Beyond the canonical PHA synthase: insights into transcriptional expression and functions of *phaC* paralogs in *Haloferax mediterranei*

Chloé Vanden Haute[1], Brendan Schroyen[1], Ulrich Hennecke[2] & Eveline Peeters[1*]

[1]Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium.

[2]Research Group of Organic Chemistry, Department of Chemistry, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium.

*Corresponding author: Eveline.Peeters@vub.be, Tel. 0032 2 6291906

Abstract

Polyhydroxyalkanoates (PHAs) are biodegradable polyesters that can be used as sustainable alternatives to petroleum-based plastics. The halophilic archaeon *Haloferax mediterranei* is a promising PHA producer with many competitive advantages, mostly linked to its extremophilic lifestyle. The PHA synthesis enzymes of *H. mediterranei* have been investigated, the most important one being a class III PHA synthase consisting of two subunits ($PhaC_{Hme}$ and $PhaE_{Hme}$). Three additional *phaC* paralogs (*phaC*$_1$, *phaC*$_2$ & *phaC*$_3$) were found in the genome sequence and were previously described as cryptic genes. In this study, we investigate these genes in greater detail. Through a targeted bioinformatics pipeline, we unravel the diversity of PHA synthases in *Halobacteriales*, which highlights the unique traits of *H. mediterranei*. Additionally, we investigated native gene expression patterns of the paralogs in different conditions relevant to PHA production: a high-growth condition, a low-growth condition and a condition in which the monomeric PHA composition is altered by feeding valeric acid. RT-qPCR assays revealed that the three *phaC* paralogs were actively transcribed and regulated under different nutritional conditions, contradicting the previous observations of the genes having a cryptic nature. The regulation of the paralogs seems to be independent on the monomeric composition of the produced polymers, but dependent on the growth phase. This suggests that all paralogs might play a role in the native PHA metabolism of *H. mediterranei*. Finally, we hypothesize about the potential functions of the paralogs and open new avenues for future characterization efforts towards the engineering of tailor-made PHAs.

Keywords: *Haloferax mediterranei*, polyhydroxyalkanoates, PHA synthase, paralogs, gene expression, archaea

Introduction

In the search for more sustainable alternatives to industrial materials, bioplastics are gaining attraction as replacements for petrochemical based polymers. A case in point are the polyhydroxyalkanoates (PHAs), which are a versatile group of biodegradable polyesters with mechanical properties similar to polypropylene and polyethylene (Laycock *et al.,* 2013). In addition to being biodegradable, PHAs can also be biobased by microbial production. The accumulation of poly(3-hydroxybutyrate) (PHB) in *Priestia megaterium* (basionym: *Bacillus megaterium*) was already discovered in 1926 (Lemoigne, 1926). Since then, many other microbial species have been found to produce these polymers. Apart from providing survival benefits through carbon and energy storage (Dawes & Senior, 1973), PHAs have also been proven to be directly involved in enhancing the stress tolerance of their producers. Examples include higher tolerance to ultraviolet irradiation, desiccation and osmotic stress (Tal & Okon, 1985). While it is widely known that carefully selected stress conditions can enhance the PHA production, the fundamental mechanisms behind the stress-protective actions of PHA have not been completely elucidated yet (Obruca *et al.,* 2018).

Despite its great potential, a large bottleneck in PHA production is the high production cost, which is still several times more expensive than the most common petrochemical based polymers (Tan *et al.* 2021). As a result, PHA research is greatly focused on the reduction of production costs and a proposed strategy is the switch to extremophilic production hosts. These extremophiles thrive in environments which are deemed to be "extreme" with regards to environmental parameters such as temperature, pH, salinity etc.

The extremely halophilic archaeon *Haloferax mediterranei* is a species that is attracting particular attention for its PHA production capabilities. Production in this extreme halophile cuts down on the energy consumption by drastically reducing sterility requirements and by allowing for cell lysis by hypo-osmotic shock in fresh water (Chen *et al.,* 2006; Hermann-Kraus *et al.,* 2013). Combined, these two factors have been indicated to reduce the cost of PHA production by more than 20% (Chen *et al.,* 2022). PHA production *H. mediterranei* and other archaea also has the advantage of being free of

endotoxins contrary to the often-used bacterial production host *Cupriavidus necator* whose polymers require extensive purification for use in biomedical applications (Koller *et al.,* 2013). Additionally, *H. mediterranei* exhibits a great metabolic versatility allowing growth on a variety of carbon sources including many industrial waste streams. Examples include whey from the dairy industry (Koller *et al.*, 2007a), glycerol from the biodiesel industry (Hermann-Krauss *et al.*, 2013), vinasse and stillage from the bioethanol industry (Bhattacharyya *et al.*, 2012; Bhattacharyya *et al.*, 2014), chitin from the seafood industry (Hou *et al.*, 2013) and agricultural waste streams such as crop waste from dates (Alsafadi *et al.*, 2020). A final advantage of *H. mediterranei* is its natural ability to produce the copolymer poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV) (Don *et al.,* 2006).

While the trait to produce PHBV copolyesters is naturally scarce among microorganisms, these copolymers have the advantage of being less crystalline and hence they are more flexible and have a larger processing window than the classically produced PHB homopolymer (Luzier, 1992). Most production organisms require the addition of costly precursor molecules to produce PHBV but *H. mediterranei,* however, has the ability to produce PHBV without addition of precursors, from unrelated carbon sources (Han *et al.*, 2013).

The PHA synthesis enzymes of *H. mediterranei* have largely been elucidated, the key enzyme being a class III PHA synthase consisting of a $PhaE_{Hme}$ and a $PhaC_{Hme}$ (E6P09_17995) subunit (Lu *et al.*, 2008). By analyzing the genome sequence, three additional *phaC* paralogs (*$phaC_1$* E6P09_13380, *$phaC_2$* E6P09_15240 & *$phaC_3$* E6P09_15825) have been discovered and they have been characterized to be cryptic genes in PHA-accumulating conditions (Han *et al.,* 2010). Each of the paralogs has previously been investigated for its PHA production ability by cloning them, together with *$phaE_{Hme}$*, into a non-PHA producing *Haloarcula hispanica* strain. While the PhaC2 subunit was found to be inactive, both PhaC1 and PhaC3 could form functional PHA synthase complexes with PhaE and additionally each of these complexes resulted in the synthesis of polymers with slightly different 3HV contents (Han *et al.*, 2010). Another study confirmed that PhaC1 and PhaC3 are capable of synthesizing PHA polymers in *H. mediterranei* by making use of multiple deletion strains (Chen *et al.,* 2020). Here, it was observed that

the monomeric composition differs depending on the PhaC paralog in use. Hence, it was hypothesized that these paralogs could be used for production of genetically tuned, tailormade PHBV copolymers in *H. mediterranei*, which would allow for the production of a range of interesting polymers, increasing its industrial versatility (Chen *et al.,* 2020).

In this study, we take a deeper look at involvement of $PhaC_{Hme}$ and its paralogs in the native PHA synthesis of *H. mediterranei*. We aim to link transcriptional gene expression patterns and chemical composition of the PHBV polymers to the structure of these enzymes, with the goal in mind to generate more fundamental knowledge that contributes to designing new strategies for production of tailormade PHAs.

Materials and methods

Bioinformatic analysis

Genomic neighbourhoods were analyzed using the accession numbers and genome assembly provided by NCBI. Paralogous protein structures were modeled and predicted using AlphaFold 3 (Abramson *et al*., 2024). Multiple sequence alignments were generated with COBALT (Papadopoulos *et al*., 2007).

To explore the diversity of PhaC paralogs within the order *Halobacteriales*, a comprehensive bioinformatics pipeline was developed. Four experimentally characterized *Haloferax mediterranei* PhaC proteins (WP_004056138.1, WP_004060157.1, WP_004061055.1, WP_004060696.1) and the associated PhaE subunit (WP_004056139.1) were retrieved from the NCBI RefSeq database and used as reference queries. A large dataset of target organisms was compiled by extracting unique scientific names from an initial set of BLAST hits. For each organism, the full proteome was downloaded using the NCBI Entrez API, and sequences were stored in FASTA format. These proteomes were concatenated to generate a single reference database representing *Halobacteriales* diversity.

A local BLAST database was constructed using the makeblastdb utility from BLAST+ (v2.16.0), enabling efficient, offline searches. Protein–protein BLAST (BLASTp) was run using each of the four PhaC

reference sequences against the custom database, with the top four hits retained per query (-max_target_seqs 4) (Camacho *et al.*, 2009). Output was formatted with extended fields (qacc sacc pident length mismatch gapopen qstart qend sstart send evalue bitscore stitle) to allow in-depth post-processing. A separate BLASTp search was conducted with PhaE (WP_004056139.1) as the query against the same database to investigate genomic colocalization of PhaC and PhaE.

BLAST results were parsed in Python using pandas. Duplicate hits were removed based on identical alignment statistics across multiple fields, and for each reference class ($PhaC_{Hme}$, $PhaC_1$, $PhaC_2$, $PhaC_3$), the highest scoring unique hit per organism was retained. The identity of the best matching reference for each hit was inferred by selecting the class with the highest percent identity. A hit was defined as weak when the percent identity of the best match was below 65%. To examine truncation patterns, the alignment length was compared to the full length of the reference sequence, expressed as a percentage of query coverage.

To determine the genomic context of each PhaC and PhaE hit, genomic coordinates were retrieved using the NCBI Entrez efetch API. These coordinates were used to identify the closest PhaC gene for each PhaE hit within the same genome, based on the smallest absolute distance between the end coordinate of PhaE and the start of the corresponding PhaC. This information was used to infer co-localized gene pairs and assign each PhaE to its most likely functional PhaC partner.

Data visualization was performed in Python using matplotlib. Each PhaC hit was represented as a colored circle according to its best matching class, plotted per organism in vertically stacked panels of 25 species. Hits with percent identity below 65% were indicated with unfilled circles. A small black dot was added to each marker when the corresponding PhaC was located in genomic proximity to a PhaE protein, indicating potential operonic organization.

To assess patterns of paralog occurrence, summary statistics were calculated from the final dataset. Species with one or more PhaC paralogs were tallied, and the most prevalent number of paralogs was identified. Among species with only one PhaC paralog, the distribution of best-matching classes was examined. Additionally, the number of species for which a PhaC-PhaE pairing could be

confidently established based on genomic colocalization was quantified, providing insight into the conservation of operonic structure across the *Halobacteriales*.

Strains, media and cultivation conditions

Haloferax mediterranei DSM1411 was cultivated at 42°C and 125 rpm in an orbital shaker in different defined cultivation media. Growth was monitored by measuring optical density at 600 nm (OD_{600}). The Hv-min minimal medium (Dyall-Smith *et al.*, 2009) contained glycerol, sodium lactate and sodium succinate as carbon sources. Briefly, 100 mL distilled water was combined with 200 mL 30% salt water (composition per liter: 240 g NaCl, 30 g $MgCl_2{\cdot}6H_2O$, 35 g $MgSO_4{\cdot}7H_2O$, 7 g KCl, 25 mL Tris-HCl pH 7.5) and autoclaved. Subsequently, the following filter-sterilized solutions were added: 8.5 mL Hv-min carbon source (for 250 mL: 41.7 mL sodium lactate solution 60% (w/w), 3.15 mL 80% glycerol 6% SW, 37.5 g sodium succinate dibasic hexahydrate in distilled water, adjusted to pH 7.5 with NaOH), 4 mL Hv-min salts (30 mL 1 M NH_4Cl, 36 mL 0.5 M $CaCl_2$, 6 mL trace elements solution containing 0.36 mg mL^{-1} $MnCl_2{\cdot}4H_2O$, 0.44 mg mL^{-1} $ZnSO_4{\cdot}7H_2O$, 2.3 mg mL^{-1} $FeSO_4{\cdot}7H_2O$ and 0.05 mg mL^{-1} $CuSO_4{\cdot}5H_2O$), 1 mL 1 M $CaCl_2$, 650 µL KPO_4 buffer (100 mL distilled H_2O, 61.5 mL 1 M K_2HPO_4 and 38.5 mL 1 M KH_2PO_4, adjusted to pH 7.0), 300 µL thiamine/biotin solution (9.6 mL 1 mg mL^{-1} thiamine and 1.2 mL 1 mg mL^{-1} biotin). When mentioned, the Hv-min medium was supplemented with 0.5 g L^{-1} valeric acid in exponential phase (OD_{600} between 0.4 and 0.6) as a precursor for the synthesis of 3-hydroxyvalerate subunits of the PHBV copolymer. The glycerol minimal medium is identical to the Hv-min medium, with the exception of the carbon source mixture being only glycerol (2.09 mL of 80% glycerol 6% SW per 300 mL of 20% SW).

Genomic DNA extraction by phenol/chloroform based protocol

Starting from an overnight culture of *H. mediterranei* on Hv-min medium, 2 mL was centrifuged at 16,000 x *g* during 2 minutes and the pellet was resuspended in 200 µL Elution Buffer ("EB" 10 mM Tris-HCl pH 8.5). An equal volume of phenol/chloroform/isoamyl alcohol solution (25:24:1) was added, after

which the solution was vortexed during 1 minute and centrifuged at 16,000 x *g* during 5 minutes. Next, the top aqueous phase was recovered and placed in a new tube. 200 µL of fresh EB was added to the first tube and the procedure was repeated, adding the aqueous phase to the previous one. To the combined aqueous phases, an equal volume of chloroform/isoamyl alcohol solution (24:1) was added and the mixture was again vortexed during 1 minute and centrifuged at 16,000 x *g* during 5 minutes. The top aqueous phase was removed and placed in a new tube to which NH_4OAc was added to a final concentration of 0.75 M as well as 20 µg of glycogen. The solution was mixed and 2.5 volumes of 100% ethanol were added and mixed again. After spinning down at 4°C and 17,000 x *g* during 20 minutes, the pellet was washed two times with 80 % ethanol followed by centrifugation at 4°C and 17,000 x *g* during 15 minutes. Residual ethanol was removed and the pellet was resuspended in EB and stored at -20°C until use.

RNA extraction and cDNA preparation

RNA was extracted using a RNeasy Mini Kit (Qiagen) with two consecutive DNase treatments for removal of genomic DNA. For the glycerol medium, RNA was extracted from 10 mL early exponential cultures. The starting volumes for extractions at late exponential and stationary phase were normalized to the OD_{600} of the cultures, in order to start from similar amounts of cells in each growth stage. The same strategy was employed for RNA extractions of cells cultivated in the Hv-min medium, using 10 mL of early exponential culture and normalizing the starting volumes for mid exponential, late exponential and stationary samples to the OD_{600}. For cells cultivated in Hv-min medium with supplementation of valeric acid, extractions were performed in mid-exponential and stationary phase. As these samples will be compared directly to those cultivated in Hv-min medium, starting volumes were normalized in order to start from similar amounts of cells as in the Hv-min medium. All extractions were performed in biological triplicates.

RNA was converted to cDNA using a GoScript™ Reverse Transcriptase kit (Promega). All RNA samples were adjusted to a concentration of 250 ng/µL, of which 4 µL were combined with 1 µL of

random primers (0.5 µg/µL). Samples were heated for 5 minutes at 70°C, followed immediately by chilling on ice for 5 minutes. After a short spin, the 5 µL samples were combined with 4 µL GoScript 5x reaction buffer, 2 µL $MgCl_2$, 0.75 µL dNTP mix (10 mM), 7.25 µL nuclease free H_2O and 1 µL reverse transcriptase. For each condition, a "no RT" control was prepared by replacing the reverse transcriptase enzyme with 1 µL of nuclease free (NF) H_2O. The samples were then placed into a thermocycler (BioRad) at a temperature program of 5 minutes at 25°C, 60 minutes at 42°C and 15 minutes at 70°C. cDNA samples were stored at -20°C until use and a tenfold dilution was used in RT-qPCR assays.

Gene expression analysis using RT-qPCR

Gene expression was monitored using Reverse Transcription quantitative PCR (RT-qPCR) using specific primers (**Supplementary Table S1**). For each of these primers, the amplification efficiencies and primer specificities were first assessed. For this purpose, a tenfold dilution series of gDNA obtained by phenol/chloroform extraction was prepared, ranging from 50 ng/µL to 5 pg/µL. For each of these dilutions, 1 µL of gDNA was combined with 10 µL GoTaq qPCR Master Mix (Promega), 7.4 µL nuclease-free H_2O and 0.8 µL of each primer (10 µM) in a 96-well plate (iCycler iQ, Bio-Rad) in three technical replicates. Two no-template-control samples were also measured for each primer pair, as a control for nucleic acid contamination. Here, the gDNA was replaced by nuclease-free H_2O. Samples were subjected to a temperature profile in the CFX96 Real-Time System (Bio-Rad) of 3 minutes at 95°C followed by 40 cycles of 10 seconds at 95°C and 30 seconds at 55°C. In a last step, a melting curve was recorded to assess primer specificity by heating the samples from 55°C to 95°C at a rate of 0.5°C/30 seconds. All primers were deemed to be specific, as a single well-defined peak appeared in the melting curve. Plotting the cycle threshold values as a function of the logarithm of the starting quantity of gDNA, allowed for the calculation of the amplification efficiencies, as described in Equation 1.

$$Amplification\ efficiency\ (\%) = \left(10^{-\frac{1}{slope}} - 1\right) \cdot 100 \qquad (1)$$

For the gene expression analysis assay, 1 µL of tenfold diluted cDNA was combined with 10 µL GoTaq qPCR Master Mix (Promega), 7.4 µL nuclease-free H2O and 0.8 µL of each primer (10 µM) in a 96-

well plate (iCycler iQ, Bio-Rad). Each condition was analyzed using three biological replicates in two technical replications. For each gene, a NTC control was also measured, in which the cDNA was replaced by nuclease-free H_2O. Samples were subjected to a temperature profile in the CFX96 Real-Time System (Bio-Rad) of 3 minutes at 95°C followed by 40 cycles of 10 seconds at 95°C and 30 seconds at 55°C.

The Pfaffl method was used to calculate relative gene expression ratios (R, Equation 3) (Pfaffl, 2001). Inputs to the Pfaffl equation are modified primer efficiencies (E, Equation 2) and Ct values of both the target genes and a housekeeping gene under a sample and a control condition. For the analysis over different growth stages in the glycerol and the Hv-min+VA media, *tbp* was used as the stable housekeeping gene whereas *ffs* was used in Hv-min medium. For comparison of the Hv-min and Hv-min+VA conditions, the geometric mean of both *tbp* and *ffs* was used according to Equation 4, as this resulted in a more accurate normalization (Vandesompele *et al.,* 2002). All input Ct values were the average of two technical replicates and three biological replicates. For the statistical analysis, technical replicates were averaged and the resulting groups of three biological replicates were compared by means of an unpaired Student's t-test in the Prism 10 software package (GraphPad).

$$E = 10^{-\frac{1}{slope}} \tag{2}$$

$$R = \frac{{E_{target}}^{\Delta Ct\ (control-sample)}}{{E_{HK}}^{\Delta Ct\ (control-sample)}} \tag{3}$$

$$R = \frac{{E_{target}}^{\Delta Ct\ (control-sample)}}{GeoMean\left[{E_{HK}}^{\Delta Ct\ (control-sample)}\right]} = \frac{{E_{target}}^{\Delta Ct\ (control-sample)}}{\prod_{i=1}^{n}\left[{E_{HK}}^{\Delta Ct\ (control-sample)}\right]_i^{\frac{1}{n}}} \tag{4}$$

Quantitative transcript levels were estimated by relating the measured Ct values of RT-qPCR assays to the standard curves recorded for each gene (**Supplementary Figure S1** and **Table S2**). The DNA quantities were then used as a proxy for transcript levels, allowing the comparison of different paralogs at different timepoints.

PHA extraction by a chloroform/methanol based protocol

H. mediterranei DSM1411 cells were cultivated in 300 mL of the different media described above, until they reached stationary phase. Cells were centrifuged at 12,100 x *g* during 15 minutes, washed with HSPB buffer (7.17 mL 1 M K_2HPO_4, 2.83 mL 1 M KH_2PO_4, 50 mL 4 M NaCl, made up to 100 mL with distilled H_2O) and snap-frozen in liquid nitrogen. When frozen, the cells were lyophilized for 16 hours in a benchtop lyophilizer (Labconco). Afterwards, the dried pellets were weighed and put in a Whatmann Extraction Thimble. This extraction thimble was suspended in 30 mL ethanol and left to shake at room temperature for 16 hours. Afterwards, the cup was transferred from the ethanol to 30 mL chloroform and left to shake at room temperature for 16 hours. The extraction thimble was removed and the chloroform was allowed to evaporate overnight. The resulting PHA was weighed and collected for further analysis.

Determination of the monomeric polymer composition by Nuclear Magnetic Resonance (NMR)

The monomeric composition of the PHA polymers was determined by ^{1}H-NMR spectroscopy. Polymers were dissolved in deuterated chloroform ($CDCl_3$) and ^{1}H-NMR spectra were measured at 400 MHz on a Bruker Avance Neo spectrometer (Bruker, Germany) using a standard pulse sequence at ambient temperature. The chemical shifts are reported in delta (δ) units in parts per million (ppm) relative to trimethylsilane. Spectra were referenced relative to the residual solvent signal. The 3-hydroxyvalerate content of the polymers was calculated from the ratio of the area under the methine proton peaks of the 3-hydroxybutyrate and 3-hydroxyvalerate subunits respectively.

Results

PhaC paralogs of H. mediterranei are encoded across the genome and megaplasmids, share a common core and are variable in their C-termini

Haloferax mediterranei contains multiple paralogs of the key PHA enzyme PhaC. While $phaC_{Hme}$ (E6P09_17995), which is considered as the main PhaC enzyme, is clustered with the gene encoding its

auxiliary protein PhaE, the other paralogs seem to be scattered randomly across the genome and megaplasmids without any PHA-related genes in their neighbourhood (**Figure 1**).

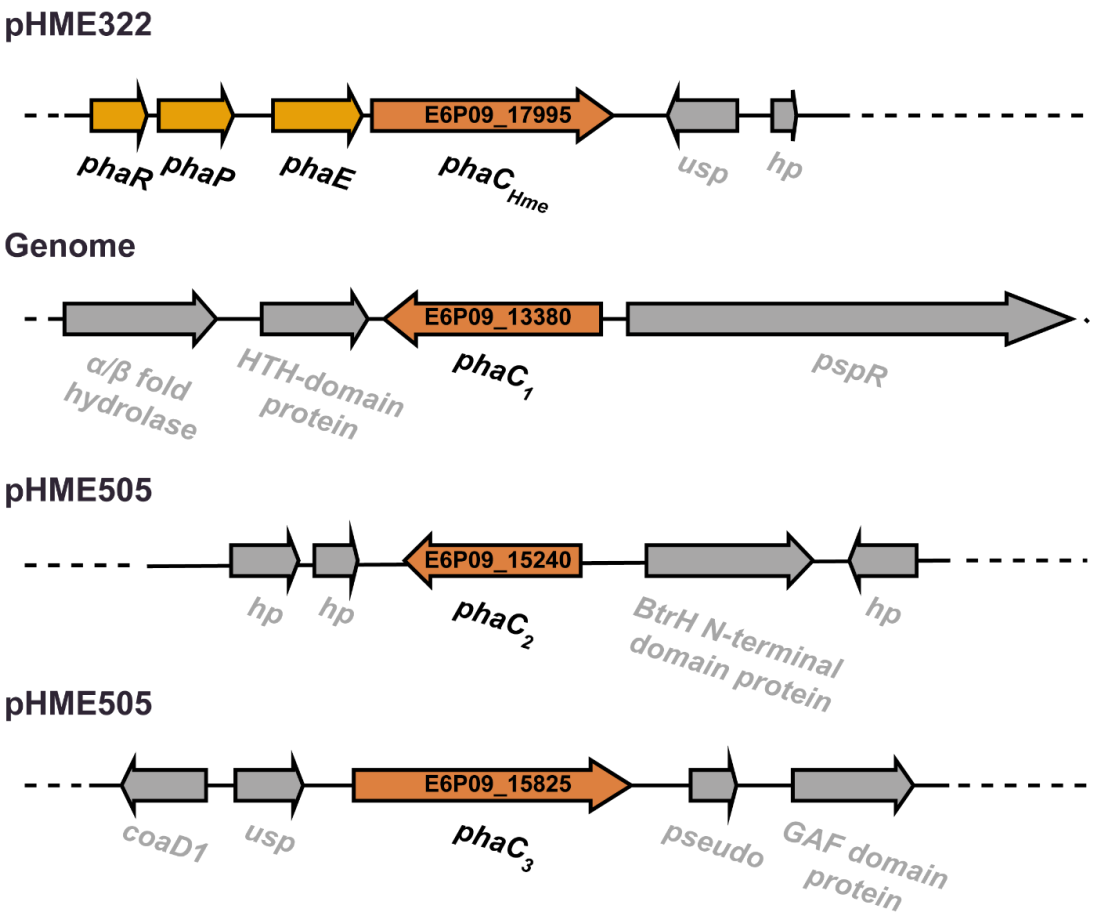

Figure 1. Graphic visualization of the genomic neighbourhoods of the *phaC* paralogs in *H. mediterranei*. All gene lengths are scaled with respect to *phaC*$_{Hme}$. Intergenic regions are not true to scale. The genes part of the core of the PHA machinery are coloured in light orange, paralogs in dark orange and unrelated genes in grey. Usp: universal stress protein, hp: hypothetical protein, HTH: helix-turn-helix, *pspR*: PPS-like regulator, BtrH: Butirosin biosynthesis protein H, *coaD1*: phosphopantetheine adenylyltransferase.

Next to the apparent distribution of the paralogs between the genome and extrachromosomal elements, we also observe a variation in gene length. The length varies between 1404 bp, with phaC$_2$ being the smallest, and the 2211 bp of phaC$_3$ which is the largest *phaC* paralog. This variation also translates to variations in the secondary and tertiary structures of the paralogs. On the protein sequence level, a multiple sequence alignment was performed of PhaC$_{Hme}$ and the three paralogs. When assessing sequence coverage, it can be observed that PhaC$_1$ and PhaC$_3$ cover most of the sequence of PhaC$_{Hme}$. In contrast, PhaC$_2$ covers only 62% of the sequence, and small deletions are found at the C- and N-termini. Percentage identity wise, all three paralogs share around 50% with PhaC$_{Hme}$. Using AlphaFold, we were

able to predict the tertiary structure of all four PHA synthases (**Figure 2**). Here, a common thread can be observed in which the core of the enzyme is predicted with a high identity while the termini are predicted with somewhat lower certainty.

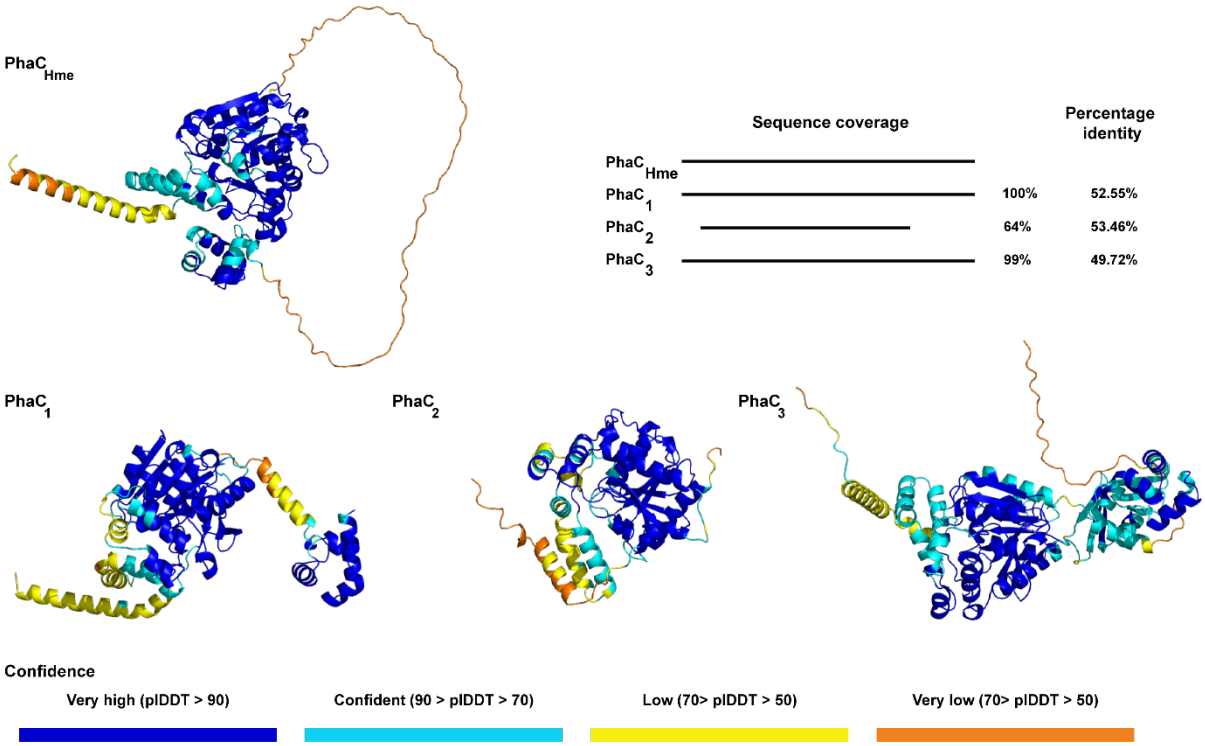

Figure 2. AlphaFold models of $PhaC_{Hme}$ and its three paralogs. The structures are coloured based on their confidence scores according to the legend below. In the upper right corner, the sequence alignments of the three paralogs can be found. All sequences were aligned against the sequence of $PhaC_{Hme}$. The black lines indicate the coverage of the paralog sequences against the sequence of $PhaC_{Hme}$.

The four structures were used to further investigate the conserved and variable structural elements by performing an alignment (**Figure 3**). Considering a superposition of the four structures, it can be observed that the α/β core, which is a signature for the α/β hydrolase superfamily, is highly conserved (Chatonnet *et al.*, 2023) (**Figure 3A**). Moreover, the indispensable catalytic triad (Histidine – Cysteine – Aspartate), is highly conserved in both sequence and structure (**Figure 3B**). In the termini, the N-terminus is quite conserved as each paralog starts with an alpha helix. In contrast, within the C-terminus a variability between the paralogs can be observed. Here, large variations are observed in the secondary structures or disordered regions but also with regards to the length of the termini. Whereas $PhaC_2$ lacks C-terminal alpha-helices, $PhaC_3$ seems to have an extended C-terminal domain with two extra β-sheets, explaining the differences in gene lengths as well. Interestingly, the N-terminal domain of

PHA synthases is hypothesized to be involved in different kinds of interactions including protein-granule interactions and the formation of (hetero)dimers (Neoh *et al.*, 2022).

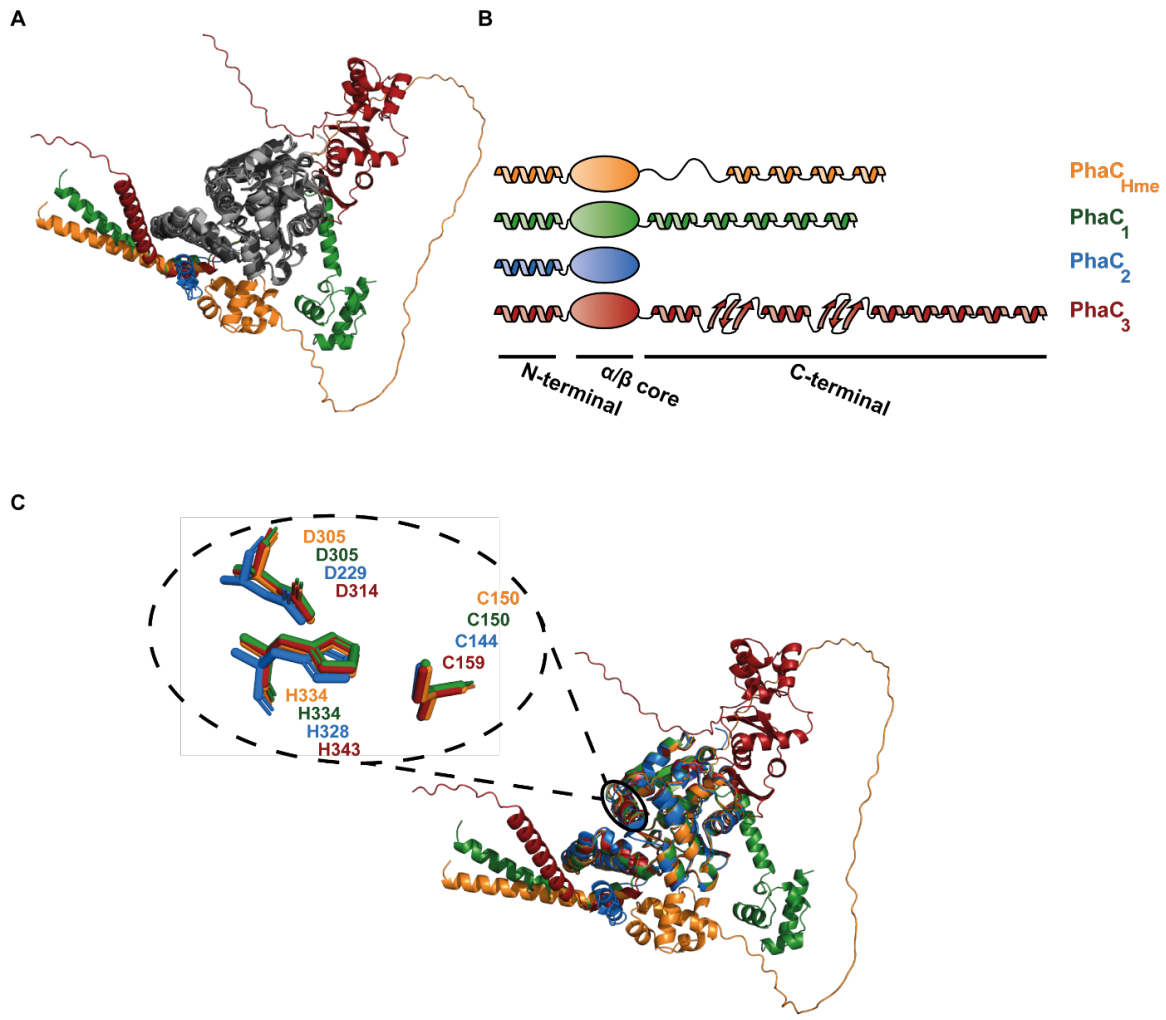

Figure 3. A: Structural superposition of the three paralogs with the main $PhaC_{Hme}$. The core of the enzyme is coloured in grey while the N- and C-terminal are coloured corresponding to the paralog according to the legend on the right. **B:** Graphical representation of the domain composition of the PhaC enzymes with a focus on the secondary structure. **C:** Superposition of the catalytic core of the enzymes. Zoom of the three catalytically indispensable residues and their position in the respective PhaC structure.

The variable C-terminal domain is often called the CAP-domain. This domain closes the active site and excludes water from it. Multiple variations of CAPs are present (Figure 3), however the absence of one, such as in $PhaC_2$, might render the enzyme non-functional as previously shown (Han *et al.*, 2010).

***Phylogenetic distribution of PhaC in* Halobacteriales**

To investigate whether the occurence of multiple PhaC paralogs is a conserved feature in polyhydroxyalkanoate (PHA)-producing *Halobacteriales*, we analyzed 225 genomes in which at least one *phaC* homolog was identified. These were considered putative PHA producers. Among them, 37 species encoded more than one PhaC paralog: 32 species contained two paralogs, four species had three, and only one species, namely *Haloferax mediterranei*, harbored four paralogs. While the presence of multiple paralogs is not uncommon, the occurrence of more than two appears to be rare.

To further understand the nature of these paralogs, each PhaC homolog was categorized based on sequence similarity relative to four reference classes defined in *H. mediterranei*: $PhaC_{Hme}$-like, $PhaC_1$-like, $PhaC_2$-like, and $PhaC_3$-like (**Figure 4**). The resulting classification revealed an absence of species that carry the full complement of paralogs found in *H. mediterranei*, indicating that this complete set is not conserved across *Halobacteriales*. Instead, varied combinations of one to three paralogs, or the main PhaC enzyme, were observed, reflecting a flexible genomic organization of the PHA biosynthetic machinery. Among genomes encoding only a single *phaC* homolog, the majority aligned with either the $PhaC_{Hme}$ or $PhaC_1$-like classes. In contrast, $PhaC_2$-like homologs were rarely detected and none of the putative hits exhibited the characteristic C-terminal truncation observed in the *H. mediterranei* paralog suggesting that this paralog may be a unique evolutionary feature.

The resulting classification revealed an absence of species that carry the full complement of paralogs found in *H. mediterranei*, indicating that this complete set is not conserved across *Halobacteriales*. Instead, varied combinations of one to three paralogs, or the main PhaC enzyme, were observed, reflecting a flexible genomic organization of the PHA biosynthetic machinery. Among genomes encoding only a single *phaC* homolog, the majority aligned with either the $PhaC_{Hme}$ or $PhaC_1$-like classes. In contrast, $PhaC_2$-like homologs were rarely detected and none of the putative hits exhibited the characteristic C-terminal truncation observed in the *H. mediterranei* paralog suggesting that this paralog may be a unique evolutionary feature.

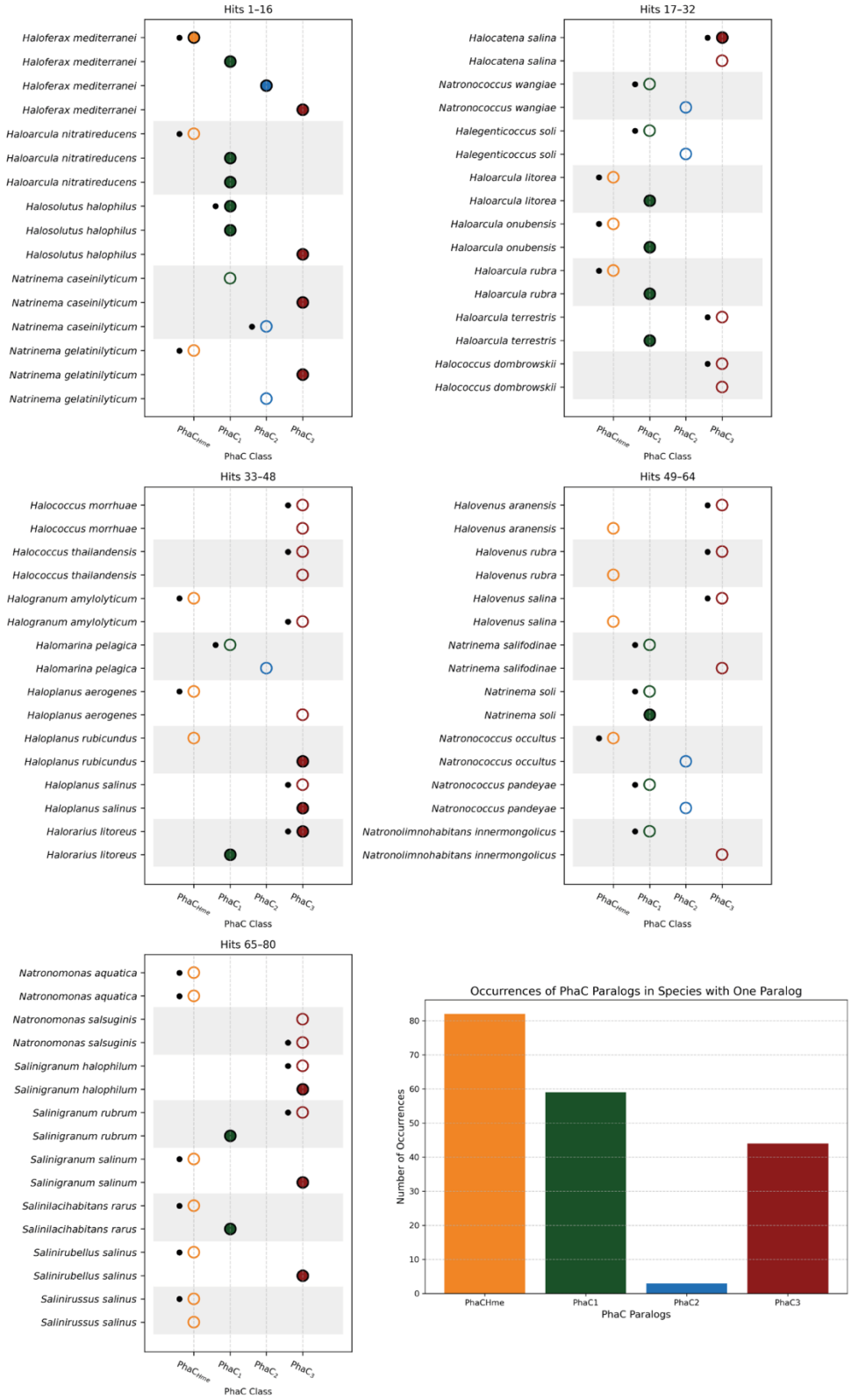

Figure 4. Graphical representation of the occurrence of PhaC paralogs in the genomes of PHA-producing *Halobacteriales*. Each PhaC homolog was classified into one of four categories—$PhaC_{Hme}$-like, $PhaC_1$-like, $PhaC_2$-like or $PhaC_3$-like —based on their sequence characteristics and genomic context in *Haloferax mediterranei*. Dots are filled when the sequence identity exceeds 65%; otherwise, only the outline is shown. A small black dot adjacent to a PhaC dot indicates that the gene encoding the PhaE subunit is located nearby in the genome, suggesting the presence of a *phaEC* operon. Only species with more than one paralog have been shown, a full table is available in **Supplementary Figure S2**. The bar chart represents the distribution of PhaC-like paralogs in microorganisms containing only one *phaC* gene.

In addition to *phaC* diversity, patterns in *phaE* co-localization were examined. It was observed that the gene encoding the PhaE subunit frequently co-occurs with various *phaC* paralogs, not exclusively with one particular class. Despite the presence of multiple *phaC* genes in some genomes, most species contained only a single copy of *phaE*. Only in a few exceptional cases, for example in *Natronomonas aquatica,* more than one *phaE* gene was identified.

All phaC paralogs are transcriptionally expressed and subject to gene regulation

To complement our comprehensive bioinformatic analysis, we also performed a series of experimental assays. Two defined media, one with glycerol as the sole carbon source (glycerol medium) and the other harboring a mixture of glycerol, sodium lactate and sodium succinate (Hv-min medium), were compared for their support of growth and PHA production in *H. mediterranei* (**Figure 5**). The glycerol medium supported growth to a maximal OD_{600} of 0.533 ± 0.062, which is relatively low for this species, while the carrying capacity of the Hv-min medium was much higher with a maximal OD_{600} of 2.220 ± 0.003 (**Figure 5A**). In terms of PHA production characteristics, both media performed similarly. The PHA content was relatively low, namely 5.2 ± 1.8% for cultivation in the glycerol medium and 6.5 ± 1.5% for the Hv-min medium. The PHA composition was determined by NMR (**Supplementary Figure S3** and **S4**) and also found to be comparable in both growth conditions with 10.4 ± 1.3 %3HV and 9.2 ± 0.6 %3HV in the glycerol and Hv-min medium, respectively (**Figure 5B**). Hence, the glycerol and Hv-min media can be definition as a low- and a high-growth condition, respectively, each resulting in a similar PHA production.

Transcriptional expression levels of $phaC_{Hme}$, $phaC_1$, $phaC_2$ and $phaC_3$ were followed over time in both the low-growth (glycerol medium) and high-growth (Hv-min medium) condition, with the aim of investigating growth-phase dependent regulatory patterns. For this purpose, early exponential, late exponential and stationary phase samples were analyzed by of RT-qPCR (**Figure 5C** and **D**).

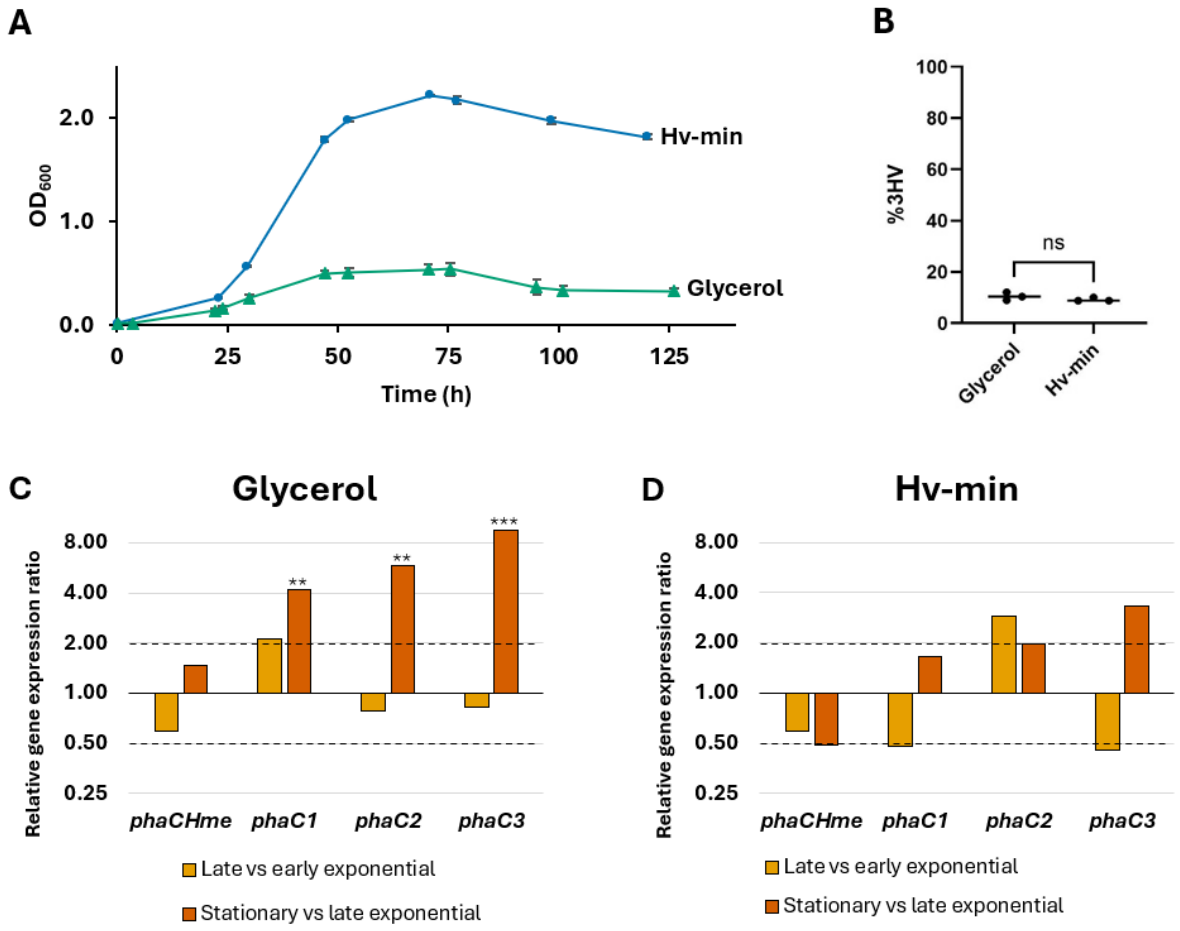

Figure 5. A. Growth curves of *H. mediterranei* on the glycerol and Hv-min defined media, measured as the optical density at a wavelength of 600 nm (OD_{600}) as a function of time. Represented values are the average of three biological replicates and the error bars depict their standard deviations. **B.** 3-Hydroxyvalerate content in the PHA polymers (%3HV) produced on the glycerol and Hv-min media, as measured by NMR. Triplicate values were compared by means of a t-test, which revealed the difference to be insignificant (ns). **C&D.** Relative gene expression ratios of the main $phaC_{Hme}$ and the *phaC* paralogs in different growth phases in the glycerol (**C**) and Hv-min (**D**) media, as determined by RT-qPCR. Underlying Ct values were compared by means of a statistical t-test. A double asterisk marks a p-value between 0.01 and 0.001, a triple asterisk marks a p-value smaller than 0.001. The dotted lines mark the levels of 2.0 (upregulation) and 0.5 (downregulation).

Transcriptional expression levels of $phaC_{Hme}$, $phaC_1$, $phaC_2$ and $phaC_3$ were followed over time in both the low-growth (glycerol medium) and high-growth (Hv-min medium) condition, with the aim of investigating growth-phase dependent regulatory patterns. For this purpose, early exponential, late exponential and stationary phase samples were analyzed by of RT-qPCR (**Figure 5C** and **D**).

In the low-growth condition, the expression levels of almost all *phaC* genes remained relatively constant when comparing late to early exponential growth phase, except for $phaC_1$, which displayed a slight upregulation (relative gene expression ratio of 2.1). Comparing stationary to late exponential growth phase, all three paralogs displayed a transcriptional upregulation (relative gene expression ratios

of 4.2, 5.8 and 9.5 for *phaC*$_1$, *phaC*$_2$ and *phaC*$_3$, respectively). In contrast, the expression level of the main PhaC subunit (*phaC*$_{Hme}$) remained rather constant over all investigated timepoints (**Figure 5C**) .

Although less pronounced, the same trends are found in the high-growth condition (**Figure 5D**). All paralogs displayed an upregulation towards the stationary phase, which was most pronounced for *phaC*$_2$ and *phaC*$_3$. Simultaneously, the main subunit *phaC*$_{Hme}$ displayed a relatively lower regulation in later growth phase stages, hinting towards an increased importance of the paralogs as growth progresses (**Figure 5D**).

Effects of feeding valeric acid on growth and PHA content

With the goal of synthesizing PHA polymers with distinct monomeric composition, valeric acid was used as a precursor for the 3-hydroxyvalerate (3HV) subunit (**Figure 3**). This precursor was added to exponentially growing Hv-min cultures and the impact on growth and PHA synthesis was monitored (**Figure 6**). The addition of valeric acid also stimulated growth, with a prolongation of the exponential phase and a higher OD_{600} that is reached (2.990 ± 0.090 compared to 2.220 ± 0.003) (**Figure 6A**). As previously described for the bacterial PHA producer *Cupriavidus necator* (Khanna & Srivastava, 2007), the addition of valeric acid increased the 3HV content in the polymer of *H. mediterranei* (**Figure 6B and Supplementary Figure S5**). Indeed, the 3HV content increased from 9.2 ± 0.6% to 53.4 ± 3.7%. Precursor feeding also significantly boosted the PHA content from 6.5 ± 1.5% to 10.5 ± 1.2%.

Gene expression of phaC paralogs seems to be independent of the monomeric composition of the polymer

A previous study showed that both $PhaC_1$ and $PhaC_3$ are able to synthesize PHA with distinct monomeric compositions when combined with PhaE in a non-PHA producing *Haloarcula hispanica* strain (Han *et al.*, 2010). Contrary to this previously report, our results indicated that the paralogs were transcriptionally expressed in the investigated conditions. This led us to investigate the potential involvement of the paralogs in native PHA synthesis, and their link to the monomeric PHA composition.

As our results indicated that the addition of valeric acid drastically increases the 3HV content of the produced polymers, we decided to investigate transcriptional gene expression levels in the Hv-min medium with and without the addition of the precursor. However, no significant differential expression levels were measured in neither mid-exponential nor stationary phase (**Figure 6C**). This clearly indicates that transcriptional regulation of the paralogs is independent of the 3HV content.

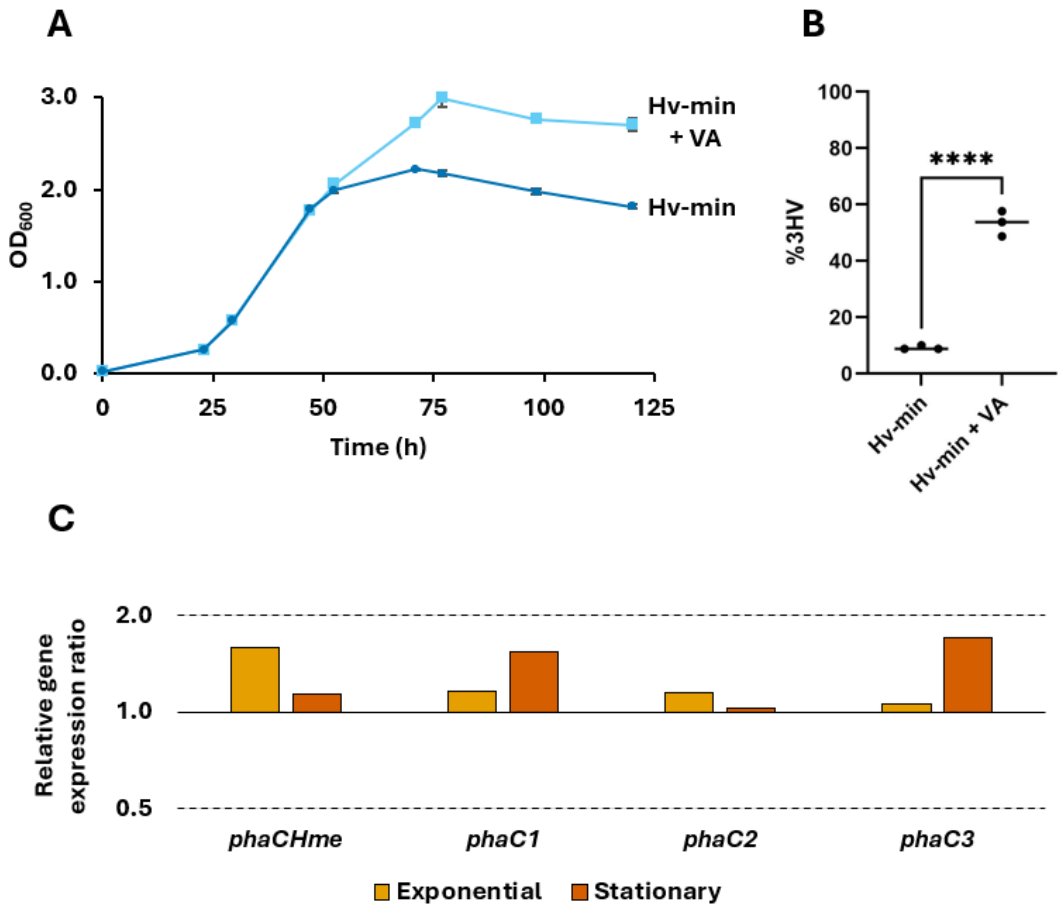

Figure 6. A. Growth curves of *H. mediterranei* on the Hv-min medium with and without the addition of valeric acid, measured as the optical density at a wavelength of 600 nm (OD_{600}) as a function of time. Represented values are the average of three biological replicates and the error bars depict their standard deviations. **B.** 3-Hydroxyvalerate content in the PHA polymers (%3HV) produced on the Hv-min and Hv-min+VA media, as measured by NMR. Triplicate values were compared by means of a t-test, which revealed the difference to be significant (**** $p<0.0001$). **C** Relative gene expression ratios of the main $phaC_{Hme}$ and the *phaC* paralogs in different growth phases in the Hv-min medium with addition of valeric acid compared to the absence of it, determined by RT-qPCR. The dotted lines mark the levels of 2.0 (upregulation) and 0.5 (downregulation).

The phaC paralogs are still subject to growth-phase dependent gene regulation when feeding valeric acid

The growth-phase dependent expression patterns of $phaC_{Hme}$ and its paralogs were also investigated in the precursor-supplemented medium by comparing stationary to mid-exponential phase.

Yet again, all three paralogs showed a significant upregulation (relative gene expression ratios of 3.69, 2.55 and 5.68 for *phaC1, phaC2 and phaC3* respectively) in stationary phase while the main subunit $phaC_{Hme}$ showed a downregulation (relative gene expression ratio of 0.33) (**Figure 7**). This further hints to an increased importance of the paralogs in later growth stages.

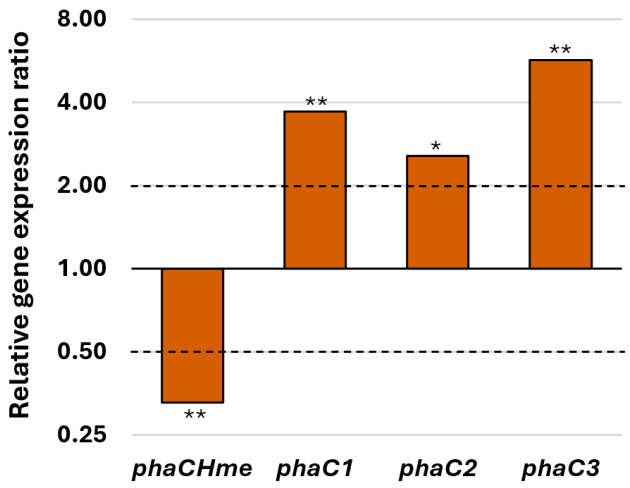

Figure 7. Relative gene expression ratios of the main $phaC_{Hme}$ and the *phaC* paralogs in stationary phase relative to mid-exponential phase in Hv-min medium with valeric acid as a 3HV precursor, as determined by RT-qPCR. Underlying Ct values were compared by means of a statistical t-test. A single asterisk marks a p-value between 0.05 and 0.01, a double asterisk marks a p-value between 0.01 and 0.001. The dotted lines mark the levels of 2.0 (upregulation) and 0.5 (downregulation).

Differences in transcription levels of between the paralogs and the main subunit

Our findings indicated that all paralogs are actively transcribed under the investigated conditions and that transcription of the paralogs increases over time. Building further on these results, we aimed to analyze potential differences in transcription levels between the main $phaC_{Hme}$ and the paralogs, as well as how these levels change over time. Making use of standard curves on gDNA, measured Ct values were converted into input DNA quantities, which were then used as an approximation of transcript levels (**Figure 8**). Relating the quantities of the main subunit $phaC_{Hme}$ to each paralog over time revealed that transcription levels differ substantially in the exponential phase across all investigated media, with $phaC_{Hme}$ being expressed at significantly higher levels. As growth progresses, the gap between the main subunit and the paralogs becomes less pronounced.

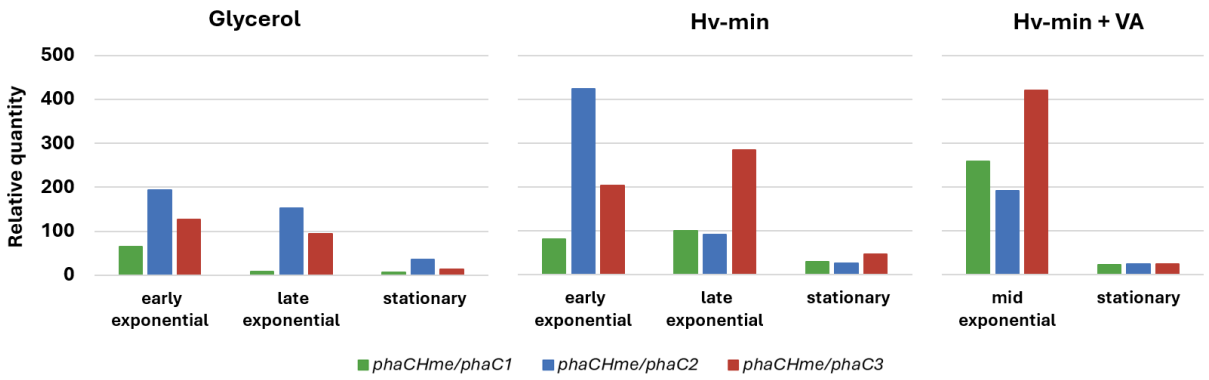

Figure 8. Relative quantity of transcript levels of the main *phaC* subunit (*phaCHme)* compared to each paralog in all investigated conditions. Values for each gene are derived by relating measured Ct values to their respective standard curves.

Discussion

As PHA research advances, it becomes clear that finding a 'one-solution-fits-all' is unlikely. Hence, research efforts are being diversified to multiple hosts and a plethora of PHA polymers, tailored to specific applications. The extreme halophilic archaeon *Haloferax mediterranei* is a prototypical example thereof. It was previously hypothesized that *H. mediterranei* could be turned into a highly specialized PHBV production platform, enabling the production of polymers with genetically tuned monomeric composition due to the presence of multiple PHA synthase PhaC subunits (Han *et al.,* 2010). In this study, we further explored the diversity and structure of these enzymes. Furthermore, we investigated the involvement of the main PhaC subunit ($PhaC_{Hme}$) and its paralogs ($PhaC_1$, $PhaC_2$ & $PhaC_3$) in the synthesis of PHBV polymers in three distinct conditions: a low growth condition, a high growth condition and a condition in which a precursor molecule was fed to alter the monomeric PHA composition.

The occurrence of multiple PhaC paralogs is not extremely rare in prokaryotes. In many cases, the paralogs themselves are redundant in the genome and contribute little to the PHA metabolism under normal conditions, as observed in *Rhodospirillum rubrum* (Jin *et al.*, 2012). In other cases, like in *Bradyrhizobium japonicum* USDA 110, five PhaC paralogs were discovered in the genome of which only two were significantly expressed (PhaC1 and PhaC2). Moreover, no PHA accumulation was observed in a ΔPhaC1 strain while there was more PHA production in a ΔPhaC2 strain, showing that while both were

expressed at the same level in wild-type cells, the interplay of the two enzymes might be more complex. In the same study, it was hypothesized that the paralogs form a heterodimer and balance out each other's activity (Quelas *et al.*, 2013).

In *H. mediterranei,* previous work has shown the presence of four PhaC paralogs of which three - $PhaC_1$, $PhaC_2$ and $PhaC_3$ - were not expressed in the investigated conditions (Han *et al.*, 2010). However, when these paralogs were heterologously expressed in the non-PHA producing *Haloarcula hispanica,* $PhaC_1$ and $PhaC_3$ were able to form a functional complex with PhaE and produce PHA with a slightly different monomer composition. $PhaC_2$ on the other hand, had a truncated C-terminal domain and was not able to yield PHA production. We found that the structural conservation across the catalytic cores of all four PhaC enzymes, combined with variation at the terminal regions, points to a shared enzymatic ancestry with potentially divergent functional capacities. The consistent presence of a conserved N-terminal α-helix across all paralogs suggests that interaction with PhaE is still feasible. However, the absence of a CAP domain in $PhaC_2$ likely renders this paralog inactive, in line with the earlier heterologous expression experiments (Han *et al*., 2010).

When analyzing the genomes of 225 PhaC containing species belonging to the class *Halobacteriales*, we found that ~83% of the species possess one PhaC enzyme in their genome, the other ~17% possess multiple paralogs with the maximum of four paralogs only seen in *H. mediterranei*. We observed that, as a consequence, the full set of PhaC paralogs is never observed in another microorganism and when there are multiple paralogs, it can be any combination of the paralog classes as defined in *H. mediterranei*. Nonetheless, species containing only a single paralog predominantly possess PhaCHme-like and PhaC1-like sequences, accounting for 44% and 31% of the one-paralog species, respectively. Finally, while we found some hits of PhaC enzymes belonging to the $PhaC_2$-like class, when investigating it further none of them lacked the C-terminal region and were thus no true $PhaC_2$ truncated enzymes. This leads us to believe that $PhaC_2$ is a unique trait of *H. mediterranei* and might have originated from some sort of gene duplication event or horizontal gene transfer.

As stated before, a previous study indicated that only *phaC*$_{Hme}$ was actively transcribed in PHA-accumulating conditions in the wild-type strain, and that the three *phaC* paralogs were cryptic genes (Han *et al.*, 2010), which matched the observation that a knockout of *phaC*$_{Hme}$ completely abolished PHA synthesis (Lu *et al.*, 2008). Our new data, however, indicates that all *phaC* paralogs are expressed, and even subjected to regulation, in all investigated native PHA production conditions. The regulation seems to be growth phase dependent, with an upregulation of *phaC*$_1$, *phaC*$_2$ and *phaC*$_3$ occuring in the stationary phase. This pattern could implicate an involvement of the paralogs in secondary metabolism. At the same time, levels of the main subunit stay constant or even undergo a slight downregulation, showing a clear distinction in the regulation of the main subunit and the paralogs. Our investigation of transcription levels showed higher levels of *phaC*$_{Hme}$ compared to the paralogs in all conditions, which is in accordance with previously reported values from an RNA-sequencing study (Chen *et al.,* 2020). This gap in transcript levels emphasizes the role of PhaC$_{Hme}$ as the primary PhaC subunit of *H. mediterranei*. This is further supported by its genomic context: *phaC*$_{Hme}$ is co-located with *phaE* in a single operon, and both are under the control of the transcription factor PhaR. PhaR itself is encoded directly upstream, along with the phasin gene *phaP*, forming a tightly regulated four-gene cluster that constitutes the core PHA biosynthetic machinery. Although transcription levels of the paralogs are always lower than that of *phaC*$_{Hme}$, our comparison of growth phases shows that the gap gradually closes towards stationary phase. The lower abundance of transcript levels of the paralogs in the exponential phase, and advancements in the sensitivity of PCR-based technologies over the past 15 years, could in part explain their initial listing as cryptic genes (Han *et al.,* 2010). Although our data suggest that the paralogs are also subject to regulation, we find it unlikely that PhaR is responsible for this. Analysis of the promoter/operator regions of the paralogs using MEME Suite did not reveal any conserved motifs or the previously described PhaR binding site found upstream of the *phaEC* operon (Cai *et al*., 2015).

Haloferax mediterranei is an exceptional PHA producer, partly due to its ability to produce PHBV copolymers, without needing any precursor molecules. This copolymer seems to naturally contain about 10% 3HV, independent of the carbon source (Chen *et al.*, 2006; Koller *et al.*, 2007a; Bhattacharyya *et al.*,

2012; Bhattacharyya *et al.*, 2014) which also matches our findings on the glycerol and Hv-min media. To create a condition with distinct PHA monomeric composition, we chose to feed valeric acid as a 3HV precursor. Addition of the precursor boosted growth and PHA content, while increasing the 3HV content of the polymer by more than five-fold to a final value of ~53%. The potential of valerate as a 3HV precursor in *H. meditteranei* had already been investigated previously (Koller *et al*., 2007b; Han *et al.*, 2015). However, our approach resulted in a much higher 3HV incorporation at the used concentration of precursor than the levels that had been reported thus far (21.8% for Koller *et al*. at 8 mM valerate; 20.8% for Han *et al*. at 6.5 mM; 53.4% in our study at 4.9 mM). Han and colleagues were able to reach 60% 3HV while feeding 17 mM valerate (unspecified anion), but with a very negative impact on growth due to precursor toxicity. A possible explanation for the discrepancy between our data and these previous studies is our use of valeric acid as the source of valerate, which might cause pH stress, which in turn could alter the PHA production characteristics. Another study showed that growth of *H. meditteranei* is capable of producing PHBV copolymers with extremely high percentages (> 90%) of 3HV when feeding valeric acid as the sole carbon source, but this approach resulted in very slow growth (Ferre-Guell & Winterburn, 2018). Hence, our study is the first to reach such high 3HV content by the feeding of valerate as a precursor, without harming growth.

As previous reports indicated that the paralogs could be involved in the synthesis of polymers with distinct monomeric compositions (Han *et al.,* 2010; Chen *et al.,* 2020*)*, we were interested to assess the effect of the addition of the precursor at the transcriptomic level. Here, we saw that the regulation of $phaC_{Hme}$ and its paralogs seems to be independent of the PHA content or its composition. Interestingly, in other PHA-producing *Halobacteriales*, we observed that different paralog classes—except for true $PhaC_2$-like variants—can be found within *phaEC*-like operons. This raises the possibility that genomic context, rather than enzyme structure alone, may influence the functional relevance of the paralogs in *H. mediterranei*. Furthermore, the regulation of the paralogs could be taking place on another level than the transcriptional level. We found, in previously published RNAseq data by Martinez Pastor *et al*., that the

phaC$_1$ paralog has a rather long 5' UTR region (~350 bp) which we do not observe in *phaC*$_{Hme}$ or the other paralogs (Supplementary Figure S6) (Martinez Pastor *et al*., 2025).

Future studies could explore whether relocating these paralogs into a more canonical genomic environment, such as the native *phaEC* operon, could alter their expression and physiological importance.

Acknowledgements

This research was funded by the Vrije Universiteit Brussel (Strategic Research Program SRP91). C.V.H. and B.S. were funded by PhD fellowships of the Research Foundation Flanders (FWO-Vlaanderen) with respective grant numbers [1SH8V24N] and [1S21424N]. U.H. is grateful for the financial support by the VUB/FWO grant OZR 3584 (NMR spectrometer).

Author contributions

C.V.H.: Conceptualization, Investigation, Formal analysis, Writing – original draft. B.S.: Conceptualization, Investigation, Formal analysis, Writing – original draft. U.H.: Resources, Writing – review & editing. E.P.: Resources, Supervision, Writing – review & editing.

References

Abramson J, Adler J, Dunger J *et al*. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500. https://doi.org/10.1038/s41586-024-07487-w

Alsafadi D, Ibrahim M I, Alamry K A *et al*. Utilizing the crop waste of date palm fruit to biosynthesize polyhydroxyalkanoate bioplastics with favorable properties. *Sci Total Environ* 2020;**737**:139716. https://doi.org/10.1016/j.scitotenv.2020.139716

Bhattacharyya A, Pramanik A, Maji S K *et al*. Utilization of vinasse for production of poly-3-(hydroxybutyrate-co-hydroxyvalerate) by *Haloferax mediterranei*. *AMB Expr* 2012;**2**:34. https://doi.org/10.1186/2191-0855-2-34

Bhattacharyya A, Saha J, Haldar S *et al*. Production of poly-3-(hydroxybutyrate-co-hydroxyvalerate) by *Haloferax mediterranei* using rice-based ethanol stillage with simultaneous recovery and re-use of medium salts. *Extremophiles* 2014;**18**:463–470. https://doi.org/10.1007/s00792-013-0622-9

Cai S, Cai L, Zhao D *et al*. A Novel DNA-Binding Protein, PhaR, Plays a Central Role in the Regulation of Polyhydroxyalkanoate Accumulation and Granule Formation in the Haloarchaeon *Haloferax mediterranei*. *Appl Environ Microbiol* 2015;**81**:373-385. https://doi.org/10.1128/AEM.02878-14

Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC bioinformatics* 2009;**10**:421. https://doi.org/10.1186/1471-2105-10-421

Chatonnet A, Perochon M, Velluet E *et al.* The ESTHER database on alpha/beta hydrolase fold proteins - An overview of recent developments. *Chem Biol Interact* 2023;**383**:110671. https://doi.org/10.1016/j.cbi.2023.110671

Chen C W, Don T M, Yen H F. Enzymatic extruded starch as a carbon source for the production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by *Haloferax mediterranei*. *Process Biochem* 2006;**41**:2289-2296. https://doi.org/10.1016/j.procbio.2006.05.026

Chen J, Mitra R, Xiang H *et al*. Deletion of the pps-like gene activates the cryptic phaC genes in *Haloferax mediterranei*. *Appl Microbiol Biotechnol* 2020;**104**:9759–9771. https://doi.org/10.1007/s00253-020-10898-0

Chen S, Cui Y, Huang M. Coupling Magnetic Field and Salinity Upshock To Improve Polyhydroxyalkanoate Productivity by *Haloferax mediterranei* Feeding on Molasses Wastewater. *Appl Environ Microbiol* 2022;**88**:e00305-22. https://doi.org/10.1128/aem.00305-22

Dawes E A, Senior P J. The Role and Regulation of Energy Reserve Polymers in Micro-organisms. *Adv Microb Physiol* 1973;**10**:135-266. https://doi.org/10.1016/S0065-2911(08)60088-0

Don T M, Chen C W, Chan T H. Preparation and characterization of poly(hydroxyalkanoate) from the fermentation of *Haloferax mediterranei*. *J Biomat Sci Polym Ed* 2006;**17**:1425–1438. https://doi.org/10.1163/156856206778937208

Ferre-Guell A, Winterburn J. Biosynthesis and Characterization of Polyhydroxyalkanoates with Controlled Composition and Microstructure. *Biomacromol* 2018;**19**:996-1005. https://doi.org/10.1021/acs.biomac.7b01788

Han J, Hou J, Zhang F *et al*. Multiple Propionyl Coenzyme A-Supplying Pathways for Production of the Bioplastic Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) in *Haloferax mediterranei*. *Appl Environ Microbiol* 2013;**79**:2922-2931. https://doi.org/10.1128/AEM.03915-12

Han J, Li M, Hou J *et al*. Comparison of four phaC genes from *Haloferax mediterranei* and their function in different PHBV copolymer biosyntheses in *Haloarcula hispanica*. *Aquat Biosyst* 2010;**6**:9. https://doi.org/10.1186/1746-1448-6-9

Han J, Wu L P, Hou J *et al*. Biosynthesis, Characterization, and Hemostasis Potential of Tailor-Made Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) Produced by *Haloferax mediterranei*. *Biomacromol* 2015;**16**:578-588. https://doi.org/10.1021/bm5016267

Hermann-Krauss C, Koller M, Muhr A *et al*. Archaeal Production of Polyhydroxyalkanoate (PHA) Co- and Terpolyesters from Biodiesel Industry-Derived By-Products. *Archaea* 2013;**2013**:129268. https://doi.org/10.1155/2013/129268

Hou J, Feng B, Han J *et al*. Haloarchaeal-Type β-Ketothiolases Involved in Poly(3-Hydroxybutyrate-*co*-3-Hydroxyvalerate) Synthesis in *Haloferax mediterranei*. *Appl Environ Microbiol* 2013;**79**:5104-5111. https://doi.org/10.1128/AEM.01370-13

Jin H, Nikolau BJ. Role of Genetic Redundancy in Polyhydroxyalkanoate (PHA) Polymerases in PHA Biosynthesis in *Rhodospirillum rubrum*. *J Bacteriol* 2012;**194**:5522-5529. https://doi.org/10.1128/jb.01111-12

Khanna S, Srivastava A K. Production of poly(3-hydroxybutyric-co-3-hydroxyvaleric acid) having a high hydroxyvalerate content with valeric acid feeding. *J Ind Microbiol Biotechnol* 2007;**34**:457–461. https://doi.org/10.1007/s10295-007-0207-7

Koller M, Hesse P, Bona R *et al*. Potential of Various Archae- and Eubacterial Strains as Industrial Polyhydroxyalkanoate Producers from Whey. *Macromol Biosci* 2007a;**7**:218-226. https://doi.org/10.1002/mabi.200600211

Koller M, Hesse P, Bona R *et al*. Biosynthesis of High Quality Polyhydroxyalkanoate Co- and Terpolyesters for Potential Medical Application by the Archaeon *Haloferax mediterranei*. *Macromol Symp* 2007b;**253**:33-39. https://doi.org/10.1002/masy.200750704

Koller M, Niebelschütz H, Braunegg G. Strategies for recovery and purification of poly[(R)-3-hydroxyalkanoates] (PHA) biopolyesters from surrounding biomass. *Eng Life Sci* 2013;**13**:549-562. https://doi.org/10.1002/elsc.201300021

Laycock B, Halley P, Pratt S *et al*. The chemomechanical properties of microbial polyhydroxyalkanoates. *Prog Polym Sci* 2013;**38**:536-583. https://doi.org/10.1016/j.progpolymsci.2012.06.003

Lemoigne M. Produits de deshydration et de polymerisation de l'acide β-oxybutyrique. *Bull Soc Chim Biol* 1926;**8**:770-782.

Lu Q, Han J, Zhou L *et* al. Genetic and Biochemical Characterization of the Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) Synthase in *Haloferax mediterranei*. *J Bacteriol* 2008;**190**:4173-4180. https://doi.org/10.1128/jb.00134-08

Luzier W D. Materials derived from biomass/biodegradable materials. *Proc Natl Acad Sci USA* 1992;**89**:839-842.

Pastor M M, Darnell C L, Vreugdenhil A, Schmid A K. Rapid rewiring of an archaeal transcription factor function via flexible cis-trans interactions. *Mol Biol Cell* 2025. https://doi.org/10.1091/mbc.E24-11-0505

Neoh S Z, Chek M F, Tan H T *et al*. Polyhydroxyalkanoate synthase (PhaC): The key enzyme for biopolyester synthesis. *Curr Res Biotechnol* 2022:**4**:87-101. https://doi.org/10.1016/j.crbiot.2022.01.002

Obruca S, Sedlacek P, Koller M *et al*. Involvement of polyhydroxyalkanoates in stress resistance of microbial cells: Biotechnological consequences and applications. *Biotechnol Adv* 2018;**36**:856-870. https://doi.org/10.1016/j.biotechadv.2017.12.006

Papadopoulos J S, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 2007;**23**:1073–1079. https://doi.org/10.1093/bioinformatics/btm076

Pfaffl M W. A new mathematical model for relative quantification in real-time RT-PCR. *NAR* 2001;**29**:e45. https://doi.org/10.1093/nar/29.9.e45

Quelas J I, Mongiardini E J, Pérez-Giménez J *et al*. Analysis of Two Polyhydroxyalkanoate Synthases in *Bradyrhizobium japonicum* USDA 110. *J Bacteriol* 2013;**195**:3145-3155. https://doi.org/10.1128/jb.02203-12

Tal S, Okon Y. Production of the reserve material poly-P-hydroxybutyrate and its function in *Azospirillum brasilense* Cd. *Can J Microbiol* 1985;**31**:608-613. https://doi.org/10.1139/m85-115

Tan D, Wang Y, Tong Y *et al*. Grand Challenges for Industrializing Polyhydroxyalkanoates (PHAs). *Trends Biotechnol* 2021;**39**:953–963. https://doi.org/10.1016/j.tibtech.2020.11.010

The Halohandbook. Protocols for haloarchaeal genetics Version 7.2 2009. https://haloarchaea.com/wp-content/uploads/2018/10/Halohandbook_2009_v7.3mds.pdf

Vandesompele J, De Preter K, Pattyn F *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;**3**: research0034.1. https://doi.org/10.1186/gb-2002-3-7-research0034