International Journal of Research (IJR) Vol-1, Issue-7, August 2014 **ISSN 2348-6848**

β-Ketothiolase Homologues in *Cupriavidus necator*

Roopesh Jain* and Archana Tiwari
School of Biotechnology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Airport Bypass Road, Gandhi Nagar, Bhopal, Madhya Pradesh, India
***E-mail:** science.roopesh@gmail.com

ABSTRACT

Polyhydroxyalkanoates (PHAs) represent a naturally occurring class of biopolyesters composed of (R)-3-hydroxy fatty acids. PHA primarily serves as carbon and energy reserves, produced by numerous bacteria in response to environmental limitation (e.g. lack of macro elements such as phosphorus, nitrogen, trace elements, or lack of oxygen) and the excess supply of carbon sources. The type of polymer produced depends on the carbon sources available, the flexibility of the organism's intermediary metabolism, and the substrate specificity of the PHA biosynthetic enzymes. Cupriavidus necator (formerly known as Ralstonia eutropha) was found to be capable of producing the P(3HB) homopolymer from even carbon numbered n-alkanoates while odd-carbon numbered n-alkanoates resulted in the accumulation of copolymers of 3HB and 3HV. Biosynthesis of PHA requires three important enzymes, β-ketothiolase, acetoacetyl-CoA reductase, and PHA synthase. The main objective of this study was to develop an understanding of β-ketothiolases in C. necator.

Keywords:

Biosynthesis, polyhydroxyalkanoates, β-ketothiolases, copolymers and protein structures

INTRODUCTION

PHAs are synthesized and deposited in diverse bacteria as spherical water-insoluble cytoplasmic inclusion bodies containing an amorphous hydrophobic polyester core surrounded by a phospholipid monolayer and associated embedded proteins (Rehm BH, 2007). PHB (Polyhydroxybutyrate), the most commonly found and isolated form of PHA from bacteria was detected in 1926 by Maurice Lemoigne as an intracellular compound of Bacillus megaterium (Girdhar et al., 2013). Accumulation of PHAs proceeds under unbalanced cultivation conditions when a carbon source is available in excess and if another macroelement like phosphorus, nitrogen, trace elements or oxygen is limiting growth at the same time (Jain et al., 2010)

The PHA operon of *C. necator* comprises three genes encoding a β-ketothiolase (phaA), an acetoacetyl-CoA-reductase (phaB), and a PHA synthase (phaC) which participate in biosynthesis of PHB (Madkour et al., 2013). In *C. Necator*, two acetyl-CoA moieties are condensed to acetoacetyl-CoA by a ß-ketothiolase (PhaA). The product then undergoes reduction by an NADPH-dependent reductase (PhaB) which produces the (R)-isomer of 3-hydroxybutyryl-CoA. Finally, the PHA synthase (PhaC) polymerizes the 3-hydroxybutyryl moieties of 3HB-CoA to poly(3HB). Besides 3HB, more than 150 different PHA constituents are currently known as components of microbial polyesters (Rehm BH, 2007).

In general, the β-ketothiolases/acetyl-CoA acetyltransferases belong to the family of transferases or acyltransferases which transfer groups other than aminoacyl groups. A β-ketothiolase catalyzes the acetylation of acetyl-CoA to acetoacetyl-CoA while one molecule of CoA is released. In addition to the fatty acid and PHA metabolism, β-ketothiolases participate in other pathways as such as ketogenesis, sterol synthesis, and propanoate and butanoate metabolism as well as in the degradation of some amino acids like valine, leucine, and isoleucine. Interestingly it was reported that many PHA-accumulating bacteria harbor more than one β-ketothiolase homologue in their genomes (Lindenkamp et al., 2010).

C. necator (formerly known as *Ralstonia eutropha* or *Alcaligenes eutrophus*) was famous as a potential single cell protein (SCP) in the 1970s and remained in the limelight as a producer of PHA (Kunasundari et al., 2013); a material that resembles commodity plastics such as polypropylene. Although tremendous achievements have been attained in the past few decades in the efficient production of PHA, but the protein structure is not fully understood. The main objective of this study was to develop an understanding of β-ketothiolases in PHA metabolism in *C. necator*.

METHODOLOGY

Data base search

β-Ketothiolase genes were searched among various data base i.e. NCBI, EMBL-EBI etc. and protein sequences were downloaded for further study. Multiple sequence alignment was carried out using ClustalW whereas motifs were searched using KEGG SSDB (Sequence Similarity Data Base).

Homology modelling

SWISS-MODEL was used for the homology modelling of protein 3D structures. Template search with Blast and HHBlits was performed against the SWISS-MODEL template library (SMTL). The target sequences were searched with BLAST against the primary amino acid sequence contained in the SMTL. An initial HHblits profile was built using the procedure outlined by Remmert et al., followed by 1 iteration of HHblits against NR20. The obtained profile was then searched against all profiles of the SMTL (Biasini et al., 2014; Arnold et al., 2006; et al., Benkert et al., 2011; Altschul et al., 1997).

Template selection

For each identified template, the template's quality was predicted from features of the target-template alignment. The templates with the highest quality were been selected for model building.

Model building

Models were built based on the target-template alignment using Promod-II. Conserved coordinates between the target and the template were copied from the template to the model. Insertions and deletions were remodelled using a fragment library. Side chains were then rebuilt. Finally, the geometry of the resulting model was regularized by using a force field. The global and per-residue model quality was assessed using the QMEAN scoring function.

Model building

Models were built based on the target-template alignment using Promod-II. Conserved coordinates between the target and the template were copied from the template to the model. Insertions and deletions were remodelled using a fragment library. Side chains were then rebuilt. Finally, the geometry of the resulting model was regularized by using a force field. The global and per-residue model quality was assessed using the QMEAN scoring function.

3D structure analysis

3D structure analysis was carried out using PDBsum database that provides an overview of the contents of each 3D macromolecular structure deposited in the Protein Data Bank (de Beer et al., 2013).

RESULTS AND DISCUSSION

Reputed public databases such as NCBI, EMBL etc. consist millions of unique protein sequences and number are growing rapidly. Eight β-Ketothiolase genes in *C. necator* were reported among these databases (Table 1).

These genes were categories as bktB, thlA, phaA, phaA1, and phaA2. Protein sequences for these genes were downloaded from NCBI and all sequences were aligned using multiple sequence alignment tool ClustalW to get pairwise alignments score (Table 2), whereas motifs were searched using KEGG SSDB that contains the information about amino acid sequence similarities among all protein-coding genes in the complete genomes (Kanehisa et al., 2002) (Table 3).

Highest sequence similarity was reported between Gene ID: 4248815 and Gene ID: 10917862. Further motif search (de Beer et al., 2013) revealed functional relation among these genes as all genes except Gene ID: 10918233 were found to possess common Thiolase, N-terminal domain, Beta-ketoacyl synthase, N-terminal domain and Thiolase, C-terminal domain. Multiple sequence alignment and motif search both indicated difference between Gene ID: 10918233 and rest

sequences. phaA2 (ID: 10918233) was found to have more resemblance with Poly-beta-hydroxybutyrate polymerase than β-ketothiolase.

Table 1. β-ketothiolases/acetyl-CoA acetyltransferases in *Cupriavidus necator*

Name/ NCBI-Gene ID	Organism name	Description	Location
bktB ID: 4248815	*Ralstonia eutropha H16*	beta-ketothiolase	Chromosome 1, NC_008313.1 (1565652 to 1566836)
thlA ID: 10920220	*Cupriavidus necator N-1*	acetyl-CoA acetyltransferase ThlA	NC_015724.1 (129331 to 130515)
bktB ID: 10917862	*Cupriavidus necator N-1*	acetyl-CoA acetyltransferase BktB	Chromosome 1, NC_015726.1 (1485802 to 1486986)
phaA ID: 4249783	*Ralstonia eutropha H16*	acetyl-CoA acetyltransferase	Chromosome 1, NC_008313.1 (1559207 to 1560388)
phaA ID: 10921806	*Cupriavidus necator N-1*	acetyl-CoA acetyltransferase PhaA	Chromosome 2, NC_015723.1 (1406745 to 1407953)
phaA ID: 10916079	*Cupriavidus necator N-1*	acetyl-CoA acetyltransferase	NC_015727.1 (1151786 to 1152964, complement)
phaA2 ID: 10918233	*Cupriavidus necator N-1*	poly(3-hydroxyalkanoate) polymerase PhaA	Chromosome 1, NC_015726.1 (1901422 to 1903230)
phaA1 ID: 10917857	*Cupriavidus necator N-1*	acetyl-CoA acetyltransferase PhaA	Chromosome 1, NC_015726.1 (1479937 to 1481118)

Table 2. ClustalW pairwise alignments

SeqA	NCBI-Gene ID	Length	SeqB	NCBI-Gene ID	Length	Score
1	4248815	394	2	10920220	394	68.78
1	4248815	394	3	10917862	394	98.98
1	4248815	394	4	4249783	393	52.16
1	4248815	394	5	10921806	402	45.94
1	4248815	394	6	10916079	392	35.71
1	4248815	394	7	10918233	602	6.6
1	4248815	394	8	10917857	393	52.16
2	10920220	394	3	10917862	394	68.27
2	10920220	394	4	4249783	393	50.89
2	10920220	394	5	10921806	402	47.72
2	10920220	394	6	10916079	392	37.76
2	10920220	394	7	10918233	602	9.14
2	10920220	394	8	10917857	393	51.15
3	10917862	394	4	4249783	393	51.91
3	10917862	394	5	10921806	402	45.94
3	10917862	394	6	10916079	392	35.46
3	10917862	394	7	10918233	602	6.6
3	10917862	394	8	10917857	393	51.91
4	4249783	393	5	10921806	402	45.04
4	4249783	393	6	10916079	392	46.68

4	4249783	393	7	10918233	602	6.62
4	4249783	393	8	10917857	393	99.75
5	10921806	402	6	10916079	392	33.42
5	10921806	402	7	10918233	602	7.21
5	10921806	402	8	10917857	393	44.78
6	10916079	392	7	10918233	602	5.61
6	10916079	392	8	10917857	393	46.68
7	10918233	602	8	10917857	393	6.36

Table 3. Motif in various *Cupriavidus necator* β-ketothiolases

NCBI-Gene ID	Motif id	From	To	Definition	E value
4248815	pf:Thiolase_N	3	263	Thiolase, N-terminal domain	3.50E-95
	pf:ketoacyl-synt	63	122	Beta-ketoacyl synthase, N-terminal domain	0.0003
	pf:Thiolase_C	271	393	Thiolase, C-terminal domain	1.70E-46
	pf:ACP_syn_III_C	302	390		0.24
10920220	pf:Thiolase_N	4	262	Thiolase, N-terminal domain	2.10E-90
	pf:ketoacyl-synt	81	122	Beta-ketoacyl synthase, N-terminal domain	0.0076
	pf:ACP_syn_III	85	129		0.31
	pf:Thiolase_C	271	393	Thiolase, C-terminal domain	1.20E-45
	pf:ACP_syn_III_C	302	390		0.24
10917862	pf:Thiolase_N	3	263	Thiolase, N-terminal domain	5.80E-95
	pf:ketoacyl-synt	63	122	Beta-ketoacyl synthase, N-terminal domain	0.00046
	pf:Thiolase_C	271	393	Thiolase, C-terminal domain	1.10E-46
	pf:ACP_syn_III_C	302	390		0.24
4249783	pf:Thiolase_N	1	262	Thiolase, N-terminal domain	5.20E-117
	pf:Ketoacyl-synt_C	19	51	Beta-ketoacyl synthase, C-terminal domain	0.79
	pf:CMD	32	62	Carboxymuconolactone decarboxylase family	0.27
	pf:ketoacyl-synt	77	119	Beta-ketoacyl synthase, N-terminal domain	0.00033
	pf:Thiolase_C	270	392	Thiolase, C-terminal domain	8.30E-55
	pf:ACP_syn_III_C	301	388		0.13
10921806	pf:Thiolase_N	1	267	Thiolase, N-terminal domain	7.60E-73
	pf:ketoacyl-synt	80	146	Beta-ketoacyl synthase, N-terminal domain	0.00051
	pf:ACP_syn_III	86	121		0.05
	pf:Thiolase_C	274	401	Thiolase, C-terminal domain	7.70E-44
	pf:SnoaL_4	300	366		0.087
10916079	pf:Thiolase_N	6	262	Thiolase, N-terminal domain	1.70E-86
	pf:ketoacyl-synt	82	128	Beta-ketoacyl synthase, N-terminal domain	0.00024
	pf:Thiolase_C	269	389	Thiolase, C-terminal domain	5.70E-43
	pf:ACP_syn_III_C	300	383		0.13
10918233	pf:PhaC_N	117	286	Poly-beta hydroxybutyrate polymerase (PhaC) N-terminus	6.60E-62
	pf:Abhydrolase_6	282	523	Alpha/beta hydrolase family	1.00E-06

	pf:Abhydrolase_1	292	530	alpha/beta hydrolase fold	7.30E-15
10917857	pf:Thiolase_N	1	262	Thiolase, N-terminal domain	7.70E-116
	pf:Ketoacyl-synt_C	19	51	Beta-ketoacyl synthase, C-terminal domain	0.79
	pf:CMD	32	62	Carboxymuconolactone decarboxylase family	0.27
	pf:ketoacyl-synt	77	119	Beta-ketoacyl synthase, N-terminal domain	0.00033
	pf:Thiolase_C	270	392	Thiolase, C-terminal domain	8.30E-55
	pf:ACP_syn_III_C	301	388		0.13

Homology models were built based on the target-template alignment whereas the templates with the highest quality were selected on the basis of various parameters such as GMQE (Global Model Quality Estimation), QMEAN (Qualitative Model Energy Analysis), sequence identity, sequence similarity and coverage etc. (Guex et al., 1997; Sali et al., 1993). Three-dimensional (3D) models for all genes are illustrated in Figure 1.

General approach for prediction of a model depends on the various parameters and analysis of identified template structures. Parameters for target-template alignment used to build suitable models are given in Table 4. PDBsum database summaries 3D structures of proteins and provide detailed structural analysis of each protein chain. Significantly much duplication of redundant information has been removed in this database therefore only a representative chain is described in details (Laskowski et al., 2009). Common sequence variation can be seen in Table 5-7.

SWISS-MODEL homo-oligomeric structure of the target protein was predicted based on the analysis of pairwise interfaces of the identified template structures. For each relevant interface between polypeptide chains (interfaces with more than 10 residue-residue interactions), the QscoreOligomer was predicted from features such as similarity to target and frequency of observing this interface in the identified templates. The prediction was performed with a random forest regressor using these features as input parameters to predict the probability of conservation for each interface. The QscoreOligomer of the whole complex was then calculated as the weight-averaged QscoreOligomer of the interfaces. The oligomeric state of the target is predicted to be the same as in the template when QscoreOligomer was predicted to be higher or equal to 0.5 (Mariani et al., 2011).

Table 4. Parameters for target-template alignment

NCBI-GeneID	Oligo-State	Ligands	GMQE	QMEAN4
4248815	Homo-tetramer	None	0.79	-3.95
10920220	Monomer	None	0.80	None
10917862	Homo-tetramer	None	0.79	-3.77
4249783	Homo-tetramer	None	0.84	-1.03
10921806	Homo-tetramer	None	0.75	-4.10
10916079	Homo-tetramer	None	0.77	-2.97
10918233	Homo-tetramer	None	0.22	-7.67
10917857	Homo-tetramer	None	0.84	-1.00

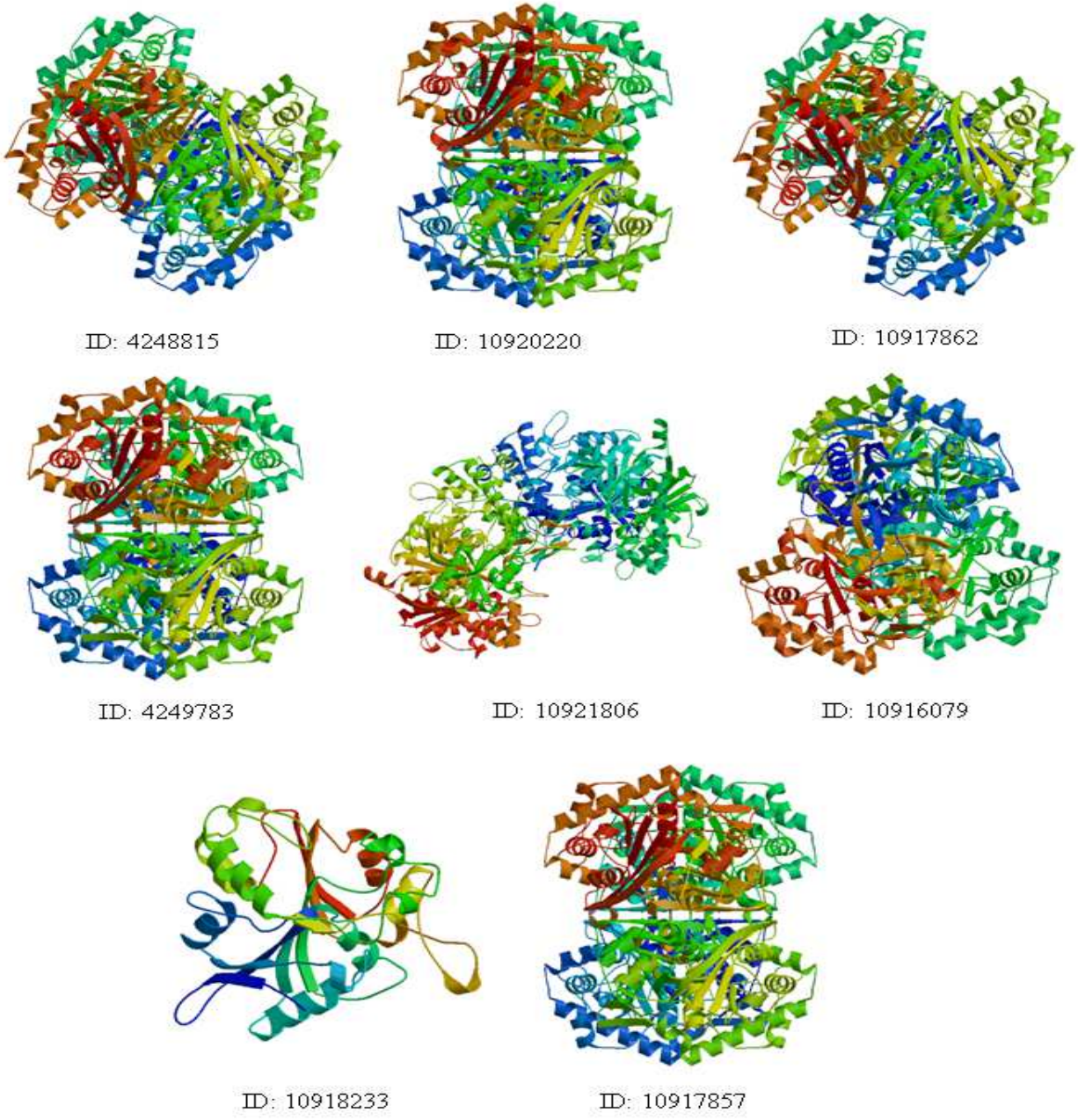

Figure 1. Models for β-ketothiolases in *Cupriavidus necator*

Table 5. Beta sheets in *Cupriavidus necator* PHB depolymerases

NCBI-Gene ID	Sheet	No. of strands	Type	Barrel	Topology
4248815	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	8	Antiparallel	Y	1 -2X -1 -2X 1 -2X -1
	C	2	Antiparallel	N	1
10920220	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	8	Antiparallel	N	-1 2X 1 2X -1 2X 1
	C	2	Antiparallel	N	1
	D	2	Antiparallel	N	1

10917862	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	8	Antiparallel	Y	1 -2X -1 -2X 1 -2X -1
	C	2	Antiparallel	N	1
4249783	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	4	Antiparallel	N	-1 2X 1
	C	2	Antiparallel	N	1
	D	2	Antiparallel	N	1
10921806	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	8	Antiparallel	Y	1 -2X -1 -3 -1 2X 1
10916079	A	14	Mixed	N	0 3X 1X -2X -1 -2X 10X -2X -1X -1X 2X 1 2X
	B	4	Antiparallel	N	-1 2X 1
	C	2	Antiparallel	N	1
	D	4	Mixed	N	-3X 1X 1
10918233	A	7	Mixed	N	-2 1X -2X -1X -1X -1X
	B	2	Antiparallel	N	1
10917857	A	20	Mixed	N	0 3X 1X -2X -1 -2X -3X 1X 1 11 -2X -1X -1X 2X 1 2X 3X -1X -1
	B	4	Antiparallel	N	-1 2X 1
	C	2	Antiparallel	N	1
	D	2	Antiparallel	N	1

Table 6. Beta-alpha-beta units in *Cupriavidus necator* β-ketothiolases

NCBI-Gene ID	Strand 1			Strand 2			No. of helices	No. of residues	
	Start	End	Length	Start	End	Length		Loop	Helix
4248815	His 51	Gly 55	5	Ala 83	Asn 87	5	1	27	11
	Val 314	Ala 317	4	Tyr 374	Ile 381	8	3	56	40
10920220	His 52	Gly 56	5	Ala 84	Asn 88	5	1	27	11
	Val 314	Ser 317	4	Tyr 374	Ile 381	8	3	56	40
10917862	His 51	Gly 55	5	Ala 83	Asn 87	5	1	27	10
	Val 314	Ala 317	4	Tyr 374	Ile 381	8	3	56	40
4249783	Glu 50	Gly 54	5	Ala 81	Asn 85	5	1	26	10
	Leu 313	Ile 316	4	Lys 373	Ile 380	8	3	56	40
10921806	Asp 50	Ala 54	5	Gly 81	Asp 85	5	1	26	10
	Leu 317	Leu 320	4	Gly 383	Ile 389	7	3	62	37
10916079	Ala 55	Gly 58	4	Ala 85	Ser 89	5	1	26	10

	Val 310	Ile 315	6	Arg 372	Leu 379	8	3	56	33
10918233	Pro 260	Val 264	5	Gln 292	Ile 295	4	1	27	9
	Val 333	Ala 338	6	Val 360	Met 366	7	1	21	12
	Lys 486	Gly 491	6	Arg 512	Ser 518	7	1	20	10
10917857	Glu 50	Gly 54	5	Ala 81	Asn 85	5	1	26	10
	Leu 313	Ile 316	4	Lys 373	Ile 380	8	3	56	40

Table 7. Beta strands in *Cupriavidus necator* β-ketothiolases

NCBI-Gene ID	Start	End	Sheet	No. of residues	Edge	Sequence
4248815	Val5	Arg12	A	8	No	VVVVSGVR
	Gly16	Thr17	A	2	Yes	GT
	His51	Gly55	A	5	No	HVVFG
	Ala83	Asn87	A	5	No	ALTVN
	Ala112	Ser119	A	8	No	AIGGGAES
	Tyr125	Ala127	B	3	No	YLA
	Ala139	Asp143	B	5	No	AGLVD
	Val203	Ser205	C	3	Yes	VVS
	Val212	Phe214	C	3	Yes	VTF
	Asn252	Glu262	A	11	No	NDAAAAVVMME
	Ala275	Gly284	A	10	No	ARLVSYGHAG
	Val314	Ala317	A	4	Yes	VIEA
	Tyr374	Ile381	A	8	No	YALVTMCI
	Gln385	Glu392	A	8	No	QGIAAIFE
10920220	Ile6	Arg13	A	8	No	IFVVGAAR
	Gly17	Thr18	A	2	Yes	GT
	His52	Gly56	A	5	No	HVVMG
	Ala84	Asn88	A	5	No	AFNVN
	Ile112	Ser120	A	9	No	IAIGAGSES

	Tyr126	Asp128	B	3	No	YFD
	Val143	Asp144	B	2	Yes	VD
	His151	Asp152	C	2	Yes	HD
	Met157	His158	C	2	Yes	MH
	Val204	Val206	D	3	Yes	VEV
	Val213	Phe215	D	3	Yes	VLF
	Asn252	Glu262	A	11	No	NDGAGAVVLAE
	Ala275	Gly284	A	10	No	ARLVGYAHAG
	Val314	Ser317	A	4	Yes	VIES
	Tyr374	Ile381	A	8	No	YALVTMCI
	Gln385	Glu392	A	8	No	QGIAAIFE
10917862	Val5	Arg12	A	8	No	VVVVSGVR
	Gly16	Thr17	A	2	Yes	GT
	His51	Gly55	A	5	No	HVVFG
	Ala83	Asn87	A	5	No	ALTVN
	Ala112	Ser119	A	8	No	AIGGGAES
	Tyr125	Ala127	B	3	No	YLA
	Ala139	Asp143	B	5	No	AGLVD
	Val203	Ser205	C	3	Yes	VVS
	Val212	Phe214	C	3	Yes	VTF
	Asn252	Glu262	A	11	No	NDAAAAVVMME
	Ala275	Gly284	A	10	No	ARLVSYGHAG
	Val314	Ala317	A	4	Yes	VIEA
	Tyr374	Ile381	A	8	No	YALVTMCI
	Gln385	Glu392	A	8	No	QGIAAIFE
4249783	Val4	Arg11	A	8	No	VVIVSAAR
	Gly15	Lys16	A	2	Yes	GK
	Glu50	Gly54	A	5	No	EVIMG
	Ala81	Asn85	A	5	No	AMTIN

	Ile109	Asn117	A	9	No	IVVAGGQEN
	His123	Leu125	B	3	No	HVL
	Val140	Asp141	B	2	Yes	VD
	Trp149	Asp150	C	2	Yes	WD
	Tyr155	His156	C	2	Yes	YH
	Val202	Leu203	D	2	Yes	VL
	Val212	Ala213	D	2	Yes	VA
	Asn251	Ser261	A	11	No	NDGAAAVVVMS
	Ala274	Gly283	A	10	No	ATIKSYANAG
	Leu313	Ile316	A	4	Yes	LMEI
	Lys373	Ile380	A	8	No	KGLASLCI
	Met384	Glu391	A	8	No	MGVALAVE
10921806	Ala4	Arg11	A	8	No	AAIVTPLR
	Gly15	Thr16	A	2	Yes	GT
	Asp50	Ala54	A	5	No	DVVFA
	Gly81	Asp85	A	5	No	GMQLD
	Val109	Ser117	A	9	No	VVIAGGVES
	Tyr123	Thr125	B	3	No	YYT
	Val137	Asp141	B	5	No	VRFFD
	Asn255	Ala265	A	11	No	NDASAACLIVA
	Ala278	Gly287	A	10	No	ASLVGWAAAG
	Leu317	Leu320	A	4	Yes	LVEL
	Gly383	Ile389	A	7	No	GLETMCI
	Gln393	Glu400	A	8	No	QGIAAVFE
10916079	Val9	Arg15	A	7	No	VIVGARR
	Gly19	Ala20	A	2	Yes	GA
	Ala55	Gly58	A	4	No	AIMG
	Ala85	Ser89	A	5	No	ATTIS
	Val113	Ser121	A	9	No	VAVAGGMES

	His127	Val129	B	3	No	HIV
	Leu144	Asp145	B	2	Yes	LD
	Glu153	Asp154	C	2	Yes	ED
	His159	Leu160	C	2	Yes	HL
	Ser250	Met259	A	10	No	SDGAAALVLM
	Ala273	Arg274	A	2	Yes	AR
	Ala279	Ala282	D	4	Yes	ATRA
	Val310	Ile315	D	6	Yes	VDLYEI
	Arg372	Leu379	D	8	No	RGLASLCL
	Glu383	Met387	D	5	No	EAVAM
10918233	Ile247	Tyr249	A	3	Yes	IQY
	Pro260	Val264	A	5	No	PQLIV
	Gln292	Ile295	A	4	No	QVFI
	Val333	Ala338	A	6	No	VNLHGA
	Val360	Met366	A	7	No	VHAATLM
	Ala465	Thr466	B	2	Yes	AT
	His475	Ala476	B	2	Yes	HA
	Lys486	Gly491	A	6	No	KYVLAG
	Arg512	Ser518	A	7	Yes	RTEFVLS
10917857	Val4	Arg11	A	8	No	VVIVSAAR
	Gly15	Lys16	A	2	Yes	GK
	Glu50	Gly54	A	5	No	EVIMG
	Ala81	Asn85	A	5	No	AMTIN
	Ile109	Asn117	A	9	No	IVVAGGQEN
	His123	Leu125	B	3	No	HVL
	Val140	Asp141	B	2	Yes	VD
	Trp149	Asp150	C	2	Yes	WD
	Tyr155	His156	C	2	Yes	YH
	Val202	Leu203	D	2	Yes	VL

Val212	Ala213	D	2	Yes	VA
Asn251	Ser261	A	11	No	NDGAAAVVVMS
Ala274	Gly283	A	10	No	ATIKSYANAG
Leu313	Ile316	A	4	Yes	LMEI
Lys373	Ile380	A	8	No	KGLASLCI
Met384	Glu391	A	8	No	MGVALAVE
Ile351	Arg369	H	19	28.48	1.53

Ligands present in the template structure are transferred by homology to the model when the following criteria are met: (a) The ligands are annotated as biologically relevant in the template library, (b) the ligand is in contact with the model, (c) the ligand is not clashing with the protein, (d) the residues in contact with the ligand are conserved between the target and the template (Schwede et al., 2003). Criteria were not satisfied therefore ligands were not included in the models (Schwede et al., 2003) (Table 4).

CONCLUSION

The results show diversity of PHA β--ketothiolase in C. necator. These β--ketothiolases are known as bktB, thlA, phaA, phaA1, and phaA2. Pairwise alignments, motif search, homology models, secondary structures, beta sheets, beta strands and beta-alpha-beta units confirms the highest sequence similarity and homology between Gene ID: 4248815 and Gene ID: 10917862.

Common Thiolase, N-terminal domain, Beta-ketoacyl synthase, N-terminal domain and Thiolase, C-terminal domain in all genes except phaA2 (Gene ID: 10918233) indicate their functional similarity. phaA2 (ID: 10918233) was found to have more resemblance with Poly-beta-hydroxybutyrate polymerase than β--ketothiolase. The diversity of β--ketothiolase is interesting to study as it triggers the first step of PHA biosynthesis, and previous reposts indicates that various β--ketothiolases are substrate specific and produce different types of PHAs. Therefore not only identification of such genes but also categorization and selection of microbes and substrate feeding on the basis of these genes required more focus research. Interesting study PHA biosynthesis ensures environment friendly & sustainable production of PHA based bioplastics and economic opportunities.

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25: 3389–3402.

Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22: 195–201.

Benkert P, Biasini M, Schwede, T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27: 343–350.

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research doi:10.1093/nar/gku340.

de Beer TA, Berka K, Thornton JM, Laskowski RA. 2014. PDBsum additions. Nucleic Acids Research 42: D292–296.

Girdhar A, Bhatia M, Nagpal S,

Kanampalliwar A, Tiwari A. 2013. Process Parameters for Influencing Polyhydroxyalkanoate Producing Bacterial Factories: An Overview. Journal of Petroleum & Environmental Biotechnology 4:155. doi: 10.4172/2157-7463.1000155.

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714–2723.

Jain R, Kosta S, Tiwari A. 2010. Polyhydroxyalkanoates: A way to sustainable development of bioplastics. Chronicles of Young Scientists 1: 10–15.

Kanehisa M, Goto S, Kawashima S, Nakaya A. 2002. The KEGG databases at GenomeNet. Nucleic Acids Research 30(1): 42–46.

Kunasundari B, Murugaiyah V, Kaur G, Maurer FHJ, Sudesh K. 2013. Revisiting the single cell protein application of *Cupriavidus necator* H16 and recovering bioplastic granules simultaneously. PLOS ONE 8(10): e78528.

Laskowski RA. 2009. PDBsum new things. Nucleic Acids Research 37: D355–359.

Lindenkamp N, Volodina E, Steinbüchel A. 2012. Genetically modified strains of Ralstonia eutropha H16 with β-ketothiolase gene deletions for production of copolyesters with defined 3-hydroxyvaleric acid contents. Applied and Environmental Microbiology 78(15):5375–5383

Madkour MH, Heinrich D, Alghamdi MA, Shabbaj, Steinbüchel A. 2013. PHA recovery from biomass. Biomacromolecules 14(9): 2963–2972.

Mariani V, Kiefer F, Schmidt T, Haas J, Schwede T. 2011. Assessment of template based protein structure predictions in CASP9. Proteins 79 (Suppl 10): 37–58.

Rehm BH. 2007. Biogenesis of microbial polyhydroxyalkanoate granules: a platform technology for the production of tailor-made bioparticles. Current Issues in Molecular Biology 9(1): 41–62.

Remmert M, Biegert A, Hauser A, Soding J. 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9: 173–175.

Sali A and Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. Journal of Molecular Biology 234: 779–815.

Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Research 31(13): 3381–338.