The polyhydroxyalkanoate genes of a stress resistant Antarctic 
Pseudomonas are situated within a genomic island

Nicolás D. Ayub, M. Julia Pettinari, Beatriz S. Méndez, Nancy I. López *

Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria,
1428 Buenos Aires, Argentina

Received 26 February 2007, revised 2 May 2007
Available online 16 July 2007
Communicated by A. Mark Osborn

Abstract

Pseudomonas sp. 14-3 is an Antarctic bacterium that shows high stress resistance in association with high polyhydroxybutyrate (PHB) production. In this paper genes involved in PHB biosynthesis (phaRBAC) were found within a genomic island named pha-GI. Numerous mobile elements or proteins associated with them, such as an integrase, insertion sequences, a bacterial group II intron, a complete Type I protein secretion system and IncP plasmid-related proteins were detected among the 28 ORFs identified in this large genetic element (32.3 kb). The G+C distribution was not homogeneous, likely reflecting a mosaic structure that contains regions from diverse origins. pha-GI has strong similarities with genomic islands found in diverse Proteobacteria, including Burkholderiales species and Azotobacter vinelandii. The G+C content, phylogeny inference and codon usage analysis showed that the phaBAC cluster itself has a complex mosaic structure and indicated that the phaB and phaC genes were acquired by horizontal transfer, probably derived from Burkholderiales. These results describe for the first time a pha cluster located within a genomic island, and suggest that horizontal transfer of pha genes is a mechanism of adaptability to stress conditions such as those found in the extreme Antarctic environment.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Polyhydroxyalkanoates; Genomic islands; Pseudomonas; Antarctica

1. Introduction

Polyhydroxyalkanoates (PHAs) are accumulated in several bacterial species as intracellular carbon and energy storage compounds under unbalanced growth conditions (Madison and Huisman, 1999). Accumulation and degradation of PHAs endow bacteria with enhanced survival, competition abilities and stress tolerance, increasing fitness in changing environments (Kadouri et al., 2005; López et al., 1995; Ruiz et al., 2001, 2004).

Polyhydroxybutyrate (PHB) is the most common PHA, and in the majority of the microorganisms that accumulate this polymer, its biosynthesis involves three enzymes: a β-ketothiolase (phaA), which condenses two acetyl-CoA into acetoacetil-CoA, a NADPH or NADH dependent reductase
(phaB), that reduces acetoacetyl-CoA into D(-)-3-hydroxybutyryl CoA, and a PHA synthase (phaC), that uses this short length monomer (C4) as a substrate for polymerization (Steinbüchel and Hein, 2001). PHAs can also be synthesized through de novo fatty acid biosynthesis and β-oxidation pathways from sugars and fatty acids (Aldor and Keasling, 2003). PHA synthases play a key role in the production of PHAs. According to their substrate specificity, they can be divided in two functional groups: classI PHA synthases, that include PHB synthases which preferentially use short length monomers (C3–C5), and classII PHA synthases, which preferentially use medium length subunits (C6–C14). Most of the PHA producing Pseudomonas has two classII PHA synthases, called PhaC1 and PhaC2, involved in medium chain length PHA production (Rehm, 2003).

Whereas many Pseudomonas species have been found to accumulate medium chain length PHAs, the inability to produce PHB by the Pseudomonas sensu stricto species has been proposed as an important taxonomic value (Kessler and Palleroni, 2000). However, classI PHA synthases have been found in some Pseudomonas strains (Ayub et al., 2006; Matsusaki et al., 1998; Solaiman and Ashby, 2005). Pseudomonas sp. 14-3 is a highly stress resistant bacterial strain isolated from Antarctic environments that, unlike other PHB producers, is able to synthesize PHB from octanoate but not from glucose (Ayub et al., 2004). Genetic analysis has demonstrated that pha genes in Pseudomonas sp. 14-3 are organized in a cluster, containing genes phaR, phaB, phaA and phaC, and that impaired PHB production from glucose is due to a defective β-ketothiolase (phaA) gene (Ayub et al., 2006).

The ability of bacteria to adapt to new environments frequently results from the acquisition of genes through horizontal transfer (Frost et al., 2005), a process that entails the incorporation into a genome of blocks of DNA with signatures of mobile genetic elements transferred from other organisms (Hacker and Carniel, 2001) often giving rise to genomic islands (GIs). GIs can be 10 kb or longer in size, and represent mosaic-like structures, because they contain elements from diverse origins. GIs are often flanked by small direct repeats, carry various mobility genes, such as integrases and transposases, and genes that can increase the adaptability and versatility of the bacterium (Dobrindt et al., 2004). GIs and large plasmids are frequently found associated to host adaptation and xenobiotic degradation in Pseudomonas strains (Gaillard et al., 2006; He et al., 2004; Ma et al., 2006; Pitman et al., 2005).

pha genes have been shown to enhance fitness and survival (Kadouri et al., 2005; López et al., 1995; Ruiz et al., 2001, 2004), so acquisition of these genes by horizontal transfer could provide recipients with advantages under unfavorable conditions. In recent studies, probable horizontal transfer events were reported for PHA proteins based on the construction of phylogenetic trees (Kadouri et al., 2005; Kalia et al., 2007). Physiological and genetic studies of Pseudomonas sp. 14-3 have shown an increased stress resistance associated with its high and uncommon PHB accumulation capability (Ayub et al., 2004, 2006). Phylogenetic analysis of its pha genes suggested that they could have been acquired by horizontal transfer. Additionally, in previous work performed in our laboratory an association between putative polyhydroxybutyrate regulatory genes and insertion sequence-like elements was reported in Azotobacter sp. FA8 (Pettinari et al., 2003) when a complete IS and a truncated one were detected in the pha region of this microorganism. In the present work, a complete analysis of the pha genetic region of Pseudomonas sp. 14-3 revealed that pha genes are included in a large genomic island element, giving further support to the hypothesis that these genes were acquired by horizontal transfer as a mechanism of adaptability to changing environments.

2. Materials and methods

2.1. Cloning and molecular analysis of the region around pha genes of Pseudomonas sp. 14-3

Two recombinant plasmids, pCT1377 and pCT1283, obtained from a genomic library of Pseudomonas sp. 14-3 containing the pha genes (Ayub et al., 2006) were used to analyze the genes present in the flanking regions. Both plasmids were digested with XhoI and/or EcoRI, subcloned into pBBR1MCS-2 and sequenced. Additional fragments for sequencing were obtained by PCR amplification. The sequencing reactions were performed by Macrogen Inc. (Korea). Sequences were aligned, assembled and analyzed using Bioedit Sequence Alignment Editor CAP (Hall, 1999). ORFs and bacterial operons were predicted by using GeneMark.hmm for Prokaryotes Version 2.4 (Lukashin and Borodovsky, 1998) and FGENESB (http://www.softberry.com). ORFs were compared with the sequences in GeneBank using the BLAST tools (http://www.ncbi.nlm.nih.gov/blast). We
scanned tRNA genes using tRNAscan-SE program (Lowe and Eddy, 1997). The G + C content variation was calculated from a 500-bp window moved along the sequence by 10-pb steps by EMBOSS FREAK (http://emboss.bioinformatics.nl/cgi-bin/emboss/freak). The Relative Synonymous Codon Usage (RSCU) distances between genes were obtained by using General Codon Usage Analysis (McInerney, 1998).

2.2. Phylogenetic analysis of sequence data

Sequence search was performed using BLASTP tools. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.0 (Kumar et al., 2004). Protein sequences were aligned using the ClustalW program. Phylogenetic trees were constructed using the neighbor-joining (NJ) method with genetic distances computed using p-distance model and bootstrap analysis of 1000 resamples and root on midpoint.

2.3. Nucleotide sequence accession number

The nucleotide sequences obtained here have been deposited in the EMBL Nucleotide Sequence Database Accession No.: AM262984.

3. Results and discussion

3.1. phaRBAC is located within a genomic island in Pseudomonas sp. 14-3

The phaRBAC cluster from Pseudomonas sp. 14-3 was found within a large genetic element (32,306 bp) that we termed pha-GI (Pseudomonas sp. 14-3 genomic island containing pha genes). The pha-GI was flanked by an 8 bp (5'-TTTTTTGA-3') direct repeat (DR), which might be part of its site-specific recombination site (Fig. 1).

The average percent G + C content for pha-GI was 58.8% (Fig. 1), close to the average for the genus Pseudomonas (58–66%), but it was not uniform, with values for individual ORFs varying between 44% and 68% (Table 1). The heterogeneity of this region was also observed within the pha gene cluster. Genes phaR, phaB and phaC had an average of 57% while the value for phaA was 67%, suggesting that they have different origins.

Analysis of pha-GI allowed the identification of 28 ORFs that showed similarity to gene products from several bacterial species (Table 1). Functional assignment by comparison with previously described proteins was possible for 23 of them (Table 1). An integrase (int) belonging to a family containing mainly phage integrases, was found at the left end of the pha-GI, followed by parA, a gene involved in the maintenance of elements after transfer, and traF, generally associated with DNA conjugal transfer. ORF7, similar to a sodium/glutamate symporter (gltS), was found next to the PHA cluster. Downstream from the phaBAC cluster and in the same orientation, a complete Type I secretion system was observed. These systems normally consist of an ATP-binding cassette (ABC) transporter protein, located within the inner membrane, a periplasmic protein and an outer membrane protein that form the secretion pore (Schmidt and Hensel, 2004). This kind of system is considered a defense mechanism which exports specific proteins such as toxins or proteases (Andersen et al., 2001). We searched for ORFs containing a glycine rich repeat characteristic of proteins secreted by Type I systems, but we could not find any within pha-GI. The secretion system found in Pseudomonas sp.

![Fig. 1](image-url)
14-3 has the same organization observed in Azotobacter vinelandii (Table 1) and is composed of 4 ORFs (Fig. 1). It has two inner membrane proteins (ORFs 13 and 14), a periplasmic protein (ORF 12) and an outer membrane protein (ORF 15). ORF12 showed high amino acid similarity (62–63% identity) with Type I antifreeze proteins belonging to the HlyD family secretion proteins (CAD85027; ABA53314; ABC35065). This ORF presents a high content of alanine residues (16.94%), characteristic of Type I antifreeze proteins (Zhang and Laursen, 1998).

Following the Type I system and in the same orientation two ORFs corresponding to a lipoprotein (ORF16) and an arginine and ornithine antiporter (arcD) were observed (Table 1). The right side of pha-GI contained a high percentage of mobile elements or proteins associated with them (Fig. 1). A complete transposase (ORF18) similar to transposases encoded by insertion sequences of the IS family (Table 1) and several incomplete transposases were observed: ORF19 belonging to the IS21 family and two ORFs, 23 and 24, belonging to the IS4 family. We did not detect insertion sequence-like elements belonging to IS3 and IS630 families, as those found in Azotobacter sp. FA8 associated with pha genes (Pettinari et al., 2003).

Another mobile element present was a bacterial group II intron. Full-length group II introns, as well as intron fragments, are often located either within plasmids, IS elements or pathogenicity islands (Dai and Zimmerly, 2002). The intron found in Pseudomonas 14-3 contained a non-coding region (422 bp) highly similar to a group II intron-encoding maturase of P. putida KT2440 (AAN66260) and an incomplete ORF similar to a reverse transcriptase found in A. vinelandii (ORF22-rvt) that is also truncated in a similar position (Table 1).

### Table 1: Localization and annotation of open reading frames (ORFs) of pha-GI from Pseudomonas sp. 14-3

<table>
<thead>
<tr>
<th>ORF No.</th>
<th>Gene name</th>
<th>G + C (%)</th>
<th>Putative product</th>
<th>Strain</th>
<th>Accession No.</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>int</td>
<td>63</td>
<td>Integrase</td>
<td>A. vinelandii</td>
<td>EAM08320</td>
<td>75</td>
</tr>
<tr>
<td>02</td>
<td>parA</td>
<td>68</td>
<td>Plasmid partition protein</td>
<td>C. oxalatus</td>
<td>CAD61131</td>
<td>81</td>
</tr>
<tr>
<td>03</td>
<td>traF</td>
<td>66</td>
<td>Plasmid conjugal protein</td>
<td>C. oxalatus</td>
<td>CAD61134</td>
<td>76</td>
</tr>
<tr>
<td>04</td>
<td></td>
<td>64</td>
<td>Hypothetical protein</td>
<td>X. campestris</td>
<td>AE012175</td>
<td>48</td>
</tr>
<tr>
<td>05</td>
<td></td>
<td>44</td>
<td>Hypothetical protein</td>
<td>S. enterica</td>
<td>AAX68249</td>
<td>29</td>
</tr>
<tr>
<td>06</td>
<td></td>
<td>52</td>
<td>Hypothetical protein</td>
<td>Mycobacterium</td>
<td>EAT01112</td>
<td>35</td>
</tr>
<tr>
<td>07</td>
<td>gltS</td>
<td>56</td>
<td>Sodium/glutamate symporter</td>
<td>R. palustris</td>
<td>CAE27703</td>
<td>53</td>
</tr>
<tr>
<td>08</td>
<td>phaR</td>
<td>57</td>
<td>Transcriptional regulator</td>
<td>A. vinelandii</td>
<td>AAK72595</td>
<td>42</td>
</tr>
<tr>
<td>09</td>
<td>phaB</td>
<td>57</td>
<td>Acetoacetyl-CoA reductase</td>
<td>A. vinelandii</td>
<td>AAK72596</td>
<td>70</td>
</tr>
<tr>
<td>10</td>
<td>phaA</td>
<td>67</td>
<td>β-ketothiolase</td>
<td>A. vinelandii</td>
<td>AAF82771</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>phaC</td>
<td>58</td>
<td>Polyhydroxybutyrate synthase</td>
<td>A. vinelandii</td>
<td>AAK72597</td>
<td>68</td>
</tr>
<tr>
<td>12</td>
<td>hlyD</td>
<td>64</td>
<td>Secretion protein</td>
<td>A. vinelandii</td>
<td>EAM07646</td>
<td>65</td>
</tr>
<tr>
<td>13</td>
<td>ABC-1</td>
<td>63</td>
<td>ABC transporter</td>
<td>A. vinelandii</td>
<td>EAM07645</td>
<td>75</td>
</tr>
<tr>
<td>14</td>
<td>ABC-2</td>
<td>62</td>
<td>ABC-2 transporter</td>
<td>A. vinelandii</td>
<td>EAM07644</td>
<td>74</td>
</tr>
<tr>
<td>15</td>
<td>tolC</td>
<td>62</td>
<td>Outer membrane efflux protein</td>
<td>P. syringae</td>
<td>AA058405</td>
<td>74</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>47</td>
<td>Lipoprotein</td>
<td>P. entomophila</td>
<td>CAK17182</td>
<td>68</td>
</tr>
<tr>
<td>17</td>
<td>arcD</td>
<td>61</td>
<td>Arginine/ornithine antiporter</td>
<td>P. putida</td>
<td>AAN66627</td>
<td>74</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>63</td>
<td>Transposase (IS4 family)</td>
<td>Azohacrus</td>
<td>CAI10600</td>
<td>80</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>64</td>
<td>Transposase (IS21 family)</td>
<td>P. alcaligenes</td>
<td>AAB48471</td>
<td>98</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>55</td>
<td>Transcriptional regulator</td>
<td>C. oxalatus</td>
<td>CAD61126</td>
<td>56</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>55</td>
<td>Conserved hypothetical protein</td>
<td>M. aquaeolii</td>
<td>EA099450</td>
<td>46</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>59</td>
<td>Reverse transcriptase (Intron)</td>
<td>A. vinelandii</td>
<td>EAM05016</td>
<td>77</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>52</td>
<td>Transposase (IS4 family)</td>
<td>P. putida</td>
<td>AAN67548</td>
<td>72</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>58</td>
<td>Transposase (IS4 family)</td>
<td>P. aeruginosa</td>
<td>AA32836</td>
<td>91</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>53</td>
<td>Hypothetical protein</td>
<td>R. etli</td>
<td>AAM55044</td>
<td>35</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>53</td>
<td>Virulence factor</td>
<td>V. cholerae</td>
<td>AAL8651</td>
<td>28</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>55</td>
<td>ATP-dependent DNA helicase</td>
<td>C. oxalatus</td>
<td>CAD61115</td>
<td>64</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>55</td>
<td>ATP-dependent endonuclease</td>
<td>C. oxalatus</td>
<td>CAD61116</td>
<td>70</td>
</tr>
</tbody>
</table>

3.2. Relationship of pha-GI to other genomic islands

The deduced amino acid sequences of several ORFs found in pha-GI showed high similarity to proteins located within genomic islands belonging to Burkholderiales species (Tn4371, found in Cupriavidus oxalaticus, GIs found in Ralstonia solanacearum and Cupriavidus metallidurans), A. vinelandii and Pseudomonas species (He et al., 2004; Pitman...
et al., 2005; Toussaint et al., 2003). These ORFs are the integrase (ORF1), parA, traF, ORF20 (transcriptional regulator), ORF27 (helicase) and ORF28 (nuclease) (Table 1). ParA, TraF, and transcriptional regulators such as the one coded by ORF20 are characteristic of IncP plasmids.

As an approach to determine the ends of pha-GI, we searched for tRNA genes, as these genes are frequently found adjacent to genomic islands (Schmidt and Hensel, 2004), but were unable to find any. However, we observed a DR-spacer-int motif at the left end. These motifs, normally situated at one end of genomic islands, are thought to be involved in the excision and integration of these elements (Pitman et al., 2005; Toussaint et al., 2003). The alignment of the 181 nucleotide sequence corresponding to the spacer region of pha-GI showed high similarity to the left ends of related GIs found in Burkholderiales and A. vinelandii (Fig. 2). The highest identity (96%) was found for A. vinelandii. Furthermore, experimental evidence has demonstrated that transposition of Tn4371 from C. oxalaticus involves an 8 bp motif (5′-TTTTTTCAT-3′) associated with the excision/integration process (Merlin et al., 1999). The eight bp-left DR from pha-GI and Tn4371 are located at the same position upstream relative to the start codon for int (Fig. 2).

The Proteobacteria species harboring broad-host-range genomic islands strongly similar to pha-GI also have PHB biosynthetic clusters: phaCAB in Burkholderiales and phaRBAC in A. vinelandii. This suggests that these gene clusters could have been transferred between different groups of Proteobacteria. However, the pha genes of the above mentioned microorganisms are not located inside their known genomic islands. In addition, analysis of the regions flanking pha genes in available bacterial genomes showed that these genes were not included in genomic islands. The genetic element presented in this paper constitutes the first description of a PHA cluster located within a genomic island.

3.3. Phylogenetic analysis supports horizontal transfer of pha genes

To investigate the possible horizontal transfer of pha genes, we analyzed the phylogenetic relationships of the proteins corresponding to the pha genes found in pha-GI (Fig. 3). In previous phylogenetic analysis of PHA proteins, trees constructed using proteins from distant taxa were congruent with 16S rRNA data only for individual clusters (Kadouri et al., 2005; Kalia et al., 2007; Rehm, 2003; Steinbüchel and Hein, 2001). The analysis performed in this work was restricted to well-characterized species belonging to the group Proteobacteria. The phylogenetic tree for PhaAs from Proteobacteria showed a clear affinity among thiolase proteins of Pseudomonadales strains, including thiolases of Pseudomonas sp. 14-3 and A. vinelandii, belonging to the phaBAC cluster, and the thiolase from P. putida KT2440 (Fig. 3). This last protein has been experimentally shown to participate in PHB biosynthesis in heterologous complementation analysis (Ayub et al., 2006), even when P. putida KT2440 does not accumulate PHB. These results strongly support the suggestion that the PhaA found in pha-GI has a Pseudomonadales origin. Contrarily, an unexpected clustering was observed for Pseudomonas sp. 14-3 within the Burkholderiales in the PhaB tree (Fig. 3). The PhaC tree shows that classI and classII PHA synthases of Proteobacteria form two distinct groups, support-

Fig. 2. Alignment of the left ends of pha-GI from Pseudomonas sp. 14-3 and related genomic islands of Azobacter vinelandii OP (Avin), Cupriavidus oxalaticus (Tn4371), Ralstonia solanacearum GMI1000 (TnRso) and Cupriavidus metallidurans CH34 (RCH4). These regions are located on the left side of the integrase genes. The left 8 bp direct repeat from pha-GI and Tn4371, and the beginning to int genes are marked.
Fig. 3. Phylogenetic relationships between representative taxa of the phylum Proteobacteria based on neighbor-joining (NJ) analysis of PhaA, PhaB and PhaC proteins. The α, β and γ subclasses of the Proteobacteria and functional classification (class I and class II) of synthases are shown. Arrows indicate proteins that fall within groups not corresponding to their 16S rRNA derived phylogenetic affiliation. Bootstrap percentages are indicated at the branch points. In all of the cases, tree topologies obtained using NJ method, Minimum evolution and Maximum parsimony methods were identical.
ing the functional classification (Fig. 3). The subtree for the classI PhaC was similar to the PhaB tree, and in both trees the proteins from *Pseudomonas* sp. 14-3 clustered within the *Burkholderiales*, suggesting that the corresponding genes were acquired by horizontal transfer. The incongruence observed between protein and 16S rRNA trees is not exclusive for this species: PhaB and PhaC of *A. vinelandii* also cluster within *Burkholderiales*, and PhaC from *Rickettsiales* does not form a group with the rest of *α-Proteobacteria*, but was found to be associated with *Legionellales* (γ-Proteobacteria) with a high bootstrap value (78%) (Fig. 3). Recent work has determined that classII PHA synthases can also be spread by horizontal gene transfer (Ciesielski et al., 2006).
To further investigate if the phaBAC genes had different phylogenetic affinities, we analyzed the extent of divergence in codon usage. We calculated the relative synonymous codon usage (RSCU) distances among genes belonging to phaBAC and Type I secretion systems. The distance matrix was used as input for program MEGA 3.0, and a dendrogram was produced by the UPGMA method. The resulting tree had two main clusters, one for phaB and phaC genes and the other for the phaA and Type I secretion system genes (Fig. 4). This type of distribution of the pha genes further supports the hypothesis that Pseudomonas sp. 14-3 pha cluster is the result of an assemblage of pha genes acquired from different origins. The same can be concluded about the A. vinelandii pha cluster, that has an identical genetic organization and similar origins for phaB, phaA and phaC.

4. Concluding remarks

The association of pha genes with fitness has been proposed almost since their discovery, but until now these genes had not been included among genes normally found in fitness islands, such as those related to the degradation of xenobiotics, or symbiotic nitrogen fixation. In this work, we identified a large genomic island (pha-GI) containing pha genes. Our results indicate that these genes, that confer high stress resistance to Pseudomonas sp. 14-3, and might contribute to survival in the specific extreme conditions encountered in Antarctica, were acquired by horizontal transfer, suggesting that horizontal transfer of pha genes is a mechanism of adaptability to changing environments.

Acknowledgments

This work was supported by grants from UBA and CONICET. B.S.M., M.J.P. and N.I.L. are career investigators from CONICET. N.D.A has a graduate student fellowship from CONICET.

References


Merlin, C., Springael, D., Toussaint, A., 1999. Tn 4371: a modular structure encoding a phage-like integrase, a Pseudo-


