



**HAL**  
open science

**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, Laure Prat, Sophie Mangenot, Stéphane Cruveiller,  
Philippe Goulas, R. Grimaud

► **To cite this version:**

Pierre-Joseph Vaysse, Laure Prat, Sophie Mangenot, Stéphane Cruveiller, Philippe Goulas, et al.. Proteomic analysis of *Marinobacter hydrocarbonoclasticus* SP17 biofilm formation at the alkane-water interface reveals novel proteins and cellular processes involved in hexadecane assimilation. *Research in Microbiology*, 2009, 160 (10), pp.829-837. 10.1016/j.resmic.2009.09.010 . hal-01614789

**HAL Id: hal-01614789**

**<https://hal.science/hal-01614789>**

Submitted on 18 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

4

5 Pierre-Joseph Vaysse<sup>a</sup>, Laure Prat<sup>a†</sup>, Sophie Mangenot<sup>b</sup>, Stéphane Cruveiller<sup>c</sup>, Philippe  
6 Goulas<sup>a</sup>, Régis Grimaud<sup>a\*</sup>

7

8 <sup>a</sup> Institut Pluridisciplinaire de Recherche en Environnement et Matériaux, Equipe

9 Environnement et Microbiologie UMR5254 CNRS, IBEAS, Université de Pau et des Pays de  
10 l'Adour, BP1155, 64013 Pau cedex, France

11

12 <sup>b</sup> CEA/DSV/IG/Genoscope, 2 rue Gaston Cremieux, 91057 Evry cedex, France

13

14 <sup>c</sup> CEA/DSV/IG/Genoscope/LGC, 2 rue Gaston Cremieux, 91057 Evry cedex, France

15

16 <sup>†</sup> Current address : EPFL, ENAC-ISTE, Laboratory of Environmental Biotechnology,

17 Lausanne, Switzerland

18

19 E-mail addresses of all authors

20 • Pierre-Joseph Vaysse: [pierre-joseph.vaysse@univ-pau.fr](mailto:pierre-joseph.vaysse@univ-pau.fr)

21 • Laure Prat: [laure.prat@epfl.ch](mailto:laure.prat@epfl.ch)

22 • Sophie Mangenot: [mangenot@genoscope.cns.fr](mailto:mangenot@genoscope.cns.fr)

23 • Stéphane Cruveiller: [scruveil@genoscope.cns.fr](mailto:scruveil@genoscope.cns.fr)

24 • Philippe Goulas: [philippe.goulas@univ-pau.fr](mailto:philippe.goulas@univ-pau.fr)

25 • Régis Grimaud: [regis.grimaud@univ-pau.fr](mailto:regis.grimaud@univ-pau.fr) \* Correspondence and reprints

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

27

28 **Abstract**

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

46

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

49

50 **1. Introduction**

51

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

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

64 Biofilms developing on hydrocarbons present two properties distinguishing them from  
65 other biofilms which are their substrate-substratum specificity and their capacity to overcome  
66 the low accessibility of hydrophobic substrates. It was observed in the case of PAH-degrading  
67 strains that the less soluble the hydrocarbon was, the more cells grew at the PAH-water  
68 interface [12,23]. Biofilm formation has been shown to promote growth on hydrocarbons by  
69 facilitating interfacial access. Kinetic studies demonstrated that growth at the interface occurs  
70 faster than mass transfer rate of hydrocarbons in the absence of bacteria would suggest  
71 [4,5,10,13]. Thus, biofilm formation constitutes an efficient adaptive strategy to assimilate  
72 hydrocarbon. However, the genetic determinants and the molecular mechanisms underlying  
73 the functioning of these biofilms remain poorly understood. In order to identify the proteins

74 and cellular functions involved, we undertook a proteomic analysis of the biofilm of  
75 *Marinobacter hydrocarbonoclasticus* SP17 growing at the hexadecane-water interface. This  
76 marine alkane degrading bacterium readily forms biofilms on metabolizable hydrophobic  
77 organic compounds including *n*-alkanes from 8 to 28 carbon atoms, whereas in the same  
78 condition this strain does not develop biofilms on an inert substratum such as plastic or glass  
79 [15]. The doubling time of biofilm cells growing on hexadecane (between 3 and 5 hours  
80 depending on the development stage of the biofilm) is similar to that observed on water-  
81 soluble substrates such as acetate, pyruvate and lactate (P-J Vaysse, unpublished results).  
82 Furthermore, the rate of hexadecane degradation dramatically decreases when the biofilm is  
83 disorganized by strong shaking, indicating a strong link between biofilm formation and the  
84 utilization of alkanes [15]. These observations indicate that the biofilm lifestyle must provide  
85 *M. hydrocarbonoclasticus* SP17 with efficient mechanisms to access the hexadecane. In the  
86 current study, we undertook a proteomic analysis of a biofilm growing at hexadecane-water  
87 interface. The results obtained indicate that adaptation to alkane utilization as carbon and  
88 energy source involves a global change in cell physiology. Novel proteins and cellular  
89 processes involved in alkane assimilation were revealed.

## 90 2. Material and methods

91

### 92 2.1 Growth conditions and preparation of protein extracts

93

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

96 All chemicals used were from Sigma-Aldrich unless otherwise specified. Biofilm cultures  
97 were inoculated with cells growing exponentially on acetate (20 mM) washed twice with one  
98 volume of SSW and resuspended to a final optical density at 600nm of 0.1 in 300 ml SSW  
99 medium supplemented with 0.2% (v/v) hexadecane in 1 l Erlenmeyer flasks. Biofilms were  
100 grown at the hexadecane-SSW interface at 30 °C, under slow shaking (50 rpm). After 35  
101 hours incubation, the medium was carefully discarded, the biofilms resuspended in 30 ml of  
102 SSW and centrifuged for 20 min at 20,000 g at room temperature. Biofilms were harvested  
103 above the supernatant while residual planktonic cells were pelleted. Cells grown on acetate  
104 (100 ml) were harvested during exponential growth phase (OD<sub>600nm</sub>= 0.3) by centrifuging  
105 20 min at 20,000 g at room temperature.

106 For protein extraction, cell pellets or biofilms were washed twice with 5 ml of acetone,  
107 resuspended in 5 ml of water containing a protease inhibitor cocktail and sonicated on ice for  
108 1 min with 500 ms/s pulses at 35 W. Sonicated cell suspensions were then incubated with 40  
109 µg of DNase I, 10 µg of RNase A plus 0.01% (v/v) Triton X100 for 30 min at room  
110 temperature. Proteins were precipitated for 30 min at 4 °C by adding 500 µl of 100%  
111 trichloroacetic acid and then centrifuged 10 min at 20,000g at 4 °C. Protein extracts were  
112 washed twice with 1 ml of trichloroacetic acid 10% (v/v) and twice with 1 ml of acetone.  
113 Proteins were air dried and dissolved in IEF buffer (urea 8 M, 3-[(3-  
114 Cholamidopropyl)dimethylammonio]-1-propanesulfonate 4% (w/v), dithiothreitol 60 mM,

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

118

## 119 *2.2 2D electrophoresis and gel analysis*

120

121 Two hundred and thirty micrograms of protein were applied to a 24 cm Immobiline  
122 Dry strip with a 3-7 non linear pH gradient (GE Healthcare). Isoelectric focusing and SDS-  
123 PAGE (12.5%) were carried out using the Multiphor II and Ettan DALTsix systems,  
124 respectively (GE Healthcare) according to the manufacturer's instructions. Gels were stained  
125 with Deep Purple Total Protein Stain (GE Healthcare) and then scanned using a Typhoon  
126 9200 fluorescent scanner (GE Healthcare). Image analysis, spot detection and matching were  
127 carried out using Image Master Platinum software (GE Healthcare) and checked manually.

128 Three replicate gels from three independent experiments were run for each growth condition.  
129 The normalized protein amount for each protein spot was defined as the fraction of that spot  
130 volume to the total spot's volume of the gel. Student t-test ( $P < 0.05$ ) and a threshold of 2-fold  
131 change were used to determine the proteins significantly differentially expressed between the  
132 two conditions.

133

## 134 *2.3 Protein identification by nanoLC-MS/MS*

135

136 Protein spots were excised from gels, destained in ammonium bicarbonate 50 mM  
137 50% acetonitrile (ACN), rinsed twice in ultrapure water and dehydrated in 100% ACN. After  
138 ACN removal by evaporation, gel pieces were dried, rehydrated with a trypsin solution (10  
139 ng/ml in 50 mM ammonium bicarbonate) at 4 °C for 10 min and finally incubated overnight

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

149

#### 150 *2.4 Database searching and sequence analysis*

151

152 Sequences similarities were searched against the translation of the non-redundant  
153 GenBank CDS database, using the BLAST program [1]. Homologs were defined as proteins  
154 giving an alignment with a bits score above 100 (scoring matrix BLOSUM62) and an  
155 expected value (E-value) below  $1.10^{-25}$ . Protein domain searches and multiple sequence  
156 alignment were carried out using the NCBI-Conserved Domain Database search program with  
157 CDD v2.16-27036 PSSMs database (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>)  
158 [18]. Synteny similarities were searched using the MAGE microbial genome annotation  
159 system from the Genoscope Evry France  
160 (<https://www.genoscope.cns.fr/agc/mage/wwwpkgdb/MageHome>).



161

162 **3. Results and discussion**

163

164 *3.1. Comparison of protein patterns of biofilm cells growing on hexadecane with planktonic*  
165 *cells growing on acetate*

166

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

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

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

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

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

215

### 216 3.2. Redirection of carbon flux in biofilm cells growing on hexadecane

217 Most of the enzymes corresponding to the CO<sub>2</sub>-releasing steps of the tricarboxylic acid  
218 cycle (TCA cycle) (encoded by MARHY0078 *idh*, MARHY2120 *sucD*, MARHY2121 *sucC*,  
219 MARHY2126 *sdhA*) were found to be down regulated in BH. In addition, the gene encoding  
220 malate synthase (MARHY1458 *glcB*), catalyzing the conversion of glyoxylate into malate,  
221 was upregulated. This modulation of TCA cycle enzymes suggests a stimulation of the  
222 glyoxylate bypass in BH. This anaplerotic pathway allows the replenishment of cells in the  
223 metabolic intermediates necessary for the synthesis of their cellular components when acetyl-  
224 CoA is the only carbon source available in the cell [16]. The activation of the glyoxylate  
225 bypass under a hexadecane diet compared to a diet of pyruvate or glucose has been already  
226 observed in proteomic analyses on *A. burkomensis* SK2 and *Geobacillus thermodenitrificans*  
227 [6,24]. In our study, stimulation of the glyoxylate bypass was observed under hexadecane  
228 conditions compared with acetate. In hexadecane and acetate conditions, the intracellular  
229 carbon source is acetyl-CoA hence the glyoxylate pathway is required in both conditions. The  
230 explanation of the glyoxylate bypass stimulation could lie in the fact that hexadecane is a  
231 more energetic substrate than acetate. In fact the breakdown of hexadecane to acetyl-CoA  
232 produces energy whereas the conversion of acetate to acetyl-CoA does not. Augmenting the  
233 flux of carbon through the glyoxylate bypass on hexadecane could allow cells to restore the  
234 balance between carbon assimilation and energy production. However interpretation of

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

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

241

### 242 3.3. A type VI secretion system is overproduced by biofilm cells growing on hexadecane

243

244 Three proteins sharing similarities with type VI secretion system (T6SS) subunits,  
245 encoded by MARHY3623, MARHY3634 and MARHY3635, were overexpressed in BH.  
246 MARHY3634 and MARHY3635 were among the most abundant protein detected in BH, with  
247 0.2 and 0.7 % of total protein respectively. MARHY3634, MARHY3635 and MARHY3623  
248 are localized within a cluster of 16 genes (from MARHY3635 to MARHY3620) that are  
249 transcribed in the same direction. All members of the *M. hydrocarbonoclasticus* SP17 cluster  
250 share similarity with genes found in T6SS clusters from other bacteria including *Vibrio*  
251 *cholerae*, *P. aeruginosa* and *Escherichia coli*. The best synteny conservation with a  
252 characterized T6SS was observed with *Vibrio cholerae* N16961 T6SS gene cluster (Fig. 2).

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

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

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

275

#### 276 3.4. Evidence for an alkane transporter family specific to marine bacteria

277

278 Two proteins produced in large quantities in BH, encoded by MARHY0478 and  
279 MARHY0477, drew our attention because their sequences were found to be conserved  
280 exclusively in marine species capable to use alkanes as a carbon source or possessing at least  
281 one copy of an alkane hydroxylase gene. The phylogenetic distribution of MARHY0478 and  
282 MARHY0477 homologs is rather restricted since they are found only in two orders of the  
283 Gammaproteobacteria: the *Oceanospirillales* and the *Alteromonadales*. The degree of peptide  
284 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

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

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

321

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.



331 **References**

- 332 [1] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman,  
333 D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search  
334 programs. *Nucleic Acids Res.* 25, 3389-3402.
- 335 [2] Baldi, F., Ivosevic, N., Minacci, A., Pepi, M., Fani, R., Svetlicic, V., Utlic, V. (1999)  
336 Adhesion of *Acinetobacter venetianus* to diesel fuel droplets studied with in situ  
337 electrochemical and molecular probes. *Appl. Environ. Microbiol.* 65, 2041-2048.
- 338 [3] Bastiaens, L., Springael, D., Wattiau, P., Harms, H., DeWachter, R., Verachtert, H., Diels,  
339 L. (2000) Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria  
340 using PAH-sorbing carriers. *Appl. Environ. Microbiol.* 66, 1834-1843.
- 341 [4] Bouchez-Naitali, M., Blanchet, D., Bardin, V., Vandecasteele, J.-P. (2001) Evidence for  
342 interfacial uptake in hexadecane degradation by *Rhodococcus equi*: The importance of cell  
343 flocculation. *Microbiology* 147, 2537-2543.
- 344 [5] Calvillo, Y.M., Alexander, M. (1996) Mechanism of microbial utilization of biphenyl  
345 sorbed to polyacrylic beads. *Appl. Microbiol. Biotechnol.* 45, 383-390.
- 346 [6] Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., et al. (2007)  
347 Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans*  
348 NG80-2 isolated from a deep-subsurface oil reservoir. *Proc. Natl. Acad. Sci. U S A* 104,  
349 5602-5607.
- 350 [7] Filloux, A., Hachani, A., Bleves, S. (2008) The bacterial type VI secretion machine: Yet  
351 another player for protein transport across membranes. *Microbiology* 154, 1570-1583.
- 352 [8] Gauthier, M.J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P., Bertrand,  
353 J.-C. (1992) *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely  
354 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., Bouriart, 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., Bangera, 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

431 **Legends to figures**

432

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

437

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

443

444 Fig. 3. Multiple sequence alignment of FadL, TbuX, TodX and translated MARHY0478.  
445 Sequence alignment was performed using the web-based tool CD search using the PSSM  
446 112176 scoring matrix. Conserved residues are shown in grey boxes and identical residues in  
447 black boxes. Secondary structures elements of FadL were retrieved from the PDB database  
448 (PDB ID:1t16) and are shown above the sequences.

449

## 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 <sup>a</sup>	Amount in planktonic/acetate condition <sup>b</sup>	Amount in biofilm/hexadecane condition <sup>b</sup>	Gene accession number <sup>c</sup>
<b>Transport and binding protein</b>							
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
<b>Cellular processes</b>							
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
<b>Information transfer</b>							
10	MARHY3200	<i>greA</i>	Transcription elongation factor	BH	ND	0.0843	FP475907
<b>Protein fate</b>							
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 <sup>a</sup>	Amount in planktonic/acetate condition <sup>b</sup>	Amount in biofilm/hexadecane condition <sup>b</sup>	Gene accession number <sup>c</sup>
<b>Stress response</b>							
15	MARHY3764	<i>ahpF</i>	Alkyl hydroperoxide reductase f-subunit	BH	ND	0.0369	FP475881
<b>Cofactor biosynthesis</b>							
16	MARHY2568	<i>coaE</i>	Dephospho-CoA kinase	PA	0.0533	ND	FP475916
<b>Aminoacid biosynthesis</b>							
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
<b>Nucleotide biosynthesis</b>							
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
<b>Central intermediary metabolism</b>							
25	MARHY3262	<i>exaC</i>	NAD+ dependent acetaldehyde dehydrogenase	-3.44	0.0968	0.0281	FP475906
<b>Energy metabolism</b>							
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 <sup>a</sup>	Amount in planktonic/acetate condition <sup>b</sup>	Amount in biofilm/hexadecane condition <sup>b</sup>	Gene accession number
<b>Fatty acid biosynthesis</b>							
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
<b>Uncertain or unknown function</b>							
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

<sup>a</sup>Positive 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. <sup>b</sup>Average 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 <sup>a</sup>	MARHY0477 homologs <sup>b</sup>	MARHY0478 homologs <sup>b</sup>	Alkane degradation genes <sup>c</sup>
<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. <sup>a</sup>ND not determined. <sup>b</sup>Number of homologous genes and their percent peptide sequence identity in brackets. <sup>c</sup>alk and P450 indicate the presence of alkane hydroxylase and cytochrome P450 alkane monooxygenase genes respectively.

Figures

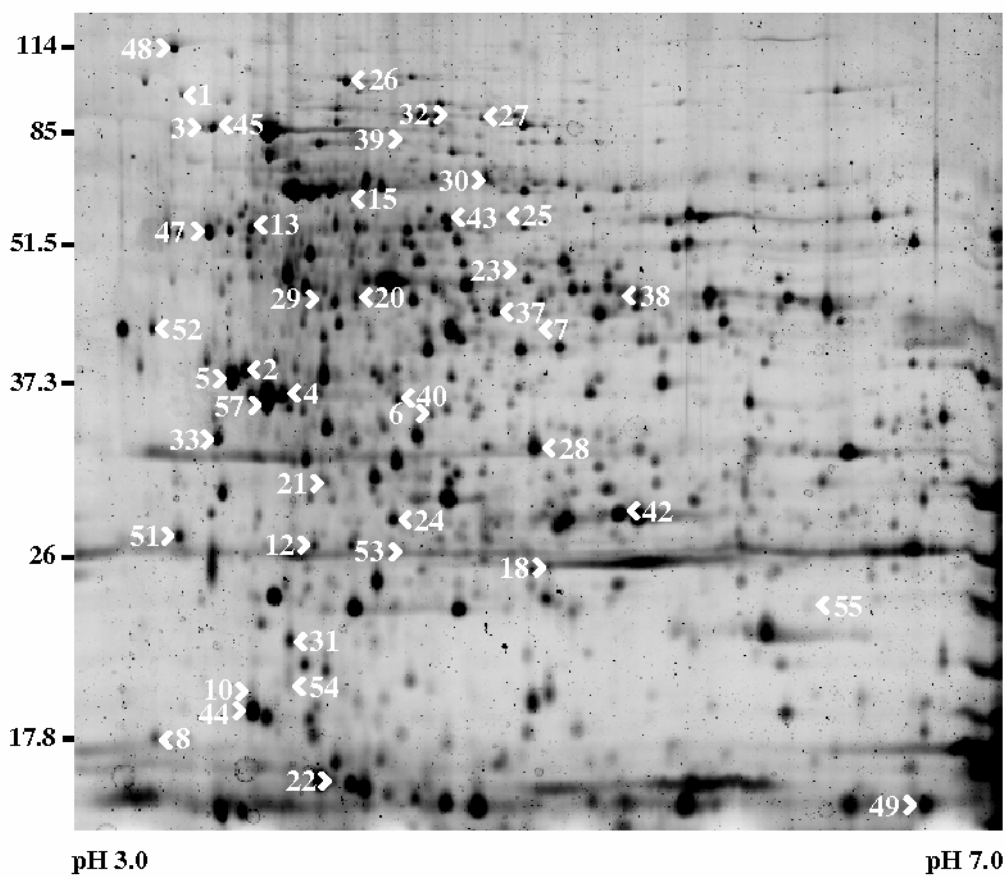
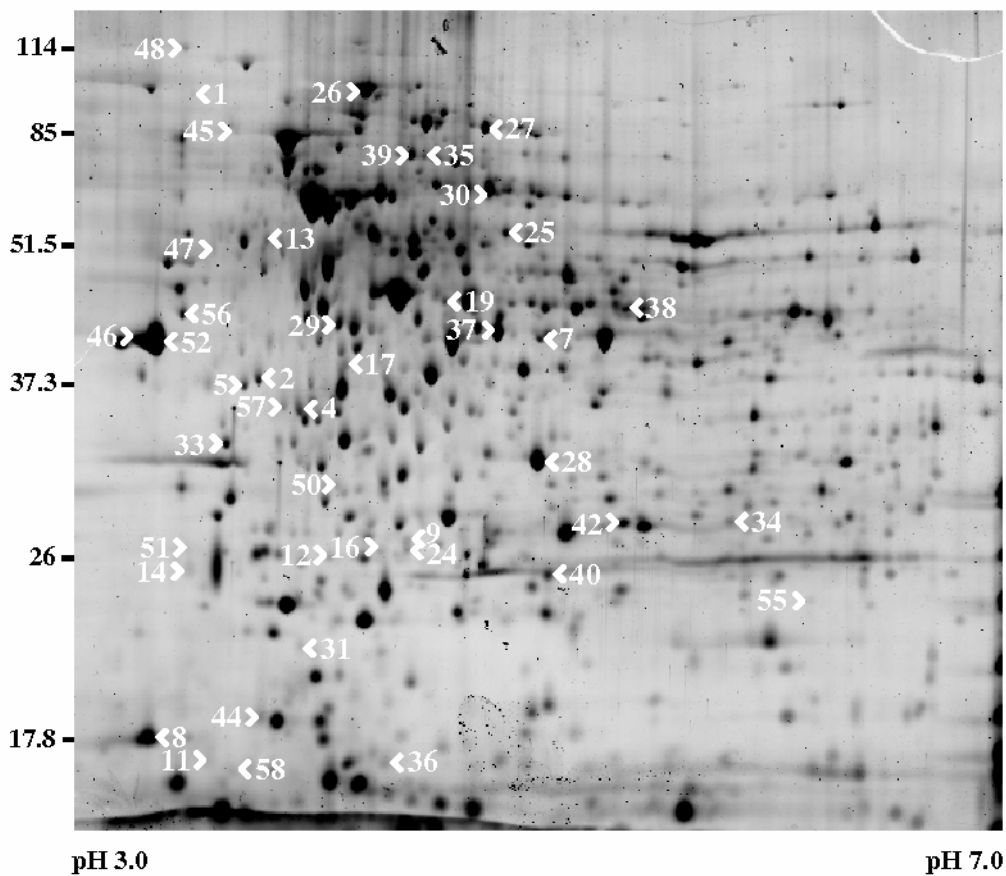


Fig. 1

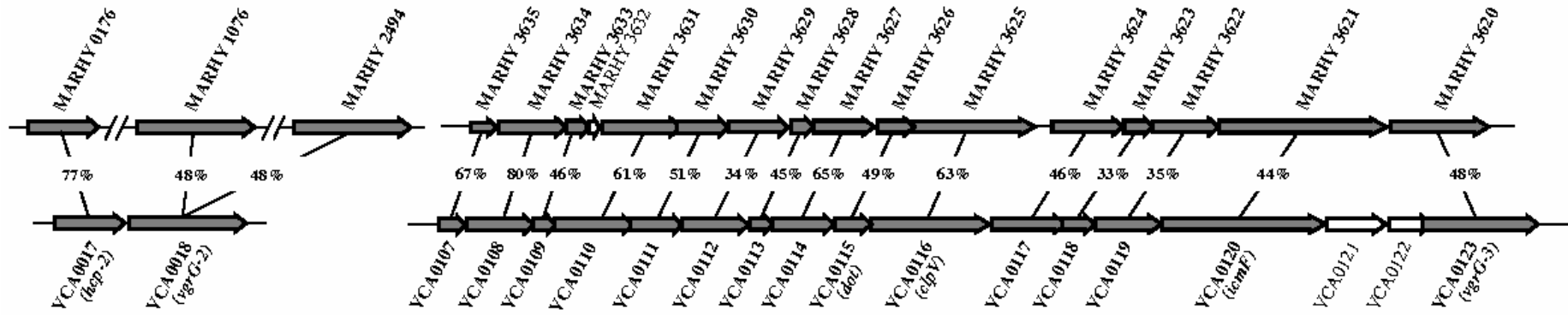


Fig. 2

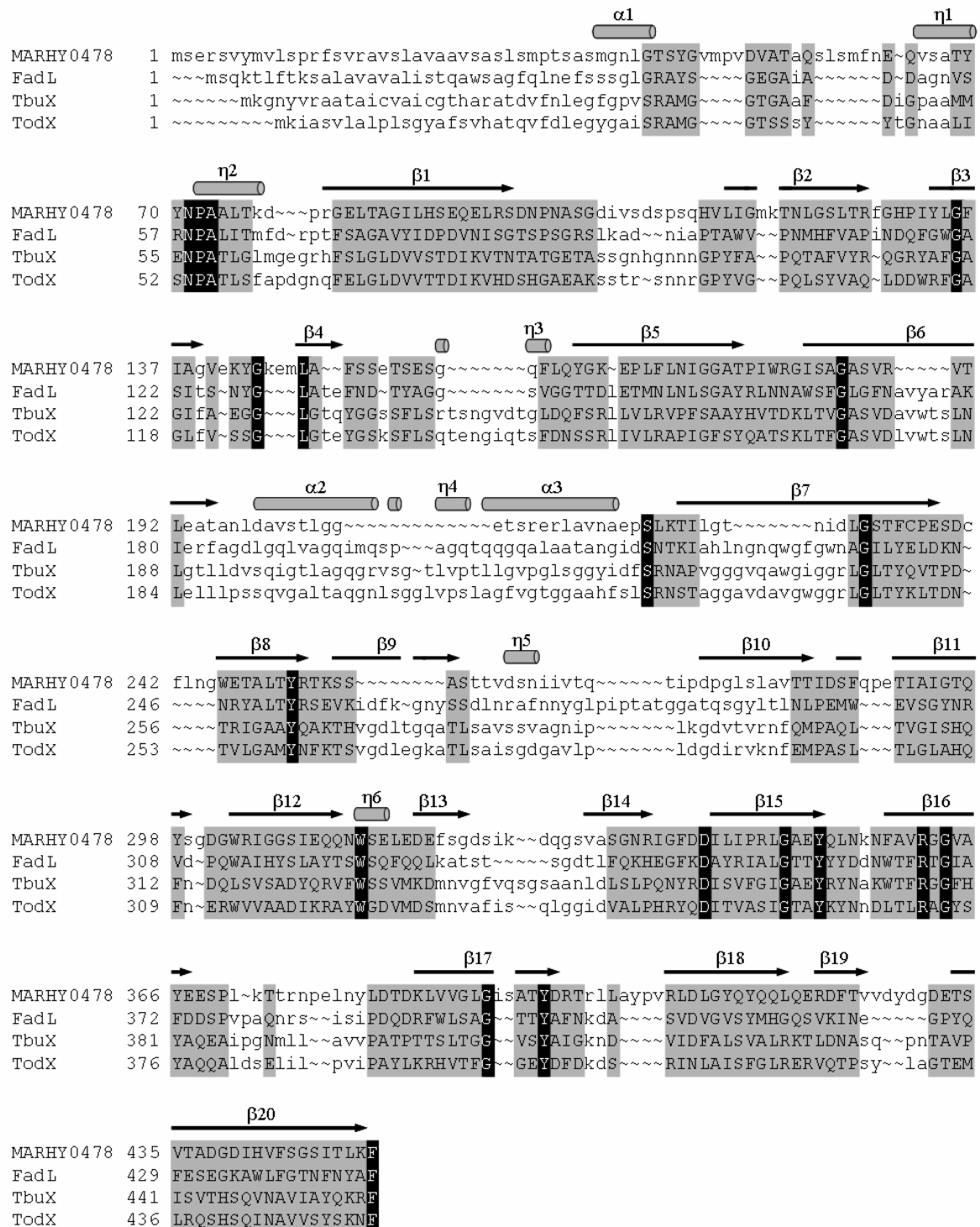


Fig. 3