Metabolism of Poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas oleovorans. Identification and Sequences of Genes and Function of the Encoded Proteins in the Synthesis and Degradation of PHA

Huisman, Gjalt W.; Wonink, Eric; Meima, Rob; Kazemier, Bert; Terpstra, Peter; Witholt, Bernard

Published in:
The Journal of Biological Chemistry

Publication date:
1991

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 4, Issue of February 5, pp. 2191-2198, 1991
Printed in U.S.A.

Metabolism of Poly(3-hydroxyalkanoates) (PHAs) by *Pseudomonas oleovorans*

IDENTIFICATION AND SEQUENCES OF GENES AND FUNCTION OF THE ENCODED PROTEINS IN THE SYNTHESIS AND DEGRADATION OF PHA*

(Received for publication, May 24, 1990)

Gjalt W. Huisman, Eric Wonink, Rob Meima, Bert Kazemier, Peter Terpstra, and Bernard Witholt‡

From the Department of Biochemistry, Groningen Biotechnology Center, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Pseudomonas oleovorans **accumulates poly(3-hydroxyalkanoates) (PHAs) after growth on medium chain length hydrocarbons. Large amounts of this polyester are synthesized when cells are grown under nitrogen-limiting conditions. When nitrogen is resupplied in the medium, the accumulated PHA is degraded. In this paper, we describe mutants which are defective in the synthesis or in the degradation of PHA. These mutants were used to select DNA fragments which encode PHA polymerases and a PHA depolymerase. A 25-kilobase (kb) DNA fragment was isolated from *P. oleovorans* that complements a *Pseudomonas putida* mutant unable to accumulate PHA. Subcloning resulted in the assignment of a 6.4-kb *Eco*RI fragment as the *pha* locus, containing genetic information for PHA synthesis. Mutants in the PHA degradation pathway were also complemented by this fragment, indicating that genes encoding PHA biosynthetic and degradative enzymes are clustered. Analysis of the DNA sequence of the 6.4-kb fragment revealed the presence of two open reading frames encoding PHA polymerases based on homology to the poly(3-hydroxybutyrate) polymerase from *Alcaligenes eutrophus*. A third open reading frame complemented the PHA degradation mutation and is likely to encode a PHA depolymerase. The presence of two PHA polymerases is due to a 2098-base pair DNA duplication. The PHA polymerases are 53% identical and show 35–40% identity to the poly(3-hydroxybutyrate) polymerase. No clear difference in specificity was found for the PHA polymerases. However, with the *pha* locus cloned on a multicopy vector, a polymer was accumulated that contains a significantly higher amount of substrate-derived monomers. An increase in the rate of polyester synthesis *versus* oxidation of the monomers in the β-oxidation explains these findings.**

Poly(3-hydroxybutyrate) (PHB)[1] is a common reserve material in both Gram-positive and Gram-negative bacteria. It serves as a carbon and/or electron sink when optimal growth conditions are not met (Dawes and Senior, 1973). Besides its physiological role for its host, this material has recently become of interest as a useful plastic for mankind (Byrom, 1987; King, 1982; Pool, 1989).

PHB is a member of the class of poly(3-hydroxyalkanoates) (PHAs). These polymers are composed of 3-hydroxy fatty acids. Until recently, the only PHAs found were formed under nonspecific growth conditions (Capon *et al.*, 1983; Findlay and White, 1983; Wallen and Rohwedder, 1989). However, recently, the synthesis by *Pseudomonas oleovorans* of a range of PHAs composed of medium chain length 3-hydroxy fatty acids was described (Brandl *et al.*, 1988; Lageveen *et al.*, 1988; Preusting *et al.*, 1990). Further studies revealed that PHAs are also formed by fluorescent pseudomonads like *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, and *Pseudomonas lemonnieri* when grown on fatty acids as the carbon and energy source (Haywood *et al.*, 1989; Huisman *et al.*, 1989).

The metabolic pathways of PHB synthesis and degradation are known. The different enzymes have all been purified (for a review, see Tomita *et al.* (1983)), and the genes encoding the biosynthetic enzymes in *Alcaligenes eutrophus* (Peoples and Sinskey, 1989b, 1989c; Schubert *et al.*, 1988; Slater *et al.*, 1988) and *Zoogloea ramigera* (Peoples *et al.*, 1987; Peoples and Sinskey, 1989a) have been cloned and characterized.

The pathways resulting in PHA synthesis and degradation have not yet been resolved. It has been proposed (Lageveen *et al.*, 1988) that PHA monomers are derived from (*S*)-3-hydroxyacyl-CoAs, which are intermediates in the fatty acid oxidation cycle. After inversion of the configuration at the asymmetric carbon atom, the resulting (*R*)-3-hydroxyacyl-CoAs are polymerized.

We initially believed that the relative incorporation of different fatty acids into PHA reflects the availability of these fatty acids or their precursor alkanes in the growth medium of *P. oleovorans* (Lageveen *et al.*, 1988). We have since studied the specificity of PHA synthesis in more detail and found this to be true only for C_6 to C_{12} fatty acids. When long chain fatty acids are supplied in the growth medium, these are taken up by *P. oleovorans* and degraded to the appropriate size range before incorporation into PHA (Huisman *et al.*, 1989). Thus the incorporation of fatty acid derivatives into PHA is restricted by the specificity of a key enzyme in PHA biosynthesis.

To further investigate the mechanism of PHA synthesis, we have cloned some of the genes which encode PHA biosynthetic enzymes from *P. oleovorans*. In this paper, we describe a 6.4-kb *Eco*RI fragment that complements mutants in the synthesis and degradation of PHA. The fragment has been

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M58445.

‡ To whom correspondence should be addressed.

[1] The abbreviations used are: PHB, poly(3-hydroxybutyrate); PHAs, poly(3-hydroxyalkanoates); kb, kilobase(s); bp, base pair(s); ORF, open reading frame; IR, inverted repeat.

TABLE I

Bacterial strains and plasmids used in this study

Strain/plasmid	Relevant genotype/phenotype[a]	Source or ref.
E. coli		
DH1	*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17, *sup*E44, *rel*A1	Maniatis *et al.* (1982)
JM109	*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17, *sup*E44, *rel*A1, (*lac-pro*AB), {F′*tra*D36, *pro*AB, *lac*I^{q}Z M15}	Maniatis *et al.* (1982)
P. oleovorans		
GPo1	OCT plasmid, PHA$^+$	Schwartz and McCoy (1973)
GPo500	NTG mutant of GPo1, no PHA degradation	This study
P. putida		
KT2442	mt2 cured of TOL, Rf^r, PHA$^+$	Bagdasarian *et al.* (1981)
GPp104	NTG mutant of KT2442, PHA$^-$	This study
Plasmids		
pLAFR1	Tc, Mob$^+$, RK2 replicon	Friedman *et al.* (1982)
pRK2013	Km, Tra$^+$, ColE1 replicon	Ditta *et al.* (1980)
pJRD215	Km, Sm, RSF1010, and pACYC177 replicon, Mob$^+$	Davison *et al.* (1987)
pGEM-7Zf(+)	Ap, *lac*I^q, T7 promoter	Promega Biotec
pGEM-5Zf(+)	Ap, *lac*I^q, T7 promoter	Promega Biotec
pGEc400	PHA$^+$, pLAFR1	This study
pGEc401	PHA$^+$, pLAFR1	This study
pGEc404	PHA$^+$, pJRD215	This study
pGEc420	3.3-kb *Eco*RI-*Cla*I fragment in pGEM-7Zf(+)	This study
pGEc422	3.1-kb *Eco*RI-*Cla*I fragment in pGEM-7Zf(+)	This study
pGEc424	6.4-kb *Eco*RI-*Eco*RI fragment in pGEM-7Zf(+)	This study
pGEc425	2.2-kb *Sal*I-*Sal*I fragment in pGEM-5Zf(+)	This study
pGEc426	2.2-kb *Sal*I-*Sal*I fragment in pGEM-5Zf(+)	This study
pGEc427	2.4-kb *Sal*I-*Sal*I fragment in pGEM-5Zf(+)	This study
pGEc428	2.4-kb *Sal*I-*Sal*I fragment in pGEM-5Zf(+)	This study
pGEc432	pGEc422 in pJRD215	This study
pGEc434	pGEc424 in pJRD215	This study
pGEc435	pGEc425 in pJRD215	This study
pGEc436	pGEc426 in pJRD215	This study
pGEc437	pGEc427 in pJRD215	This study
pGEc438	pGEc428 in pJRD215	This study

[a] Rf, rifampicin; Tc, tetracycline; Km, kanamycin; Sm, streptomycin; Ap, ampicillin.

sequenced and subcloned to determine the functions of the different open reading frames. Three activities were identified, corresponding to two polymerases and a depolymerase.

Closely related results have recently been obtained by Peoples and Sinskey.[2]

MATERIALS AND METHODS

Bacterial Strains and Plasmids—The *Pseudomonas*, *Alcaligenes*, and *Escherichia coli* strains as well as the plasmids used in this study are listed in Table I.

Isolation and Analysis of Plasmid DNA—Small-scale plasmic isolations were done according to a modified procedure (Ish-Horowitz and Burke, 1981) of the method of Birnboim and Doly (1979). Large-scale preparations were done as described by Maniatis *et al.* (1982). Plasmids were digested with restriction endonucleases under the manufacturer's conditions and analyzed by electrophoresis on 0.7–1.5% agarose gels in Tris-acetate buffer. DNA fragments were isolated using the GeneClean method (BIO 101 Inc., La Jolla, CA). Where relevant, restricted DNA was ligated as described by Maniatis *et al.* (1982). Restriction endonucleases were extracted with phenol/chloroform prior to ligation.

Transformation—Preparation of competent *E. coli* cells and transformation were done according to the polyethylene glycol method described by Chung *et al.* (1989).

Conjugations—Mating of the recipient strains with the *E. coli* donor strains was achieved using the helper plasmid pRK2013 (Ditta *et al.*, 1980). Transconjugants were isolated on minimal medium with the appropriate antibiotic (tetracycline, 12.5 μg/ml; kanamycin, 50 μg/ml; ampicillin, 50 μg/ml).

Cloning of PHA Biosynthetic Locus—Genes encoding enzymes involved in the synthesis of PHA were isolated by complementation of PHA-free mutants.[3] A *P. oleovorans* gene library (Eggink *et al.*, 1984) was transferred to GPp104 via triparental mating using pRK2013 as helper plasmid. Complementation was tested by screening for colonies that regained a white phenotype when grown on 2-fold diluted E_2 medium (Lageveen *et al.*, 1988) (0.5 N E_2 medium) on 10 mM octanoate. The presence of PHA was confirmed by phase-contrast microscopy.

Identification of PHA Depolymerase Locus—*P. oleovorans* GPo500 is a PHA degradation mutant which cannot degrade PHA.[3] Complementation of *P. oleovorans* GPo500 was determined by growing conjugants on E_2 medium containing 10% of the usual amount of nitrogen source (0.1 N E_2 medium) on 10 mM octanoate. After PHA had accumulated, extra nitrogen was added to 4 times the normal amount, and plates were further incubated at 30 °C. Colonies were subsequently screened for the disappearance of PHA due to the degradation of the previously stored polymer.

Determination of Nucleotide Sequence of PHA Biosynthetic Locus—The nucleotide sequence of the 6.4-kb *Eco*RI-*Eco*RI fragment encoding the PHA biosynthetic locus was determined from a set of unidirectional overlapping deletion clones generated by exonuclease III digestion of clones carrying (parts of) the fragment in the vector pGEM-7Zf(+) according to the method of Henikoff (1984). The two *Eco*RI-*Cla*I fragments were cloned into this vector to give pGEc420 (the 3.3-kb *Eco*RI-*Cla*I fragment as a *Bam*HI-*Hin*dIII fragment from pGEc404) and pGEc422 (the 3.1-kb *Eco*RI-*Cla*I fragment from pGEc401). The complete 6.4-kb fragment was inserted as an *Eco*RI fragment to give pGEc424. The dideoxy sequencing method of Sanger *et al.* (1977) with T7 DNA polymerase was used, replacing GTP with 7-deaza-GTP to reduce the formation of secondary structure during the sequence reactions. The nucleotide sequences of both strands of the 6.4-kb fragment were determined. Amino acid sequences derived from the nucleotide sequence were compared with known sequences in the SWISSPROT (release 12, October 1989) protein library with the program FASTA (Pearson and Lipman, 1988).

RESULTS

In our study, to clone the genes encoding enzymes involved in PHA metabolism, we first isolated mutants in PHA bio-

[2] O. P. Peoples and A. J. Sinskey, personal communication.

[3] G. W. Huisman, E. Wonink, G. de Koning, H. Preusting, P. J. Lemstra, and B. Witholt, manuscript in preparation.

synthesis and PHA degradation. These were subsequently used to identify the *pha* locus.

Mutants in PHA Metabolism—Mutant strains unable to synthesize or degrade PHA were isolated after *N'*-methyl-*N'*-nitro-*N*-nitroguanidine mutagenesis of *P. oleovorans* GPo1 or *Pseudomonas putida* KT2442.[3] Colonies of strains unable to accumulate PHA when grown on 0.1 N E_2 medium containing 10 mM octanoate appear translucent on this medium and do not stain with Sudan Black. After screening 1.5×10^4 colonies, seven mutants in PHA biosynthesis were isolated. One such mutant, *P. putida* GPp104, was used for further complementation studies.

The isolation method for biodegradation mutants involves a first step in which PHA is accumulated under low nitrogen conditions. When excess nitrogen is added to restore growth, there is PHA degradation, and such colonies no longer stain with Sudan Black. Mutants unable to degrade their stored polymer still retain Sudan Black after the addition of nitrogen. One mutant (GPo500) was isolated from 2.5×10^3 colonies by this approach.

Cloning of PHA Biosynthetic Genes—A *P. oleovorans* GPo1 gene library in the broad host-range cosmid vector pLAFR1 (Eggink *et al.*, 1984) was conjugated into *P. putida* GPp104. Out of 400 recombinant clones, one PHA-synthesizing colony was identified. The relevant plasmid (pGEc400) was isolated from the corresponding *E. coli* clone. It consisted of five *Eco*RI fragments comprising 25 kb in total, which were subcloned by *Eco*RI digestion and religation. Of these subclones, pGEc401, consisting of a single 6.4-kb *Eco*RI insert in pLAFR1, was still able to complement the PHA$^-$ mutation of GPp104.

Further subcloning (Fig. 1*A*) was achieved by ligating *Eco*RI-*Cla*I fragments into the broad host range vector pJRD215. Recombinant plasmids were conjugated into *P. putida* GPp104, and eight complementing plasmids were isolated from the corresponding *E. coli* clones. All plasmids contained a 3.3-kb *Eco*RI-*Cla*I insert (pGEc404).

The complementing plasmids described above were also used for complementation studies of the depolymerization-negative mutant *P. oleovorans* GPo500. Degradation of the accumulated PHA was restored with pGEc401, but not with pGEc404. The genetic locus of the depolymerization activity was further determined by subcloning the 3.1-kb *Eco*RI-*Cla*I fragment into pGEM-7Zf(+) (pGEc422) and the two *Sal*I fragments in both orientations into pGEM-5Zf(+) (pGEc425/426 and pGEc427/428) and increasing the host range of these plasmids by ligating them into the *Nsi*I site of pJRD215 (pGEc432, pGEc435/436, and pGEC437/438, respectively).

Plasmids pGEc432, pGEc435, and pGEc436 complemented the mutation in GPo500. Surprisingly, introduction of these plasmids in the PHA$^-$ mutant *P. putida* GPp104 resulted in complementation of the PHA$^-$ phenotype by pGEc432. The other plasmids did not complement this mutation (Fig. 1*A*). It is not clear why the 3.1-kb *Eco*RI-*Cla*I fragment of pGEc432 was not identified earlier as complementing the PHA$^-$ phenotype during subcloning of the 6.4-kb *Eco*RI fragment.

It is clear, however, that the 6.4-kb insert in pGEc401 contains genes involved in PHA biosynthesis and in PHA degradation. The 3.3-kb *Eco*RI-*Cla*I fragment complements only the PHA$^-$ mutant GPp104, whereas the 3.1-kb *Eco*RI-*Cla*I fragment encodes both PHA polymerizing and depolymerizing activities.

Nucleotide Sequence of 6.4-kb EcoRI-EcoRI Fragment—The 6.4-kb *Eco*RI-*Eco*RI fragment (Figs. 1*B* and 2) was found to contain three open reading frames that could code for proteins according to the rules of Fickett (1982) (Fig. 1*C*). ORF1 runs from positions 554 to 2230 and encodes a gene product of 559 amino acids (62.4 kDa). ORF2 starts at position 2297, ends at position 3145, and encodes a 31.5 kDa gene product (283 amino acids). ORF3 begins at one of two ATG codons in the same frame at positions 3217 and 3268 and ends at position 4947. The resulting polypeptide of 560 or 577 amino acids has a calculated molecular mass of 62.6 or 64.5 kDa.

Comparison of ORF1- and ORF3-encoded Amino Acid Sequences to Known Protein Sequences—The proteins encoded by the ORFs identified in the 6.4-kb sequence were compared with known sequences to establish possible functions for these hypothetical proteins. No significant amino acid sequence homology to known proteins was found, except to the recently reported PHB polymerase encoded by the *phb*C gene of *A. eutrophus* (Peoples and Sinskey, 1989c). Based on deduced amino acid sequences, PHB polymerase showed a 37.8% identity in a 511-amino acid overlap with the protein encoded by ORF1 and a 39.5% identity in a 488-amino acid sequence overlap in ORF3 as calculated with the program FASTA (Pearson and Lipman, 1988). These observations suggest that both ORF1 and ORF3 encode PHA polymerases. The three polymerases were aligned with the program CLUSTAL (Higgins and Sharp, 1989) (Fig. 3) and show regions of significant similarity.

The overall identity between the ORF1 and ORF3 gene products is 53.6%. In fact, the entire 6.4-kb sequence contains a 2098-bp duplication (Fig. 4). The alignment of these DNA stretches according to FASTA runs from positions 536 to 2593 and from positions 3251 to 5314 with 38 gaps, all smaller than six nucleotides, and shows 65.4% identity. Although the first sequence includes 300 base pairs that encode the first 100 amino acids of ORF2, the second sequence does not contain a similar fourth open reading frame downstream of ORF3. Based on this duplication, it is assumed that ORF3 starts with the ATG triplet at position 3268. Alignment of the sequences which surround the start codons at positions 3217 and 3268 with the start of the coding region of ORF1

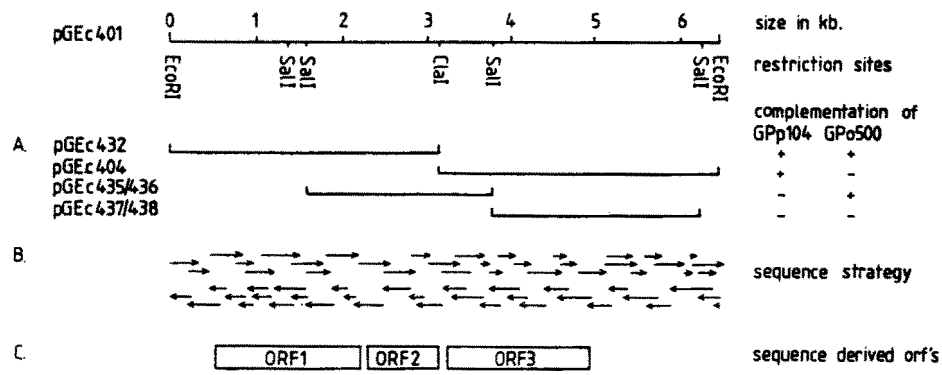

FIG. 1. **Genetic structure of 6.4-kb *Eco*RI fragment that complements mutations in PHA biosynthesis and degradation.** *A*, several subclones were prepared and tested for their ability to complement mutants in PHA synthesis (GPp104) and degradation (GPo500). *B*, sequence strategy for determination of the nucleotide sequence. *C*, after the nucleotide sequence of the entire 6.4-kb fragment was determined, three open reading frames were assigned according to the rules of Fickett (1982).

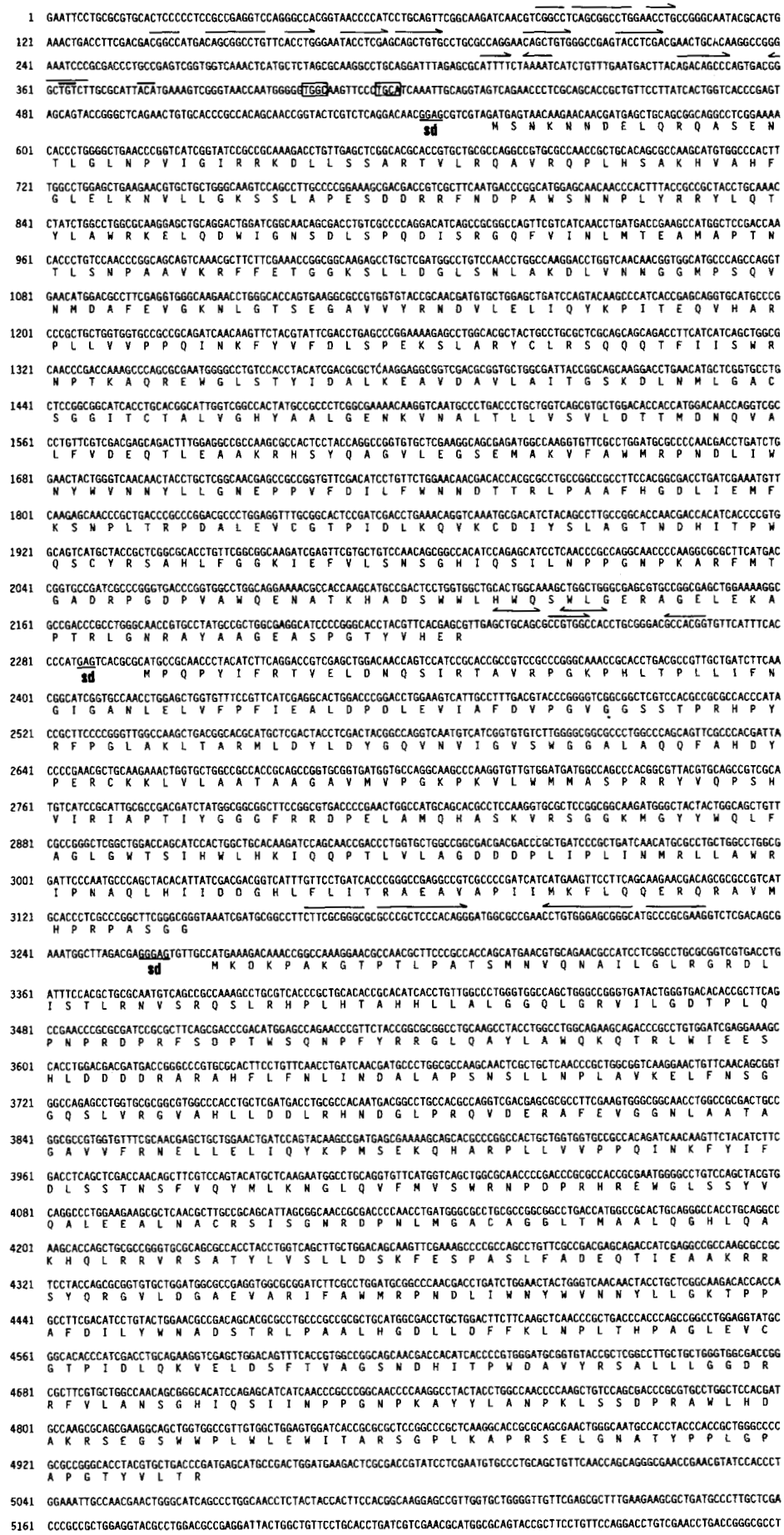

FIG. 2. **Nucleotide sequence of 6.4-kb *Eco*RI-*Eco*RI fragment.** The noncoding strand is shown. The amino acid sequences of the three open reading frames indicated in Fig. 1 are shown. The ORFs are preceded by a Shine-Dalgarno box (*SD*) (Shine and Dalgarno, 1975). Inverted and directed repeats are indicated by *arrows*. A possible −24/−12 promoter region (Dixon, 1986) preceding ORF1 is *boxed*; a sequence resembling the NifA box (Morett and Buck, 1988; Ow *et al.*, 1985) is *overlined*.

FIG. 2—*continued*

```
5281  GCCCAAACTGGCCCGCGGCATGCGCAACCTGATCAACGCGCTCAAGCGCACACTGGCGGCGTTGCTGGCCAGCCTCAAGGGCCAAGGGTTGGTAGAGAGCGAGACCCAGGCGCTGGGGCA
5401  ACTGGTGGAGCAGATCACCCTGACACTGATGTTCTCGCTGGATTATCAGCGGGTACTGGGGCGCGAGGGGGATGTGGGGATTGTGGTGTACCAGGTGATGATGCTGGTGGCGCCGCATCT
5521  GCAGGCCCAGGCGCGGGGGGCGGCGGAGCAATTGGCGGTGCGGTATCTGGAGGGGTAAGCCTGTTGATTCGGTGTCGCGGCTTTCGCGGGCATGCCCGCTCCCACAGGTGAAATGCAGTG
5641  CTCGAGTGCACACAGGACCTGTGGGAGCGGGCAAGCCCGCGAAGATGGCGACGCGGTATCAGATCAGGGTACCGGTGCCTGTCTGTGCCGAAGGCGGGTTGCTGGCCGGAGCCGCAGTGG
5761  GCGCCGAAGCTGCGCTAGCGGCCGGAGCGGCGCTGGCCGCTGGAGCTGCCGGCTTGGCGGCTGCCGGCTTGGCCGGTGCTTTCTTCACTGCAGGCTTCTTCGCCACAGCCGGTTTGGCCG
5881  CTGCCGGTTTGCCGCAGGCTTGGCCGCCGCAGTCTTGGCAGCAGGTTTAGCCGCTGCGGCCTTGGCTGCCGGCTTGGCTGCCGCGGTTTTGCCGCAGGCTTGGCCGCTGCCGTCTTGGCC
6001  AGTGGCTTGGCCGCCGCCTTGCTCGCAGCCGGTTTGGCTGCAGTGCGCGACGAAATCGGGGTAACCGAAGCGCCGGTGAGTTTCTCGATCTGCTTGGTCAGGCTGTCCACCTGCTGGTGC
6121  AGGGCCTTGATCTCGTTGCGGCTCGGCACGCCAAGGCGCGAGATGGCACTGTTCAGGCGCTTGTCGAAGGCCTCTTCGAGTTCGCTCCACTTGCCTAGCGCACGGTCCTTCACGCCCGAC
6241  ACACGCGAAGTGGTCGACGACTTGGCAGTTTCAGCAACATCTTCTGCGGTCTTCTTCGCCTGTTTCTCGGCCTCTCGCCATCCTTTACCAGCGAGTCGAACAGCTTCGGGCCGTCCTGG
6361  TCGATCTTCGAATAGATACCAAGCCCCGCCAGCCAGATCTTGCGGGAGTACTTCTCGATCCCGCCGCCCAGGAGCTGCCTTCTTTCTCGGAATTC
```

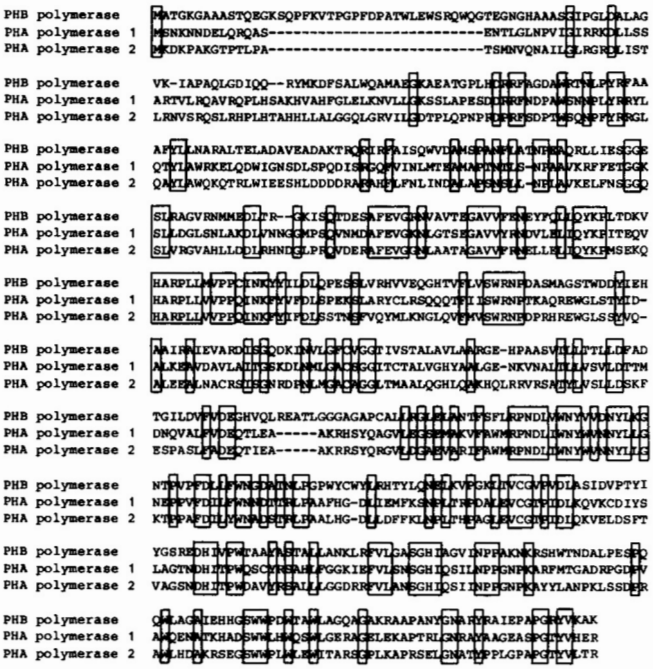

FIG. 3. **Comparison of amino acid sequences of ORF1 and ORF3 and PHB polymerase.** The amino acid sequences of the two PHA polymerases and the PHB polymerase were aligned according to the program CLUSTAL (Higgins and Sharp, 1989). Identical residues are *boxed*. Two gaps were introduced, one of which is located in the nonhomologous N terminus. In the alignment shown, PHA polymerase 1 is 53.6% identical to PHA polymerase 2. PHB polymerase from *A. eutrophus* is 37.8 and 39.5% identical to PHA polymerases 1 and 2, respectively.

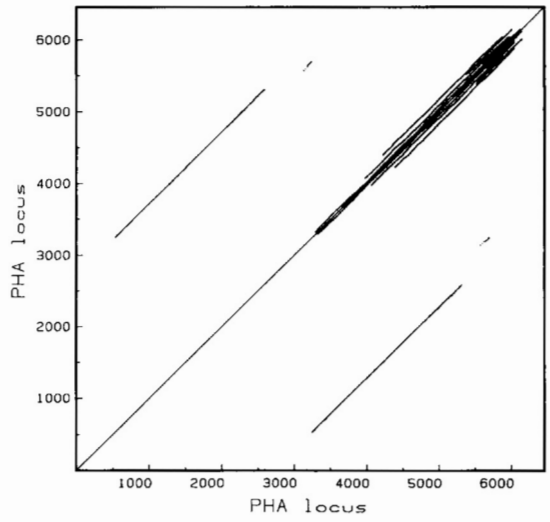

FIG. 4. **Duplication of segments within 6.4-kb *Eco*RI-*Eco*RI fragment.** The nucleotide sequence of the 6.4-kb *Eco*RI fragment was compared with itself (FASTX) (Pearson and Lipman, 1988). Lines which are not located on the main diagonal indicate repeated sequences.

```
ORF3 (3268) AATGGCTTAGACGAG GGAG TGTTG-CC ATG AAAGACAA
                ::: :  :::: :: :    ::: :   ::::
ORF1        TCGTCTCAGGACAAC GGAG CGTCGTAG ATG AGTAACAA
                :  :    :::: :: :    ::: :   ::
ORF3 (3217) GGCGCCGAACCTGTG GGAG CG--G-GC ATG CCCGCGAA
```

FIG. 5. **Comparison of regions containing two potential start condons of ORF3 with translation initiation region of ORF1.** The nucleotide sequences surrounding the start codon of ORF1 and the two possible start condons of ORF3 were aligned. The start codon ATG is printed in *italics*, and the ribosome-binding sites are in *boldface*. Gaps are indicated (-). The spacing between the ribosome-binding site and the start codon at position 3268 is 7 nucleotides, whereas the distance between the ribosome-binding site and ATG (position 3217) is only five nucleotides.

shows that the conservation of ribosome-binding sites, the spacing between the ribosome-binding site and ATG codon, and homology of the downstream nucleotides are the best for the region at position 3268 (Fig. 5). Moreover, assigning this ATG as the start codon results in PHA polymerases of identical molecular mass encoded by completely homologous genes.

Comparison of ORF2-encoded Amino Acid Sequence to Known Proteins—The amino acid sequence of ORF2 does not show any significant homology to proteins from the SWISSPROT protein library, except for a decapeptide around position 100, which resembles the lipase fingerprint (Persson *et al.*, 1989). The decapeptide sequence VNVIGVSWGG contains the lipase consensus sequence GX_1SX_2G. However, the X residues differ from those of other lipases where X_1 is either H or Y and X_2 is preferentially an L (Fig. 6*a*). We therefore defined an alternate lipase box (Fig. 6*b*) consisting of 10 amino acids based on the known 14 lipase boxes (Fig. 6*c*) and the ORF2 sequence. Screening of the SWISSPROT protein library with this alternate lipase box yielded ORF2 and the lipases of Fig. 6*C*, with a maximal score of 10; no other proteins were identified that contained this lipase box. Thus, it is likely that ORF2 encodes a protein which is related to known lipases in at least some respects.

Specificity of PHA Polymerases—To compare the two PHA polymerases *in vivo*, various GPp104 recombinants were analyzed for the composition of accumulated PHA. Both GPp104/pGEc404 and GPp104/pGEc432 were found to accumulate polymer as seen by light microscopy. Analysis of the polymer formed in liquid cultures by these strains indicated little, if any, difference in the specificity of the two polymerases (Table II). Both polymerases produced PHAs that contained 45% 3-hydroxydecanoate, 49–51% 3-hydroxyoctanoate, and 4–7% 3-hydroxyhexanoate. Under the same conditions, GPp104/pGEc434 accumulated a polymer in which 3-hydroxydecanoate was the major monomer (Table II).

These results suggest that the introduction of the *pha* locus in a multiple copy vector influences the metabolism of PHA in such a way that 3-hydroxyacyl-CoA intermediates are more

a	ORF2 gene product	V	N	V	I	G	V	S	W	G	G
b	Alternative lipase box	V	H	L	I	G	H	S	L	G	A
		I	V	Y	L		Y		Q		G
		L	F	V	V		D		M		T
			Q	M	F		V		W		S
			N	F							C
c	human LPL-1	V	H	L	L	G	Y	S	L	G	A
	mouse LPL	V	H	L	L	G	Y	S	L	G	A
	bovine LPL	V	H	L	L	G	Y	S	L	G	A
	guinea pig LPL	V	H	L	L	G	Y	S	L	G	A
	human hepatic TAG lipase	V	H	L	I	G	Y	S	L	G	A
	rat hepatic TAG lipase-1	V	H	L	I	G	Y	S	L	G	A
	pig pancreatic TAG lipase	V	H	V	I	G	H	S	L	G	S
	canine pancreatic TAG lipase	V	Q	L	I	G	H	S	L	G	A
	S. aureus lipase	V	H	L	V	G	H	S	M	G	G
	S. hyicus lipase	V	H	F	I	G	H	S	M	G	G
	P. fragi lipase	V	N	L	I	G	H	S	Q	G	A
	rat lingual lipase	I	H	Y	V	G	H	S	Q	G	T
	human gastric lipase	L	H	Y	V	G	H	S	Q	G	T
	human lecithin-cholesterol acyl transferase	V	F	L	I	G	H	S	L	G	C

FIG. 6. **Alignment of possible lipase box from ORF2 with consensus lipase box.** The possible lipase box of ORF2 (*a*) was used to define an alternate lipase box (*b*), which is based on sequences (*c*) from several lipoprotein lipases (*LPL*) and triacylglycerol lipases (*TAG*) (Persson *et al.*, 1989) and ORF2. Scanning of the protein data bank with this new lipase box did not result in identification of enzymes other than lipases.

TABLE II

Formation of PHA by P. oleovorans GPo1, P. putida GPp104, and recombinant strains

Strains were grown on 10 mM decanoate, and cultures were supplied with an additional 10 mM decanoate after 24 h of growth. After 40 h of growth, 30-ml samples were collected, and cells were lyophilized and analyzed for PHA accumulation and composition.

Strain/plasmid	PHA polymerases[a]		PHA[b]	PHA composition[c]		
	Host	Plasmid		F_{C6}	F_{C8}	F_{C10}
			%			
GPp104	0	0	ND[d]	—[e]	—	—
GPp104/pGEc432	0	1	29	0.07	0.49	0.45
GPp104/pGEc404	0	2	36	0.04	0.51	0.45
GPp104/pGEc434	0	1, 2	33	0.04	0.42	0.54
GPo1	1, 2	0	39	0.05	0.61	0.34
GPo1/pGEc434	1, 2	1, 2	28	0.06	0.47	0.47

[a] PHA polymerases may be encoded by the host and/or the plasmid. 1 = PHA polymerase 1 (ORF1 gene product); 2 = PHA polymerase 2 (ORF3 gene product).
[b] PHA is the percentage of polymer relative to the total cell mass.
[c] F_{Cx} is the mass fraction of the C_x monomer.
[d] ND, not detectable ($p < 0.05$).
[e] —, not relevant.

rapidly incorporated into PHA, compared to being processed via β-oxidation.

DISCUSSION

Genes involved in PHA biosynthesis and degradation were cloned from a gene library of *P. oleovorans* GPo1 in pLAFR1 by complementation of the PHA⁻ mutant *P. putida* GPp104 and the degradation mutant *P. oleovorans* GPo500. A 6.4-kb *Eco*RI-*Eco*RI fragment encodes enzymes involved in both PHA biosynthesis and degradation.

From the determined nucleotide sequence, the presence of three open reading frames was deduced. The open reading frames are preceded by the ribosome-binding site GGAG (ORF1 and ORF3, independent of the proposed ATG start codons) or GAG (ORF2). A putative promoter analogous to the *ntr*A-dependent −24/−12 consensus promoter (Dixon, 1986) was found 124 bp upstream of the ribosome-binding site of ORF1. This sequence, TGGCAAN$_5$CTGCA, differs from the consensus *nif*- and *ntr*-like promoters only in the last A of the −24 box and the first C of the −12 box (Dixon, 1986).

Secondary Structure in 6.4-kb EcoRI-EcoRI Fragment—The noncoding regions in the 6.4-kb fragment contain several sequences which might form secondary structures. These regions are indicated by *arrows* in Fig. 2.

A very large inverted repeat with a 26-bp stem and a 12-bp loop is located between ORF2 and ORF3 ($\Delta G = -71.7$ kcal). The stem contains a 14-bp palindromic sequence. A similar inverted repeat was found in the DNA downstream of ORF3. The stem of this inverted repeat contains an almost identical palindromic sequence (Fig. 7*A*). The loop of the structure consists of 28 bases, a number of which can form additional secondary structure, resulting in a free energy of −80.1 kcal. Comparison of these homologous inverted repeats with known DNA sequences identified two other sequences able to form similar inverted repeats (Fig. 7*B*). The first is located downstream of the *trp*BA operon of *P. putida* encoding the two subunits of tryptophan synthase (Crawford and Eberly, 1989), whereas the second was found between the *cat*BC genes, also of *P. putida* (Aldrich and Chakrabarty, 1988). Although the tryptophan synthases in *P. putida* and *P. aeruginosa* are over 70% identical (Crawford and Eberly, 1989), no such inverted repeat was found in the latter strain.

The plot in Fig. 4 depicts the numerous direct repeats in the second half of the 6.4-kb fragment. This sequence contains 54 direct repeats with a minimum repeat length of 8 bases. Due to these structures, long stretches of identity were calculated (>50% identity for stretches longer than 1000 nucleotides, >55% for stretches longer than 400 nucleotides), and these give rise to many diagonals parallel to the main diagonal in Fig. 4.

ORF1 and ORF3 Both Encode PHA Polymerases—The gene products of ORF1 and ORF3 were found to share 35–40% identity with the PHB polymerase encoded by the *phb*C gene cloned from *A. eutrophus* (Peoples and Sinskey, 1989c) (Fig. 3). Both ORFs were found to encode functional PHA polymerases. The polymer composition was only slightly dependent on the type of PHA polymerase present in the cell. When the two polymerases were encoded on multicopy plasmids, the amount of substrate-derived monomers in the polymer increased significantly, but the total amount of polymer accu-

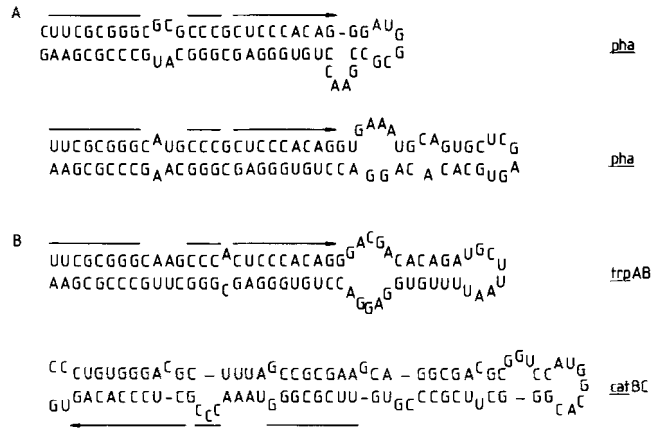

FIG. 7. **Structure of inverted repeats upstream and downstream of ORF3.** *A*, analysis of the nucleotide sequence of the *pha* locus indicated the presence of two large homologous inverted repeats upstream (IR1: $\Delta G = -1.7$ kcal) and downstream (IR2: $\Delta G = -80.1$ kcal) of ORF3. *B*, comparison of one part of the stem of IR1 with the DNA sequence library identified similar stem-loop structures in other *P. putida* regulons. Identical nucleotides in one part of the stem are indicated by *arrows* in the four inverted repeats. Downstream of the *trp*BA operon (Crawford and Eberly, 1989), an IR in the same orientation was found ($\Delta G = -65.6$ kcal), whereas an IR in the opposite orientation is localized between the *cat*B and *cat*C genes ($\Delta G = -60.6$ kcal) (Aldrich and Chakrabarty, 1988).

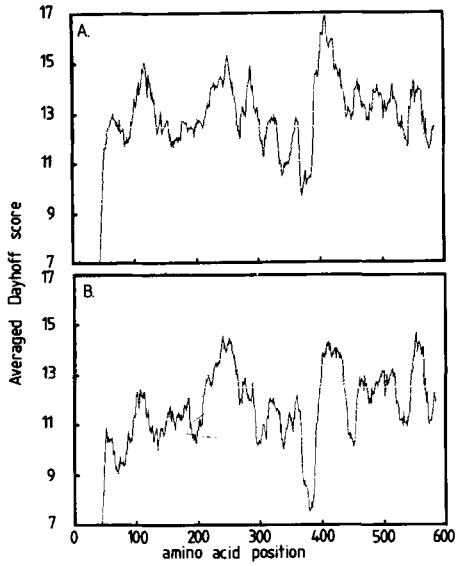

FIG. 8. **Comparison of two PHA polymerases and PHB polymerase using averaged Dayhoff score.** The amino acid sequences of PHA polymerases 1 and 2 were aligned, and the averaged Dayhoff score (Dayhoff *et al.*, 1979) was calculated over a window of 21 residues (*A*). The same was done with PHA polymerase 1 and PHB polymerase (*B*). Gaps were introduced in the PHA polymerase sequences to make all the sequences equally long. Peaks indicate a high homology between the polymerases. These residues are candidates for involvement in the polymerization of 3-hydroxy fatty acids.

mulated did not change (Table II).

Previous studies have demonstrated the importance of the substrate (Brandl *et al.*, 1988; Lageveen *et al.*, 1989) and the strain (Haywood *et al.*, 1989; Huisman *et al.*, 1989) for the composition of the accumulated PHA. The results presented here suggest that the polymer composition is in fact determined by the specific cellular PHA polymerase activity because with high polymerase levels, substrate is drawn into polymer elongation rather than to complete β-oxidation. Thus, bacteria which contain high levels of PHA polymerase activity are expected to accumulate homopolymers. This is in fact found when octane or octanoate is used as growth substrate rather than decane or decanoate; the polymer-synthesizing system shows greater activity for the shorter substrate, and the resulting PHA consists of 90% 3-hydroxyoctanoate (Brandl *et al.*, 1988; Huisman *et al.*, 1989; Lageveen *et al.*, 1988; Preusting *et al.*, 1990). Since high polymerase levels did not lead to higher cellular PHA contents, polymer accumulation apparently depends on other factors which remain to be defined. One possibility is that the formation of acyl-CoA intermediates is limiting. Another possibility is that polymer synthesis is followed by granule assembly processes and that these are limiting.

Comparison of PHA Polymerases—Although the two PHA polymerases show similar levels of homology to the PHB polymerase from *A. eutrophus*, they are mutually different in some respects. Their isoelectric points as calculated according to the PC/GENE program CHARGEPRO are pI 6.53 for PHA polymerase 1 and pI = 9.55 for PHA polymerase 2. Their amino acid sequences show two highly conserved segments: amino acids 180–234 share a 75.9% identity over a 54-amino acid stretch, and amino acids 337–492 have a 73.5% identity over a 155-amino acid stretch. The 100 amino acids between these stretches are 40% identical, whereas the 60 N-terminal amino acids of the two PHA polymerases are only 31.6% identical (Fig. 8*A*).

Similarities between the amino acid sequences of PHA and PHB polymerases might identify segments important for enzyme structure and catalysis. The conserved amino acid stretches at portions 180–233 and 367–406 in the PHA polymerases show 57.4 and 60.0% identity to the PHB polymerase, respectively, and may thus be important for the polymerization reaction itself. Amino acids between these segments are less conserved, and the N termini of the three enzymes show very little homology (Figs. 3 and 8*B*). Consequently, these segments may be involved in substrate binding and define the enzyme specificity.

It has been proposed that the first step in the synthesis of PHB is the formation of an acyl-*S*-enzyme intermediate followed by transesterification to a primer acceptor (Griebel and Merrick). Peoples and Sinskey (1989c) have identified 2 cysteine residues in the PHB polymerase that appear to be conserved in PHA polymerase. It is noteworthy that these conserved cysteine residues in the two PHA polymerases are at positions 296 and 430 and therefore occur in the less homologous segments of the polymerases.

ORF2 Encodes Enzyme Involved in PHA Degradation—The two PHA polymerases are separated by an ORF that encodes a protein involved in PHA degradation. It restores the ability to degrade the accumulated PHA in *P. oleovorans* GPo500. The protein encoded by ORF2 showed no significant homology to the extracellular PHB depolymerase cloned from *Alcaligenes faecalis* (Saito *et al.*, 1989) or to other known protein sequences. A decapeptide was identified in ORF2 that resembles the fingerprint encountered in triacylglycerol lipases and lipoprotein lipases (Persson *et al.*, 1989). The presence of such a lipase fingerprint is consistent with the assignment of a PHA depolymerase activity for the ORF2 gene product because a PHA depolymerase is in fact a lipase.

PHA Biosynthetic Pathway—The biosynthesis of PHB is known to require two enzymes in addition to the PHB polymerase (Dawes and Senior, 1973). In contrast, the number of enzymes involved in PHA biosynthesis remains unknown. It is clear from our results that PHA monomers can be polymerized by either one of two PHA polymerases. Future studies will show how the (*S*)-3-hydroxyacyl-CoA intermediates in the β-oxidation are inverted to the (*R*)-3-hydroxyalkanoates found in PHA.

Acknowledgments—We thank H. Enequist for technical assistance and Gerrit Eggink, Menno Kok, and Hans Preusting for helpful discussions.

REFERENCES

Aldrich, T. L., and Chakrabarty, A. M. (1988) *J. Bacteriol.* **170,** 1297–1304
Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J., and Timmis, K. N. (1981) *Gene (Amst.)* **16,** 237–247
Birnboim, H. C., and Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523
Brandl, H., Gross, R. A., Lenz, R. W., and Fuller, R. C. (1988) *Appl. Environ. Microbiol.* **54,** 1977–1982
Byrom, D. (1987) *Treads Biotechnol.* **5,** 246–250
Capon, R. J., Dunlop, R. W., Ghisalberti, E. L., and Jefferies, P. R. (1983) *Phytochemistry (Oxf.)* **22,** 1181–1184
Chung, C. T., Niemala, S. L., and Miller, R. H. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 2172–2175
Crawford, I. P., and Eberly, L. (1989) *Biochimie (Paris)* **71,** 521–531
Davison, J., Heusterspreute, M., Chevalier, N., Ha-Thi, V., and Brunel, F. (1987) *Gene (Amst.)* **51,** 275–280
Dawes, E. A., and Senior, P. J. (1973) *Adv. Microb. Physiol.* **10,** 135–266
Dayhoff, M. O., Schwartz, R. M., and Orcutt, B. C. (1979) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed) Vol. 5, pp. 345–363, National Biomedical Research Foundation, Washington, D. C.
Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. (1980) *Proc.*

Natl. Sci. U. S. A. **77,** 7347–7351
Dixon, R. (1986) *Mol. Gen. Genet.* **203,** 129–136
Eggink, G., Van Lelyveld, P. H., and Witholt, B. (1984) *Prog. Ind. Microbiol.* **20,** 373–380
Fickett, J. W. (1982) *Nucleic Acids Res.* **10,** 5303–5318
Findlay, R. H., and White, D. C. (1983) *Appl. Environ. Microbiol.* **45,** 71–78
Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J., and Ausubel, F. M. (1982) *Gene (Amst.)* **18,** 289–296
Griebel, R. J., and Merrick, J. M. (1971) *J. Bacteriol.* **108,** 782–789
Haywood, G. W., Anderson, A. J., and Dawes, E. A. (1989) *Biotechnol. Lett.* **11,** 471–476
Henikoff, S. (1984) *Gene (Amst.)* **28,** 351–359
Higgins, D. G., and Sharp, P. M. (1989) *Comput. Appl. Biosci.* **5,** 151–153
Huisman, G. W., de Leeuw, O., Eggink, G., and Witholt, B. (1989) *Appl. Environ. Microbiol.* **55,** 1949–1954
Ish-Horowitz, D., and Burke, J. F. (1981) *Nucleic Acids Res.* **9,** 2989–2998
King, P. P. (1982) *J. Chem. Technol. Biotechnol.* **32,** 2–8
Lageveen, R. G., Huisman, G. W., Preusting, H., Ketelaar, P., Eggink, G., and Witholt, B. (1988) *Appl. Environ. Microbiol.* **54,** 2924–2932
Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Morett, E., and Buck, M. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 9401–9405
Ow, D. W., Xiong, Y., Gu, Q., and Shen, S.-C. (1985) *J. Bacteriol.* **161,** 868–874
Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 2444–2448
Peoples, O. P., and Sinskey, A. J. (1989a) *Mol. Microbiol.* **3,** 349–357
Peoples, O. P., and Sinskey, A. J. (1989b) *J. Biol. Chem.* **264,** 15293–15297
Peoples, O. P., and Sinskey, A. J. (1989c) *J. Biol. Chem.* **264,** 15298–15303
Peoples, O. P., Masamume, S., Walsh, C. T., and Sinskey, A. J. (1987) *J. Biol. Chem.* **262,** 97–102
Persson, B., Bengtsson-Olivecrona, G., Enerback, S., Olivecrona, T., and Jornvall, H. (1989) *Eur. J. Biochem.* **179,** 39–45
Pool, R. (1989) *Science* **245,** 1187–1189
Preusting, H., Nijenhuis, A., and Witholt, B. (1990) *Macromolecules* **23,** 4220–4224
Saito, T., Suzuki, K., Yamamoto, J., Fukui, T., Miwa, K., Tomita, K., Nakanishi, S., Odani, S., Suzuki, J.-I., and Ishikawa, K. (1989) *J. Bacteriol.* **171,** 184–189
Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
Schubert, P., Steinbuchel, A., and Schlegel, H. G. (1988) *J. Bacteriol.* **170,** 5837–5847
Schwartz, R. D., and McCoy, C. J. (1973) *Appl. Microbiol.* **26,** 217–218
Shine, J., and Dalgarno, L. (1975) *Nature* **254,** 34–38
Slater, S. C., Voige, W. H., and Dennis, D. E. (1988) *J. Bacteriol.* **170,** 4431–4436
Tomita, K., Saito, T., and Fukui, T. (1983) in *Biochemistry of Metabolic Processes* (Lennon, D., Stratman, F., and Zahlten, R., eds) Elsevier/North-Holland Biomedical Press, Amsterdam
Wallen, L. L., and Rohwedder, W. K. (1984) *Environ. Sci. Technol.* **8,** 576–579
Witholt, B. (1972) *J. Bacteriol.* **109,** 350–364