Homology Modeling of Bifunctional Enzyme Alanine Racemase from *Taibaiella chishuiensis*

Gulsanga lemar^{1*} and Farid Ahmad Danish Far²

¹Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Malaysia and Faculty of Biology, Kabul University, Afghanistan. ²Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Malaysia and Faculty of Science, Herat University, Afghanistan.

http://dx.doi.org/10.13005/bbra/2864

(Received: 02 September 2020; accepted: 22 September 2020)

Alanine Racemase from Taibaiella chishuiensis bacteria is one of the bifunctional enzymes that catalyze the L- and D-alanine racemization of peptidoglycan biosynthesis in bacteria and ligation (UDP-N-acetylmuramoyl-Tripeptide-D-alanyl-D-alanine ligase). It had two EC numbers 5.1.1.1 and 6.3.2.10 respectively. This enzyme is an important target for antimicrobial drug productions or inhibitor design. However, the 3D structure of Alanine Racemase from Taibaiella or UDP-N-acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase/ alanine racemase has remained unknown. Thus, this study modeled and validated the 3D structure of the enzyme in the query. The bioinformatics tools/databases and software such as BRENDA, NCBI, Uni Prot, Clustal Omega, Prot Param, Swiss model, Phyre2, GOR, PROCHECK, and PyMOL were used for modeling, validation, and structural comparison. From the sequence and 3D structure analysis, it is indicated that Alanine racemase from Taibaiella had the same active and binding sites with the reference enzymes. Thus, we were able to study the similarities and differences in the sequence and structural properties of alanine racemase in two different bacteria. Finally, it was found that our enzyme has two parts for two different functions (racemization and ligation). The predicted model of alanine racemase of T. chishuiensis from this study could serve as a useful model for further study regarding the other bifunctional enzymes structure and function as well as drug design projects.

Keywords: Alanine Racemase; *Taibaiella chishuiensis*; Bifunctional enzyme; Bioinformatics tools; Predicted model; Active site; Binding site.

Alanine racemase, which belongs to the isomerases family, is a bacterial enzyme with Enzyme Classification (EC) number: 5.1.1.1 (Muhammad and Zhao, 2019). It catalyzes the racemization of L- and D- alanine (Figure1),and pyridoxal5'- phosphate (PLP) is required as a cofactor (Muhammad *et al.*,2019). Woodand Gunsalus discovered and isolated this enzyme in Streptococcus faecalis, for the first time (Wood and Gunsalus, 1959).

This enzyme has a significant role in the growth of the bacteria by providing D-alanine, the

peptidoglycan layer constitutive of the bacterial cell wall. (Liu *et al.*, 2018). In both gram-positive and Gram-negative bacteria, the peptidoglycan layer of the bacterial cell wall prepares resistivity to osmotic lysis (Islam *et al.*, 2017). Alanine racemase is special to bacteria. There are few exceptions to the fact that certain eukaryotes have the enzyme D- alanine-containing peptides biosynthesis in fungi, D- alanine metabolism in yeast, as well as osmotic regulationin a crayfishplusa bivalve mollusk (Nomura *et al.*, 2001).

This is an ⁽²⁾ Open Access article licensed under a Creative Commons license: Attribution 4.0 International (CC-BY). Published by Oriental Scientific Publishing Company © 2020



^{*}Corresponding author E-mail: gulsangalemar@gmail com

This enzyme is absent in humans and is aunique enzyme that transforms L-alanine and D- alanine in most bacteria thus, it is an important target for antimicrobial drug productions (Im et al., 2011; Wei et al., 2016). For instance, alanine racemase was chosen as a target to discover new drugs to treat tuberculosis or for the inhibitor design (Anthony et al., 2011). Based on previous studies, most of previously studied inhibitors suffer from the lack of specificity, and further investigation is required to solve this issue (Azam and Jayaram, 2016). To target an enzyme for effective antibacterial drugs, and to develop new and selective inhibitors, which improves treatment in public health systems, it is important to know about the structure and characteristics of the enzyme in various living organisms. The 3D structure of the alanine racemase of several bacteria is available in the PDB. But there is no 3D structure of the Taibaiellachi shuiensis alanine racemase. Strain AY17 T of Taibaiella chishuiensis contained MK-7 as the predominant respiratory quinone, plus the main polar lipid phosphatidylethanolamine, and hydrolyses the aesculin, casein, and gelatin (Tan et al., 2014). Therefore, in the current study, we perform modelling of alanine racemase in the mentioned bacteria. The present investigation is the first study of the sequence and 3D structural characterization of the *Taibaiella chishuiensis* alanineracemase enzyme. This study modeled the structure of the novel bifunctional, *Taibaiella chishuiens*-based alanine racemase, which belongs to the family of isomerases. A bacterial strain, AY17 T of *Taibaiella chishuiensis*, was isolated from the Chishui River in Guizhou Province, Southwest China.

Based on phenotypic, phylogenetic, plus genetic evidence, theAY17T strain was categorized as a novel representative species of the *Taibaiella* genus for which the *Taibaiella chishuiensis* sp.Nov., the name has been suggested. This bacteria belongs to the Chitinophagaceae family (Wei *et al.*, 2016). The structure of the alanine racemase was first performed on an enzyme isolated from Bacillus stearothermophilus. This enzyme is a homodimeric enzyme with each monomer consisted of á/â barrel domain at the N-terminus and a C-terminal domain consisting mainly of â strands. The location of the active site is at the interface of the á/â barrel andthe â domain, close to the PLP cofactor, forming



Fig. 1. The reaction catalyzed by alanine racemase (5.1.1.1) (Schomburg *et al.*, 2004)

The amino acids number	831	
Molecular weight	93816.05	
Theoretical pI	6.42	
Total amount of positively	92	
charged residues (Arg + Lys)		
Total amount of negatively	101	
charged residues (Asp + Glu)		
Atomic composition	Hydrogen	Н 6656
	Carbon	C 4197
	Oxygen	O 1222
	Nitrogen	N 1144
	Sulfur	S 35
Formula	Formula: $C_{4107}H_{6656}N_{1144}O_{1222}S_{25}$	
Total amount of atoms	13254	
The estimated half-life	30 hours (mammalian reticulocytes, in vitro)	
	>10 hours (Escherichia coli, in vivo)	
	>20 hours (yeast, in vivo)	
Aliphatic index	97.3	
Grand average of hydropathicity	-0.16	
(GRAVY)		

Table 1. Calculation of some physical and chemical parameters of alanine racemase from *T. chishuiensis* (Gasteiger *et al.*, 2003)

an internal aldimine connection to the lysine residue (Im *et al.*, 2011). Crystal structure study has shown that almost all alanine racemases are dimeric (Dong *et al.*, 2018). Mutagenic, modeling and structural analyzes indicate Tyr265 and Lys39 of G.Stearothermophilus alanine racemase are residues involved in the movement of protons to the racemase reaction. Lys39 and Tyr265 residues are composed of two similar polypeptides. These two residues are well preserved in other Alanine racemases (Ju *et al.*, 2011).

To clarify the bifunction potency of the alanine racemase enzyme, a comparative study was performed between the primary sequence and the predicted structure of the new bifunctional alanine racemase enzyme and similar, clostridium difficile (PDB: 4LUT) from the family, Peptostreptococcaceaeas as well as T. maritime (PDB: 3z18) from the family, thermotogaceae (Couñago *et al.*, 2009; Tan *et al.*, 2014; Dong *et al.*, 2018).

WP_106523916.1	MNLNDILKATKAEAIPEVKDIDIAELLIDSRKIVNPARAMFIAIKGSHRDGHQFIDD	57
WP_118951764.1	MNLAEIQKATKAEVRPGVEDRVINELLIDSRKIINPSRAMFIAIKGSHRDGHEFIAD	57
DJW29663.1	MNLAEIQEITEAEVEPDIEDRTINELLIDSREIINFTEAMFIAIESBROGHSFIAD	22
WP_110999378.1	MLAAQQICAWSQGTWMAFHADIVIEELIIDSRKINEPETALFIALKAPRRDGHSFIAS	5.8
WP_086102057.1	MQRPGAEQLERUCNGTMLLHENDIAIEELVIDTREVAEPENALFIALQTSRRDGHSFIPQ	60
WP_106523916.1	AVERGVRTFLISEETDTAKYKDAQFLKVTDTLTALHQLATAHRNQFHIPVIGITGSNGRT	3.3
WP_118951764.1	AYKKOVRTFLVSSDTO-DMYPDAIFLOVODTLTALHOLAAHRNOFHFPVIGITDSNGKT	11
03879003.1	AVERGYRNFLYAGDTDTTLYEGANFIRVEDTLOGLOGIAETYREGEDIFVIGITDSNORT	3.3
WP_110999378.1	AYDKGVRNFLVQEEINTSDYPGANFILVKDTLKALGLFVAQVRKQFSIPVIGITGSNGKT	11
WP_086102057.1	AYERGIRNFLVSQDVDVQPYQGACFIRVNDTLAALQQIAAGYRRQFDIPVIGITSSNGRT	3.2
WP_106523016.1	TKEWLFOLLSENVKIIRSPKSVNSOIGVPLSVWRMOAHHTLGIFEAGISOPGEMSRSER	17
WP 118973406.1	TPREWLPGLLSONYRTTRSPRSYNSOTOVPLSYWRMOEIHITLGTPEAGTSOPGEMSHSER	17
03979663.1	I/KEWLYQLLSPDHNMVRSPKSYNSQLGVPLSVWLLEPQHQIAVFEAGISQPGEMEQLER	3.2
WP_110999378.1	VEREWLFOLLSNSFSIARSPRSYNSOTOVPLSVWLLEPENELGIFEAGISOPPEMENLER	17
		1.00
WP 118951764.1	TUMPETGTFTNIGEAMSEGFUNLROKINERUVUFRMSKDUTVCKDMHELNECTVOVCHOV	23
WP_118973406.1	ILHPEIGIFTNIGEAHSEGFLNLRÖKINEKLVLFRHSKELIVCKDHHELNECIVQVCHÖV	28.28
WP 110000170.1	TREPTION THOSE AND GET LIKENOK INERLELEF HHALOL VYCROVHELNDCI LOVING	
WP_086102057.1	I IOPT IGVF TN IGEAHSEGF LNMROK INEKLGLFRHVOOMVVC IDVPELNEVIVHVK SNV	2.4
	···· ·································	
WP_106523916.1	RSG-EEDDLKVFTWSGKS-EATLRIGGIDKSSNSTTIRGVYEEREIAILIPFSDNASIE	29
WP_118951764.1	RNA-PONPLQLFTWSQRS-EATLRITQVERGSGATQIGGVVREQEMSIRIPFTDNASIEN	20
03079663.1	RNS-PODPLQLFTWSQRS-EATLRILRIERGTGSTQITGIYRERIMSIRIPFIDAASIEN BSAGNGOOLTLFTWSQRS-EATLRILRIERGTGSTQITGIYRERIMSIRIPFIDAASIEN	29
WP_110999378.1	KONGSEPPLOLTTWSORH-EADLRITRLEGIDSKTKIEGOVOGVDLAITIPFTDPASVEN	29
WP_086102057.1	RNHDGEQPLELFTWSQKNVEADLYVENTSSAAGKTTINAIYEQQPLSITIPENDAASIEN	34
WP_106523916.1	AIHCHCVLLLFKIEDSQIAERMARLHPVAMID ELLHGVDNCTIINDTVN5DLTSLQIALD	39
WP_118951764.1	ATHEWELMLLLHIDAETTAERMERLHEVAMELELLHEMDNETTINDTYN DU TSLNIALD	3.5
03879663.1	ATHCWSVLLLGVSADNIKORMOOLHAVSMELELKHGINGCTVINDTYNDDLTSFRIALD	35
WP_110999378.1	AIHCWCTLLLLKVPQQEIEKQMEQLHTVAMRLELRHGINDSTLINDTVNEDLTSLYIALD	35
WP_086102057.1	AIHCWAVMLLLALPOETIAERMORLOPVSMILELRHGINDCTIINDTYNIDMISLLIALD	36
WP_100523910.1	FLEQOROHPRETVILSOMLQIGRPORDLYEEVAELVSQKNIYRIICIOVALSKHKPTFRK	
WP_118973466.1	FLEGORGHPARTVILSDMLGIGKPDGDLYTEVAELVSGRNIVRIICIGPALTRHRAVFRK	43
03979663.1	FLOQOOGHSOHTVILSDIHGGVATDTDLYDEVAQLVAQHNIKRFIGIGPALYKNKASFRK	-9.3
WP 086302057.3	VLEQGROHOBTVILSDILGIARPDALLVEEVAEALARRNIGRFIGIGPALVESKSLFRA VLEQGROHODTVILSDILGIARPDALLVEEVAEALARRNIGRFIGIGPALVENKALFRO	43
	isissis iisses i a second i i i i i i i i i i i i i i i i i i i	
WP 106523916.1	HKKLBSTFFKSTEFFLKKMMATTFENEATLLEGABAFBFEKTEKLLEGKTHKTVLEINLS	47
WP_118951764.1	HKKLRSIFFKSTEDLLKKLHMITFENEAILL <mark>K</mark> GARAFRFEKIEKLLEØKIHKTVLEINLS	4.7
WP_118973406.1	YKKLRSIFFKSTEDLLKKLHMITFENEAILL DARAFRFEKIEKLLEOKIHKTVLEINLS	-8.2
WP 330999378.3	HERLESS FEESTEEF IRRUHLVOF ORDATLL BOARST REEKVSTLLEOEVHOTVLSTOLS	
WP_086302057.1	HERLESIFFESTODFLEEFHLISFDNEAILL <mark>E</mark> GARVFTFEEIDMLLEQQVHQTVLSINLS	-9-23
WP_106523916.1	AIQNNLNVYRALLPAPVKMMAM	5.3
WP_118951764.1	ATOMNESTYRSELEPPRVETMANNE VYSYGSGSYFTANLEGRADDVLTVAVVDEGVALRE	2.2
03979663.1	ALAHNLDAVRSYLEPOVKTMATY AFSYGHOSYEIANLLQYSRVDYLAVAYTDEGIALRK	53
WP_110999378.1	AMMNNLNVFRTRVAPGVKVMAM AF SYGSGSYEVAVALOVAGVDYLTVAYTDEGIALRK	5.0
wp_006102057.1	St Innenvriskt provident of stososyte and European States and States	5-4
		-
WP 118951764.1	AGITAPIMVMSPESGAFORMISWELEPEIPSITESLIKETEIAHTLNARNYPVHILELOTGA	5.0
WP_118073406.1	AGITAPIMVMSPEAGAFDRMISWKLEPEIFSIESLHQFTGIAHTLNTHHYPVHIKLDTGM	5.0
03679663.1	TGITMPIMVMSPDLGSFDRMIAWKLEPELFNMRSVRAFMEIARMLDVRDVPIHIKLDTGM	5.0
WP_086102057.1	AGITMPIMVMSPDATSFDRMIAMRLEPEIVNVRSLELFTSIAQTLQVRNVPIHIRLDTGM	60
	···· ······· ······ ··················	
WP_106523016.1	HRLGFMPHELPELMAELKVNTATKVVSTFSHLAGSDSADFDGFTAEQARNFDNMSQELMQ	-
WP_118951764.1	HRLGFMPQELDGLIEELKENDAAKVVSIFSHLAGSDAADFDGFTKEGAAAFTQMSNRLIA	65
WP_118973406.1	HRLGFMPHELDALIELLKVNDTTKVV5IFSHLAGSDTADFDGFTQEQATAFTQMSNRLIA	65
WP_110999378.1	HREGENVHEEPDEFTELAQNEHVKVASVESHLAGSEDAGEDEFTRQQAAAFRSMSDSLTE	65
WP_086102057.1	HRLGFMPHELEALMEHCDNNPHIQIASVFSHLAASEDPALDHFTAHOKEQFEQMSHALLT	66
WP_106523916.1	VLPTRPLRHLVNSTGITRHPELVYDMVRLGIGMYGFDG5DKV9H0LQNIGTLKTTIAQIK	73
WP 118973406.1	ALPORPLIAL CNSTGTTRAPEL VEDWALGVGAVGYGSGKEGSGLAHIGTLKTTIAGIK	23
03879663.1	GL SYRPL RHLCN SAGVVRHKOLHF DMVRLGIGL YGVDSSNAMHGELKQIGTLKTTIAQIK	23
WP_110999378.1	QLQYKPYLHLCNSAGIARHPELHFDMYRLGLGLYGYDGSGYLHDELIQYGTLKTTIAQIK BI PNKPI BHI CNSAAIABHPDMRYDM/BI GI GI YGTDGSGYLHDELIQYGTLKTTIAQIK	73
	· · · · · · · · · · · · · · · · · · ·	
WP 106523916-1	ELKAGETVERBERKERLERDSRIATVSIGVADGVFRDFGNGNAHMLVHSKPAPVVGAVEMD	
WP_118951764.1	DLPAGETVCV RKGKLKKNSRIATVSIGVADGVFRDFGNGQAVMLLHGKQAPVLGSVCMD	77
WP 118973486.1	ELPAGETWINGOLTROSKTATVSTOVADOVPROPONGOAVMLLIGKOAPVLGAVCMD	77
WP 110999378.1	ELDPGET V WRRGEVONPREIATISIGVADGVPRALGNGTANVL VIGEPACI VGVVAD	22
WP_086102057.1	EIPAGETVCTBRKGFVDKDTHIATISIGVADGYPRALGNGKAHVLVQGKPAKLVGVVCHD	78
WP_106523916.1	MCMLDITGIDNAAVGDEVIVEGNELPIGQLADMAKTIPVEVLTGISQRVNRIVVNE- 831	
WP 118051764.1	MCMLDITGIEQAKVGDEVTVFGDELPISLLASWAKTIPVEIMTSISGRVNRIVVNE - 838	
03879663.1	MCMVDVTNIPEAQEGDEVIVFGEQLPLEQLAAMDGTIAVEIMTGISGRVKRVVTQED 833	
WP_110000378.1	MCMVDVTGTPEAKEGDEVIVFGRELPLTQLSQHADTIPVEM4TGISQRVKRVVVHEG 834	
W086102057.1	ACTION IN THE REPORT OF A STATE AND A STAT	

Fig. 2. Multiple sequences alignment of 5 bifunctional enzymes against the enzyme alanine racemase from *T. chishuiensis*. Yellow highlight box: active site for alanine racemase, green highlight: binding site for Dalanyl-D-alanine ligase and boxes: binding region for Dalanyl-D-alanine ligase

WP_10652391 6.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase WP_11895176 4.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase WP_109937 8.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase WP_08610205 7.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase	Access ion	Enzyme				Or	ganism		Identity	(%)
WP_11897340 6.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Taibaiella sp. F-4 82 Dalanyl-D- alanine ligase/alanine racemase WP_11895176 4.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Taibaiella sp. F-4 82 Dalanyl-D- alanine ligase/alanine racemase WP_08610205 7.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Taibaiella soli 66 Dalanyl-D- alanine ligase/alanine racemase UP_08610205 7.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Dalanyl-D- alanine ligase/alanine racemase Dalanyl-D-10 Alanyl-D-10 Alanyl	WP_10652391 6.1	Bifunctional U	DP-N- acet	ylmuramoyl	- tripeptide:	Ta	ibaiella		100	
WP_11895176 4.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Taibaiella sp. F-4 82 Dalanyl-D- alanine ligase/alanine racemase WP_08610205 7.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Taibaiella soli 66 Ualanyl-D- alanine ligase/alanine racemase UJW79663.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Dalanyl-D- alanine ligase/alanine racemase Ccccchhanhunhanhunche/cheeseeeee crocchhanhunhanhunch/cheeseeeeee crocchhanhunhanhunch/cceeseeeeeeececccchhanhunhanhunch/cceeseeeeee crocchanyanhunhanhunhunch/cceeseeeeeeeeeeececchhanhunhanhunch/cceeseeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee	WP_11897340 6.1	Bifunctional U	DP-N- acet	ylmuramoyl	- tripeptide:	Ta	ibaiella		82	
WP_11099937 8.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalaryl-D- alanine ligase/alanine racemase Bifunctional Distribution Comparison of the second secon	WP_11895176 4.1	Bifunctional U	DP-N- acet	ylmuramoyl	- tripeptide:	Та	ibaiella sp. F-4	4	82	
WP_08610205 7.1 Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Bifunctional UDP-N- alanine ligase/alanine racemase Bifunctional UDP-N- acetylmuramoyl- tripeptide: Dalanyl-D- alanine ligase/alanine racemase Bifunctional UDP-N- alanine ligase/alanine racemase Bifunctional UDP-N-alanine l	WP_11099937 8.1	Bifunctional U	DP-N- acet	ylmuramoyl	- tripeptide:	Та	ibaiella soli		66	
OJW79663.1 Bitunctional UDP-N-acetylmuramoyl-trippetidie: bacterium DV 0501 Bacteroidet es bacterium A6-16 61	WP_08610205 7.1	Bifunctional U	DP-N- acet	ylmuramoyl	- tripeptide:	Cł	nitinopha gace	ae	64	
10 20 30 40 50 60 70 The Not Clear KARA LEP KYKD LID S.R.KT VIN YAAWF LALK GAM BOG KG LID KYKS WR TH. LISE Cocce high high high high high high high hig	OJW79663.1	Bifunctional U Dalanyl-D- ala	DP-N- acet	ylmuramoyl alanine race	- tripeptide: mase	Ba ba	cteroidet es cterium 46-16	CDI	61	
MULNDILLATAEAIPEVKU DIAELLIDSKITVNPAANFIAIKGSHBGAGFIDDVKEGORTFLISE ccccchhhhhhhhhccAteAIPEVKU TALIALATAIRAGHITYUGTGSNGKITVKILIQLISUNVKIISEKSU ccccchhhhhhhhhhhhhhhhhhhhhhhhhhhcccceecee	10	20 30	40	50	60	78				
	chhhhnhhhhnhhhhhh NSQIGVPLSVWRMQAH ccceeeceehhhhhh FRHAKELIYCKOHHEL hhhhnhhhhnccccc IAILIPFSDNASIENA heeecccccceeee QIALDFLEQQKQHPRR hhnhnhnhhhhhhhhhhhhhhhhhhhhhhhhhhhhh	hhhhhhhhhhhhccc HTLGIFEAGISQPGEMS ceeceeeecccccc IHCWCVLLLFKIEDSQL ecceeeeeecccchhh TVILSDHLQIGKPDGDL VVILSDHLQIGKPDGDL HHH Hhhhhhhhhhhhhhhhhhh SYEIANLLQFAGVDYLT chhhhhhhhhcccceee AHTLNVSHYPVHIKLDT hhhhhhhhhhccccee hhhhhhhhhccccee AMGEVIVFGNELPIG cccceeeecccccc	ceeeeeeccc RSERILHPDI ceeeeeccc AERMARLHPV/ hhhhhhhhhh PEKIEKLEQ hhhhhhhhhh FEKIEKLEQ hhhhhhhhhh GHIRLGFMPHI chhhccccc HLVNSTGITR eeecccccc QLADWAKIP hhhhhcccc	ccchhhhhhhh GIFTNIGEAHSI Ceeeecccccci SEATLRIQGID ccchhhhcccc AMRLELLHGVD hhhhhhhccccc KNIYRIICIGV Ccceeeeeeee KIHKTVLEINL hhhchCccccee ELPELMAELKV Cchhhhhhhhhh HPELYYDMVRLC CCCEEEEEE YADGYFRDFGNC CCCCEEEEEE	hhnhnhnhnhnhn GELNNRQKINE GELNNRQKINE GESNSTIRGVY CCCCCERECCH CTIINDTYNSD CCREECCCCCC CLSKHKPTFRKH CCCCCCCChhnh GAIQNNLNVYRA hhnhnhnhnhnhnhnhnhnhnhnhnhnhnhnhnhnhnh	CCCC KLIL hhhh EERE ERRE hhhhh LTSL chhh KKLR KKLR KKLR KKLR CCCC SWKL LLPA hccc SWKL LLPA hccc SWKL CCCC VQHQ hhhh CCCC VQHQ hhhhh CCCC	R4 : Alpha helix 3g helix Pi helik Beta bridge Extended strand Bend region Randon coll Ambiguous state Other states	(Hb) (Gg) (II) (Ee) (Tt) (Cc) (Cc) (Cc)	355 is 0 is 0 is 0 is 0 is 0 is 300 is 0 0 is 0 0 is	42,73 8,04 0,04 21,14 0,04 0,04 0,04 0,04 0,04
	· ··• • •		+ 18 +	+ + + (- 111	ŧ.	***	∎₩		++
		200	300				508	700		

Table 2. The results and Alanine racemase enzymes obtained from BLASTp

Prediction result file (text): [GOR4 (/tmp/ebb716a41afe.gor4.mpsa)]

Fig. 3. Secondary structure prediction of alanine racemase using the GOR database

Finally, several analyses were performed to provide useful details on the bifunctionality of this enzyme. This study wishes to provide useful information for modeling/designing other unknown structures of the bifunctional enzymes as well as drug design projects.

Research methods

Sequence retrieval and multiple sequence alignment(MSA)

The information about the enzyme alanine racemase of *T.chishuiensis* is obtained by searching the EC number (5.1.1.1) in BRENDA (Schomburg

et al., 2004). So, we found out the optimum temperature, PH, metabolic pathway reaction. Also, the FASTA format and other informationregarding the PDB structure is received from UNIPROT (Ju *et al.*, 2011). Then, the Blastp from NCBI database was done to obtain a similar sequence of alanine racemase from different organism to compare and analyze the structure (Asojo *et al.*, 2014). Based on the data, it was found that our enzyme is very similar to bifunctional enzymes "UDP-N-acetylmuramoyl-tripeptide-D-alanyl-Dalanine ligase/alanine racemase" of different bacteria



Fig. 4. A. The cartoon 3D structure of the template model of UDP-N-acetylmuramoyl-tripeptide-D- alanyl-Dalanineligase (3zl8.1.A) (which i ssimilar to N-terminal half of our sequence). B.The cartoon 3D structure of the template model of alanineracemase (4lut.1.A) (identical to C terminal half of our sequence), and both were modeled using the SWISS-MODEL database. C. The cartoon 3D structure of alanine racemase from *T. chishuiensis* was modeled using the Phyre2 modeler database



Fig. 5. Ramachandran plot of alanine racemase of T. chishuiensis predicted 3D structure

especially genus *Taibaiella* and Chitinophagaceae bacterium. But all of them do not have a PDB structure (Chen *et al.*,2019).

The primary structure analysis model of enzyme alanine racemase from *T. chishuiensis* was done using the ProtParam tool (Gasteiger *et al.*, 2003;Hassan *et al.*,2020).This step was carried out to identify the estimated molecular weight, the theoretical pI value, the composition of the amino acid, the total amount of positively charged plus negatively charged residues, the atomic composition, the chemical formula, the aliphatic index, and the hydropathicity value of the model protein.

Secondary structure prediction plus 3D structure-modeling

The secondary structure of enzyme alanine racemase in *Taibaiella chishuiensis* was

predicted using the GOR database (Hassan *et al.*, 2020). The 3D structure of the alanine racemase of *T. chishuiensis* was modeled using the SWISS-MODEL (Biasini *et al.*, 2014). The best model generated from SWISS-MODEL was selected on the basis of a high sequence identity score. Our enzyme (alanine racemase of *Taibaiella chishuiensis*) was similar to two different enzymes: UDP-N-acetylmuramoyl-tripeptide-Dalanyl-D-alanine ligase (PDB ID: 3zl8) plus monomer of alanine racemase (PDB ID: 4lut). Also, PHYRE2 modeler was used for our sequence to build the 3D model (Kelley *et al.*,2015).

Homology model validation

The predicted 3D structure was validated using the PROCHECK software and Ramachandran plot to investigate the psi-phi angles to determine the accuracy of the predicted structure. Verify 3D





Fig. 6. Verify 3D structure of alanine racemase of T. chishuiensis from verify 3D prediction

is also used to evaluate the 3D protein structure model (Meo and Cozzetto, 2006).

Structural comparison

The structure of alanine racemase from *T. chishuiensis* was analyzed and compared with alanine racemase from *C. difficile*. PyMol was used to superimpose the modeled structure of the alanine racemase of *T. chishuiensis* and the known structure of the alanineracemase of *C. difficile*. Since our enzyme is bifunctional, act also as a ligase, thus, we compared with the known structure of UDPN-acetylmuramoyl- tripeptide-D-alanyl-D-alanine ligase (PDB ID: 3zl8). For comparison of the surface structure and determining active site and binding site, we used also the PyMol (Meo and Cozzetto,2006).

RESULT AND DISCUSSION

Primary sequence

The amino acids sequence (FASTA format) of alanine racemase of *T. chishuiensis* retrieved from the UniProt database (Consortium, 2015).

>tr|A0A2P8D088|A0A2P8D088_9BACT Alanine racemase OS=*Taibaiella chishuiensis* OX=1434707 GN=B0I18_10740 PE=3 SV=1 MNLNDILKATKAEAIPEVKDIDIAELLIDSR KIVNPARAMFIAIKGSHRDGHQFIDDAYEK GVRTFLISEETDTAKYKDAQFLKVTDTLTA LHQLATAHRNQFHIPVIGITGSNGKTIVKE WLFQLLSENYKIIRSPKSYNSQIGVPLSVWR MQAHHTLGIFEAGISQPGEMSRSERILHPDI GIFTNIGEAHSEGFLNNRQKINEKLILFRHA



Fig. 7. A. Right: superimposed diagram of reference UDP-N-acetylmuramoyl-tripeptide-D- alanyl-D-alanine ligase (3zl8) (green) with Phyre2 model (blue) and left: reference enzyme of Alanine racemase (4lut) (green) with Phyre2 model (blue). B. Superimposed of Phyre 2 model of alanineracemase of *T.chishuiensis* (green) with reference enzyme UDPN- acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase (3zl8, pink) plus reference enzyme of monomer Alanine racemase (4lut,blue).



Fig. 8. A. Superimposed amino acid involved in the active site of the Phyre2 model of alanine racemase of *T. chishuiensis* (green) with reference enzyme monomer of alanine racemase (4lut) (blue). Active sites show in the boxes (red for reference and pink modeled). B. Superimposed of amino acids involved in the binding site of the Phyre2 model of alanine racemase of *T. chishuiensis* (green) with reference enzyme UDP-N- acetylmuramoyl-tripeptide-D-alanyl-D-alanineligase (3zl8)(blue). Binding sites are shown in the boxes (red for reference and pink is formodeled)

KELIYCKDHHELNECIVQYCHQIRSGEEDD LKVFTWSQKSEATLRIQGIDKSSNSTTIRGV YEEREIAILIPFSDNASIENAIHCWCVLLLFK IEDSQIAERMARLHPVAMRLELLHGVDNCT IINDTYNSDLTSLQIALDFLEQQKQHPRRTV ILSDMLQIGKPDGDLYEEVAELVSQKNIYRI ICIGVALSKHKPTFRKHKKLRSIFFKSTEEFL KKMHMITFENEAILLKGARAFRFEKIEKLL EQKIHKTVLEINLSAIQNNLNVYRALLPAPV KMMAMVKAYSYGSGSYEIANLLQFAGVD YLTVAYVDEGVALRKAGISAPIMVMSPEA GAFDRMISWKLEPEIFSIYSLQQFTEIAHTL NVSHYPVHIKLDTGMHRLGFMPHELPELM AELKVNTAIKVVSIFSHLAGSDSADFDGFT AEQARNFDNMSQELMQVLPTRPLRHLVNS TGITRHPELYYDMVRLGIGMYGFDGSDKV QHQLQNIGTLKTTIAQIKELKAGETVGYSR KGKLERDSRIATVSIGYADGYFRDFGNGNA HMLVHGKPAPVVGAVCMDMCMLDITGID NAAVGDEVIVFGNELPIGQLADWAKTIPYE VLTGISQRVNRIYVNE



Fig. 9. Superimposed of amino acids involved in the binding zone of the Phyre2 model of alanine racemase of *T. chishuiensis* (green) with reference enzyme UDP-N-acetylmuramoyl-tripeptide-D-alanyl

4LUTIA POBID CHAIN SEQUENCE Er AAA2PBDEBB AAA2PBDEBB_9BACT	MNLNDILKATKAEAIPEVKDIDIAELLIDSRKIVNPARAMFIAIKGSHRDGHQFIDDAVE	e.
4LUT:A[POBID[CHAIN]SEQUENCE tr]A0A2P8D088[A0A2P8D088_98ACT	KGVRTFLISEETDTAKYKDAQFLKVTDTLTALHQLATAHRNQFHIPVIGITGSNGKTIVK	0 120
4LUT:A PDBID CHAIN SEQUENCE tr A@A2P8D@88 A@A2P8D@88_98ACT	EWLFQLLSENYXIIRSPKSYNSQIGVPLSVWRMQAMHTLGIFEAGISQPGEMSRSERILH	0 180
4LUT:A[PDBID[CHAIN]SEQUENCE tr[A0A2P8D088]A0A2P8D088_98ACT	PDIGIFTNIGEAHSEGFLNNRQKINEKLILFRHAKELIYCKD8HELNECIVQYCHQIR5G	0 240
ALUT:A[POBID[CHAIN]SEQUENCE tr]ABA2P8D088[A6A2P8D088_98ACT	EEDOLKVFTWSQKSEATLRIQGIDKSSNSTTIRGVYEEREIAILIPFSONASIENAIHCW	0 300
4LUTIA POBID CHAIN SEQUENCE tr A0A2P8D088 A0A2P8D088_98ACT	CVLLLFKIEDSQIAERMARLHPVAMRLELLHOVDNCTIINDTYNSDLTSLQIALDFLEQQ	9 360
4LUT:A]POBID[CHAIN]SEQUENCE tr]ABA2PBD058[A6A2PBD058_98ACT	KQHPRRTVILSOMLQIGKPDGDLYFEVAELVSQKNIVRIICIGVALSKHKPTFRKHKKLR	0 420
4LUT:A PDBID CHAIN SEQUENCE tr A0A2PBD0888 A0A2PBD088_98ACT	MQKITVPTHAEINLONLAPN STPPKSTEEFLKKNHHITPENFAILLKGARAPRFEKTEKLLEQKTHKTVLEINLSAIQNN	20 480
ALUT: A POBID CHAIN SEQUENCE tr AAA2PBD088 AAA2PBD088_98ACT	LNNIKNLLEEDIKICGVIR DAYGHGAVEVAKLLEKEKVDYLAVARTAEGIELRQNGITL LNVYRALLPAPVKNAANKKNYSYGSGSYIITANLLQPAGVDYLTVAYVDEGVALRKAGISA	80 540
4LUT:A POBID CHAIN SEQUENCE tr AAA2P8D088 AAA2P8D088_98ACT	PILNLGYTPDEAFEDSIKNKITMTVYSLETAQKINEIAKSLGE-KACVHVKIDSGMTRIG PIMMMSPE-AGAFORNISMKLEPEIFSIYSLQQFTEIAHTLNYSHYPVHIKLDTGMHRLG	139 599
ALUTIA POBID CHAIN SEQUENCE tr Adazpedobe Adazpedobe_98ACT	FQPNEESVQEIIELNKLEVIDLEGMFTHFATADEV-SKEVTYKQANNYKFMSDKLDEAGV FMPHELPE-LMAELKVNTAIKVVSIFSHLAGSDSADFDGFTAEQARNFDNMSQELMQVLP	198 658
4LUTIA POBID CHAIN SEQUENCE tr Adazpedoes Adazpedoes _98ACT	KIAIKHVSNSAAIMDCPDLRLNMVRAGIILYGHYPSDDVFKDRLELRPANKLKSKIGHIK TRPLRHLVNSTGITRHPLLYDMVRLGIGHYGFDGSDKVQHQLQNIGILKTTIAQIK	258 715
4LUT:A PDBID CHAIN SEQUENCE tr A0A2P8D0888 A0A2P8D0888_98ACT	QVEPGVG1575LKYTTTGKETIATVPIGYADG#TRIQKNPKVLIKGEVFDVVGRICMD ELKAGETVCVFRKGKLERDSRTATV5IGYADGYFRDFGNGNAUMLVHGKPAPVVGAVCMD	316 775
4LUT:A[PDBID]CHAIN]SEQUENCE tr]ABA2PBDBBB[ABA2PBDBBB_9BACT	QIMWRIDKDIDIKVGDEVILFGEGEVTAERIAKDLGTINYEVLCHISRRVDRVYMENNEL MCMLDITGIDNAAVGDEVIVFGNEL-PIGQLADWAKTIPYEVLTGISGRVNRIYVNE	376 831
4LUT:A PDBID CHAIN SEQUENCE tr A8A2PBDE88 A8A2PBDE88_9BACT	VQINSYLLK 385 831	

Fig. 10. Sequence alignment of alanine racemase of *T. chishuiensis* (down sequence) with reference enzyme monomer of alanine racemase of *C. difficile* (4lut) (up sequence). Active sites are shown in the boxes

Primary structure analysis of the alanine racemase sequence was done using of Protpram tool. The protein informational analysis is in table 1.

Multiple Sequence Alignment

From the BLAST results obtained from NCBI, alanineracemase was chosen from the five different species to perform the multiple sequences alignment (Table 2). These 5 enzymes were aligned by using Clustal Omega which is shown in Figure 2. Positions with a single, completely conserved residue have been demonstrated by an asterisk (*), conservation between groups with highly similar characteristics were indicated by a colon(:), conservation between groups with less similar characteristics were indicated by a period (.) and the variable regions were indicated by a gap in between. From the obtained result, it can be revealed that the active sites of all of 5 enzymes were highly conserved (Mcwilliam et al., 2013). **Prediction of the Secondary Structure**

Analysis of secondary structure prediction from GOR on alanine racemase showed that there are 335 amino acid residues involving in the formation of the helix, 176 amino acids for extended strands (beta-sheet) formation and 300 amino acid residues in the formation of the coil, consisting of 42.72 %, 21.18 %, and 36.10 %, respectively (Figure 3) (Sen *et al.*, 2005; Consortium, 2015; Hassan *et al.*, 2020).

The 3D Structure Modeling

The 3D structure of the alanine racemase was modeled using the SWISS-MODEL database. Our enzyme (alanine racemase of *Taibaiella chishuiensis*) was similar to two different enzymes: UDPN-acetylmuramoyl- tripeptide-D-alanyl-Dalanine ligase (3zl8.1.A) and monomer of alanine racemase (4lut.1.A). So that our sequence from residue 18 up to 468 is similar to a template model of 3zl8, and from 468 upto 829 residues are very similar to the template model of 4lut (Figure 4).

At which about half of our sequence (from N-terminal) identity was 27.10% (with 3zl8.1.A model template), and another half of sequence (from C-terminal) identity was 39.33% (with 4lut.1.A model template).So when we combine these two template models, it will complete all residues of our sequence and it will reveal a reliable

tr A0A2P8D088 A0A2P8D088_9BACT	MNLNDILKATKAEAIPEVKDIDIAELLIDSRKIVHPARAMFIAIKGSHRDGHDFIDDAYE	60
3ZLB:A PDBID CHAIN SEQUENCE		41
tr[A0A2P8D088]A0A2P8D088_98ACT	KGVRTFLISEETDTAKYKDAQFLKVTDTLTALHQLATAHRNOFHIPVIGITDSNGKTIVK	120
32L8:A[PDBID[CHAIN[SEQUENCE	NGAVAITAERTTVNSDRIFLVESSVDTLARLAREKLGAFSGTVVGVTDSSGKTTTK	97
tr A0A2P8D088 A0A2P8D088_9BACT	EHLFQLLSENYKIIRSPKSYNSQIGVPLSVWRMQAUHTLGIFEAGISQPGEMSRSERILH	180
3ZL8:A PDBID CHAIN SEQUENCE	EILVNLLKNRSVFKTPGNWNTFYGLPLSILNDYKGHEILVLEMAASRPGDTAHLCKIAP	157
LP[A0A2P8D088]A0A2P8D088_9BACT 3ZL8:A[PDBID]CHAIN[SEQUENCE	PDIGIFTNIGEAHSEGFLNNRQKINEKLILFRHAKELIYCKDHHELNECIYQYCHQI PDVAVLLNYGSAHLEFFGTRENIMETKMEIIXHSKENAIAYTLFDDPDLRKEVPRYNNTL **1.11 *1*******	237 217
tr[A0A2P8D088]A0A2P8D088_98ACT	RSGEEDD-LKVFTWSQKSEAT-LRIQGIDKSSNSTTIRGVYEEREIAILIPFSDNASILR	295
3ZL8:A[PDBID[CHAIN]SEQUENCE	FFGKEGGDSVLKDWAVYEGSTIAEFEAFD-SLFTVKLSGVHVGQLLW	264
tr A0A2P8D088 A0A2P8D088_9BACT	ATHCWCVLLLFKIEDSQTAERMARLHPVAMILELLHGVDNCTIINOTYNEDLTSLQIALD	355
3ZL8:A PDBID CHAIN SEQUENCE	TAASLCVMRTLGETVDIFDLASLFVVCGFNV-REKEGVLIVDOTYNEDBEAFQTSIE	321
tr[A0A2P8D088]A0A2P8D088_9BACT	FLEQQKQHPRRTVILSDMLQIGKPDGDLYEEVAELVSQ-KNIYRIICIGVALSKHKPTFR	414
JZL8:A[PDBID[CHAIN[SEQUENCE	ALLRFPGKFAVVGAMKELGRSKEFHEELGRRNVLDGVVVFLSEPRAEWI	373
tr[A0A2P8D088]A0A2P8D088_90ACT	KHKKLRSIFFKSTEEFLKKM®MITFENCAILLKGARAFRFEKIEKLLEQKIHKTVLEI	472
3ZL8:A[PDBID]CHAIN[SEQUENCE	KSKKIILKSDOPEKTAKDLATRVKKODVVLKASRAVNIERVLEMFEKELEKRAAL	430
tr ABA2PRDBRB ABA2PRDBR8_BBACT 32LBIA PDBID CHAIN SEQUENCE	NL SATQNNLNVVRAL LPAPVKMMAMVKAVSYGSGSYETANL LQFAGVDYLTVAVVDEGVA EHRHORH : ::	532 437
tr[A0A2P8D088]A0A2P8D088_98ACT 32L8:A[PDBID]CHAIN]SEQUENCE	LRKAGISAPIMVMSPEAGAFDRMISWKLEPEIFSIYSLQQFTEIAHTLNVSHYPVHIKLD	592 437
tr A0A2P8D088 A0A2P8D088_98ACT 3ZL8:A PDBID CHAIN SEQUENCE	TGPHRLGFMPHELPELMAELKVNTAIKVV5IF5HLAG5D5ADFDGFTAEQARNFDNMSQE	652 437
tr[A0A2P8D088[A0A2P8D088_98ACT 32L81A[PD8ID[CHAIN[SEQUENCE	LMQVLPTRPLRHLVNSTGITRHPELYVDHVRLGIGMVGPDGSDKVQHQLQNIGTLKTTIA	712 437
Tr[A0A2P8D088]A0A2P8D088_9BACT 3ZL8:A[PDBID[CHAIN]SEQUENCE	QIKELKAGETVGYSRKGKLERDSRIATVSIGYADGYFRDFGNGNAHMLVHGKPAPVVGAV	772 437
tr A0A2P8D088 A0A2P8D088_9BACT 3ZL8: A POBID CHAIN SEQUENCE	CMDMCMLDITGIDNAAVGDEVIVFGNELPIGQLADWAKTIPVEVLTGISQRVNRIVVNE 831 437	

Fig. 11. Sequence alignment of alanine racemase of *T. chishuiensis* (top sequence) with reference enzyme UDP-Nacetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase of *T. maritima* (3zl8) (down sequence). Binding sites are shown in the highlighted boxes. Also, the boxes without highlight show the binding region

model for our sequence. Also, we modeled our sequence using Phyer2 (Figure 4). The generated structure, which was consists of one chain is very similar to the combined model of SWISS-MODELER

Homology Model Validation Ramachandran Plot

Ramachandran plot was used to validate the accuracy of the 3D structure by visualizing the phi (\hat{O}) and psi (q) dihedral angles of the residues of amino acids in the protein structure (Mcwilliam *et al.*, 2013) (Figure 5). In this case, there are 568 amino acid residues in the preferred zone, which provide 76.6 % accuracy, 127 amino acid residues (16.9 %) in the additional allowed regions, as well as 30 amino acid residues (4.0 %) in the generously allowed regions. There are 26 (3.5 %) residues of amino acid in the disallowed zones.

Verify 3D

The Verify3D software has been used to assess the consistency of the degree of every amino acid in a structure to the environment which includes the polarity. The alanine racemase

Table 3	The	active	sites	of the	model	enzyme	and	reference	enzymes.

The active sites of Alanine racemase, <i>T. chishuiensis</i> and alanine racemase, <i>C. difficile</i>	The active sites of Alanine racemase, <i>T. chishuiensis</i> plus UDP-N-acetylmuramoyl-tripeptide-D-alanyl- D-alanine ligase, T. maritime
Alanine racemase, <i>T. chishuiensis</i> Lys 499	Alanine racemase, <i>T. chishuiens is</i> Asn 295, UDP-N-
Alanine racemase, <i>T. chishulensis</i> Tyr 499	maritima Asn 264
Alanine racemase, C. difficile Tyr 725	Alanine racemase, <i>T. chishuiensis</i> Arg 326, UDP-N- acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase, <i>T. maritima</i> Arg 293
	Alanine racemase, <i>T. chishuiensis</i> Lys 447, UDP-N- acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase, <i>T. maritima</i> Lys 405

Table 4. The Sarface of modeled enzyme alanine racemase of T. chishuiensis and reference enzymes of alanine racemase C. difficile as well as T. maritima



616

predicted structure possesses an 84% of the residues with the averaged 3D-1D score $\geq =0.2$.

From the obtained results of validation software, it is revealed that the model generated possesseda high resolution and high accuracy. Thus, the 3D model generated for alanine racemase is highly reliable (Figure6).

3D structure comparison

A comparative study was performed to investigate the structural differences of alanine racemase and the contribution to the ligand-binding cofactor on its function, based on the sequence alignment with different alanine racemase. The generated 3D structure from Phyre2 for alanine racemase of T.chishuiensis was super imposed with the structure of the templates obtained from SWISS-Model (3zl8, UDP-N-acetylmuramoyltripeptide-D-alanyl-D-alanine ligase, also 4lut, alanine racemase of C. difficile) (Figure7). From the superimposed structure shown in Figure 8, A it can be evealed that the catalytic residues of alanine racemase of T. chishuiensis is conserved with alanine racemase of C. difficile (PDB ID: 4lut). Also, the superimposed structure is shown in Figure 8,B revealed that the binding sites of UDP- N-acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase of Thermotoga maritima (PDB ID:3zl8) are conserved with alanineracemase of T. chishuiensis. Besides, the binding region of UDP-N- acetylmuramoyl- tripeptide- D-alanyl-D-alanine ligase of T. maritima (PDB ID: 3zl8) is conserved with alanine racemase of T. chishuiensis (Figure 9). The enzyme UDP-Nacetylmuramoyltripeptide- D-alanyl-D-alanine ligase function is: ATP+UDP-N-acetylmuramoyl-L-alanyl-gamma-D- glutamyl-L-lysine +

D-alanyl-D-alanine = ADP + phosphate + UDP-N-acetylmuramoyl-L-alanyl- gamma-D-glutamyl-L-lysyl-D-alanyl-D-alanine (KEGG-Enzyme) (Bhardwaj *et al.*,2018).

These observations show that the enzyme alanine racemase of *T. chishuiensis* is a bifunctional enzyme. This enzyme works with racemization of L-alanine and D-alanine as well as ligation of amino acid (Favini-stabile *et al.*, 2013). Therefore, the enzyme alanine racemase of *T. chishuiensis* functionally is the same with alanine racemase of *C. difficile* and also has the same function with UDP-Nacetylmuramoyl-tripeptide- D-alanyl-Dalanine ligase of *T. maritima*. So, this enzyme has two functions that we can call it a bifunctional enzyme.

Structurally, there are differences between alanine racemase of *T. chishuiensis* and alanine racemase of *C. difficile* as well as some difference between alanineracemase and UDP-Nacetylmuramoyl-tripeptide-D-alanyl-D-alanine ligaseof *T.maritima*. These two parts of our enzyme are connected at the loop region: I465, H466, K467, and T468.

The superimposed structure of alanine racemase with binding site labeled with the reference structure (UDP-N-acetylmuramoyltripeptide-D-alanyl-D-alanine ligase (3zl8)) shows that the binding site was found at the similar region, overlapping each other (Figure 8, B and Figure 11). This specifies that the binding site for both alanine racemase and UDP-N- acetylmuramoyltripeptide-D-alanyl-D-alanine ligase (3zl8) was conserved and possessed similar functions.

Also, there are two binding regions. The first binding region shows less conservation and it



Fig. 12. Major differences between modeled enzyme alanine racemase of *T. chishuiensis* (green) and a reference model of alanine racemase *C. difficile* (blue)



Fig. 13. Major differences between modeled enzyme alanine racemase of *T. chishuiensis* (green) and a reference model of enzyme UDP- Nacetylmuramoyl-tripeptide-D-alanyl-D- alanine ligase of *T. maritima* (3zl8, blue), highlighted by red color

is because of some changes of amino acids such as, the change of amino acid S to amino acid N, but both of them are small and polar so functionally both of them are the same, and also the change of amino acid T to amino acid I, which have the same characteristics, they are hydrophobic. The second binding region is strongly conserved.

The major different spots between alanine racemase of *T. chishuiensis* and alanine racemase of *C. difficile* were highlighted by using PYMOL (Figure 12, Table3 and Table4) and shown in red color.

The main different spots between alanine racemase of *T.chishuiensis* and reference enzyme UDP-Nacetylmuramoyl-tripeptide-D-alanyl-Dalanine ligase of *T. maritima* (3zl8) are highlighted by using PYMOL (Figure 13 and Table 4). Major different spots between two sequences are highlighted in redcolor.

The predicted model of alanine racemase of T. chishuiensis from this study could serve as a useful model for further study of the structure and function of other bifunctional enzymes. Based on this study, the name of alanineracemase of T.chishuiensis can be bifunctional UDP-Nacetylmuramoyl- tripeptide: D-alanyl-D-alanine ligase/alanine racemase in the UniProt database. Besides, the structural information about alanine racemase of T. chishuiensis especially the active sites and binding sites could help drug designers to do further research. However, many scientists have been studied for the design of more specific alanine racemase inhibitors by targeting the active sites and binding sites and mutations in the enzyme strain in various microorganisms (Anthony et al., 2011; Azam and Jayaram, 2016). Due to the lack of specificity of most of the previous inhibitors, further study is required for highly selective alanine racemase inhibitors (Azam and Jayaram, 2016). Since the current study is the first study of 3D structural characterization of bifunctional alanine racemase in T. chishuiensis, it may provide useful information about the structure and function of this enzyme for drug designers to do more investigation for more specific inhibitors of alanine racemases and provide effective treatment.

CONCLUSION

In the currentstudy, the bioinformatics

approaches have been used to predict the 3D structure of alanineracemase (EC:5.1.1.1) from T. chishuiensis. Due to the absence of a 3D structure of the bifunctional enzyme (UDP-Nacetylmuramoyl-tripeptide: D-alanyl-D- alanine ligase/alanine racemase) in PDB database, we have done the homology modeling with template models of SWISS- MODELLER database (UDP-N- acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase (3zl8) plus alanine racemase (4lut)).Our enzyme (alanine racemase of Taibaiella chishuiensis) was similar to two different enzymes: UDP-N-acetylmuramoyl-tripeptide- D-alanyl-Dalanine ligase (PDB ID: 3zl8, EC: 6.3.2.10) and of alanine racemase (PDB ID: 4lut, EC: 5.1.1.1). From the sequence and 3D structure analysis, we realized that the first half of the sequence of our modeled enzyme, from residue 18 upto 468 was 27.10% identical with the existing template 3zl8, and the identity of the second half of our query from 468 upto 829 residues was 39.33%, with the template 4lut. Hence, the combination of these two template models has made a reliable model for our sequence. Both, alanine racemase from Taibaiella chishuiensis and alanine racemase of C.difficile (4lut) have the same active site at Lys (499), Lys (39), Tyr (499), Tyr (725)(Tan etal., 2014). On the other hand, bothalanine racemase from Taibaiella chishuiensis and UDP-N-acetylmuramoyltripeptide-D-alanyl- D-alanine ligase (3zl8) from T. maritima also have the same binding site at Asn (295), Asn (264), Arg(326), Arg(293), Lys(447), Lys (405) (Dong et al., 2018). Besides, verify 3D shows good quality with a percentage of 80. Based on these observations, our enzyme (alanine racemase from T. chishuiensis) is bifunctional. Therefore, it has two commission numbers (EC: 5.1.1.1 and EC: 6.3.2.10). Findings from this study may help to contribute to the rational design of the remaining unknown structure of the same enzyme in other organisms and assist in designing proteins model with enhanced properties.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my main lecturer, Prof. Dr. Mohd Shahir Shamsir Omar, for encouragement, guidance, critics, and friendship. Without his continued support and interest, this study would not have been the same as presented here. I am also indebted to the Ministry of Higher Education Afghanistan and Higher Education Development Program (HEDP), for supporting and funding this study. Special thanks to the University Technology Malaysia (UTM) for providing the opportunity to study at this prestigious university.

REFERENCES

- M. Muhammad and B. Zhao, "Purification , Characterization and Inhibition of Alanine Racemase from a Pathogenic Strain of Streptococcus iniae," *Polish J. Microbiol.*, 2019; 68(3): pp. 331–341.
- M.Muhammad,Y.Li,S.Gong,Y.Shi,andJ.Ju, "Cloning, Biochemical Characterization and Inhibition of Alanine racemase from Streptococcusiniae,"*bioRxiv*, 2019; no. 611251: pp.1–25.
- A. Wood and C. Gunsalus, "d-Alanine Formation: A Racemase In Streptococcus Faecalis," *ournal Biol.Chem.*, 1959; 190(1): pp. 403–416.
- S.Liu*etal.*, "Function of alanineracemase in the physiological activity and cariogenicity of Streptococcus mutans," *Sci. Rep.*, 2018; 8(30): pp. 1–8.
- M. S. Islam, A. Aryasomayajula, and P. R. Selvaganapathy, "A Review on Macroscale and Microscale Cell Lysis Methods," *Micromachines*, 2017; 8(83): pp. 1–27.
- T. Nomura, I. Yamamoto, F. Morishita, Y. Furukawa, and O. Matsushima, "Purification and Some Properties of Alanine Racemase From a Bivalve Mollusc Corbicula japonica," *J. Exp. Zool.*, 2001; 9(289): pp. 1–9.
- I. Schomburg *et al.*, "BRENDA, the enzyme database/: updates and major new developments," *Nucleic Acids Res.*, 2004; **32**(4): pp. 431–433.
- Y. Wei *et al.*, "Alanine racemase is essential for the growth and interspecies competitiveness of Streptococcus mutans," *Int. J. Oral Sci.*, 2016; 8(4): pp. 231–238.
- H.Im, M.L.Sharpe, U.Strych, M. Davlieva and K. L. Krause, "The crystal structure of alanine racemase from Streptococcus pneumoniae, a target for structure-based drug design," *BMC Microbiol.*, 2011; **11**: pp. 1–15.
- K. G. Anthony *et al.*, "New classes of alanine racemase inhibitors identified by high- through put screening show antimicrobial activity against mycobacterium tuberculosis," *PLoS One*, 2011; 6(5): pp. 1–9.
- 11. M. A. Azam and U. Jayaram, "Inhibitors of

alanine racemase enzyme: A review," J.Enzyme Inhib. Med. Chem., 2016; **31**(4): pp. 517–526.

- X.Tan, R. Zhang, T. Meng,H.Liang,andJ.Lv, "Taibaiella chishuiensis sp. nov., isolated from freshwater," *Int. J. Syst. Evol. Microbiol.*, 2014; 64: pp. 1795–1801.
- R. M. Couñago, M. Davlieva, U. Strych, R. E. Hill, and K. L. Krause, "from Bacillus anthracis (Ames)," *BMC Struct. Biol*, 2009; 15: pp. 1–15.
- H. Dong *et al.*, "Biochemical and Biophysical Research Communications Enzymatic characterization and crystal structure of biosynthetic alanine racemase from Pseudomonas aeruginosa PAO1," *Biochem. Biophys. Res. Commun.*, 2018; **503**(4): pp. 2319–2325.
- J. Ju *et al.*, "Correlation between catalytic activity and monomer — dimer equilibrium of bacterial alanine racemases," *J. Biochem.*, 2011; **149**(1): pp. 83–89.
- O. A. Asojo, D. A. Hyman, and M. Monica, "Structural and biochemical analyses of alanine racemase from the multidrugresistant Clostridium difficile strain 630," *Biol. Crystallogr.*, 2014; **70**(7): pp. 1922–1933.
- P. Chen *et al.*, "Structure of the full-length Clostridium difficile toxin B," *Nat. Struct. Mol. Biol.*, 2019; 26(8): pp. 712–719.
- S. Favini-stabile, C. Contreras-martel, N. Thielens, and A. Dessen, "MreB and MurG as scaffolds for the cytoplasmic steps of peptidoglycan biosynthesis," *Environ. Microbiol.*, 2013; 15(12): pp. 3218–3228.
- E. Gasteiger *et al.*, "ExPASy/ : the proteomics server for in-depth protein knowledge and analysis,"*Nucleic Acids Res.*, 2003; **31**(13): pp. 3784–3788.
- S. Hassan, J. Javid, A. Sadiqua, N. Fatima, A. Mohsin, and H. Akram, "Thyroid Cancer TC-1/: an insight from 3D structure prediction to virtual screening for computational drug design," *Biomed. Lett.*, 2020; 6(1): pp. 127–137.
- T. Z. Sen, R. L. Jernigan, J. Garnier, and A. Kloczkowski, "Structural bioinformatics GOR V server for protein secondary structure prediction," *Bioinformatics*, 2005; **21**(11): pp. 2787–2788.
- M. Biasini *et al.*, "SWISS-MODEL/: modelling protein tertiary and quaternary structure using evolutionary information," *Nucleic Acids Res.*, 2014; **42**: W1, pp. W252–W258.
- L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, "The Phyre2 web portal for protein modeling, prediction and analysis," *Nat. Protoc.*, 2015; **10**(6): pp. 845–858.
- 24. P. D. O. De Meo and D. Cozzetto, "The PMDB Protein Model Database," *Nucleic Acids Res.*,

2006; **34**(suppl_1): pp. 306–309.

- 28. M. Kouza and E. Faraggi, "Chapter 2," *Humana Press. New York, NY.*, 2017; **1484**: pp. 7–24.
- G. Janson, C. Zhang, M. G. Prado, and A. Paiardini, "Structural bioinformatics PyMod 2 . 0/ : improvements in protein sequence-structure analysis and homology modeling within PyMOL," *Bioinformatics*, 33(3): pp.444–446,2017.
- U. Consortium, "UniProt/ : a hub for protein information,"*Nucleic Acids Res.*, 2015; 43(D1), pp. D204–D212.
- H. Mcwilliam *et al.*, "Analysis Tool Web Services from the EMBL-EBI," *Nucleic Acids Res.*, 2013; 41(W1): pp. 597–600.
- 29. M. Wiltgen and G. G. Hospital, *Algorithms for Structure Comparison and Analysis/: Homology Modelling of Proteins*, vol.1.ElsevierLtd., 2018.
- T. Bhardwaj and P.Somvanshi, "Gene Reports A computational approach using mathematical modeling to assess the peptidoglycan biosynthesis of Clostridium botulinum ATCC 3502 for potential drug targets," *Gene Reports*, 2018; 12: pp. 179–186.