Effect of *In Silico* Site-Directed Mutagenesis of Trp164 and Ser165 on the Predicted 3D structure of the C-Terminal Domain of Glutathione s-transferase from *Acidovoras sp.* KKS102

*[1,2]Murtala, Y., [2]Nwanguma, B. C. and [2]Ezeanyika, L. U. S.

[1]Department of Biochemistry, Faculty of Basic Medical Sciences, Bayero University, Kano

[2]Department of Biochemistry, Faculty of Biological Sciences, University of Nigeria, Nsukka

*Corresponding author: myau.bch@buk.edu.ng

ABSTRACT

In silico site-directed mutagenesis of two evolutionarily conserved amino acids, Trp164 and Ser165 in the C-terminal domain of cytosolic glutathiones transferase from *Acidovorax sp.* KKS102, a biphenyl/polychlorobiphenyl degrading bacterium was carried out using a SWISS MODEL DeepView/Swiss-Pdb viewer molecular graphics program. Of all the proteinogenic amino acids used for the mutation, only substitutions for Gly, W164G and S165G were found to likely affect the stability of the C-terminal domain of the both protein models. However, S165P and S165W substitutions at position 165 were theoretically predicted to affect the 3D model of the GSTKKS102 by either introducing a sharp turn in the peptide as proposed for Pro substitution, or by replacing the predicted hydrogen bond interactions contributed by Ser side chain with proposed hydrophobic interactions by the nonpolar aromatic Trp side chain. This hydrophobic interaction may result in the formation of a hydrophobic region in the C-terminal domain of GSTKKS102 models and consequently changes the stability of the protein structure and perhaps, its function.

Keywords: 3D structure, a*cidovoras sp.* KKS102, glutathione s-transferase, site-directed mutagenesis, Trp164, Ser165

INTRODUCTION

The glutathione s-transferases (GSTs) are composed of ubiquitous and versatile proteins that play some vital roles in detoxification of both endogenous and exogenous substrates [1-4]. They normally catalyse the nucleophilic attack of the sulfhydryl group of glutathione (GSH) on the electrophilic centres of a wide range of substrates, making them more soluble and easily removable from the biological systems [3, 5]. In bacteria, some GSTs were involved in reductive dehalogenation reactions, where a chlorine atom from a chlorinated substrate is replaced by the thiol group from the reduced glutathione [5-7]. A homologous gene of BphK (biphenyl upper

pathway K) from *Acidovorax* sp. KKS102 was reported and designated as BphK-KKS [6]. This bacterial gene product was named GST-KKS102 and identified as a beta class GST from *Acidovoras sp.* KKS102, a biphenyl/polychlorobiphenyl degrading organism. The GST-KKS102 was recently documented to show a dechlorination function on various organochlorines. It reacted towards 1 -chloro -2, 4 dinitrobenzene (CDNB), ethacrynic acid, hydrogen peroxide and cumene hydroperoxide [6,9]. *Acidovorax* sp. KKS102 is a biphenyl/polychlorobiphenyl degrading organism isolated from the soil near a refinery in Japan. Sources from the database suggest that *Acidovorax* sp. KKS102 contained at least eleven putative GSTs [10]. The bacterial GSTs are dimeric proteins, which composed of two domains: the N-terminal and C-terminal domains. They have an average molecular weight of 25,000 Da. The N-terminal thioredoxin-like domain serves as a GSH binding pocket termed as G site, while the hydrophobic electrophilic substrate binding site termed H site is located at the C-terminal α-helix domain [11, 12].

Many studies involving mutagenesis were reported on bacterial GSTs in order to identify the role of specific amino acids involved in protein stability and catalysis. Indeed, numerous site-directed mutagenesis studies have shown that some conserved residues in both N-terminal and C-terminal domains of various GSTs have been involved in the stabilization of the GST proteins and/or their co-substrates, which ultimately aids catalysis [12-18].

In recent years, *in-silico* side-directed mutagenesis has been proved to be vital for finding a direction for the intended protein engineering work or to confirm the observed effect of *in-vitro* side-directed mutagenesis on a particular protein entity. *In-silico* side-directed mutagenesis is an essential tool for predicting the effect of mutation on the theoretical 3D model in terms of protein stability and function prior to the actual *in-vitro* side-directed mutagenesis.

This study is aimed to predict the *in silico* effect on substitution of the conserved amino acids, Trp164 (W164) and Ser165 (S165) in the C-terminal domain of the theoretical 3D model of GST-KKS102 in an attempt to understand roles these amino acids in the protein stability.

MATERIALS AND METHODS

Protein Homology Studies

The protein sequence homology study of cytosolic GST-KKS102 was carried out using the European Molecular Biology Laboratory and European Bioinformatics Institute (EMBL-EBI)

Clustal Omega Online Resources (https://www.ebi.ac.uk/Tools/services/web). Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. Four protein sequences, *Bphk*LB400 (PDB file ABE37052.1), *Ralstonia sp.* B15 (PDB file Q9RAF0), *Pseudomonas pseudoalcaligenes* (PDB file Q52037) and *Paraburkholderia xenovoras* LB400 (PDB file Q59721) were used as templates for this sequence homology study.

3D Structure Prediction

The GST-KKS102 is a protein with 202 amino acid residues. Its theoretical 3D model was predicted using the SWISS-MODEL homology modelling alignment interface approach mode [19, 20]. Homology modelling determines structure based on the target sequence, GST-KKS102, possessing homology with template sequence in the structural database. The predicted theoretical 3D model was viewed and manipulated using DeepView/Swiss-Pdb viewer molecular graphics program version 4.1 (SP5) [19].

***In Silico* Site-Directed Mutagenesis**

Prediction of the effect of *in silico* site-directed mutagenesis of the evolutionarily conserved W164 and S165 on the predicted 3D model of GST-KKS102 was carried out using the DeepView/Swiss-Pdb Viewer molecular graphics program version 4.1 (SP5). New amino acid side chains were selected by opting for the mutate tool in accordance with DeepView/Swiss-Pdb viewer user guide protocols.

RESULTS AND DISCUSSION

Sequence Homology Studies

The sequence homology studies revealed at least 40% sequence identity between the cytosolic GST from *Acidovoras sp.* KKS102 and the four template sequences used (Table 1). In bacteria, one of the most important criterion for classifying cytosolic GSTs is the use of percentage sequence similarity, which posits that proteins that shared 40% and above of the sequence similarity belongs to the same class, while less than 20% are classified into a different class [21]. All the four templates used in this analysis belong to beta class GSTs. Thus, going by the above

criterion, cytosolic GST-KKS102 belongs to the beta class, as previously reported by Shehu and Zazali [9].

Table 1: Percentage Sequence Similarity of Cytosolic GST-KKS102 and some Bacterial GST Templates

Template	**% Sequence similarity with GST-KKS102**
*Bphk*LB400	47.26
Ralstonia sp. B15	44.28
Pseudomonas pseudoalcaligenes	46.77
Paraburkholderia xenovoras LB400	47.26

The two highly conserved amino acid residues, W164 and S165 (Figure 1), selected in the current study were located in the C-terminal domain of GST-KKS102 protein. This domain contained a region that is responsible for electrophilic substrate binding [12]. These neighbouring residues selected are hypothetically thought to likely contribute to the structural stability of the domain.

```
kks102                  MKLYYAPGACSLAVHIALREVGVAFDLVKVDLVRHTTETGANYLDISPRGYVPLLELADQ      60
sp|Q52037|Q52037_PSEPS  MKLYYSPGACSLSPHIALREAGLNFELVQVDLASKKTASGQDYLEINPAGYVPCLQLDDG      60
ABE37052.1              MKLYYSPGACSLSPHIALREAGLNFELVQVDLASKKTASGQDYLEINPAGYVPCLQLDDG      60
sp|Q9RAF0|Q9RAF0_9RALS ---------CSLSPHIALREAGLNFELVQVDLASKKTASGQDYLEINPAGYVPCLQLDDG      51
sp|Q59721|Q59721_BURXL  MKLYYSPGACSLSPHIALREAGLNFELVQVDLASKKTASGQDYLEINPAGYVPCLQLDDG      60
                                ***: ******.*: *:**:***. :.* :* :**:*.* **** *:* *

kks102                  SRHTEAAALLQYVADLDPARALIGQPGSSERLAVLEWLTFVSTELHKGFSPWLWHKETAD      120
sp|Q52037|Q52037_PSEPS  RTLTEGPAIVQYVADQVPGKQLAPANGSFERYHLQQWLNFISSELHKSFSPLFNP-ASSD      119
ABE37052.1              RTLTEGPAIVQYVADQVPGKQLAPANGSFERYHLQQWLNFISSELHKSFSPLFNP-ASSD      119
sp|Q9RAF0|Q9RAF0_9RALS RTLTEGPAIVQYVADQVPGKQLAPANGSFERYHLQQWLNFISSELHKSFSPLFNP-ASSD      110
sp|Q59721|Q59721_BURXL  RTLTEGPAIVQYVADQVPGKQLAPANGSFERYHLQQWLNFISSELHKSFSPLFNP-ASSD      119
                           **. *::***** *.: *    ** **  : :**.*:*:****.*** :    ::*

kks102                  STRQAVKAKLAVRFAEMEAVLSRSDFL-AGGYSVADAYGFTIVNWSHLLGIPLTAYPHLQ      179
sp|Q52037|Q52037_PSEPS  EWKNAVRQSLNTRLGQVARQLEHAPYLLGDQLSVADIYLFVVLGWSAYVNIDLSPWPSLQ      179
ABE37052.1              EWKNAVRQSLNTRLGQVARQLEHAPYLLGDQLSVADIYLFVVLGWSAYVNIDLSPWPSLQ      179
sp|Q9RAF0|Q9RAF0_9RALS EWKNAVRQSLNTRLGQVARQLEHAPYLLGDQLSVADIYLFVVLGWSAYVNIDLSPWPSLQ      170
sp|Q59721|Q59721_BURXL  EWKNAVRQSLNTRLGQVARQLEHAPYLLGDQLSVADIYLFVVLGWSAYVNIDLSPWPSLQ      179
                        . ::**: .* .*:.::   *.:: :* ..  **** * *.::.**  :.* *: :* **

kks102                  AWMARVAARPQVQAALRAEGLLS-    202
sp|Q52037|Q52037_PSEPS  AFQGRVGGREAVQSALRAEVLIKE    203
ABE37052.1              AFQGRVGGREAVQSALRAEGLIKE    203
sp|Q9RAF0|Q9RAF0_9RALS AFQGRVGGREAVQSALR-------    187
sp|Q59721|Q59721_BURXL  AFQGRVGGREAVQSALRAEGLIKE    203
                        *: .**..*  **:***
```

Figure 1: Sequence Homology of Cytosolic GST-KKS102 and various GST Templates, Indicating Evolutionarily Conserved W164 and S165 Residues marked by Reddish Bars.

Predicted 3D Model of GSTKKS102

The predicted 3D models of GSTKKS102 were perfectly modelled using DeepView/Swiss-PdbViewer molecular graphics program. The quality of the model was ascertained using a Ramachandran plot [5]. The models are depicted in figures 2 and 3, in which the wild type of GSTKKS102 and its mutants were presented. The high rotamer score values (in positive integer) of some few amino acid substitutions of the conserved serine side chain at position 165 (Table 3) showed possibility of some effect on the GSTKKS102 protein structure. Conversely, substitutions of the conserved tryptophan side chain at position 164 with exception of glycine did not affect the 3D structure of the protein as indicated by the very low rotamer scores (Table 2).

***In silico* Site-Directed Mutagenesis**

Table 2: Rotamer and Score Values of *In-Silico* Site-Directed Mutagenesis of W164

Amino Acid Residue	**W164 Mutation**	**Rotamer**	**Score**
Alanine	Ala	1/1	-3
Arginine	Arg	19/26	-5
Asparagine	Asn	1/5	-3
Aspartic acid	Asp	1/4	-3
Cysteine	Cys	3/3	-4
Glutamine	Gln	5/14	-4
Glutamic acid	Glu	1/14	-3
Glycine	Gly	1/1	0
Histidine	His	2/6	-4
Isoleucine	Ile	1/5	-3
Leucine	Leu	1/4	-3
Lysine	Lys	1/16	-4
Methionine	Met	1/12	-3
Phenylanine	Phe	2/5	-3
Proline	Pro	1/2	-3
Serine	Ser	2/3	-4
Threonine	Thr	2/2	-4
Tryptophan	Trp	-	-
Tyrosine	Tyr	2/5	-5
Valine	Val	1/3	-3

Score value with a negative index shows no significant effect on the 3D structure of the protein

The *in silico* substitutions of W164 and S165 residues in the C-terminal domain of cytosolic GST-KKS102 were carried out using all the proteinogenic amino acids. Rotamer and score values were allocated for each substitution depending on how much disruption the new amino acid was predicted to affect the theoretical 3D model of the cytosolic GST-KKS102 (Tables 2 and 3). DeepView/Swiss-Pdb viewer molecular graphics program selects the best rotamer or conformation of the new amino acid residue from libraries of rotamers within the program. The best rotamer is defined as one that produces the lowest score indicating the least amount of effect on the structure of the protein and is based on minimization of energy. However, any high score have a tendency to affect the energy level which creates an effect on the protein's 3D model.

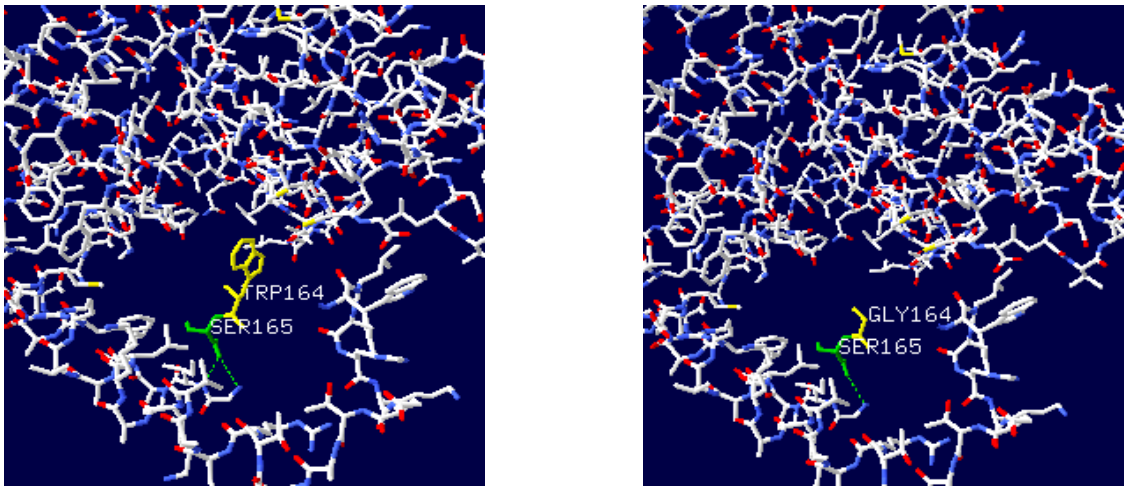
Table 3: Rotamer and Score Values of *In-Silico* Site-Directed Mutagenesis of S165

Amino Acid Residue	**S165 Mutation**	**Rotamer**	**Score**
Alanine	Ala	1/1	-3
Arginine	Arg	4/26	-2
Asparagine	Asn	2/5	-5
Aspartic acid	Asp	2/4	-3
Cysteine	Cys	3/3	-4
Glutamine	Gln	2/14	-2
Glutamic acid	Glu	2/14	-1
Glycine	Gly	1/1	0
Histidine	His	2/6	-3
Isoleucine	Ile	2/5	-3
Leucine	Leu	1/16	-3
Lysine	Lys	1/16	-3
Methionine	Met	3/12	-3
Phenylanine	Phe	4/5	-3
Proline	Pro	1/2	+1
Serine	Ser	-	-
Threonine	Thr	2/2	-4
Tryptophan	Trp	4/6	+5
Tyrosine	Tyr	4/5	-3
Valine	Val	2/3	-3

Score value with a negative index shows no significant effect on the 3D structure of the protein

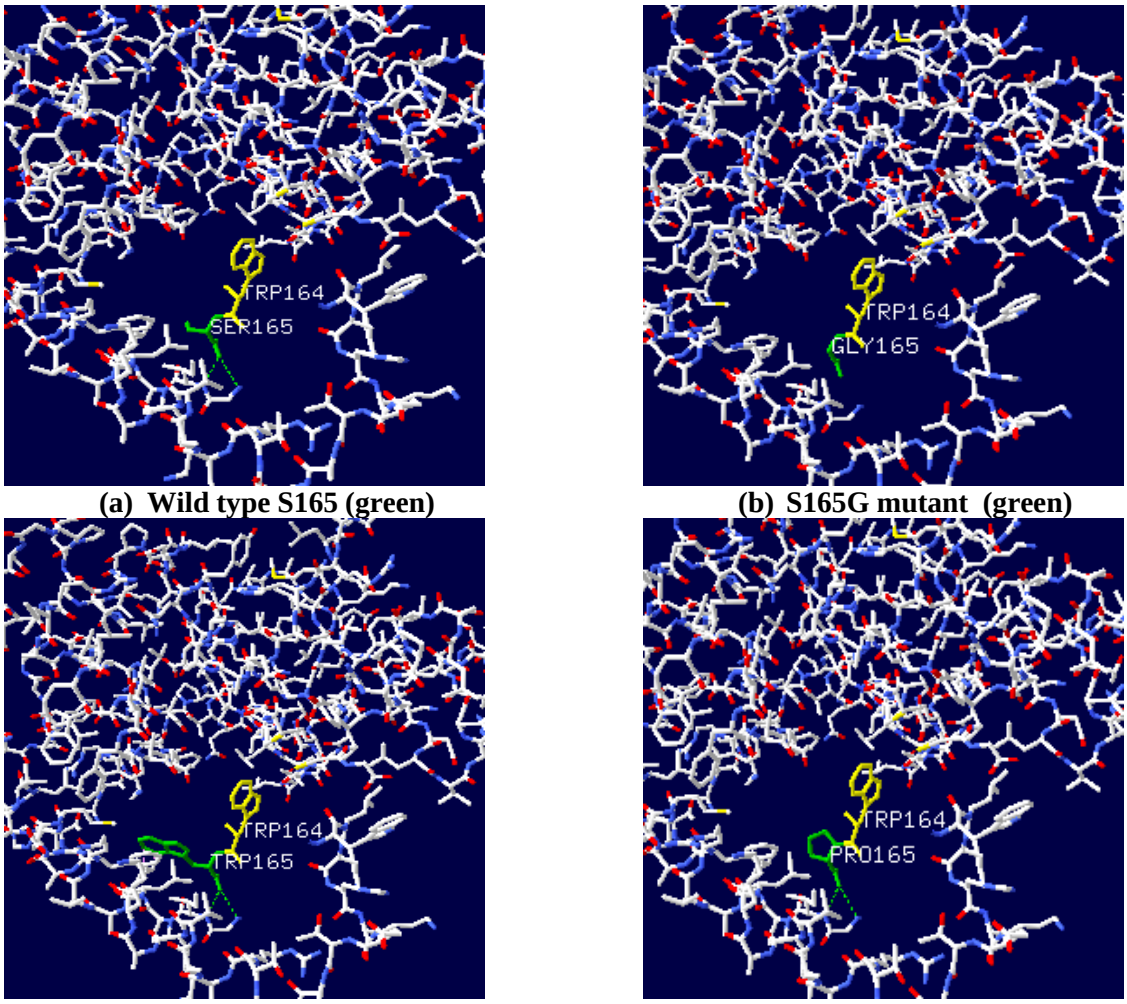
The predicted effects of *in silico* Site-Directed Mutagenesis on the 3D Model of GST-KKS102

The *in silico* substitution with glycine in the two targeted positions (W164G and S165G mutants) was likely to affect the 3D structure of GST-KKS102 as indicated by the relatively high score values of 0 in both cases (Tables 2 and 3). This structural effect could be attributed to the nature of the shorter side chain of glycine which contains only a hydrogen atom a side chain. This structural uniqueness of glycine side chain made it possible to acquire the ability to introduce kinks into the peptide chain, possibly affecting the substituted regions and indeed, the entire C-terminal domain and perhaps, the whole protein model. Additionally, in S165G mutant, there is possibility of disruption of hydrogen bonds due to the absence of hydroxyl group (OH) in glycine which may interfere with the protein stability induced by serine side chain in the wild type model. Similar scenario was reported by Gilmartin *et al.* [18] that S152G mutant showed a decrease in hydrogen bond capacity and consequently decreased structural and thermal stability and drop in the activity of *Bphk* LB400 GST protein. However, in W164G mutant, with exception of glycine substitution, all other proteinogenic amino acids substitutions showed no significant effect on the C-terminal domain of GSTKKS102 protein model as indicated by the lower scores in negative index as shown in Table 2.



(a) Wild type W164 (yellow) **(b) W164G mutant (yellow)**

Figure 2: The 3D models of tryptophan wild type (a) and W164G substituted mutant of GSTKKS102. Position of the mutation is shown in yellow



(a) Wild type S165 (green) **(b) S165G mutant (green)**

(c) S165W mutant (green) **(d) S165P mutant (green)**

Figure 3: The 3D models of serine wild type (a) and S164G, S164W and S164P substituted mutants of GSTKKS102. Position of the mutation is shown in green

In the *in silico* substitution of Ser at position 165 with all the proteinogenic amino acids, only Pro and Trp have higher rotamer scores which are likely to affect the stability of the C-terminal domain of GSTKKS102 model. In Ser165 substitution for Pro, S165P (Figure 3d), Pro lacks side chain hydroxyl group found in the side chain of Ser. This may likely alter the predicted hydrogen bond interactions contributed by the side chain of Ser165. Indeed, Pro is known to induce sharp turn in the peptide chain because its side chain is bonded to the peptide backbone nitrogen and carbon atoms [22]. These properties may alter the stability of the affected region and consequently affect the conformation of the C-terminal domain of the GSTKKS102 3D model.

On the other hand, in the S165W substitution (Figure 3c), a polar side chain of Ser was replaced by the hydrophobic side chain of Trp. Unlike the polar Ser residue which is normally found in the exposed region of protein, Trp is a large hydrophobic amino acid often found stacked and buried in the hydrophobic region of the protein domain [22]. In the S165W mutant, Trp was likely to form hydrophobic interactions with other hydrophobic amino acids. In this case, Val52, Pro53, Lue69, Lue70, Val101, Lue105 Phe109, Ile161, Leu170 and Pro171 were proposed to form a hydrophobic region with Trp165 in the S165W 3D model. This predicted hydrophobic interaction in the S165W model may replace the initial Ser side chain contributory hydrogen bond interactions thought to occur in the wild type 3D model (Figure 3a). This may create changes in the structure and stability of C-terminal domain of the S165W model.

CONCLUSION

In silico site-directed mutagenesis of two evolutionarily conserved amino acids, W164 and S165 was carried out using DeepView/Swiss-Pdb Viewer molecular graphics program. Of all the proteinogenic amino acids used for the mutation, only the substitutions for Gly, W164G and S165G were found to likely affect the stability of the C-terminal domain of the both protein models. However, S165P and S165W substitutions at position 165 were theoretically predicted to affect the 3D model of GSTKKS102 by replacing the predicted hydrogen bond interactions contributed by Ser side chain with proposed hydrophobic interactions by the nonpolar aromatic Trp side chains. This may result in the formation a hydrophobic region in the C-terminal domain of GSTKKS102 models and consequently alters the stability of the protein structure and perhaps, its function.

REFERENCES

[1] Habig, W.H., Pabst, M.J. & Jakoby, W.B. (1974). Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *Journal of Biological Chemistry,* 249,7130–7139.

[2] Simarani, K., Yusoff, W.H.A.C. & Alias, Z. (2016). Purification of Glutathione Transferases (GSTs) from Identified Rhizospheric Bacteria. *Sains Malays*,45(3), 1057-1062

[3] Wilce, M.C.J. & Parker, M.W. (1994). Structure and function of glutathione S-transferases.

Biochimica et Biophysica Acta 1205, 1-18.

[4] Venu, V. & Zazali A. (2018) Substrate specificities and kinetics parameters of recombinant Drosophila melanogaster glutathione s-transferase E6 and E7. *Drosophila melanogaster-Model for Recent Advances in Genetics and Therapeutics.* In Tech. Pp. 227-239

[5] Edwina, B., McGuinness, M. & Dowling, D. (2007). *In silico* site-directed mutagenesis and the predicted 3d structure of *BphK*LB400. *Proceedings of ESAI Environ,* 2007. 42-44 p, Ireland.

[6] Shehu, D. & Zazali, A. (2018a). Functional Role of Tyr12 in the Catalytic Activity of Novel Zeta-like Glutathione *S*-transferase from *Acidovorax* sp. KKS102. *The Protein Journal*. 37(3), 261-269.

[7] Fortin, P.D., Horsman, G.P., Yang, H.M. & Eltis, L.D. (2006). A glutathione S-transferase catalyzes the dehalogenation of inhibitory metabolites of polychlorinated biphenyls. *Journal of bacteriology,*188, 4424-4430.

[8] Warner, J.R., Lawson, S.L. & Copley, S.D.A. (2005). Mechanistic Investigation of the Thiol-Disulfide Exchange Step in the Reductive Dehalogenation Catalyzed by Tetrachlorohydroquinone Dehalogenase. *Biochemistry-US*. 44(30), 10360-10368

[9] Shehu, D. & Zazali, A. (2018b). Dechlorination of polychlorobiphenyl (PCB) degradation metabolites by a recombinant glutathione transferase from Acidovorax sp. KKS102. *FEBS Open Bio.* Accepted article

[10] Ohtsubo, Y., Maruyama, F., Mitsui, H., Nagata, Y. & Tsuda, M. (2012). Complete genome sequence of *Acidovorax* sp. strain KKS102, a polychlorinated-biphenyl degrader. *Journal of Bacteriology* 194, 6970–6971

[11] Allocati, N., Federici, L., Masulli, M. & Di Ilio, C. (2012). Distribution of glutathione transferases in Gram-positive bacteria and Archaea. *Biochimie,* 94(3), 588-596.

[12] Federici, L., Masulli, M., Bonivento, D., Di Matteo, A., Gianni, S., Favaloro, B., Di Ilio, C. & Allocati, N. (2007). Role of Ser(11) in the stabilization of the structure of Ochrobactrum anthropi glutathione transferase. *Biochemical Journal* 403(2), 267-74

[13] Nishida, M., Kong, K.H., Inoue, H. & Takahashi, K. (1994). Molecular cloning and site-directed mutagenesis of glutathione Stransferase from Escherichia coli. The conserved tyrosyl residue near the N terminus is not essential for catalysis. *Journal of Biological*

Chemistry, 269, 32536–32541

[14] Favaloro, B., Tamburro, A., Angelucci, S., Luca, A.D., Melino, S., Di Ilio, C. & Rotilio, D. (1998). Molecular cloning, expression and site-directed mutagenesis of glutathione S-transferase from Ochrobactrum anthropi. *Journal of Biochemistry* 335 (Pt 3), 573–579.

[15] Allocati, N., Masulli, M., Casalone, E., Santucci, S., Favaloro, B., Parker, M.W. & Di Ilio, C. (2002). Glutamic acid-65 is an essential residue for catalysis in Proteus mirabilis glutathione Stransferase B1-1. *Journal of Biochemistry* 363, 189–193.

[16] Armstrong, R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chemical Research in Toxicology,* 10(1), 2-18.

[17] Inoue, H., Nishida, M. & Takahashi, K. (2000). Effects of Cys10 mutation to Ala in glutathione transferase from *Escherichia coli. Journal of Organometallic Chemistry* 611, 593-595.

[18] Gilmartin, N., Ryan, D. & Dowling, D. (2005). Analysis of the C-terminal domain of *Burkholderia sp.* strain LB400 BphK reveals a conserved motif that affects catalytic activity. *FEMS Microbiologyletters* 249, 23-30.

[19] Guex, N. & Peitsch, M.C. (1997). SWISSMODEL and Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*18, 2714-2723.

[20] Schwede, T. Kopp, J., Guex, N. & Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology modelling server. *Nucleic AcidsResearch* 31, 3381-3385.

[21] Brennan, E., Mcguinness, M. &Dowling, D.N. (2009). Bioinformatic analysis and in vitro site-directed mutagenesis of conserved amino acids in BphK LB400, a specific bacterial glutathione transferase. *International Biodeterioration & Biodegradation,*63,928-932.

[22] Betts, J.M. & Russell, R.B. (2003) Amino acids properties and consequences of substitutions. *Bioinformatics for Geneticists*. John Wiley and Sons, Ltd P. 297