Proteomic analysis of *Marinobacter hydrocarbonoclasticus* SP17 biofilm formation at the alkane-water interface reveals novel proteins and cellular processes involved in hexadecane assimilation.

Pierre-Joseph Vaysse^a, Laure Prat^{a†}, Sophie Mangenot^b, Stéphane Cruveiller^c, Philippe Goulas^a, Régis Grimaud^{a*}

^a Institut Pluridisciplinaire de Recherche en Environnement et Matériaux, Equipe Environnement et Microbiologie UMR5254 CNRS, IBEAS, Université de Pau et des Pays de l'Adour, BP1155, 64013 Pau cedex, France

^b CEA/DSV/IG/Genoscope, 2 rue Gaston Cremieux, 91057 Evry cedex, France

^c CEA/DSV/IG/Genoscope/LGC, 2 rue Gaston Cremieux, 91057 Evry cedex, France

[†] Current address : EPFL, ENAC-ISTE, Laboratory of Environmental Biotechnology, Lausanne, Switzerland

E-mail addresses of all authors

- Pierre-Joseph Vaysse: pierre-joseph.vaysse@univ-pau.fr
- Laure Prat: laure.prat@epfl.ch
- Sophie Mangenot: mangenot@genoscope.cns.fr
- Stéphane Cruveiller: scruveil@genoscope.cns.fr
- Philippe Goulas: philippe.goulas@univ-pau.fr
- Régis Grimaud: regis.grimaud@univ-pau.fr * Correspondence and reprints

Tel: 33 (0)5 59 40 74 86, Fax: 33 (0)5 59 40 74 94

Abstract

Many hydrocarbon degrading bacteria form biofilms at the hydrocarbon-water interface to overcome the low accessibility of these poorly water-soluble substrates. In order to gain insight into the cellular functions involved, we undertook a proteomic analysis of *Marinobacter hydrocarbonoclasticus* SP17 biofilm developing at the hexadecane-water interface. Biofilm formation on hexadecane led to a global change of the cell physiology involving modulation of the expression of 573 out of 1144 detected proteins when compared with planktonic cells growing on acetate. Biofilm cells overproduced a protein, encoded by MARHY0478 that contains a conserved domain belonging to the family of the outer membrane transporters of hydrophobic compounds. Homologs of MARHY0478 were exclusively found in marine bacteria degrading alkanes or possessing alkane degradation genes and hence presumably constitute a family of alkane transporter specific to marine bacteria. Interestingly, we also found that sessile cells growing on hexadecane overexpressed type VI secretion system components. This secretion system has been identified as a key factor in virulence and in symbiotic interaction with host organisms. This observation is the first experimental evidence of the contribution of a type VI secretion system to environmental adaptation and raises the intriguing question about the role of this secretion machine in alkane assimilation.

Keywords: *Marinobacter hydrocarbonoclasticus* SP17; biofilm; alkane degradation; proteomic; type VI secretion system

1. Introduction

Biofilm formation at the hydrocarbon-water interface has been observed with various alkane degrading strains, e.g. *Rhodococcus* sp. Q15 [34], *Acinetobacter venetianus* RAG-1 [2] and *Oleiphilus messinensis* [9] as well as polycyclic aromatic hydrocarbons (PAHs) degrading strains, e.g. *Pseudomonas* sp. strain 8909N [19], *Sphingomonas* sp. CHY-1 [35] and *Mycobacterium frederiksbergense* LB501T [3].

The ecological significance of these biofilms has been demonstrated by a study devoted to the diversity of biofilm communities developing on PAHs [27]. It has been shown that biofilm communities contained a greater diversity of active species and of PAH degradation genes than the planktonic communities enrichment. Furthermore the diversity of active species found in the biofilm closely matched that of the PAH-contaminated soil used as inoculum. These findings suggest that biofilm formation on hydrocarbons is a probable lifestyle in natural ecosystems.

Biofilms developing on hydrocarbons present two properties distinguishing them from other biofilms which are their substrate-substratum specificity and their capacity to overcome the low accessibility of hydrophobic substrates. It was observed in the case of PAH-degrading strains that the less soluble the hydrocarbon was, the more cells grew at the PAH-water interface [12,23]. Biofilm formation has been shown to promote growth on hydrocarbons by facilitating interfacial access. Kinetic studies demonstrated that growth at the interface occurs faster than mass transfer rate of hydrocarbons in the absence of bacteria would suggest [4,5,10,13]. Thus, biofilm formation constitutes an efficient adaptive strategy to assimilate hydrocarbon. However, the genetic determinants and the molecular mechanisms underlying the functioning of these biofilms remain poorly understood. In order to identify the proteins

and cellular functions involved, we undertook a proteomic analysis of the biofilm of *Marinobacter hydrocarbonoclasticus* SP17 growing at the hexadecane-water interface. This marine alkane degrading bacterium readily forms biofilms on metabolizable hydrophobic organic compounds including *n*-alkanes from 8 to 28 carbon atoms, whereas in the same condition this strain does not develop biofilms on an inert substratum such as plastic or glass [15]. The doubling time of biofilm cells growing on hexadecane (between 3 and 5 hours depending on the development stage of the biofilm) is similar to that observed on water-soluble substrates such as acetate, pyruvate and lactate (P-J Vaysse, unpublished results). Furthermore, the rate of hexadecane degradation dramatically decreases when the biofilm is disorganized by strong shaking, indicating a strong link between biofilm formation and the utilization of alkanes [15]. These observations indicate that the biofilm lifestyle must provide *M. hydrocarbonoclasticus* SP17 with efficient mechanisms to access the hexadecane. In the current study, we undertook a proteomic analysis of a biofilm growing at hexadecane-water interface. The results obtained indicate that adaptation to alkane utilization as carbon and energy source involves a global change in cell physiology. Novel proteins and cellular processes involved in alkane assimilation were revealed.

2. Material and methods

2.1 Growth conditions and preparation of protein extracts

The bacterial strain used in this study was *Marinobacter hydrocarbonoclasticus* SP17 (ATCC 49840). Cultures were carried out in synthetic sea water (SSW) [8].

All chemicals used were from Sigma-Aldrich unless otherwise specified. Biofilm cultures were inoculated with cells growing exponentially on acetate (20 mM) washed twice with one volume of SSW and resuspended to a final optical density at 600nm of 0.1 in 300 ml SSW medium supplemented with 0.2% (v/v) hexadecane in 1 l Erlenmeyer flasks. Biofilms were grown at the hexadecane-SSW interface at 30 °C, under slow shaking (50 rpm). After 35 hours incubation, the medium was carefully discarded, the biofilms resuspended in 30 ml of SSW and centrifuged for 20 min at 20,000 g at room temperature. Biofilms were harvested above the supernatant while residual planktonic cells were pelleted. Cells grown on acetate (100 ml) were harvested during exponential growth phase (OD_{600nm}= 0.3) by centrifuging 20 min at 20,000 g at room temperature.

For protein extraction, cell pellets or biofilms were washed twice with 5 ml of acetone, resuspended in 5 ml of water containing a protease inhibitor cocktail and sonicated on ice for 1 min with 500 ms/s pulses at 35 W. Sonicated cell suspensions were then incubated with 40 µg of DNase I, 10 µg of RNase A plus 0.01% (v/v) Triton X100 for 30 min at room temperature. Proteins were precipitated for 30 min at 4 °C by adding 500 µl of 100% trichloroacetic acid and then centrifuged 10 min at 20,000g at 4 °C. Protein extracts were washed twice with 1 ml of trichloroacetic acid 10% (v/v) and twice with 1 ml of acetone. Proteins were air dried and dissolved in IEF buffer (urea 8 M, 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate 4% (w/v), dithiothreitol 60 mM,

Pharmalyte 3-10 2% (v/v) (Amersham Biosciences) and bromophenol blue 0.0002% (w/v)). Protein concentration was estimated using Biorad Protein assay and finally adjusted to 1 µg/µl with IEF buffer.

2.2 2D electrophoresis and gel analysis

Two hundred and thirty micrograms of protein were applied to a 24 cm Immobiline Dry strip with a 3-7 non linear pH gradient (GE Healthcare). Isoelectric focusing and SDS-PAGE (12.5%) were carried out using the Multiphor II and Ettan DALTsix systems, respectively (GE Healthcare) according to the manufacturer's instructions. Gels were stained with Deep Purple Total Protein Stain (GE Healthcare) and then scanned using a Typhoon 9200 fluorescent scanner (GE Healthcare). Image analysis, spot detection and matching were carried out using Image Master Platinum software (GE Healthcare) and checked manually. Three replicate gels from three independent experiments were run for each growth condition. The normalized protein amount for each protein spot was defined as the fraction of that spot volume to the total spot's volume of the gel. Student t-test ($P < 0.05$) and a threshold of 2-fold change were used to determine the proteins significantly differentially expressed between the two conditions.

2.3 Protein identification by nanoLC-MS/MS

Protein spots were excised from gels, destained in ammonium bicarbonate 50 mM 50% acetonitrile (ACN), rinsed twice in ultrapure water and dehydrated in 100% ACN. After ACN removal by evaporation, gel pieces were dried, rehydrated with a trypsin solution (10 ng/ml in 50 mM ammonium bicarbonate) at 4 °C for 10 min and finally incubated overnight

at 37 °C. The supernatant was collected by two successive extractions with H₂O/ACN/HCOOH (47.5/47.5/5), pooled, and concentrated in a vacuum centrifuge to a final volume of 25 µl. The peptide mixture was analyzed by on-line capillary HPLC (LC Packings, Amsterdam, The Netherlands) coupled to a nanospray LCQ Deca XP Ion Trap mass spectrometer (Thermo-Finnigan, San Jose, CA). MS/MS data were acquired using a 2 *m/z* units ion isolation window and a 35% relative collision energy. Data were searched using SEQUEST through Bioworks 3.3.1 interface (ThermoFinnigan) against the *M. hydrocarbonoclasticus* SP17 whole genome sequence. Gene numbers, gene function, and functional category are presented according to the unpublished annotated genome.

2.4 Database searching and sequence analysis

Sequences similarities were searched against the translation of the non-redundant GenBank CDS database, using the BLAST program [1]. Homologs were defined as proteins giving an alignment with a bits score above 100 (scoring matrix BLOSUM62) and an expected value (E-value) below 1.10^{-25} . Protein domain searches and multiple sequence alignment were carried out using the NCBI-Conserved Domain Database search program with CDD v2.16-27036 PSSMs database (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) [18]. Synteny similarities were searched using the MAGE microbial genome annotation system from the Genoscope Evry France (<https://www.genoscope.cns.fr/agc/mage/wwwpkgdb/MageHome>).

3. Results and discussion

3.1. Comparison of protein patterns of biofilm cells growing on hexadecane with planktonic cells growing on acetate

In order to characterize the molecular mechanisms involved in the development of biofilm on alkanes, we compared the proteomes of biofilm growing on hexadecane and planktonic cells growing on acetate, hereafter referred to as BH and PA respectively (Fig. 1). A total of 1144 spots appearing on all three replicates of at least one growth condition were detected. 576 spots (50%) were found to change significantly between PA and BH. Among these proteins, 245 were overexpressed in BH and 331 in PA. 81% of the differentially expressed protein had an induction ratio above 10. This high proportion of differentially expressed proteins is unusual when compared to other similar proteomic analyses. For instance, a study of *Alcanivorax burkomensis* SK2 comparing growth on hexadecane with growth on pyruvate revealed 97 proteins differentially expressed [24]. However, *A. burkomensis* SK2 protein extracts were prepared from whole cultures consisting of a mixture of biofilm cells and detached planktonic cells. Therefore, the expression protein profiles might reflect the average protein expression levels from different cellular states, thus lessening the actual expression fold-change of one specific condition. Furthermore, in our study, proteins were revealed using a fluorescent dye with a higher sensitivity and larger dynamic range than Coomassie blue.

The high number of differentially expressed proteins between the biofilm cells growing on hexadecane and those growing on acetate signify that growth on hexadecane involve a global change in the cell physiology, requiring numerous cellular functions. This

could be explained by the change of two major factors between the compared conditions: the lifestyle (biofilm or planktonic) and the carbon source (hexadecane or acetate). In many species, biofilm formation has been shown to require a large number of protein functions. The biofilm lifestyle induces changes in the environmental conditions encountered by cells, such as the formation of nutrient microgradients. The adaptation to these new conditions requires most likely a great number of proteins [28]. On the other hand, the response to alkanes involves the modulation of the expression of numerous proteins required for the transport, the metabolism or to avoid toxic effects of these compounds [31]. A total of 58 spots, that could unambiguously be assigned to a single protein, were kept for further analysis (Table 1). The major type IV pilus subunit, PilA (encoded by MARHY2564) was under-produced in BH. Type IV pili have been shown to be involved in host cell adhesion, biofilm formation, DNA uptake and twitching motility [21]. A DNA microarray analysis showed that pilin genes were repressed in the biofilm of *Pseudomonas aeruginosa* [33]. Furthermore, the type IV pilus of *P. aeruginosa* has been shown to have multiple effects on biofilm formation mainly through twitching motility and adhesion [11,20].

Proteins involved in nutrient transport across cellular envelopes constitute a large class of proteins overproduced in BH. This includes components of phosphate (encoded by MARHY3535 *pstS*) and thiosulfate (encoded by MARHY2019 *cysP*) ABC transporters. Iron uptake and transport proteins were represented by FhuE, FbpA and CirA encoded by MARHY1035, MARHY2192 and MARHY3135 respectively. Four other proteins coding for putative ABC transporters and porins (encoded by MARHY0256, MARHY0299, MARHY3277, MARHY34332 and MARHY3166) were also found, although their actual function remain uncertain. Increased capacities in solute transport may reflect an adaptation to the constraints imposed by the biofilm lifestyle. Indeed, the increase in biofilm thickness is thought to hinder nutrient penetration into the deepest layers of the biofilm. Thus, cells would

require an increased capacity in nutrient uptake and transport. This is particularly important in seawater where iron and phosphate are in limiting concentrations. Micronutrients such as iron and inorganic phosphate are also known to strongly influence biofilm development by acting as environmental cues regulating biofilm formation [25].

3.2. Redirection of carbon flux in biofilm cells growing on hexadecane

Most of the enzymes corresponding to the CO₂-releasing steps of the tricarboxylic acid cycle (TCA cycle) (encoded by MARHY0078 *idh*, MARHY2120 *sucD*, MARHY2121 *sucC*, MARHY2126 *sdhA*) were found to be down regulated in BH. In addition, the gene encoding malate synthase (MARHY1458 *glcB*), catalyzing the conversion of glyoxylate into malate, was upregulated. This modulation of TCA cycle enzymes suggests a stimulation of the glyoxylate bypass in BH. This anaplerotic pathway allows the replenishment of cells in the metabolic intermediates necessary for the synthesis of their cellular components when acetyl-CoA is the only carbon source available in the cell [16]. The activation of the glyoxylate bypass under a hexadecane diet compared to a diet of pyruvate or glucose has been already observed in proteomic analyses on *A. burkomensis* SK2 and *Geobacillus thermodenitrificans* [6,24]. In our study, stimulation of the glyoxylate bypass was observed under hexadecane conditions compared with acetate. In hexadecane and acetate conditions, the intracellular carbon source is acetyl-CoA hence the glyoxylate pathway is required in both conditions. The explanation of the glyoxylate bypass stimulation could lie in the fact that hexadecane is a more energetic substrate than acetate. In fact the breakdown of hexadecane to acetyl-CoA produces energy whereas the conversion of acetate to acetyl-CoA does not. Augmenting the flux of carbon through the glyoxylate bypass on hexadecane could allow cells to restore the balance between carbon assimilation and energy production. However interpretation of

proteomic data with regard to metabolic flux should be taken with caution, since metabolic pathways can be regulated at the level of enzyme activity, in addition to gene expression.

Fatty acid biosynthesis genes *fabA*, *fabB* and *fabF* (MARHY3086, MARHY3087 and MARHY1438, respectively) were found to be down regulated in BH. This corroborates the fact that in cells growing on hexadecane the main fatty acids of cellular lipids were derived from the oxidation of alkanes [26].

3.3. A type VI secretion system is overproduced by biofilm cells growing on hexadecane

Three proteins sharing similarities with type VI secretion system (T6SS) subunits, encoded by MARHY3623, MARHY3634 and MARHY3635, were overexpressed in BH. MARHY3634 and MARHY3635 were among the most abundant protein detected in BH, with 0.2 and 0.7 % of total protein respectively. MARHY3634, MARHY3635 and MARHY3623 are localized within a cluster of 16 genes (from MARHY3635 to MARHY3620) that are transcribed in the same direction. All members of the *M. hydrocarbonoclasticus* SP17 cluster share similarity with genes found in T6SS clusters from other bacteria including *Vibrio cholerae*, *P. aeruginosa* and *Escherichia coli*. The best synteny conservation with a characterized T6SS was observed with *Vibrio cholerae* N16961 T6SS gene cluster (Fig. 2).

Homologs of component of T6SS already characterized, i.e. Dot, IcmF, and ClpV (a subfamily of ClpB ATPase), were also found in *M. hydrocarbonoclasticus* SP17 cluster (MARHY3626, MARHY3621 and MARHY3625). Vgr and Hcp are two proteins secreted through T6SS. One homolog of Vgr is found (MARH3620) within this cluster while other homologs of Vgr and Hcp are found elsewhere on the *M. hydrocarbonoclasticus* SP17 chromosome [7] (Fig. 3). Experimental evidence of the functionality of homologs of MARHY3635 and 3634 was provided by studies of the pathogen *Edwardsiella tarda* [22].

Deletion mutants of *evpA* and *evpB* (MARHY3635 and MARHY3634, respectively) were impaired in protein secretion and showed reduced virulence in blue gourami fish. Based on the high degree of sequence similarity and synteny conservation between characterized T6SS clusters, we propose that the genes from MARHY3635 to MARHY3620 constitute a T6SS gene cluster. The functionality of this T6SS is supported by the detection of three proteins encoded by this cluster.

Although type VI secretion systems have been identified as key factors in virulence of pathogenic bacteria and in symbiotic interaction with host organisms, *in silico* analyses have revealed their presence in many species that are not considered as pathogens or symbionts. This led to the hypothesis that T6SS may also contribute to environmental adaptation [29]. Our data provide the first experimental evidence of the production of a T6SS during biofilm development by an environmental bacterium. The role of T6SS in biofilm formation during growth on hexadecane by *M. hydrocarbonoclasticus* SP17 remains puzzling. T6SS may be required to secrete proteins needed for biofilm formation or for efficient assimilation of carbon source difficult to access.

3.4. Evidence for an alkane transporter family specific to marine bacteria

Two proteins produced in large quantities in BH, encoded by MARHY0478 and MARHY0477, drew our attention because their sequences were found to be conserved exclusively in marine species capable to use alkanes as a carbon source or possessing at least one copy of an alkane hydroxylase gene. The phylogenetic distribution of MARHY0478 and MARHY0477 homologs is rather restricted since they are found only in two orders of the Gammaproteobacteria: the *Oceanospirillales* and the *Alteromonadales*. The degree of peptide sequence identity varies between 29% and 98%. Interestingly, these proteins are found in

285 *Alcanivorax borkumensis* SK2, a key alkane degrader in polluted seawater. Several strains
 286 possess multiple homologs of MARHY0478 and MARHY0477 with a maximum of six in
 287 *Alcanivorax* sp DG881 (Table 2).

288 In *M. hydrocarbonoclasticus* SP17, MARHY0478 and MARHY0477 are separated by only
 289 11 base pairs and are transcribed in the same direction. This spatial organization is indicative
 290 of a putative operon. Such an operon would suggest that the products of these genes are
 291 involved in the same function. An operon structure is likely to be conserved in *M.*
 292 *hydrocarbonoclasticus* VT8 and *A. burkumensis* SK2 since these strains contain homologs of
 293 MARHY0477 and MARHY0478 that are adjacent and transcribed in the same direction.

294 The strong expression of MARHY0478 and MARHY0477 in biofilm on hexadecane indicates
 295 that they might play an important role in alkane assimilation. In order to gain more insight
 296 into their function, conserved domains in the two proteins were searched against the NCBI
 297 Conserved Domain Database using RPS-BLAST [18]. Only the protein encoded by
 298 MARHY0478 gave a positive hit (E-value = 10^{-13}) with the pfam03349 domain family. This
 299 conserved domain is found in outer membrane proteins transporting hydrophobic compounds
 300 out of Gram negative bacteria. This family includes the monoaromatic hydrocarbon transport
 301 proteins TodX from *Pseudomonas putida* F1 and TbuX from *Ralstonia pickettii* PKO1, and
 302 the long chain fatty acid transporter FadL from *Escherichia coli* [14,30,32]. Fig. 3 shows the
 303 multiple peptide sequence alignment for the conserved domain including the amino acid
 304 sequence deduced from MARHY0478, TodX, TbuX and FadL. This alignment is based upon
 305 the frequency of each amino acid substitution in the set of protein used to create the conserved
 306 domain model [17]. Peptide encoded by MARHY0478 exhibited a clear sequence similarity
 307 with the four hydrophobic compounds transporters. In view of this significant domain
 308 conservation, it is very likely that MARHY0478 encodes for a hydrophobic compound
 309 transporter. Given that MARHY0478 was induced during growth on hexadecane, and that

homologs are only found in marine alkane degrading strains or putative alkane degrading strains, we hypothesized that MARHY0478 and its homologs would constitute a family of alkane transporter specific to marine bacteria. The production in large quantity of an alkane transporter in biofilm cells would contribute to the high growth rate observed on alkane despite its low solubility.

Polypeptides encoded by MARHY0477 and their homologs define a protein family whose function is likely to be related to MARHY0478, as members of the two families co-occur in the same strains and occasionally form a putative operon. The ecological significance of the MARHY0478 and MARHY0477 families certainly deserve closer inspection as they are present in genera known to be the main actors in hydrocarbon degradation in marine environments.

322

323 **Acknowledgments**

324 We thank Dr Rizard Lobinsky, Dr Hugues Prudhomme and the Pôle Protéomique of
325 the Bordeaux 2 University for protein identification. We also thank Dr Pierre Sivadon and Dr
326 Anne Fahy for useful discussions and critical reading of the manuscript. We gratefully
327 acknowledge the 6th European Framework Programme, Contract 018391 FACEIT, the
328 National Program ANR “ECCO” INDHYC project, the CNRS program Ingénierie
329 Ecologique, the région Aquitaine and the département des Pyrénées Atlantiques for financial
330 support.

References

- [1] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
- [2] Baldi, F., Ivosevic, N., Minacci, A., Pepi, M., Fani, R., Svetlicic, V., Utlic, V. (1999) Adhesion of *Acinetobacter venetianus* to diesel fuel droplets studied with in situ electrochemical and molecular probes. *Appl. Environ. Microbiol.* 65, 2041-2048.
- [3] Bastiaens, L., Springael, D., Wattiau, P., Harms, H., DeWachter, R., Verachtert, H., Diels, L. (2000) Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using PAH-sorbing carriers. *Appl. Environ. Microbiol.* 66, 1834-1843.
- [4] Bouchez-Naitali, M., Blanchet, D., Bardin, V., Vandecasteele, J.-P. (2001) Evidence for interfacial uptake in hexadecane degradation by *Rhodococcus equi*: The importance of cell flocculation. *Microbiology* 147, 2537-2543.
- [5] Calvillo, Y.M., Alexander, M. (1996) Mechanism of microbial utilization of biphenyl sorbed to polyacrylic beads. *Appl. Microbiol. Biotechnol.* 45, 383-390.
- [6] Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., et al. (2007) Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *Proc. Natl. Acad. Sci. U S A* 104, 5602-5607.
- [7] Filloux, A., Hachani, A., Bleves, S. (2008) The bacterial type VI secretion machine: Yet another player for protein transport across membranes. *Microbiology* 154, 1570-1583.
- [8] Gauthier, M.J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P., Bertrand, J.-C. (1992) *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int. J. Syst. Bacteriol.* 42, 568-576.

- 355 [9] Golyshin, P.N., Chernikova, T.N., Abraham, W.-R., Lünsdorf, H., Timmis, K.N.,
356 Yakimov, M.M. (2002) Oleiphilaceae fam. nov., to include *Oleiphilus messinensis* gen. nov.,
357 sp. nov., a novel marine bacterium that obligately utilizes hydrocarbons. Int. J. Syst. Evol.
358 Microbiol. 52, 901-911.
- 359 [10] Harms, H., Zehnder, A.J.B. (1995) Bioavailability of sorbed 3-chlorodibenzofuran. Appl.
360 Environ. Microbiol. 61, 27-33.
- 361 [11] Heydorn, A., Ersbøll, B., Kato, J., Hentzer, M., Parsek, M.R., Tolker-Nielsen, T.,
362 Givskov, M., Molin, S. (2002) Statistical analysis of *Pseudomonas aeruginosa* biofilm
363 development: Impact of mutations in genes involved in twitching motility, cell-to-cell
364 signaling, and stationary-phase sigma factor expression. Appl. Environ. Microbiol. 68, 2008-
365 2017.
- 366 [12] Johnsen, A.R., Karlson, U. (2004) Evaluation of bacterial strategies to promote the
367 bioavailability of polycyclic aromatic hydrocarbons. Appl. Microbiol. Biotechnol. 63, 452-
368 459.
- 369 [13] Johnsen, A.R., Wick, L.Y., Harms, H. (2005) Principles of microbial PAH-degradation
370 in soil. Environ. Pollut. 133, 71-84.
- 371 [14] Kahng, H.Y., Byrne, A.M., Olsen, R.H., Kukor, J.J. (2000) Characterization and role of
372 *tbuX* in utilization of toluene by *Ralstonia pickettii* PKO1. J. Bacteriol. 182, 1232-1242.
- 373 [15] Klein, B., Grossi, V., Bouriat, P., Goulas, P., Grimaud, R. (2008) Cytoplasmic wax ester
374 accumulation during biofilm-driven substrate assimilation at the alkane-water interface by
375 *Marinobacter hydrocarbonoclasticus* SP17. Res. Microbiol. 159, 137-144.
- 376 [16] Kornberg, H.L. (1966) The role and control of the glyoxylate cycle in *Escherichia coli*.
377 Biochem. J. 99, 1-11.

- 378 [17] Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales,
379 N.R., Gwadz, M., Hao, L., He, S., et al. (2007) CDD: a conserved domain database for
380 interactive domain family analysis. *Nucleic Acids Res.* 35, D237-D240.
- 381 [18] Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., Deweese-Scott, C.,
382 Fong, J.H., Geer, L.Y., Geer, R.C., et al. (2009) CDD: Specific functional annotation with the
383 Conserved Domain Database. *Nucleic Acids Res.* 37, D205-D210.
- 384 [19] Mulder, H., Breure, A.M., Van Honschooten, D., Grotenhuis, J.T.C., Van Andel, J.G.,
385 Rulkens, W.H. (1998) Effect of biofilm formation by *Pseudomonas* 8909N on the
386 bioavailability of solid naphthalene. *Appl. Microbiol. Biotechnol.* 50, 277-283.
- 387 [20] O'Toole, G.A., Kolter, R. (1998) Flagellar and twitching motility are necessary for
388 *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* 30, 295-304.
- 389 [21] Proft, T., Baker, E.N. (2009) Pili in Gram-negative and Gram-positive bacteria -
390 structure, assembly and their role in disease. *Cell. Mol. Life Sci.* 66, 613-635.
- 391 [22] Rao, P.S., Yamada, Y., Tan, Y.P., Leung, K.Y. (2004) Use of proteomics to identify
392 novel virulence determinants that are required for *Edwardsiella tarda* pathogenesis. *Mol.*
393 *Microbiol.* 53, 573-586.
- 394 [23] Rodrigues, A.C., Wuertz, S., Brito, A.G., Melo, L.F. (2005) Fluorene and phenanthrene
395 uptake by *Pseudomonas putida* ATCC 17514: Kinetics and physiological aspects. *Biotechnol.*
396 *Bioeng.* 90, 281-289.
- 397 [24] Sabirova, J.S., Ferrer, M., Regenhardt, D., Timmis, K.N., Golyshin, P.N. (2006)
398 Proteomic insights into metabolic adaptations in *Alcanivorax borkumensis* induced by alkane
399 utilization. *J. Bacteriol.* 188, 3763-3773.
- 400 [25] Singh, P.K., Parsek, M.R., Greenberg, E.P., Welsh, M.J. (2002) A component of innate
401 immunity prevents bacterial biofilm development. *Nature* 417, 552-555.

- 402 [26] Soltani, M., Metzger, P., Largeau, C. (2004) Effects of hydrocarbon structure on fatty
403 acid, fatty alcohol, and beta-hydroxy acid composition in the hydrocarbon-degrading
404 bacterium *Marinobacter hydrocarbonoclasticus*. *Lipids* 39, 491-505.
- 405 [27] Stach, J.E., Burns, R.G. (2002) Enrichment versus biofilm culture: a functional and
406 phylogenetic comparison of polycyclic aromatic hydrocarbon-degrading microbial
407 communities. *Environ. Microbiol.* 4, 169-182.
- 408 [28] Stewart, P.S., Franklin, M.J. (2008) Physiological heterogeneity in biofilms. *Nat. Rev.*
409 *Microbiol.* 6, 199-210.
- 410 [29] Tseng, T.-T., Tyler, B.M., Setubal, J.C. (2009) Protein secretion systems in bacterial-host
411 associations, and their description in the Gene Ontology. *BMC Microbiology* 9 (Suppl 1):S2,
412 1-9.
- 413 [30] van den Berg, B., Black, P.N., Clemons, W.M., Jr., Rapoport, T.A. (2004) Crystal
414 structure of the long-chain fatty acid transporter FadL. *Science* 304, 1506-1509.
- 415 [31] Van Hamme, J.D., Singh, A., Ward, O.P. (2003) Recent advances in petroleum
416 microbiology. *Microbiol. Mol. Biol. Rev.* 67, 503-549.
- 417 [32] Wang, Y., Rawlings, M., Gibson, D.T., Labbe, D., Bergeron, H., Brousseau, R., Lau,
418 P.C. (1995) Identification of a membrane protein and a truncated LysR-type regulator
419 associated with the toluene degradation pathway in *Pseudomonas putida* F1. *Mol. Gen.*
420 *Genet.* 246, 570-579.
- 421 [33] Whiteley, M., Banger, M.G., Bumgarner, R.E., Parsek, M.R., Teitzel, G.M., Lory, S.,
422 Greenberg, E.P. (2001) Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature* 413,
423 860-864.
- 424 [34] Whyte, L.G., Slagman, S.J., Pietrantonio, F., Bourbonniere, L., Koval, S.F., Lawrence,
425 J.R., Inniss, W.E., Greer, C.W. (1999) Physiological adaptations involved in alkane

426 assimilation at a low temperature by *Rhodococcus sp.* strain Q15. Appl. Environ. Microbiol.
427 65, 2961-2968.

428 [35] Willison, J.C. (2004) Isolation and characterization of a novel sphingomonad capable of
429 growth with chrysene as sole carbon and energy source. FEMS Microbiol. Lett. 241, 143-150.

430

Legends to figures

Fig. 1. Two-dimensional gel electrophoresis of protein extracts of *M. hydrocarbonoclasticus* SP17. Planktonic cells grown on acetate (top image) and biofilm cells grown on hexadecane (bottom image). Spots of identified proteins are numbered. Molecular weights (in kDa) are indicated on the left side.

Fig. 2. Organization of the genes encoding for type VI secretion system of *M. hydrocarbonoclasticus* SP17 (top) and *Vibrio cholerae* N16961 (bottom). The values between homologous genes refer to the percent peptide sequence identity of their products. Accession number : MARHY3635 to MARHY3620 (FP475883 to FP47598), MARHY0176 (FP475953), MARHY1076 (FP475938) and MARHY2494 (FP475918)

Fig. 3. Multiple sequence alignment of FadL, TbuX, TodX and translated MARHY0478. Sequence alignment was performed using the web-based tool CD search using the PSSM 112176 scoring matrix. Conserved residues are shown in grey boxes and identical residues in black boxes. Secondary structures elements of FadL were retrieved from the PDB database (PDB ID:1t16) and are shown above the sequences.

Tables

Table 1

Identification of proteins differentially expressed in biofilm on hexadecane compared to planktonic cells on acetate.

Spot NO.	CDS	Gene	Product	Differential abundance ^a	Amount in planktonic/acetate condition ^b	Amount in biofilm/hexadecane condition ^b	Gene accession number ^c
Transport and binding protein							
1	MARHY1035	<i>fhuE</i>	Outer membrane receptor for ferric iron uptake	+5.01	0.0102	0.0513	FP475940
2	MARHY2192	<i>fbpA</i>	Iron(III) ABC transporter, periplasmic iron(III)-binding protein	+3.89	0.1233	0.4793	FP475923
3	MARHY3135	<i>cirA</i>	Ferric iron-catecholate outer membrane transporter	BH	ND	0.079	FP475909
4	MARHY2019	<i>cysP</i>	Thiosulfate transporter subunit	+3.01	0.1170	0.3525	FP475928
5	MARHY3535	<i>pstS</i>	ABC phosphate transporter, periplasmic component	+10.74	0.1018	1.0934	FP475901
6	MARHY3277	–	ABC-type metal ion transporter, periplasmic component	BH	ND	0.0831	FP475905
7	MARHY1478	–	ABC-type branched-chain amino acid transporter, periplasmic component	-4.32	0.0882	0.0204	FP475934
Cellular processes							
8	MARHY2564	<i>pilA</i>	Fimbrial protein precursor	-2.44	0.6983	0.2861	FP475917
9	MARHY2994	<i>tpm</i>	Thiopurine methyltransferase	PA	0.0551	ND	FP475913
Information transfer							
10	MARHY3200	<i>greA</i>	Transcription elongation factor	BH	ND	0.0843	FP475907
Protein fate							
11	MARHY0922	–	FKBP-type peptidyl-prolyl cis-trans isomerase	PA	0.0418	ND	FP475944
12	MARHY0942	<i>dsbC</i>	Protein disulfide isomerase II	+4.44	0.0363	0.1611	FP475943
13	MARHY0958	<i>mucD</i>	Serine protease	+4.05	0.0517	0.2091	FP475942
14	MARHY1798	<i>slyD</i>	FKBP-type peptidyl-prolyl cis-trans isomerase	PA	0.0584	ND	FP475931

Spot NO.	CDS	Gene	Product	Differential abundance ^a	Amount in planktonic/acetate condition ^b	Amount in biofilm/hexadecane condition ^b	Gene accession number ^c
Stress response							
15	MARHY3764	<i>ahpF</i>	Alkyl hydroperoxide reductase f-subunit	BH	ND	0.0369	FP475881
Cofactor biosynthesis							
16	MARHY2568	<i>coaE</i>	Dephospho-CoA kinase	PA	0.0533	ND	FP475916
Aminoacid biosynthesis							
17	MARHY1723	<i>asd</i>	Aspartate-semialdehyde dehydrogenase	PA	0.0655	ND	FP475932
18	MARHY3112	<i>hisH</i>	Glutamine amidotransferase	BH	ND	0.1733	FP475910
19	MARHY0406	<i>lysA</i>	Diaminopimelate decarboxylase	PA	0.0901	ND	FP475949
20	MARHY2259	<i>serC</i>	3-phosphoserine/phosphohydroxythreonine aminotransferase	BH	ND	0.1327	FP475920
Nucleotide biosynthesis							
21	MARHY2255	<i>cmk</i>	Cytidylate kinase	BH	ND	0.0691	FP475921
22	MARHY2158	<i>ndk</i>	Multifunctional nucleoside diphosphate kinase and apyrimidinic endonuclease and 3'-phosphodiesterase	BH	ND	0.0836	FP475924
23	MARHY3664	<i>pyrC</i>	Aspartate carbamoyltransferase, non-catalytic chain	BH	ND	0.0533	FP475882
24	MARHY2249	<i>pyrF</i>	Orotidine-5'-phosphate decarboxylase	+3.35	0.0582	0.1948	FP475922
Central intermediary metabolism							
25	MARHY3262	<i>exaC</i>	NAD+ dependent acetaldehyde dehydrogenase	-3.44	0.0968	0.0281	FP475906
Energy metabolism							
26	MARHY1444	<i>acnB</i>	Bifunctional aconitate hydratase 2 and 2-methylisocitrate dehydratase	-2.89	0.9104	0.3147	FP475936
27	MARHY0078	<i>idh</i>	Isocitrate dehydrogenase	-3.13	0.2014	0.0643	FP475954
28	MARHY2120	<i>sucD</i>	Succinyl-CoA synthetase, NAD(P)-binding, alpha-subunit	-3.2	0.8657	0.2708	FP475927
29	MARHY2121	<i>sucC</i>	Succinyl-CoA synthetase, beta-subunit	-6.97	0.1894	0.0272	FP475926
30	MARHY2126	<i>sdhA</i>	Succinate dehydrogenase, flavoprotein subunit	-4.35	0.3639	0.0837	FP475925
31	MARHY0774	<i>petA</i>	Ubiquinol-cytochrome c reductase iron-sulfur subunit	+6	0.0449	0.2696	FP475945
32	MARHY1458	<i>glcB</i>	Malate synthase G	BH	ND	0.0203	FP475935
33	MARHY1802	<i>etfA</i>	Electron transfer flavoprotein subunit, FAD-binding	+3.13	0.1707	0.5347	FP475930
34	MARHY1815	<i>paaG</i>	Enoyl-CoA hydratase	PA	0.073	ND	FP475929

Spot NO.	CDS	Gene	Product	Differential abundance ^a	Amount in planktonic/acetate condition ^b	Amount in biofilm/hexadecane condition ^b	Gene accession number
Fatty acid biosynthesis							
35	MARHY1009	<i>acs</i>	Acetyl-CoA synthetase	PA	0.0435	ND	FP475941
36	MARHY3086	<i>fabA</i>	Beta-hydroxydecanoyl thioester dehydrase	PA	0.1247	ND	
37	MARHY3087	<i>fabB</i>	3-oxoacyl-[acyl-carrier-protein] synthase I	-2.54	0.5215	0.2056	FP475911
38	MARHY1438	<i>fabF</i>	3-oxoacyl-[acyl-carrier-protein] synthase II	-4.04	0.1566	0.0388	FP475937
39	MARHY1579	<i>prpE</i>	Propionate-CoA ligase	-5.75	0.1203	0.0209	FP475933
Uncertain or unknown function							
40	MARHY0256	—	Probable ABC transporter, ATPase subunit	PA	0.2227	ND	FP475952
41	MARHY0299	—	Probable TRAP dicarboxylate transporter	BH	ND	0.123	FP475951
42	MARHY3613	—	Probable PspA protein	+3.5	0.1716	0.6	FP475899
43	MARHY3634	—	Probable type VI secretion system subunit	BH	ND	0.2073	FP475884
44	MARHY3635	—	Probable type VI secretion system subunit	+6.68	0.1118	0.7467	FP475883
45	MARHY3432	—	Probable TonB-dependent receptor	+10.99	0.0193	0.2118	FP475902
46	MARHY3289	—	Probable porin	PA	0.4185	ND	FP475904
47	MARHY0478	—	Probable hydrophobic compounds transporter	+11.23	0.0369	0.4144	FP475946
48	MARHY0477	—	Conserved protein	+5.62	0.0369	0.2076	FP475947
49	MARHY0460	—	Conserved protein	BH	ND	0.696	FP475948
50	MARHY1073	—	Conserved protein	PA	0.0385	ND	FP475939
51	MARHY2963	—	Conserved protein	+9.21	0.0326	0.3001	FP475914
52	MARHY3166	—	Conserved protein	-14.81	3.612	0.2445	FP475908
53	MARHY3295	—	Conserved protein	BH	ND	0.0961	FP475903
54	MARHY0333	—	Conserved protein	BH	ND	0.1249	FP475950
55	MARHY3623	—	Probable type VI secretion system subunit	+4.71	0.0142	0.0671	FP475895
56	MARHY3550	—	Conserved protein	PA	0.1351	ND	FP475900
57	MARHY2686	—	Conserved protein	+22.01	0.0924	2.0331	FP475915
58	MARHY2432	—	Conserved protein	PA	0.0642	ND	FP475919

^aPositive values represent overexpression in biofilm/hexadecane condition. Negative values represent underexpression in biofilm/hexadecane condition. BH means that the protein is solely detected in biofilm/hexadecane condition. PA means that the protein is solely detected in planktonic/acetate condition. ^bAverage percent of spot volume relative to total spot volume. ND, not detected.

Table 2

Phylogenetic distribution of MARHY0477 and MARHY0478 homologs.

Strain	Provenance	Alkane degradation ^a	MARHY0477 homologs ^b	MARHY0478 homologs ^b	Alkane degradation genes ^c
<i>Marinobacter hydrocarbonoclasticus</i> SP17	Marine sediment, Mediterranean Sea	yes	-	-	2 alk, 2 P450
<i>Marinobacter hydrocarbonoclasticus</i> VT8	Oil well off the Vietnamese coast	yes	2 (74%, 33%)	3 (98%, 74%, 74%)	3 alk, 1 P450
<i>Marinobacter algicola</i> DG893	Culture of a dinoflagellate	yes	1 (59%)	1 (83%)	2 alk
<i>Alcanivorax borkumensis</i> SK2	Sea water, North Sea	yes	2 (34%, 32%)	1 (40%)	2 alk, 1 P450
<i>Alcanivorax</i> sp. DG881	Culture of marine algae	yes	6 (35%, 32%, 33%, 31%, 31%, 29%)	6 (54%, 54%, 50%, 48%, 46%, 42%)	2 alk, 2 P450
<i>Bermanella marisrubri</i>	Sea water, Red Sea	ND	1 (31%)	1 (45%)	1 alk
<i>Moritella</i> sp. PE36	Deep sea floor, Pacific Ocean	ND	1 (29%)	1 (39%)	1 alk

Sequences similar to translated MARHY0477 and MARHY0478 were searched against the non-redundant GenBank CDS translations database, using the BLAST program. ^aND not determined. ^bNumber of homologous genes and their percent peptide sequence identity in brackets. ^calk and P450 indicate the presence of alkane hydroxylase and cytochrome P450 alkane monooxygenase genes respectively.

Figures

Fig. 1

Fig. 2

Fig. 3