarensis* sp. nov., an obligately mixotrophic,
609 moderately thermophilic, thiosulfate-oxidizing bacterium. Int J Syst Evol
610 Microbiol. 2009;59: 2171-2175. doi: 10.1099/ijs.0.007120-0.
611 46. Hallberg KB, Hedrich S, Johnson DB. *Acidiferrobacter thiooxydans*, gen. nov.
612 sp. nov.; an acidophilic, thermo-tolerant, facultatively anaerobic iron- and
613 sulfur-oxidizer of the family Ectothiorhodospiraceae. Extremophiles. 2011;15:
614 271-279. doi: 10.1007/s00792-011-0359-2.

615 47. Johnson MDL, Garrett CK, Bond JE, Coggan KA, Wolfgang MC, Redinbo
616 MR. *Pseudomonas aeruginosa* PilY1 binds integrin in an RGD-and calcium-
617 dependent manner. PloS One. 2011;6: e29629. doi:
618 10.1371/journal.pone.0029629
619 48. Kadnikov VV, Ivasenko DA, Beletsky AV, Mardanov AV, Danilova EV,
620 Pimenov NV, et al. A Novel uncultured bacterium of the family
621 Gallionellaceae: Description and genome reconstruction based on the
622 metagenomic analysis of microbial community in acid mine drainage.
623 Microbiology. 2016;85: 421-435. doi: 10.1134/S002626171604010X.
624 49. Jones DS, Roepke EW, Hua AA, Flood BE, Bailey JV. Complete genome
625 sequence of *Sulfuriferula* sp. strain AH1, a sulfur-oxidizing autotroph isolated
626 from weathered mine tailings from the Duluth Complex in Minnesota.
627 Genome Announc. 2017;5: e00673-17. doi: 10.1128/genomeA.00673-17.
628 50. Krishnan V, Dwivedi P, Kim BJ, Samal A, Macon K, Ma X, et al. Structure of
629 *Streptococcus agalactiae* tip pilin GBS104: A model for GBS pili assembly
630 and host interactions. Acta Crystallogr Sect D: Biol Cristallogr. 2013;69:
631 1073-1089. doi:10.1107/S0907444913004642.
632 51. Zhang K, Chen J. The regulation of integrin function by divalent cations. Cell
633 Adhes Migr. 2012;6: 20-29. doi: 10.4161/cam.18702.
634 52. Lee JO, Rieu P, Arnaout MA, Liddington R. Crystal structure of the A domain
635 from the *a* subunit of integrin CR3 (CD11 b/CD18). Cell. 1995;80: 631-638.
636 doi: 10.1016/0092-8674(95)90517-0.
637 53. Broeker NK, Gohlke U, Müller JJ, Uetrecht C, Heinemann U, Seckler R, et al.
638 Single amino acid exchange in bacteriophage HK620 tailspike protein results

639 in thousand-fold increase of its oligosaccharide affinity. Glycobiology.
640 2012;23: 59-68. doi: 10.1093/glycob/cws126.
641 54. Anthony C. Pyrroloquinoline quinone (PQQ) and quinoprotein enzymes.
642 Antioxid Redox Signal. 2001;3: 757-774. doi: 10.1089/15230860152664966
643 55. Craig L, Taylor RK, Pique ME, Adair BD, Arvai AS, Singh M, et al. Type IV
644 pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-
645 coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. Mol Cell.
646 2003;11: 1139-1150. doi: 10.1016/S1097-2765(03)00170-9
647 56. Wang F, Coureuil M, Osinski T, Orlova A, Altindal T, Gesbert G, et al.
648 Cryoelectron microscopy reconstructions of the *Pseudomonas aeruginosa*
649 and *Neisseria gonorrhoeae* type IV pili at sub-nanometer resolution. Struct.
650 2017;25: 1423-1435. doi: 10.1016/j.str.2017.07.016.
651 57. Dalrymple B, Mattick JS. An Analysis of the organization and evolution of
652 type 4 fimbrial (MePhe) subunit proteins. J Mol Evol. 1987;24: 261-269. doi:
653 10.1007/BF02100020.
654 58. Mattick JS. Type IV pili and twitching motility. Annu Rev Microbiol. 2002;56:
655 289-314. doi: 10.1146/annurev.micro.56.012302.160938.
656 59. Hobbs M, Mattick JS. Common components in the assembly of type 4
657 fimbriae, DNA transfer systems, filamentous phage and protein-secretion
658 apparatus: a general system for the formation of surface-associated protein
659 complexes. Mol Microbiol. 1993;10: 233-243. doi: 10.1111/j.1365-
660 2958.1993.tb01949.x.
661

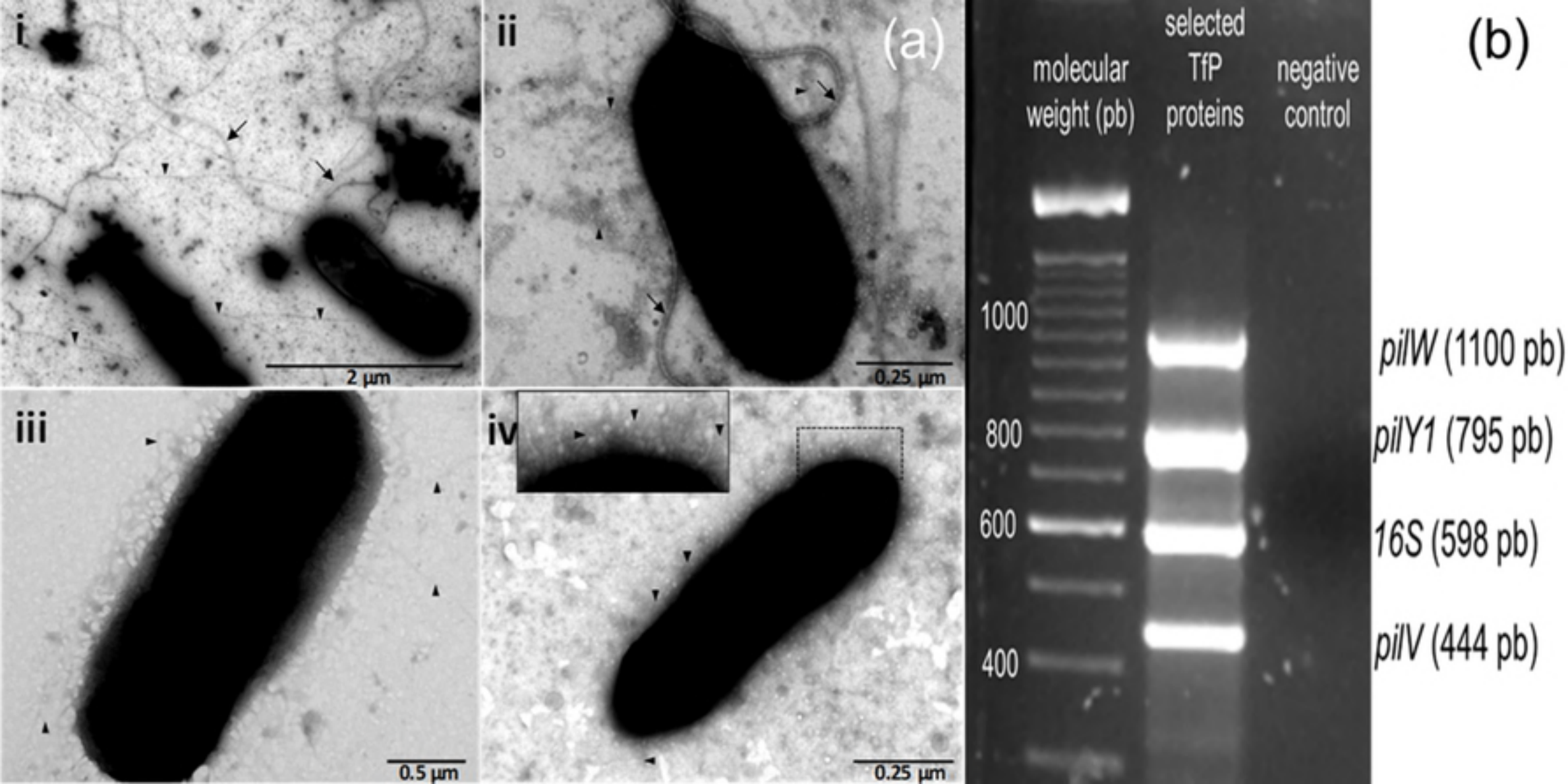
Supporting information

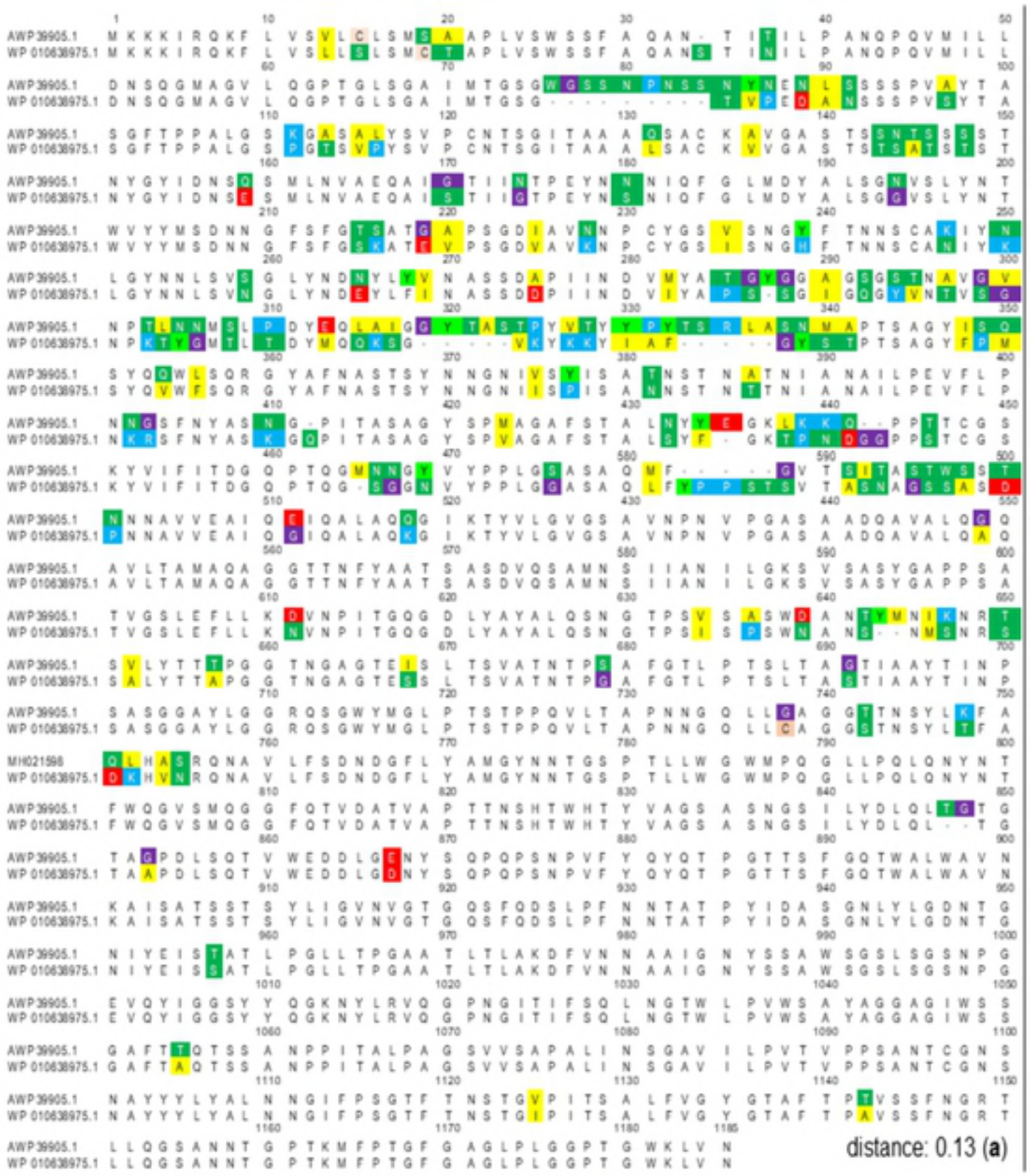
S1 Table

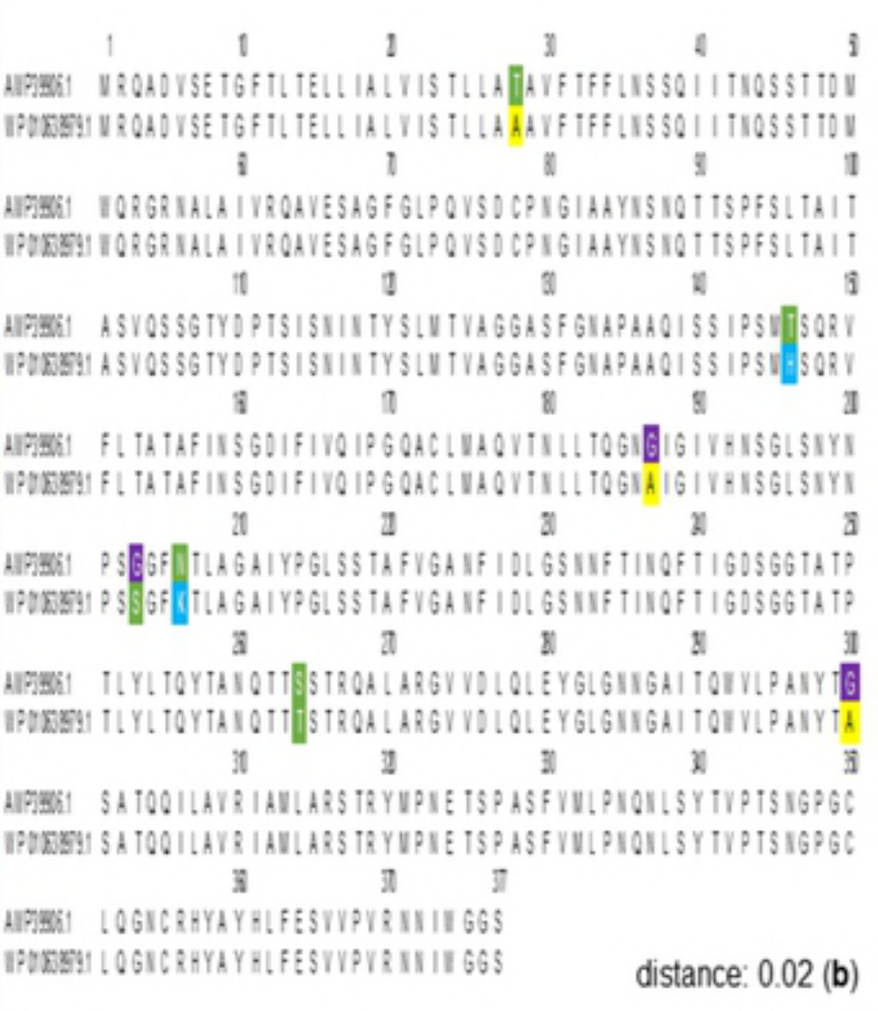
S2 Table

Authors' contributions

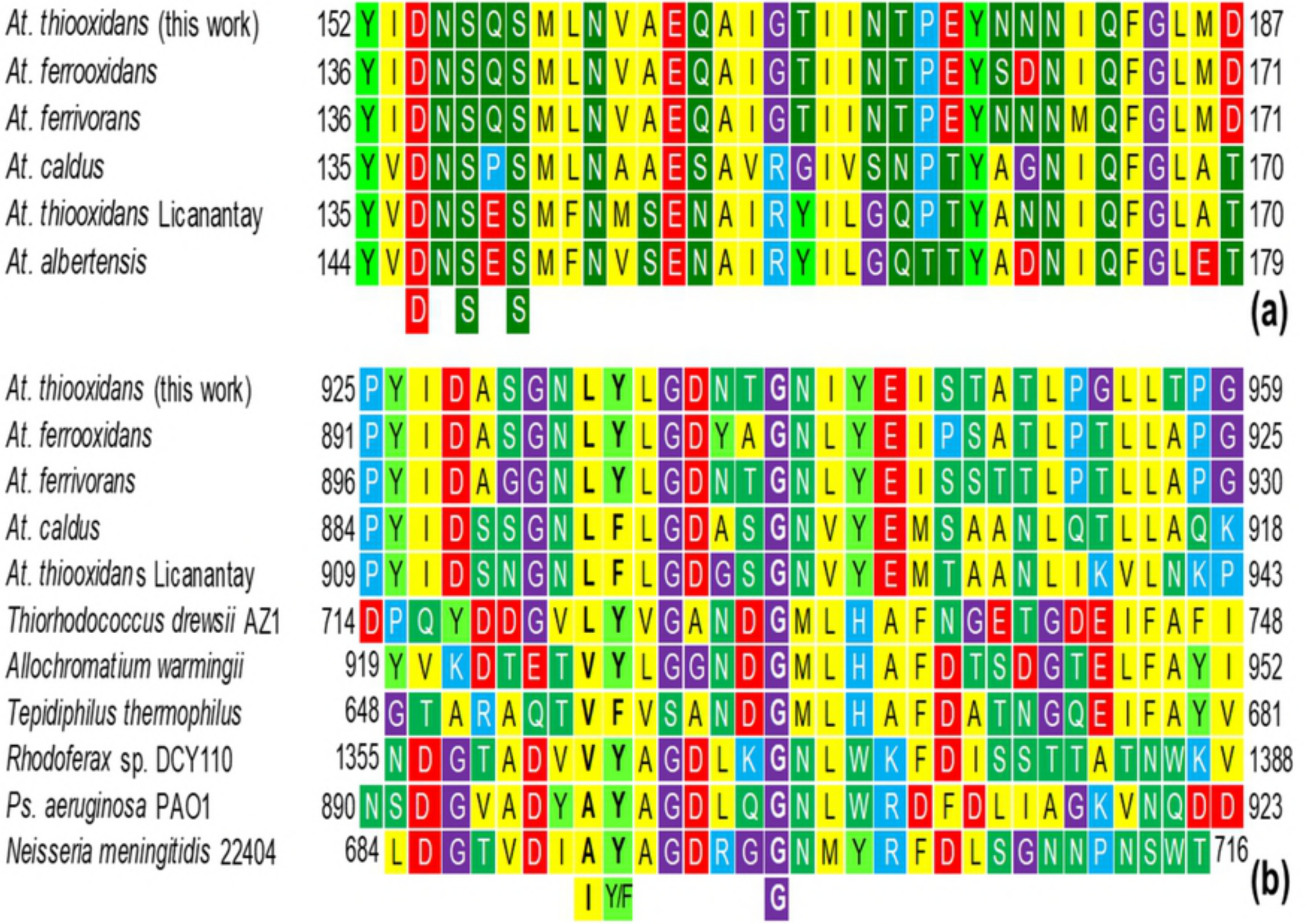
JVGM and JAMC conceived and designed the experiments. EAS and AHS performed, described and discuss the sequencing. OASP conceived, designed, performed, described and discuss the TEM analyses. JVGM performed and discussed the bioinformatics analysis. MAG discussed the overall results. JAMC, JVGM and OASP contributed with reagents/materials/analysis tools. JVGM is senior author.

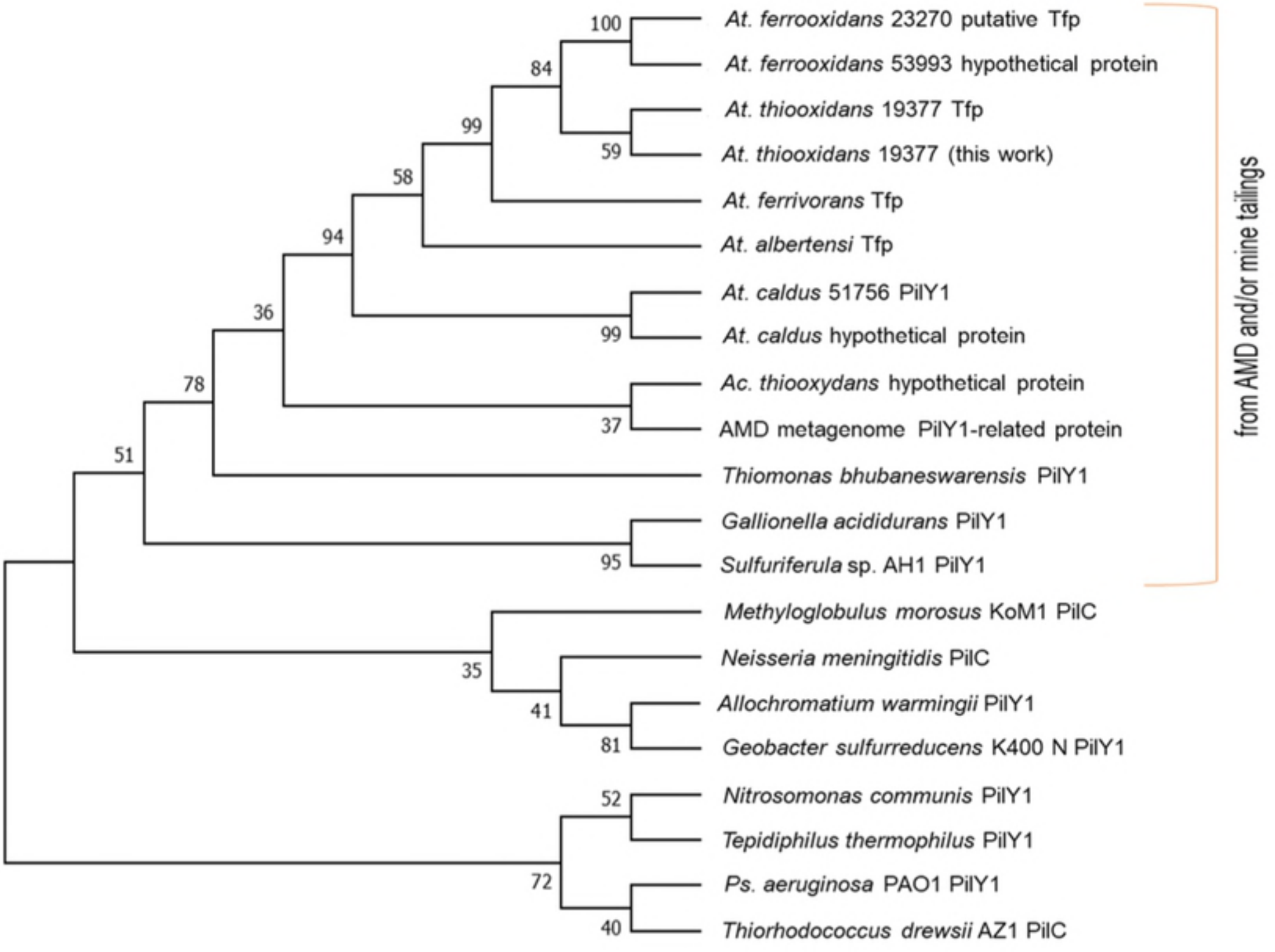

i
2 µm
ii
(a)
0.25 µm
iii
0.5 µm
iv
0.25 µm
molecular weight (pb)
selected TfP proteins
negative control
(b)
1000
800
600
400
pilW (1100 pb)
pilY1 (795 pb)
16S (598 pb)
pilV (444 pb)


AWP39905.1
WP 010638975.1
MIH021598
distance: 0.13 (a)


AWP39905.1
WP 010638975.1
distance: 0.02 (b)


AWP39907.1
WP 010638981.1
distance: 0.0 (c)


At. thiooxidans (this work) 152 Y I D N S Q S M L N V A E Q A I G T I I N T P E Y N N N I Q F G L M D 187
At. ferrooxidans 136 Y I D N S Q S M L N V A E Q A I G T I I N T P E Y S D N I Q F G L M D 171
At. ferrivorans 136 Y I D N S Q S M L N V A E Q A I G T I I N T P E Y N N N M Q F G L M D 171
At. caldus 135 Y V D N S P S M L N A A E S A V R G I V S N P T Y A G N I Q F G L A T 170
At. thiooxidans Licanantay 135 Y V D N S E S M F N M S E N A I R Y I L G Q P T Y A N N I Q F G L A T 170
At. albertensis 144 Y V D N S E S M F N V S E N A I R Y I L G Q T T Y A D N I Q F G L E T 179
D S S
(a)
At. thiooxidans (this work) 925 P Y I D A S G N L Y L G D N T G N I Y E I S T A T L P G L L T P G 959
At. ferrooxidans 891 P Y I D A S G N L Y L G D Y A G N L Y E I P S A T L P T L L A P G 925
At. ferrivorans 896 P Y I D A G G N L Y L G D N T G N L Y E I S S T T L P T L L A P G 930
At. caldus 884 P Y I D S S G N L F L G D A S G N V Y E M S A A N L Q T L L A Q K 918
At. thiooxidans Licanantay 909 P Y I D S N G N L F L G D G S G N V Y E M T A A N L I K V L N K P 943
Thiorhodococcus drewsii AZ1 714 D P Q Y D D G V L Y V G A N D G M L H A F N G E T G D E I F A F I 748
Allochromatium warmingii 919 Y V K D T E T V Y L G G N D G M L H A F D T S D G T E L F A Y I 952
Tepidiphilus thermophilus 648 G T A R A Q T V F V S A N D G M L H A F D A T N G Q E I F A Y V 681
Rhodoferax sp. DCY110 1355 N D G T A D V V Y A G D L K G N L W K F D I S S T T A T N W K V 1388
Ps. aeruginosa PAO1 890 N S D G V A D Y A Y A G D L Q G N L W R D F D L I A G K V N Q D D 923
Neisseria meningitidis 22404 684 L D G T V D I A Y A G D R G G N M Y R F D L S G N N P N S W T 716
I Y/F G
(b)


100
84
99
59
58
94
36
99
78
37
51
95
35
41
81
52
72
40
At. ferrooxidans 23270 putative Tfp
At. ferrooxidans 53993 hypothetical protein
At. thiooxidans 19377 Tfp
At. thiooxidans 19377 (this work)
At. ferrivorans Tfp
At. albertensi Tfp
At. caldus 51756 PilY1
At. caldus hypothetical protein
Ac. thiooxydans hypothetical protein
AMD metagenome PilY1-related protein
Thiomonas bhubaneswarensis PilY1
Gallionella acididurans PilY1
Sulfuriferula sp. AH1 PilY1
Methyloglobulus morosus KoM1 PilC
Neisseria meningitidis PilC
Allochromatium warmingii PilY1
Geobacter sulfurreducens K400 N PilY1
Nitrosomonas communis PilY1
Tepidiphilus thermophilus PilY1
Ps. aeruginosa PAO1 PilY1
Thiorhodococcus drewsii AZ1 PilC
from AMD and/or mine tailings

a

b

T320

N156

b'

c'

c

S77

M78

G776

L774

L43

I634

d

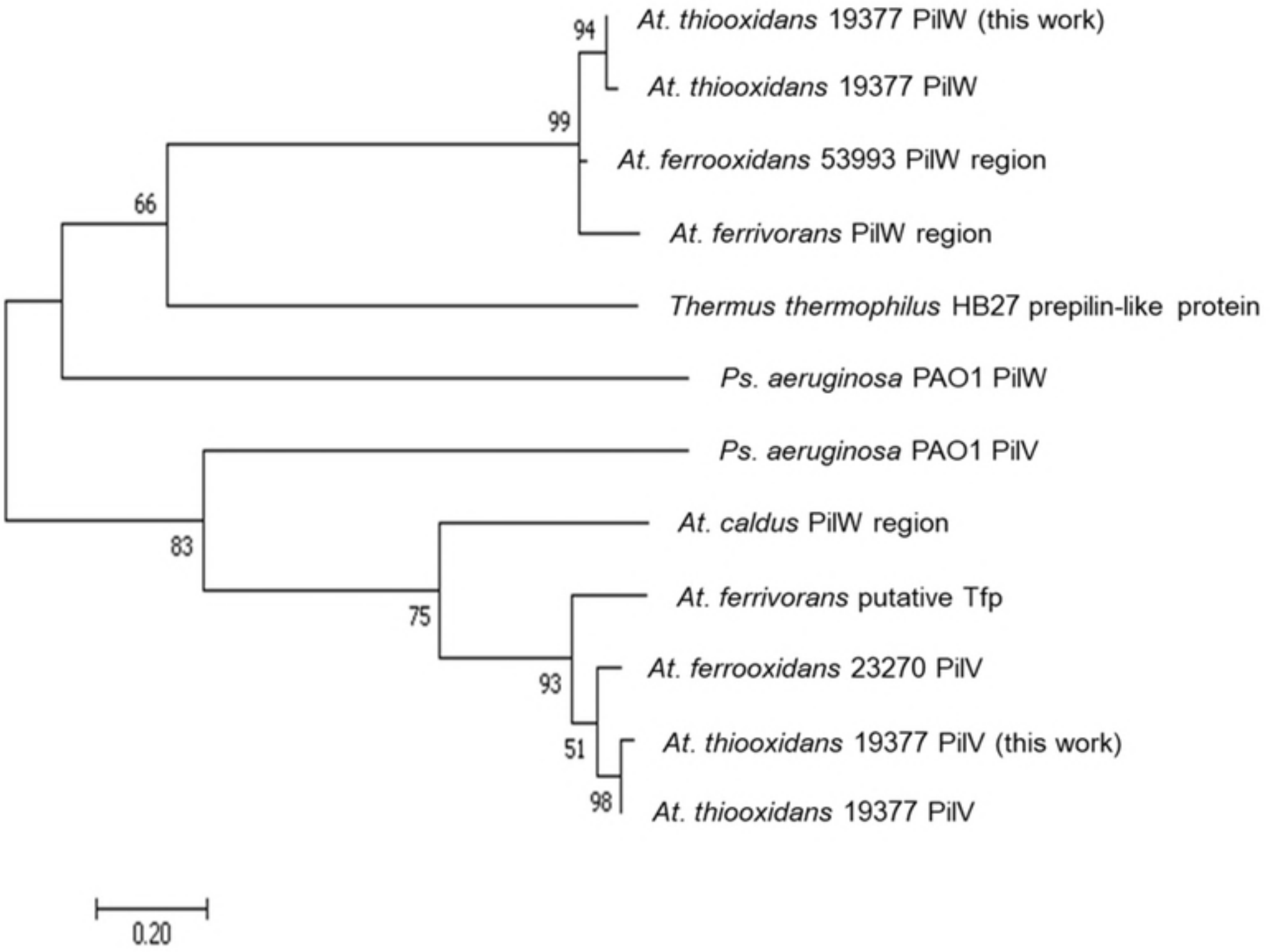
Y-152

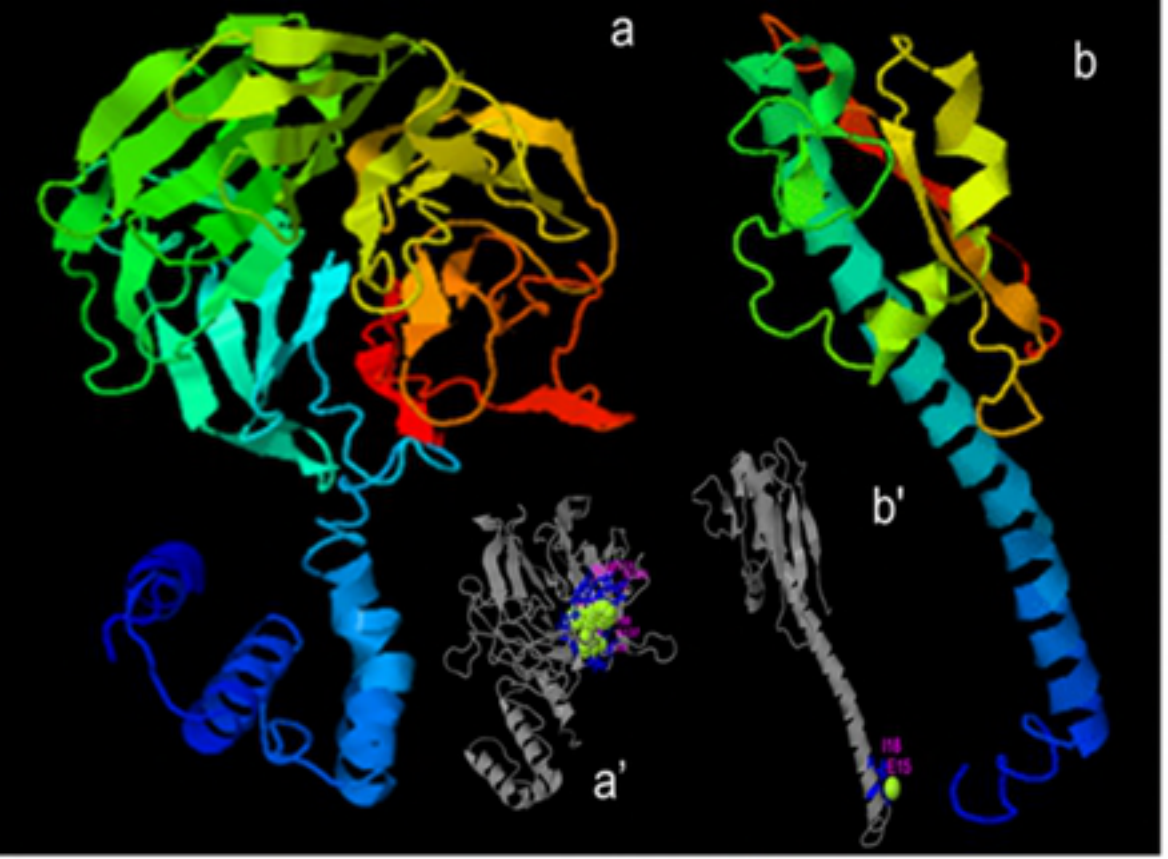
L-151

d'

T-157

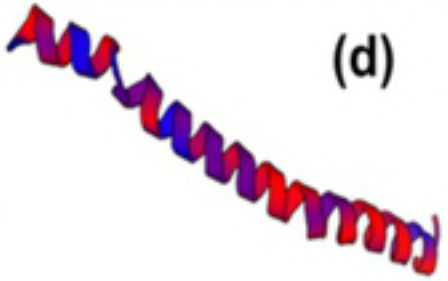
G-158

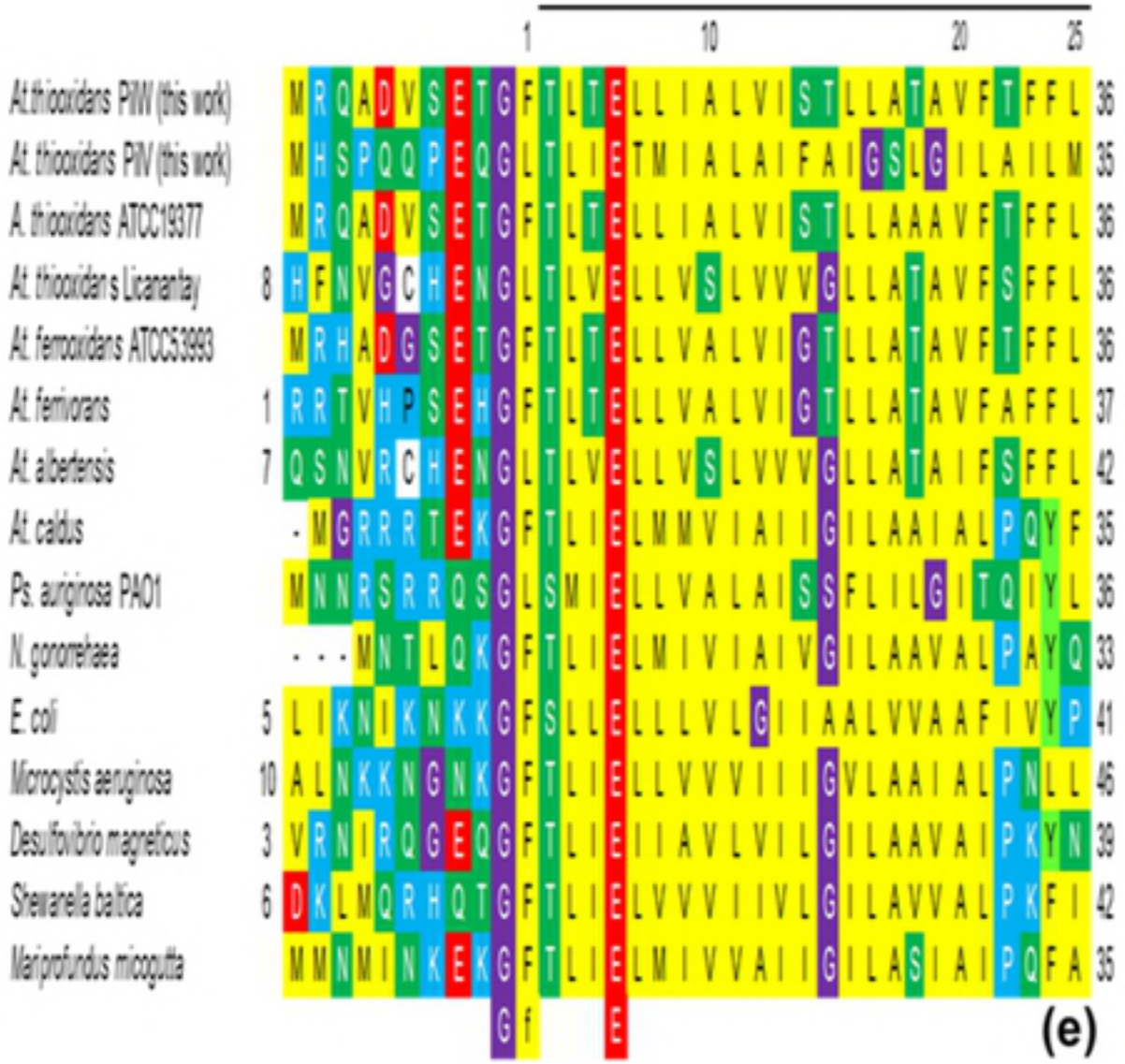

94
99
66
83
75
93
51
98
At. thiooxidans 19377 PilW (this work)
At. thiooxidans 19377 PilW
At. ferrooxidans 53993 PilW region
At. ferrivorans PilW region
Thermus thermophilus HB27 prepilin-like protein
Ps. aeruginosa PAO1 PilW
Ps. aeruginosa PAO1 PilV
At. caldus PilW region
At. ferrivorans putative Tfp
At. ferrooxidans 23270 PilV
At. thiooxidans 19377 PilV (this work)
At. thiooxidans 19377 PilV
0.20


a
b
b'
a'

(c)

(d)




1
10
20
25
At.thiooxidans PIIW (this work)
At. thiooxidans PIIV (this work)
A. thiooxidans ATCC19377
At. thiooxidans Licanantay
At. ferrooxidans ATCC53993
At. ferrivorans
At. albertensis
At. caldus
Ps. auriginosa PAO1
N. gonorrheaea
E. coli
Microcystis aeruginosa
Desulfovibrio magneticus
Shewanella baltica
Maripofundus micoquita

(e)