J. Mater. Environ. Sci. 6 (3) (2015) 686-692
ISSN : 2028-2508
CODEN: JMESCN

Acetoacetyl-CoA reductases in *Cupriavidus necator*

Roopesh Jain*, Archana Tiwari
School of Biotechnology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India.

Received 03 June 2014; Revised 04 September 2014; Accepted 26 September 2014.
**Corresponding Author. E-mail: science.roopesh@gmail.com ; Tel: (+91-9819496231)*

Abstract
Polyhydroxyalkanoates (PHAs) have been recognized as good substitutes for the non-biodegradable and petrochemically produced polymers. NADPH-dependent acetoacetyl-coenzyme A (acetoacetyl-CoA) reductase (PhaB) is one of the key enzyme in the synthesis of poly(3-hydroxybutyrate) [P(3HB)], along with β-ketothiolase (PhaA) and polyhydroxyalkanoate synthase (PhaC). The PhaB-encoding gene has been reported in many bacteria, including *Cupriavidus necator*, and is usually located in the phb operon together with β-ketothiolase (PhaA) and PHA synthase (PhaC). In present study PHA-specific acetoacetyl-CoA reductase (PhaB) was studied on the basis of available sequences in various database to understand their diversity.

Key words: Polyhydroxyalkanoates, Biosynthesis, Biodegradation, *Cupriavidus necator*, Acetoacetyl-CoA reductase (PhaB)

Introduction
Conventional plastics have been important materials for the society since its development not only because of their low reactivity and higher molecular weight but also for their cost efficiency and durability [1]. On the other side, their ubiquitous persistence in the environment has upraised worries about their end-of-life disposal. In spite of their excellent stability, durability and weight saving properties; an environmental dilemma with more far-reaching implications is climate change. The need for rapid and deep greenhouse gas (GHG) emissions cuts is one of the drivers for the search for bio-based plastic. The search for biodegradable plastics from sustainable resources and their comprehensive introduction to the marketplace has been considered a suitable approach [1, 2]. A special class of optically active biopolymers called Polyhydroxyalkanoates (PHAs) have showed some of the extraordinary similarities to the well-known synthetic polymers i.e. polypropylene and polyethylene. In addition, their disposal as bio-waste made them increasingly attractive in the pursuit of sustainable development of bioplastics [3]. Poly-3-hydroxybutirate (PHB) and its copolymers with 3-hydroxyvalerate (3HV) known as poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV), have been documented as the best known representatives of PHA family [4].
In *Cupriavidus necator*, the biosynthesis of the (*R*)-3-hydroxybutyrate coenzyme A [(*R*)-3-HB-CoA] precursor is conducted by multiple steps [5]. First, two acetyl-CoA molecules are condensed into one acetoacetyl-CoA molecule by the enzyme β-ketothiolase (PhaA). The acetoacetyl-CoA is then reduced to (*R*)-3-HB-CoA by a PHA-specific acetoacetyl-CoA reductase (PhaB) [6]. The resulting (*R*)-3-HB-CoA is subsequently incorporated into PHB, catalysed by PHA synthases [4,6]. PhaB is members of a vast protein family, the SDR superfamily, which catalyse a wide variety of NAD(P)(H)-dependent oxidation/reduction reactions [7]. NADPH-dependent acetoacetyl-coenzyme A (AcAc-CoA) reductase (PhaB) stereo-selectively reduces the 3-ketone group of acetoacetyl-CoA to synthesize (R)-3-hydroxybutyryl(3HB)-CoA, which is known to be a monomer precursor of microbial polyester polyhydroxyalkanoate (PHA) [6,8]. In bacteria, the PHA-specific acetoacetyl-CoA reductase has two other conserved residues (Val and Ile) in the cofactor binding sequence (ValThrGlyXXXGlyIleGly). Most bacterial PhaB proteins use NADPH as the cofactor, except that PhaB from *Chromatium vinosum* has been described to be NADH dependent. The main objective of this study was to develop an understanding of PHA-specific acetoacetyl-CoA reductase and their diversity in *Cupriavidus necator*.

2. Materials and methods
2.1. Data base search
Acetoacetyl-CoA reductase 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 (SSDB).

2.2. 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 [9-12].

2.3. 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 selected for model building.

2.4. 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.

2.5. 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 [13].

3. Results and discussion
Reputed public databases such as NCBI, EMBL etc. consist millions of unique protein sequences and number are growing rapidly. Seven acetoacetyl-CoA reductases genes in *Cupriavidus necator* were reported among these databases (Table 1).

Table 1: Acetoacetyl-CoA reductases in *Cupriavidus necator*

Name/ NCBI-Gene ID	Organism name	Description	Location
phbB ID: 10915757	*Cupriavidus necator* N-1	Acetoacetyl-CoA reductase PhbB	NC_015727.1 (803923 to 804666)
phaB1 ID: 10917858	*Cupriavidus necator* N-1	Acetoacetyl-CoA reductase PhaB	Chromosome 1, NC_015726.1 (1481195 to 1481935)
phbB ID: 10920675	*Cupriavidus necator* N-1	Acetoacetyl-CoA reductase	Chromosome 2, NC_015723.1 (208876 to 209607)
phaB3 ID: 4250155	*Ralstonia eutropha* H16	Acetoacetyl-CoA reductase	Chromosome 1, NC_008313.1 (2364912 to 2365622, complement)
phaB2 ID: 4249785	*Ralstonia eutropha* H16	Acetoacetyl-CoA reductase	Chromosome 1, NC_008313.1 (2174303 to 2175049)
phaB1 ID: 4249784	*Ralstonia eutropha* H16	Acetyacetyl-CoA reductase	Chromosome 1, NC_008313.1 (1560463 to 1561203)
NEWENTRY	*Cupriavidus necator*	Record to support submission of GeneRIFs for a gene not in Gene (*Alcaligenes eutrophus*; *Cupriavidus necator; Hydrogenomonas eutropha; Ralstonia eutropha; Wautersia eutropha*. Use when strain, subtype, isolate, etc. is unspecified, or when different from all specified ones in Gene.).	

These genes were categories as phbB, phbB1, phbB2, phbB3 and 'NEWENTRY'. Protein sequences for these genes were downloaded from NCBI except 'NEWENTRY' as this entry doesn't have sequences in database. 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 [14].
Highest sequence similarity was reported between ID: 10917858 and ID: 4249784 followed by ID: 10915757 and ID: 4250155. Further motif search revealed functional relation among these genes with common motifs (Table 3) such as short chain dehydrogenase, KR domain and NAD dependent epimerase etc.

Table 2: ClustalW pairwise alignments

SeqA	NCBI-Gene ID	Length	SeqB	NCBI-Gene ID	Length	Score
1	10915757	247	2	10917858	246	56.91
1	10915757	247	3	10920675	243	32.1
1	10915757	247	4	4250155	236	84.75
1	10915757	247	5	4249785	248	63.56
1	10915757	247	6	4249784	246	56.91
2	10917858	246	3	10920675	243	33.74
2	10917858	246	4	4250155	236	52.12
2	10917858	246	5	4249785	248	52.44
2	10917858	246	6	4249784	246	97.56
3	10920675	243	4	4250155	236	32.63
3	10920675	243	5	4249785	248	36.63
3	10920675	243	6	4249784	246	34.16
4	4250155	236	5	4249785	248	64.41
4	4250155	236	6	4249784	246	52.12
5	4249785	248	6	4249784	246	52.44

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. [15,16]. Three-dimensional (3D) models for all genes have been illustrated in Figure 1.

Table 3:Motifs in various acetoacetyl-CoA reductases

NCBI-Gene ID	Motif id	From	To	Definition	E value
4250155	pf:adh_short	1	160	short chain dehydrogenase	7.70E-36
	pf:KR	1	159	KR domain	8.80E-21
	pf:adh_short_C2	3	234		2.00E-29
4249785	pf:adh_short	6	172	short chain dehydrogenase	8.70E-42
	pf:F420_oxidored	6	46	NADP oxidoreductase coenzyme F420-dependent	0.32
	pf:KR	7	179	KR domain	1.10E-24
	pf:Epimerase	7	201	NAD dependent epimerase/dehydratase family	5.30E-07
	pf:NAD_binding_10	8	187		1.40E-05
	pf:3Beta_HSD	8	76	3-beta hydroxysteroid dehydrogenase/isomerase family	0.041
	pf:adh_short_C2	14	246		1.40E-24
4249784	pf:adh_short	5	171	short chain dehydrogenase	2.50E-41
	pf:Epimerase	6	141	NAD dependent epimerase/dehydratase family	5.80E-08
	pf:KR	8	178	KR domain	6.80E-22
	pf:NAD_binding_10	8	188		5.30E-05
	pf:3Beta_HSD	8	142	3-beta hydroxysteroid dehydrogenase/isomerase family	0.024
	pf:NmrA	8	76	NmrA-like family	0.17
	pf:F420_oxidored	11	49	NADP oxidoreductase coenzyme F420-dependent	0.081
	pf:adh_short_C2	13	244		2.30E-26
10917858	pf:adh_short	5	171	short chain dehydrogenase	2.40E-41
	pf:Epimerase	6	141	NAD dependent epimerase/dehydratase family	5.50E-08
	pf:KR	8	178	KR domain	1.80E-22
	pf:NAD_binding_10	8	189		4.40E-05
	pf:3Beta_HSD	8	142	3-beta hydroxysteroid dehydrogenase/isomerase family	0.032
	pf:F420_oxidored	11	49	NADP oxidoreductase coenzyme F420-dependent	0.2

	pf:adh_short_C2	13	244		4.20E-27
	pf:Snurportin1	39	58	Snurportin1	0.057
10920675	pf:Eno-Rase_NADH_b	4	20	NAD(P)H binding domain of trans-2-enoyl-CoA reductase	0.24
	pf:adh_short	7	161	short chain dehydrogenase	5.40E-22
	pf:KR	7	150	KR domain	1.70E-07
	pf:Epimerase	8	139	NAD dependent epimerase/dehydratase family	7.60E-08
	pf:NAD_binding_10	8	179		1.70E-07
	pf:3Beta_HSD	8	89	3-beta hydroxysteroid dehydrogenase/isomerase family	0.0081
	pf:Polysacc_synt_2	8	77	Polysaccharide biosynthesis protein	0.088
	pf:adh_short_C2	13	236		1.00E-18
10915757	pf:IlvN	3	42	Acetohydroxy acid isomeroreductase, catalytic domain	0.19
	pf:adh_short	5	171	short chain dehydrogenase	2.10E-39
	pf:KR	6	168	KR domain	4.30E-22
	pf:Epimerase	6	141	NAD dependent epimerase/dehydratase family	9.60E-06
	pf:NAD_binding_10	7	190		0.0053
	pf:adh_short_C2	13	245		2.10E-28

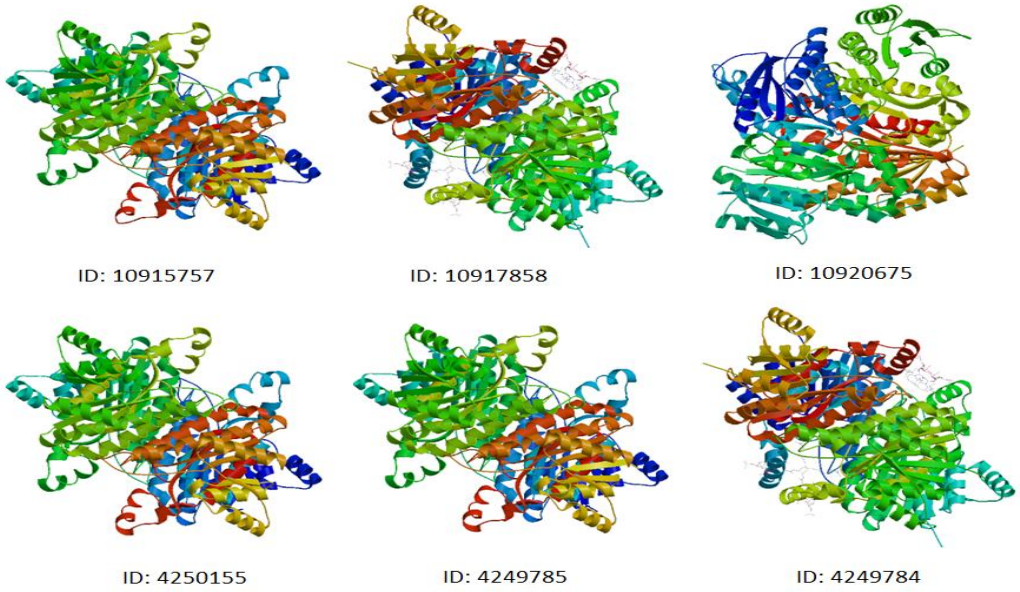

Figure 1: Models for acetoacetyl-CoA reductases in *Cupriavidus necator*

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 have been given in Table 4.

Table 4: Parameters for target-template alignment

NCBI-Gene ID	Oligo-State	Ligands	GMQE	QMEAN4
10915757	Homo-tetramer	None	0.96	0.47
10917858	Homo-tetramer	4 x CAA	0.99	-0.45
10920675	Homo-tetramer	None	0.71	-4.85
4250155	Homo-tetramer	None	0.92	0.20
4249785	Homo-tetramer	None	0.87	-0.17
4249784	Homo-tetramer	4 x CAA	0.99	-0.21

PDBsum database hassummarized 3D structures of proteins and provided structural analysis of each protein chain. Significantly much duplication of redundant information has been removed in this database, therefore only a representative chain was described in details [17]. Common sequence variation can be seen in Table 5-7.

Table 5: Beta sheets in *Cupriavidus necator*acetoacetyl-CoA reductases

NCBI-Gene ID	Sheet	No. of strands	Type	Barrel	Topology
10915757	A	7	Parallel	N	-1X -1X 3X 1X 1X1X
10917858	A	7	Parallel	N	-1X -1X 3X 1X 1X1X
10920675	A	7	Parallel	N	-1X -1X 3X 1X 1X1X
4250155	A	2	Parallel	N	1X
	B	4	Parallel	N	1X 1X1X
4249785	A	7	Parallel	N	-1X -1X 3X 1X 1X1X
4249784	A	7	Parallel	N	-1X -1X 3X 1X 1X1X

Table 6: Beta-alpha-beta units in *Cupriavidus necator*acetoacetyl-CoA reductases

NCBI-Gene ID	Strand 1			Strand 2			No. of helices	No. of residues	
	Start	End	Length	Start	End	Length		Loop	Helix
10915757	Ile 5	Thr 9	5	Asn 29	His 34	6	1	19	14
	Asn 29	His 34	6	Arg 56	Ala 59	4	1	21	11
	Ile 84	Asn 87	4	Gly 133	Ile 138	6	2	45	31
	Gly 133	Ile 138	6	Val 176	Pro 183	8	2	37	30
	Val 176	Pro 183	8	Asn 237	Ile 240	4	3	53	29
10917858	Ile 5	Val 8	4	Arg 29	Cys 34	6	1	20	14
	Arg 29	Cys 34	6	Val 56	Glu 59	4	1	21	12
	Val 82	Asn 87	6	Gly 133	Ile 138	6	3	45	31
	Gly 133	Ile 138	6	Val 176	Pro 183	8	2	37	30
	Val 176	Pro 183	8	Asp 236	Leu 239	4	3	52	32
10920675	Thr 7	Thr 10	4	Val 32	Asp 36	5	1	21	14
	Arg 75	Asn 78	4	Gly 124	Met 129	6	2	45	31
	Gly 124	Met 129	6	Val 166	Pro 173	8	1	36	23
	Val 166	Pro 173	8	Val 229	Val 232	4	3	55	30
4250155	Val 20	His 23	4	Arg 45	Glu 48	4	1	21	11
	Ile 73	Asn 76	4	Gly 122	Ile 127	6	2	45	31
	Gly 122	Ile 127	6	Val 165	Pro 172	8	2	37	30
	Val 165	Pro 172	8	Asn 226	Ile 229	4	3	53	29
4249785	Ile 6	Thr 10	5	Arg 30	His 35	6	1	19	14
	Arg 30	His 35	6	Thr 57	Pro 60	4	1	21	11
	Ile 85	Asn 88	4	Gly 134	Ile 139	6	2	45	31
	Gly 134	Ile 139	6	Ile 177	Pro 184	8	2	37	30
	Ile 177	Pro 184	8	Asn 238	Ile 241	4	3	53	29
4249784	Ile 5	Val 8	4	Arg 29	Cys 34	6	1	20	14
	Arg 29	Cys 34	6	Ile 56	Glu 59	4	1	21	12
	Val 82	Asn 87	6	Gly 133	Ile 138	6	3	45	31
	Gly 133	Ile 138	6	Val 176	Pro 183	8	2	37	30
	Val 176	Pro 183	8	Asp 236	Leu 239	4	3	52	32

The most successful general approach for predicting the structure of proteins involves the detection of homologs of known three-dimensional (3D) structure that is called template-based homology modelling. These methods work on the observation that one can compare protein sequence of interest with the sequences of proteins of experimentally determined structures. If a homolog can be found, an alignment of the two sequences is generated

and used to build a 3D model of the sequence of interest [18]. In present study SWISS-MODEL was used for homology modelling (Figure 1) whereas detailed 3D structural analysis was carried out using PDBsum (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 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 was predicted to be the same as in the template when QscoreOligomer was predicted to be higher or equal to 0.5 [19].

Table 7: Beta strands in *Cupriavidus necator* acetoacetyl-CoA reductases

NCBI-GeneID	Start	End	Sheet	No. of residues
10915757	Ile5	Thr9	A	5
	Asn29	His34	A	6
	Arg56	Ala59	A	4
	Ile84	Asn87	A	4
	Gly133	Ile138	A	6
	Val176	Pro183	A	8
	Asn237	Ile240	A	4
	Ile5	Val8	A	4
	Arg29	Cys34	A	6
	Val56	Glu59	A	4
	Val82	Asn87	A	6
	Gly133	Ile138	A	6
	Val176	Pro183	A	8
	Asp236	Leu239	A	4
10920675	Thr7	Thr10	A	4
	Val32	Asp36	A	5
	Thr48	Gln51	A	4
	Arg75	Asn78	A	4
	Gly124	Met129	A	6
	Val166	Pro173	A	8
	Val229	Val232	A	4
4250155	Val20	His23	A	4
	Arg45	Glu48	A	4
	Ile73	Asn76	B	4
	Gly122	Ile127	B	6
	Val165	Pro172	B	8
	Asn226	Ile229	B	4
4249785	Ile6	Thr10	A	5
	Arg30	His35	A	6
	Thr57	Pro60	A	4
	Ile85	Asn88	A	4
	Gly134	Ile139	A	6
	Ile177	Pro184	A	8
	Asn238	Ile241	A	4
4249784	Ile5	Val8	A	4
	Arg29	Cys34	A	6
	Ile56	Glu59	A	4
	Val82	Asn87	A	6
	Gly133	Ile138	A	6
	Val176	Pro183	A	8
	Asp236	Leu239	A	4

Ligands present in the template structure were transferred by homology to the model upon meeting following criteria: (a) The ligands were annotated as biologically relevant in the template library, (b) the ligand was in

contact with the model, (c) the ligand was not clashing with the protein, (d) the residues in contact with the ligand were conserved between the target and the template [20]. Criteria were satisfied for Gene ID:10917858 and Gene ID:4249784 therefore ligands for these two were included during homology modelling while rest were excluded (Table 4).

Conclusion

In this study homologues of acetoacetyl-CoA reductases involved in polyhydroxyalkanoate biosynthesis were studied in detail at single platform. The results showed diversity of acetoacetyl-CoA reductases in *Cupriavidus necator*. These acetoacetyl-CoA reductases homologues were identified as phbB, phbB1, phbB2, phbB3 and 'NEWENTRY'. Multiple sequence alignment and homology models confirmed highest similarities between ID: 10917858 and ID: 4249784 followed by ID: 10915757 and ID: 4250155. Common motifs such as short chain dehydrogenase, KR domain and NAD dependent epimerase etc. indicated functional similarity among homologues and suggested their associate role in biopolymer production. NADPH-dependent acetoacetyl-CoA (AcAc-CoA) reductase (PhaB) was stereoselectively reported to reduces 3-ketone group of acetoacetyl-CoA to synthesize (R)-3-hydroxybutyryl(3HB)-CoA; a monomer precursor of microbial polyester polyhydroxyalkanoate (PHA). The PhaB, PhaA and PHA PhaC; these three enzymes have been catalysed successive reactions synthesizing P(3HB) from acetyl-CoA [21]. To use 3HB-copolymer industrially, it has beensuggested to improve the enzymatic activities of genes involved in biosynthesis operon of bioplastics producing microorganism. Structure based design may be one of the major strategiesto engineer the enzyme activities. In several studies, enhanced expression of either PhaB by itself or both PhaA and PhaB has been reported toincrease P(3HB) production as the result of an increase in gene dosage and codon optimization. It is on the microbial level where the tools of genetic engineering can be most readily applied. Characterization of these genes homologues involved in production of bioplastics will help scientists in modification of recombinants as per requirement of substrate specificity and expression of genes.

References

1. Jain R., Kosta S., Tiwari A., *Chron. Young. Sci.* 1 (2010) 10-15.
2. Madkour M.H., Heinrich D., Alghamdi M.A., Shabbaj II., Steinbüchel A., *Biomacromolecules* 14(9) (2013) 2963-2972.
3. Rehm B.H., *Curr. Issues Mol. Biol.* 9(1) (2007) 41-62.
4. Jain R., Kosta S., Tiwari A., *Indian J. Cancer* 47 (2010) 87-88.
5. Tiwari A., Ramirez A.M., Jain R., Saxena A., *Key Engineer. Mater.* 517 (2012) 755-762.
6. Mullaney J.A., Rehm B.H., J Biotechnol. 147(1)(2010)31-36.
7. Jain R., Tiwari A., *Int. J. Univ. Pharm. Bio. Sci. (IJUPBS)* 3(3) (2014) 214-223.
8. Lau N.S., Sudesh K., *AMB Express* 2(1) (2012) 41.
9. Biasini M., Bienert S., Waterhouse A., Arnold K., Studer G., Schmidt T., Kiefer F., Cassarino TG., Bertoni M., BordoliL.,and Schwede T., *Nucleic Acids Res.* (2014) doi:10.1093/nar/gku340
10. Arnold K., Bordoli L., Kopp J., Schwede T., *Bioinformatics* 22 (2006) 195-201.
11. Benkert P., Biasini M., Schwede T., *Bioinformatics* 27 (2011) 343-350.
12. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., *Nucleic Acids Res.* 25 (1997) 3389-3402.
13. de Beer T.A., Berka K, Thornton J.M., Laskowski R.A., *Nucleic Acids Res.* 42 (2014) D292-296.
14. Kanehisa M., Goto S., Kawashima S., Nakaya A., *Nucleic Acids Res.* 30(1) (2002) 42-46.
15. Guex N., Peitsch M.C., *Electrophoresis* 18 (1997) 2714-2723.
16. Sali A., Blundell T.L., *J. Mol. Biol.* 234 (1993) 779-815.
17. Laskowski R.A., *Nucleic Acids Res.* 37 (2009) D355-D359.
18. Remmert M., Biegert A., Hauser A., Soding J., *Nat. Methods* 9 (2012) 173-175.
19. Mariani V., Kiefer F., Schmidt T., Haas J., Schwede T., *Proteins* 79 (Suppl 10) (2011) 37-58.
20. Schwede T., Kopp J., Guex N., Peitsch M.C., *Nucleic Acids Res.* 31(13) (2003) 3381-3385.
21. Matsumoto K., Tanaka Y., Watanabe T., Motohashi R., Ikeda K., Tobitani K., Yao M., Tanaka I., Taguchi S., *Appl. Environ. Microbiol.* 79(19) (2013) 6134-613.