Research Article
Malays. j. med. biol. res.

Homology Modelling and in *Silico* Substrate Binding Analysis of a *Rhizobium* sp. RC1 Haloalkanoic Acid Permease

Muhammed Adamu Musa[1], Fahrul Zaman Huyop[2*]

[1]Department of Biotechnology and Medical Engineering, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor, **MALAYSIA**
[2]Professor, Department of Biotechnology and Medical Engineering, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor, **MALAYSIA**

[*]Email for Correspondence: fahrul@utm.my

ABSTRACT

Rhizobium sp. RC1 grows on haloalkanoic acid (haloacid) pollutants and expresses a haloacid permease (DehrP) which mediates the uptake of haloacids into the cells. For the first time, we report the homology model and docking analysis of DehrP and proposed its putative binding residues. The Protein Data Bank for protein of similar sequence. Ligand structures were retrieved from the ChemSpider database. The 3-dimensional (3-D) structure of DehrP was modelled based on the structure of Staphylococcus epidermidis glucose: H+ symporter (GlcPse) by Phyre2, refined by 3Drefine and evaluated by ProSA z-score, ERRAT and RAMPAGE. Docking of monobromoacetate, monochloroacetate, dibromoacetate, dichloroacetate, trichloroacetate, and 2,2-dichloropropionate ligands was done with AutoDock vina1.1.2. The 3-D structure of DehrP protein has twelve transmembrane helices. The overall quality factor of the model is ~91%, with 93.6% of the residues in the favored region and the z-score is within the ≤ 10 limit. The putative H+ binding site residues are Gln133, Asp36, and Arg130. Docking analysis showed that Glu33, Trp34, Phe37, Phe38, Gln165, and Glu370 are potential haloacid interacting residues. DehrP-haloacid complexes had a binding affinity between -2.9 to -4.0 kcal/mol. DehrP has both putative H+ and haloacid binding sites that are most likely involved in the co-transport of H+ and haloacids. DehrP interacts with haloacids majorly through van der Waals and halogen bond interactions and has greater affinity for 2,2-dichloropropionate and could be a specialized chloropropionate uptake system. Site-directed mutagenesis of DehrP binding residues could improve its haloacid binding affinity.

Key words: Docking, Haloacids, Homology modelling, Permease, Rhizobium sp. RC1

Manuscript Received: 02 January 2020 - Revised: 02 March 2020 - Accepted: 18 March 2020

INTRODUCTION

Microbial metabolism of haloalkanoic acid (haloacid) pollutants depends on haloacid transport (Su *et al.*, 2012, Su *et al.*, 2013) and dehalogenation (Slater, 1979, Allison N., 1983, Tsang *et al.*, 1988, Stringfellow *et al.*, 1997). Some haloacids are difficult to be degraded by microorganisms (Effendi *et al.*, 2000, Berthiaume *et al.*, 2014) because they do not support cell growth (van der Ploeg and Janssen, 1995). Cases of haloacid selectivity by cells (Berry *et al.*, 1979, Janssen *et al.*, 1985, Chaudhry and Chapalamadugu, 1991) and haloacid cytotoxicity (Strotmann *et al.*, 1990, Plewa *et al.*, 2010) have been reported. Dehalogenase-associated inducible haloacid transport proteins are involved in the uptake of haloacids into the cytoplasm of bacteria (Higgins *et al.*, 2005, Yu *et al.*, 2007, Jing *et al.*, 2010). Kinetics studies have shown that *Burkholderia caribensis* MBA4 haloacid transporter (Dehp2) has preference for chloropropionate over

monochloroacetate (Su and Tsang, 2013). Therefore, it is important to understand the structure and mechanism of haloacid transport proteins for efficient haloacid degradation. At present, the three-dimensional (3-D) structure of haloacid transport proteins and their mechanism of action is unknown.

Rhizobium sp. RC1, can use haloacids as carbon and energy source by producing three distinct dehalogenases (D-2-haloacid dehalogenase; DehD, L-2-haloacid dehalogenase; DehL and dual isomeric haloacid dehalogenase; DehE) to cleave the carbon-halogen (C···X) bonds (Leigh *et al.*, 1986, Leigh *et al.*, 1988, Huyop, 2003). *Rhizobium* sp. RC1 haloacid permease gene *dehrP* is found 511 bases upstream of the D-2-haloacid dehalogenase gene *dehD*, and encodes the haloacid transport protein (DehrP) (Jing *et al.*, 2010). DehrP has the Major Facilitator Superfamily (MFS) transport protein conserved domain with a sugar signature (Jing *et al.*, 2010) and is closely related to the Metabolite: H^+ Symporter family of haloacid transporters (Musa, 2017) (Deh4p and Dehp2) from *Burkholderia caribensis* MBA4, a subfamily of the MFS (Tse *et al.*, 2009, Saier *et al.*, 2016).

Proteins from the six subfamilies of the MFS show low sequence similarities with unique substrate specificities, and distinct transport coupling mechanisms, but they all have a common structural organisation, known as the MFS fold (Abramson *et al.*, 2004). Therefore, these transporters conform to the same three-dimensional (3-D) structural arrangement with twelve (12) transmembrane helices (TMs) that are arranged into two folded domains (N and C), each domain having six TMs (Abramson *et al.*, 2003, Huang *et al.*, 2003, Yin *et al.*, 2006, Dang *et al.*, 2010, Newstead *et al.*, 2011, Solcan *et al.*, 2012, Sun *et al.*, 2012, Yan, 2013).

In this study, for the first time, we report the homology modelling and haloacid binding analysis of *Rhizobium* sp. RC1 haloacid permease (DehrP). The three-dimensional (3-D) structure of DehrP was modelled based on the crystal structure of *Staphylococcus epidermidis* glucose: H^+ symporter (GlcPse) (Iancu *et al.*, 2013). We conducted a comparative binding analysis of the putative haloacid binding site of DehrP with six haloacid ligands (monobromoacetate; MBA, monochloroacetate; MCA, dibromoacetate; DBA, dichloroacetate; DCA, trichloroacetate; TCA, and 2,2-dichroropropionate; 2,2-DCP) to identify the putative residues that could be involved in haloacid transport.

MATERIALS AND METHODS

Sequence Retrieval and Alignment

Rhizobium sp. RC1 haloacid permease (DehrP) (Jing *et al.*, 2010) amino acid sequence was downloaded from the UniProtKB Database (Consortium, 2015) under the accession number Q1M2W6. NCBI-Blastp (Altschup *et al.*, 1990), PSI-Blast (Schaffer *et al.*, 2001) and UniProt-Blastp (Consortium, 2015) were used to analyze the amino acid sequence of DehrP. DehrP sequence was aligned with *Staphylococcus epidermidis* glucose: H^+ symporter (GlcPse, Uniprot number: Q5HKL0) (Iancu *et al.*, 2013) using Clustal omega (Sievers *et al.*, 2011).

Homology Modelling and Model Evaluation

The three-dimensional structure of DehrP was modelled based on the crystal structure of *Staphylococcus epidermidis* glucose: H^+ symporter (GlcPse, PDB accession number: 4LDS B) by Phyre^2 (Kelley *et al.*, 2015) server, refined by 3D^{refine} (Bhattacharya *et al.*, 2016) and evaluated using ProSA-web (Wiederstein and Sippl, 2007), ERRAT (Colovos and Yeates, 1993), and RAMPAGE (Lovell *et al.*, 2003). DehrP 3-D structure was then superimposed with GlcPse (template) 3-D structure by the Chimera 1.11.2rc (Pettersen *et al.*, 2004) software. Based on the longer length of GlcPse structure (425 residues) the stretch of residues from the C-terminal of *Agrobacterim* sp. NHG3 (with 98% sequence identity, see supplementary Figure 1) without corresponding residues in DehrP was added to DehrP C-terminal before modelling.

MOLECULAR DOCKING

Ligand and Receptor Preparation

The 3-D structures of D-glucose (DGlc), monobromoacetate (MBA), monochloroacetate (MCA), dichloroacetate (DCA), dibromoacetate (DBA), trichloroacetate (TCA), and 2,2-dichroropropionate (2,2-DCP) were downloaded from ChemSpider (Pence and Williams, 2010) ligand database under the following ChemSpider ID; 96749, 5991, 452706, 24195, 5602522, 106518, and 2815701, respectively. The ligand was used as input files for AutoDock Tools 1.5.6 (Morris *et al.*, 2009). The ligand rigid roots were automatically set and all possible rotatable bonds and torsions were defined as active. The modelled DehrP 3-D structure (receptor) was used for docking simulation. The protein consists of one single chain B (419 residues). DehrP in pdb format was inputted into AutoDock Tools 1.5.6 (Morris *et al.*, 2009), after which polar hydrogen atoms were added and Kollman charges, atomic solvation parameters and fragmental volumes were assigned to the protein. The Chain B (425 residues) of the 3-D structure of *Staphylococcus epidermidis* glucose transporter (GlcPse; PDB: 4LDS) downloaded from the Protein Data Bank was equally prepared for docking the docking procedure as a control.

```
DehrP       MTTTLVARTSSAGRMTREERKVIFASSLGTVFEWYDFFLYGSLAAIIGATFFKDFPPATQ    60
DehrP*      MTTTLVARTSSAGRMTREERKVIFASSLGTVFEWYDFFLYGSLAAIIGATFFKDFPPATQ    60
            ************************************************************

DehrP       AIFALLAFAAGSLVRTFGALIFGRLGDMIGRKYTFLVTILIMGLSTFVVGLLPGSDTIGL    120
DehrP*      AIFALLAFAAGSLVRTFGALIFGRLGDMIGRKYTFLVTILIMGLSTFVVGLLPGSDTIGL    120
            ************************************************************

DehrP       AAPTILILLRLLQGLALGGEYGGAAVYEAEHAPPGRRGFYTSWIQTTATGGLFLSLLVIL    180
DehrP*      AAPTILILLRLLQGLALGGEYGGAAVYEAEHAPPGRRGFYTSWIQTTATGGLFLSLLVIL    180
            ************************************************************

DehrP       GTRSLLGEESFTSWGWRVPFLLSVVLLGISVWIRMQLNESPVFQRMKAEGKASKAPLREA    240
DehrP*      GTRSLLGEESFTSWGWRVPFLLSVVLLGISVWIRMQLNESPVFQRMKAEGKASKAPLREA    240
            ************************************************************

DehrP       FAHWPNARLALVALFGMVAGQAVVWYTGQFYVLFFLQSILKVDGFTTTLLICWSLLLGSG    300
DehrP*      FAHWPNARLALVALFGMVAGQAVVWYTGQFYVLFFLQSILKVDGFTTTLLICWSLLLGSG    300
            ************************************************************

DehrP       FFVFFGWLSDRIGRKPIMIAGCLLAVVTYFPIFEAITERANPTLAKAISEVKVTRGLPTR    360
DehrP*      FFVFFGWLSDRIGRKPIMIAGCLLAVVTYFPIFEAITERANPTLAKAISEVKVTRGLPTR    360
            ************************************************************

DehrP       SNAGIYSIRSEPACSHRLVTLPEPILLRVRCDMIVRQVLQESQLACSSMAPR--------    412
DehrP*      SNAGIYS------------IPEPILLRVRCDMIVRQVLQESQLACSSMAPRVLVNGTEV    407
            *******            :*******************************

DehrP       ------------------------------------------------------------    412
DehrP*      KFDTAQLKESQTRVTAALQAAGYPKPGQSAAVHMTHAFDFGKPQVLPLIGLLFILAIYVT    467

DehrP       ------------------------------------------------------------    412
DehrP*      MVYGPMAAALVELFPARIRYSGLSLPYHIGNGWFGGLLPAAAFAMVAQTGDIYFGLWYPI    527

DehrP       ----------------------------    412
DehrP*      VIAAVTVVVGLIWLPETKDRDIHALD      553
```

Supplementary Figure 1: Pairwise sequence alignment of *Rhizobium* sp. RC1 haloacid transport (DehrP, Uniprot accession number: Q1M2W6,(Jing *et al.*, 2010) with the new sequence (DehrP*). The deleted 13 amino acids (highlighted in red) and the additional 154 amino acids (highlighted in cyan) is from the C-terminal of a 98% identical *Agrobacterium* sp. NHG3 haloacid transporter (DehP, Uniprot accession number: Q8KLT0, (Higgins *et al.*, 2005) after DehrP sequence was aligned with DehP sequence. The asterisks (*) indicates conserved residues as generated by Clustal Omega (Sievers *et al.*, 2011) sequence alignment server.

Docking Procedure

AutoDock vina 1.1.2 (Trott and Olson, 2010) was used for docking simulation which employs the preparation of receptor by adding hydrogens and assigning Kollman charges, followed by conversion of PDB file to pdbqt. Ligands were assigned with Gasteiger charges and non-polar hydrogens. Docking simulations were run using Lamarckian Genetic Algorithm (LGA) which is known to be the most efficient and reliable method of AutoDock. The ligand centered maps were generated by AutoGrid program with a spacing of 0.200 Å and dimensions of 100 x 100 x 100 points. The grid box center was set to coordinate 3.01841, 48.2993, and 67.4807 in x, y, and z respectively. The default settings were used for all other parameters. For each docking simulation, nine (9) different conformers were generated. Chimera 1.11.2 (Pettersen *et al.*, 2004) software was then used to visualize the obtained docking conformations. The AutoDock vina 1.1.2 (Trott and Olson, 2010) binding free energies (ΔGb in kcal/mol) for the seven (7) ligand-receptor complexes for DehrP and GlcPse were then compared. The results were visualized in UCSF Chimera v.1.11.2rc (Pettersen *et al.*, 2004) and Discovery studio v.16.1.0 (BIOVIA, 2016).

RESULTS AND DISCUSSION

Sequence Alignment

Comparative analysis of protein sequences by sequence alignment provides useful information during structural and functional analysis by revealing sequence-structure-function relationships (Liò and Bishop, 2008, Shenoy and Jayaram, 2010). PSI-Blast of *Rhizobium* sp. RC1 haloacid permease sequence in the NCBI database shows that DehrP has 30%

sequence identity with *Staphylococcus epidermidis* glucose: H^+ symporter (GlcPse, NCBI accession number: WP_002486092.1). An alignment of the two sequences shows the DehrP and GlcPse share some common amino acids at specific locations (Figure 1), including some glucose and H^+ binding site residues.

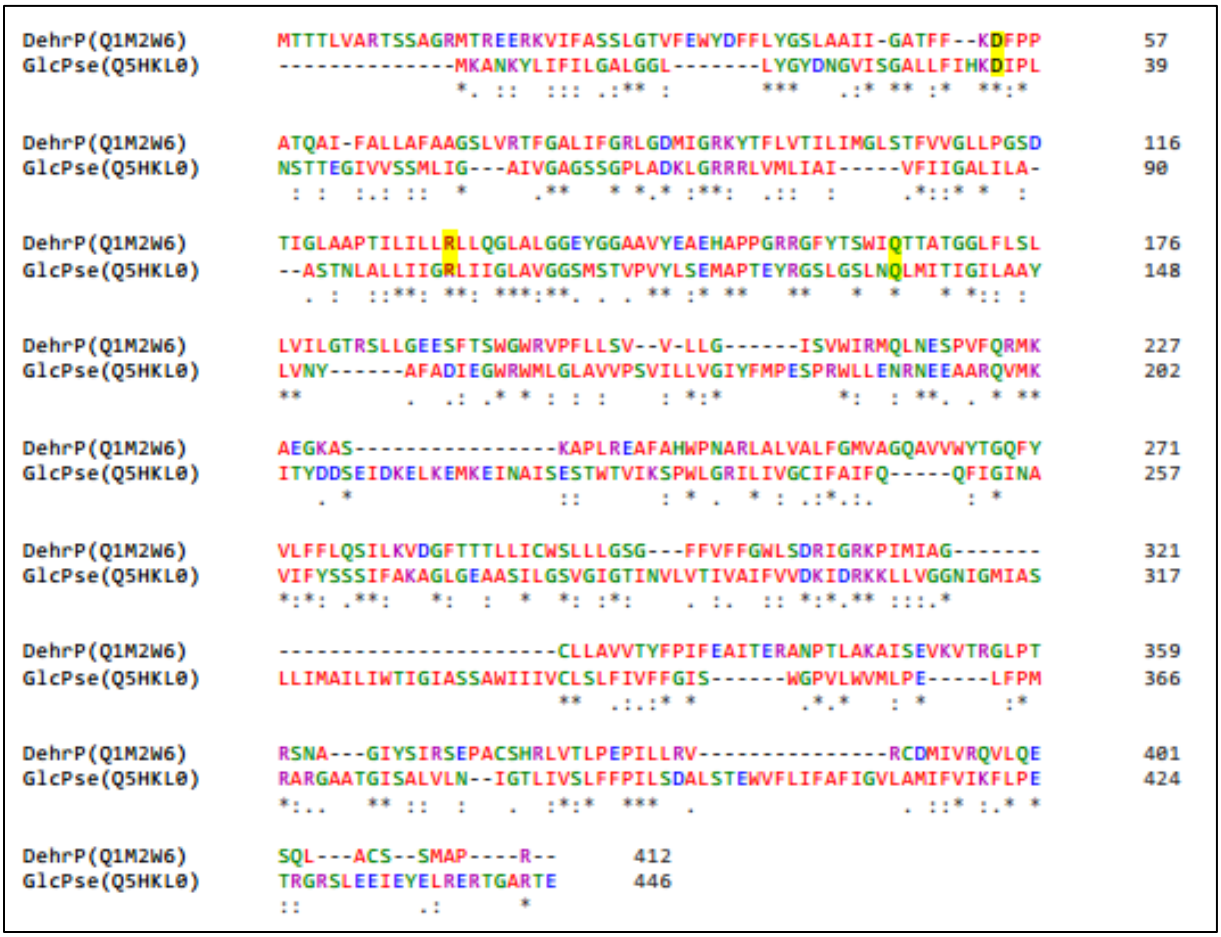


Figure 1: Alignment of DehrP sequence with glucose symporter (GlcPse) template. Asterisks (*) indicate highly conserved amino acid residues, double dots (:) and single dot (.) indicates moderate and low conservation respectively. The conserved glucose binding residue (Gln^{137}) and H^+ binding residues (Asp^{22} and Arg^{102}) are highlighted in yellow colour and the Uniprot accession number for the transporters are in bracket. The alignment was generated by Clustal Omega (Sievers *et al.*, 2011).

Homology Modelling

As suggested by Abramson *et al.* (2004), transporter proteins that belong to the major facilitator Superfamily (MFS) may agree to the same three-dimensional (3-D) structure. Since DehrP is a member of the MFS (Jing *et al.*, 2010, Saier *et al.*, 2016), it is expected to have a similar structure to the experimentally determined 3-D structures of microbial MFS transporters. DehrP sequence was threaded onto the crystal structure (at 3.2-Å resolution) of GlcPse (PDB accession number: 4LDS) by Phyre[2] software.

The homology model (Figure 2) of DehrP has twelve (12) transmembrane helices (12TMs) that are interlinked by six periplasmic loops on the extracellular face and five intracellular loops on the cytoplasmic face of the membrane. Although the large periplasmic loop between TMs 9 and 10 proposed by Tse *et al.* (2009), was not modelled by the Phyre[2] (Kelley *et al.*, 2015) server, obviously because the template does not have such a loop, the model has the 12 TMs typical of a Major Facilitator Superfamily (MFS) transporter. Therefore, removal of the large loop could allow for the modelling of more haloacid transporters transmembrane helices based on the existing experimentally determine 3-D structures of MFS transports in the Protein Data Bank. The relatively large cytoplasmic α-helical loop connecting TM6 and TM7 (Figure 3) serves as a linker between the two domains (N and C) of the 3-D structure of DehrP and is within the 30-100 residues that connect the two domains of MFS transporters (Abramson *et al.*, 2003, Huang *et al.*, 2003, Yin *et al.*, 2006, Dang *et al.*, 2010, Newstead *et al.*, 2011, Solcan *et al.*, 2012, Sun *et al.*, 2012, Iancu *et al.*, 2013). Like all the known 3-D structures of MFS transporters, the intracellular loops between TMs 2 and 3, and 8 and 9 of DehrP appears to be short. The short loops in MFS are presumed to provide restraint to the loops in the cytoplasmic face of the MFS proteins

(Yan, 2013). The two DehrP domains (N and C) are superimposable (Figure 4) with the inward facing structure of GlcPse (template) and the common binding site residues in the cavity of the domains are well aligned.

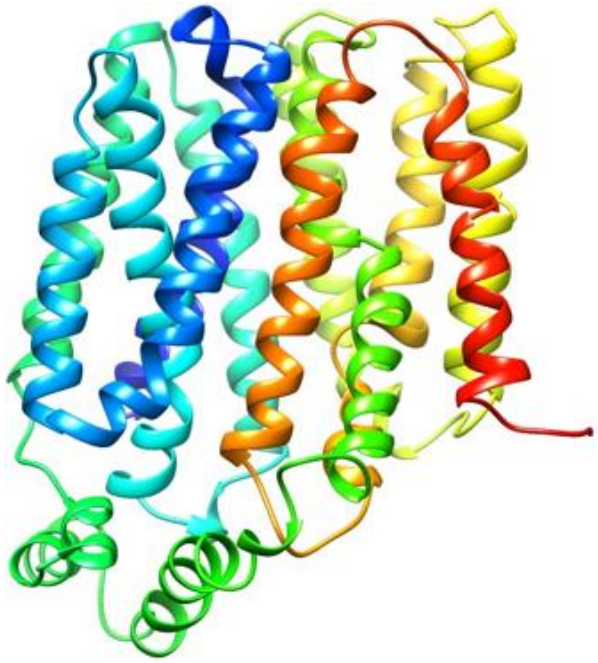


Figure 2: Homology model of DehrP. There are 12 transmembrane α-helices and the rainbow-coloured spectrum was used for the distinct N (blue) and C (brown) domains starting from the left in structures. The structure was prepared using UCSF Chimera 1.11.2rc (Pettersen *et al.*, 2004) structural visualization software.

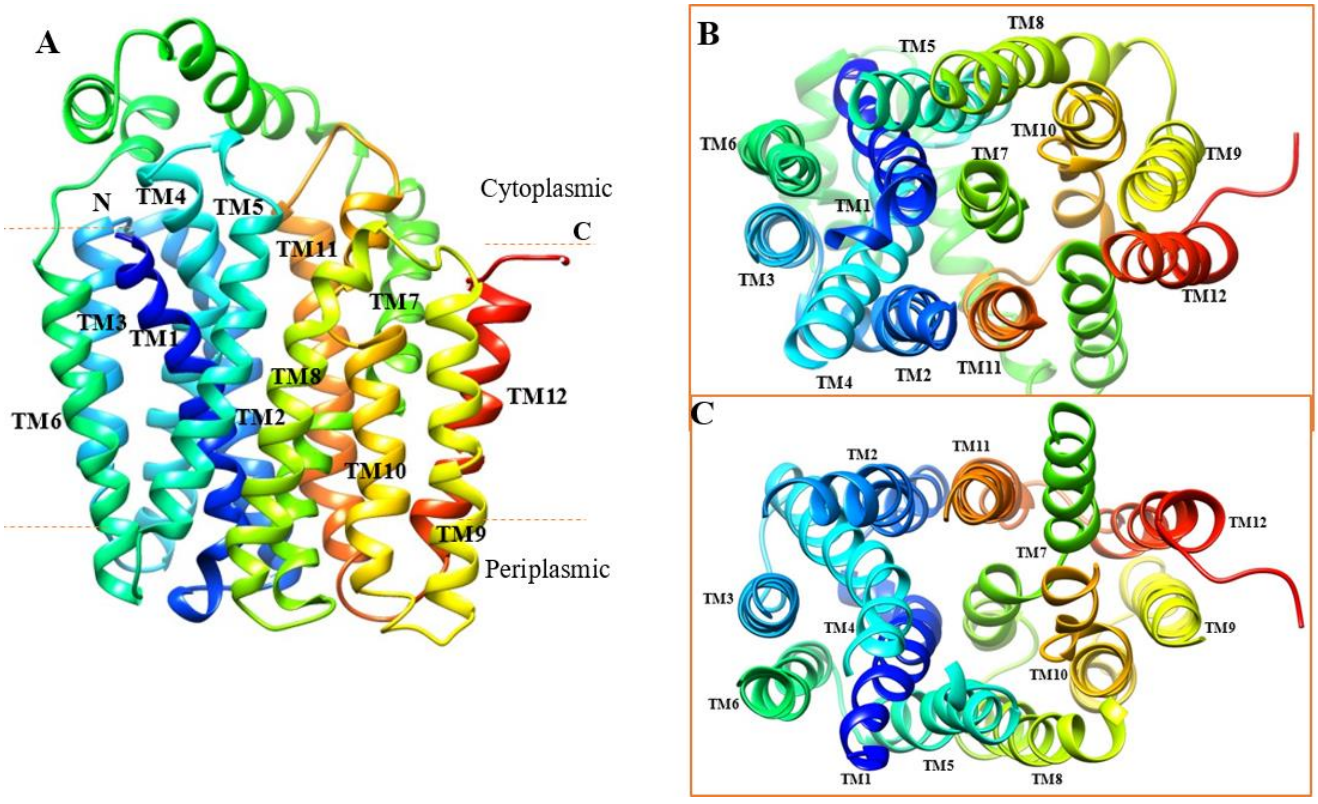


Figure 3: Overview of DehrP structural arrangement. (A) Side-view representation of the 12 TMs as two-fold pseudosymmetrical domains connected by a relatively large cytoplasmic α-helical loop between TM6 and TM7. (B) Periplasmic-view showing the cavity that leads into the periplasmic space with TM7 at the middle while TMs 3, 6, 9, and 12 farther away from the cavity. (C) Cytoplasmic-face showing the cavity that leads into the cytoplasm with TM7 at the middle while TMs 3, 6, 9, and 12 farther away from the cavity.

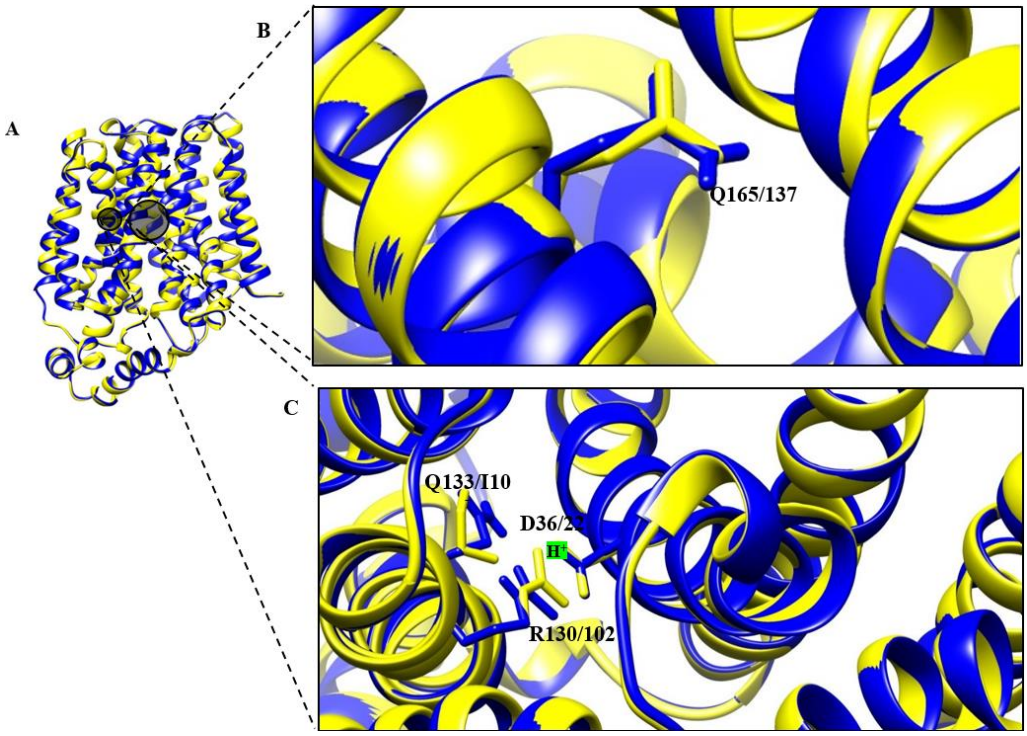
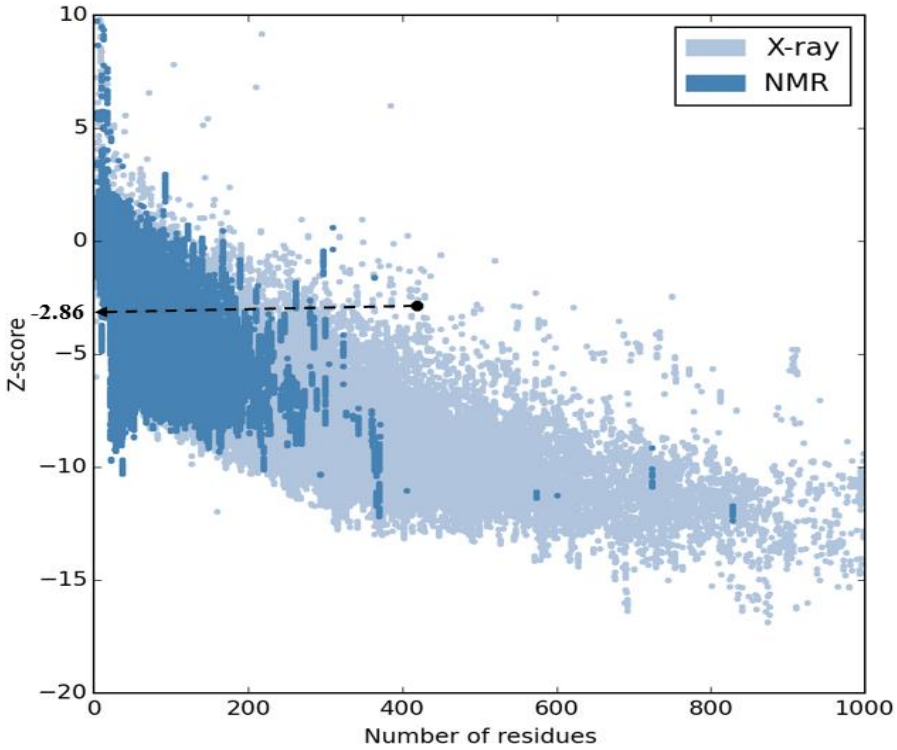


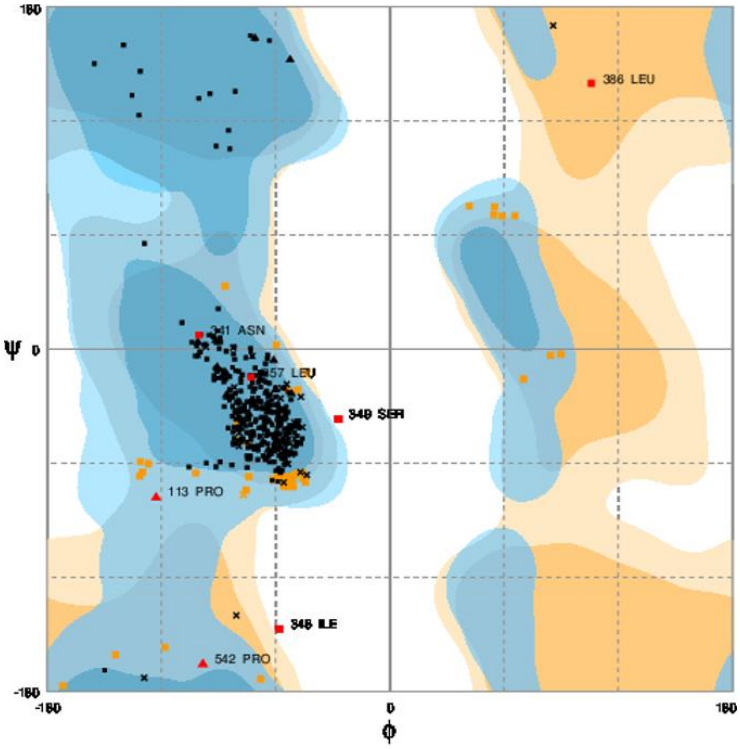
Figure 4: Structural alignment of DehrP. (A) Superposition of DehrP (blue) and GlcPse (yellow) in a cartoon representation. (B) Gln^{165} of DehrP is superposed on the glucose binding site residue (Gln^{137}) of GlcPse. (C) Gln^{133}, Asp^{36}, and Arg^{130} of DehrP superposition on the proposed H^+ binding site residues (Ile^{105}, Asp^{22}, and Arg^{102}) of GlcPse. The grey box shows the domain region including the catalytic region. Structural alignment and visualization of binding residues was done using UCSF Chimera 1.11.2rc (Pettersen *et al*., 2004) software.



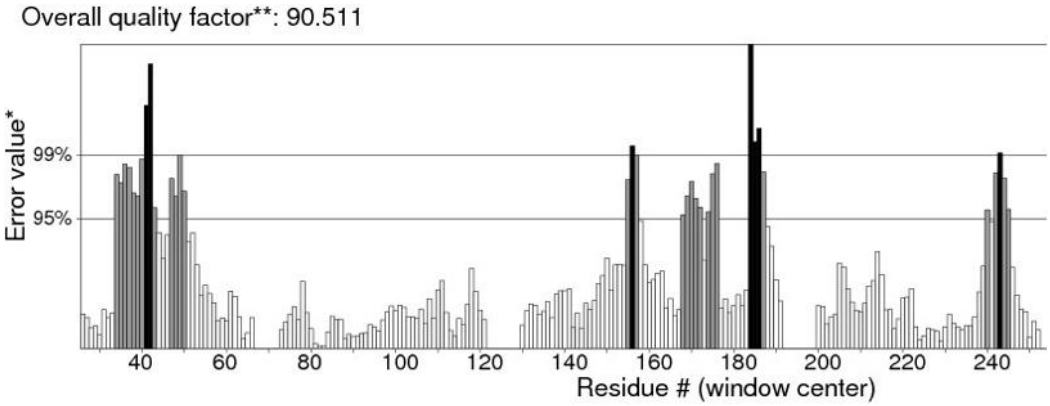
Supplementary Figure 2: The overall model quality (OMQ) of DehrP. The model z-score (highlighted as large dot) is -2.86 according to the ProSA-web (Wiederstein and Sippl, 2007), and it is within the z-score ≤ 10 limit for residues below 1000 (Wiederstein and Sippl, 2007). This shows the ProSA-web z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length with domain chains less than 1000 residues.

Validity of the 3-D Model

The quality and reliability of the DehrP structure was assessed using z-score ProSA-web (Wiederstein and Sippl, 2007), RAMPAGE (Lovell *et al.*, 2003) Ramachandran plot and ERRAT (Colovos and Yeates, 1993). The z-score of the overall DehrP structure (Supplementary Figure 2) shows that it is within the range of the usual scores (z-score ≤ 10) for native proteins of similar size (Wiederstein and Sippl, 2007). Further, the stereochemical quality (phi; Φ) and psi; Ψ angles) of the model by Ramachandran plot reveals that most of the residues (93.6%) are in favoured region (Supplementary Figure 3) while 3.9% and 2.4% allowed and outlier regions respectively. To check the reliability of the model, a statistical analysis of the nonbonded interactions between different atom types was done by ERRAT (Colovos and Yeates, 1993). The ERRAT plot (Supplementary Figure 4) shows that the overall quality factor (OQF) of the model is ~91%. The average OQF of low resolutions (2.5 to 3 Å) structures is put at 91% (Colovos and Yeates, 1993), and therefore the 3-D structure is a reliable model.



Supplementary Figure 3: Ramachandran plot showing residues in the most favourable region and disallowed regions by RAMPAGE (Lovell *et al.*, 2003). The number of residues in the favoured, allowed, and outlier regions is 383 (93.6%), 16 (3.9%), and 10 (2.4%) residues, respectively.



Supplementary Figure 4: The graphic of ERRAT (Colovos and Yeates, 1993) program for DehrP after refinement by $3D^{refine}$ (Bhattacharya *et al.*, 2016) server. The overall quality factor (OQF) of the model is ~91%. ∗On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value. ∗∗Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit. Low resolutions (2.5 to 3 Å) the average OQF is around 91% (Colovos and Yeates, 1993).

PUTATIVE BINDING RESIDUES AND DOCKING ANALYSIS

Transporters are membrane proteins that mediate the uptake of substrates into and out of the cells. The transport process is possible because of the presence of key residues that coordinate the passage of substrates across the membrane (Hashiramoto, 1992, Mueckler and Makepeace, 2009). Based on sequence similarity, some of these essential residues have been shown to be conserved among the characterized major facilitator superfamily (MFS) transporters (Sanderson *et al.*, 1998, Dang *et al.*, 2010). Superposition of the 3-D structure of DehrP on the 3-D structure of GlcPse reveal the presence of common binding site residues (Figure 4). Asp^{36}, Arg^{130} and Gln^{133} in DehrP overlapped, respectively, with the H^+-binding site residues; Ile^{105}, Asp^{22}, and Arg^{102} in GlcPse, while Gln^{165} overlapped the glucose-binding site residue; Gln^{137}. Studies has shown that these residues play important roles in the transport of H^+ into the cells (Dang *et al.*, 2010, Iancu *et al.*, 2013). *Burkholderia caribensis* MBA4 haloacid transporters (Deh4p and Dehp2) have been reported to be dependent on pH and transmembrane electrochemical gradient for transport activities (Su *et al.*, 2012). Further, previous comparative analysis show that DehrP is a member of the Metabolite: H^+ symporter (MHS) based on its sequence similarity with known members of MHS (Musa, 2017). Perhaps, these amino acids residues serve as H^+-binding site during haloacid transport into the cell since Asp^{36}, Arg^{130}, and Gln^{133} are part DehrP sequence motifs (Musa, 2017). On the other hand, biochemical studies have shown that Glutamine is important for transport activity and exofacial ligand binding (Mueckler *et al.*, 1994). Therefore, the Gln^{165} of DehrP may be important in haloacid binding as well.

Docking analysis of the 3-D structure of DehrP with DGlc, MBA, MCA, DBA, DCA, TCA, and 2,2-DCP shows that in addition to the Gln^{165} there are at least four other residues that might be involved in haloacid binding. This was revealed by the docking simulation performed by AutoDock vina 1.1.2 (Trott and Olson, 2010). AutoDock vina 1.1.2 uses the Lamarckian Genetic algorithm (LGA), which is known to be the most efficient and reliable method for docking (Morris *et al.*, 1998). From the docking simulation, the best conformation complexes were selected from 9 docking poses for each of the ligand. Visual screening of the docking poses shows that Glu^{33}, Trp^{34}, Phe^{37}, Phe^{38}, Gln^{165}, and Glu^{370} are potential binding site residues in DehrP cavity (Figure 5) and these residues correspond respectively to Thr^{19}, Gly^{20}, Asn^{23}, Gly^{24}, Gln^{137}, and Asn^{256} in GlcPse. Both Gln^{137} and Asn^{256} are among the five glucose binding site residues in GlcPse (Figure 6) (Iancu *et al.*, 2013). Glu^{33}, Trp^{34}, Phe^{37} and Phe^{38} are found on TM1. Gln^{165} and Glu^{370} are located on TMs 5 and 7, respectively. Residues on TMs 1, 4, 5, 7, and 10 have been reported to contribute to substrate binding in Major Facilitator Superfamily (MFS) transporters (Abramson *et al.*, 2003, Huang *et al.*, 2003, Yin *et al.*, 2006, Dang *et al.*, 2010, Newstead *et al.*, 2011, Solcan *et al.*, 2012, Sun *et al.*, 2012, Iancu *et al.*, 2013). Glu^{33}, Trp^{34}, Phe^{37}, Phe^{38} and Gln^{165} are part of DehrP sequence motifs (Musa, 2017).

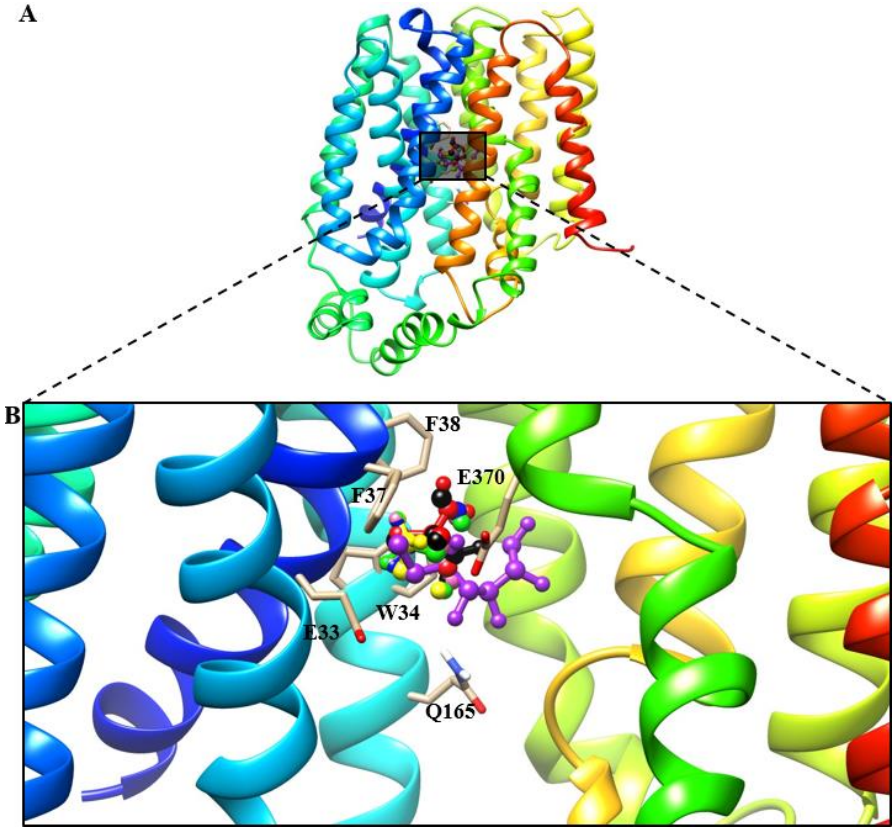


Figure 5: Docked conformation of DehrP (model). with DGlc. (A) Putative binding site DehrP binding site in the inward-facing conformation (rainbow) complexed with D-glucose (DGlc), monobromoacetate (MBA),

monochloroacetate (MCA), dibromoacetate (DBA), dichloroacetate (DCA), trichloroacetate (TCA), and 2, 2-dichloropropionate (2, 2-DCP) shown as rectangular gray shading. (B) Expanded view of the binding site residues; Glu^{33}, Trp^{34}, Phe^{37}, Phe^{38}, Gln^{165}, and Glu^{370} (stick model) and DGlc (purple), MBA (yellow), MCA (blue), DBA (green), DCA (pink), TCA (black), and 2, 2-DCP (red) in the ball and stick representation. The figure was prepared using the UCSF Chimera 1.11.2rc (Pettersen *et al.*, 2004) software.

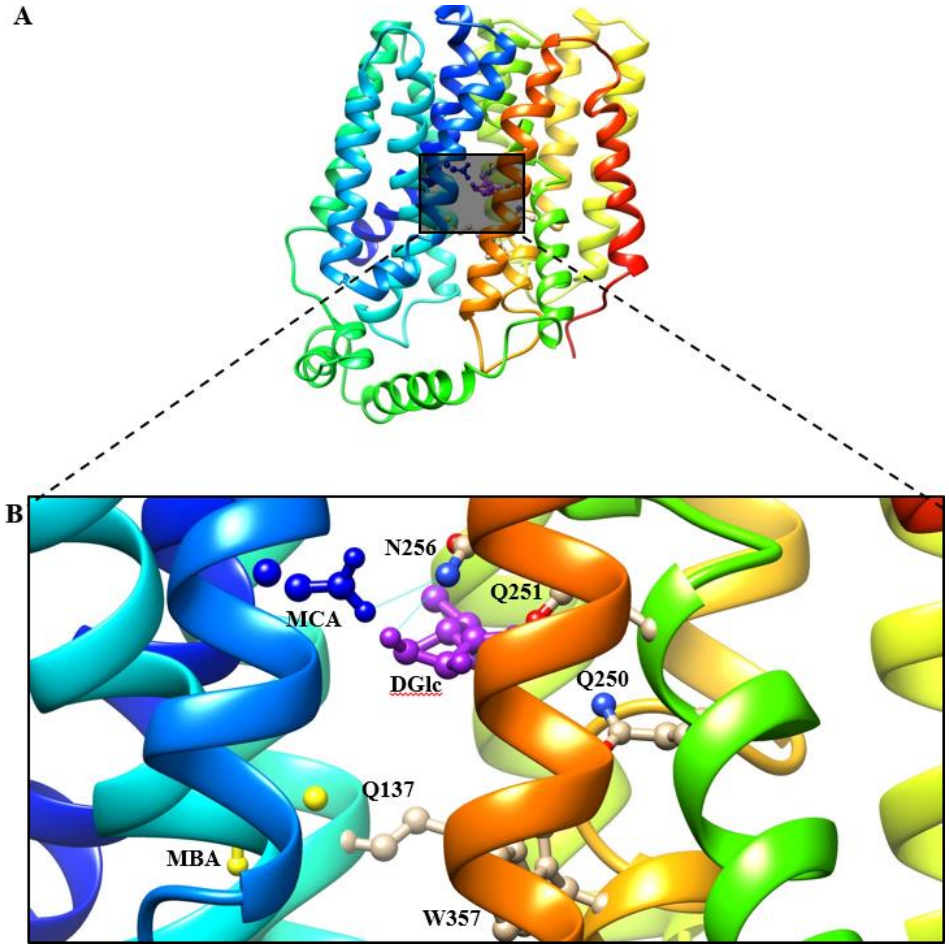


Figure 6: Docked conformation of GlcPse (template). (A) The glucose binding site in the inward-facing conformation (rainbow) complexed with D-glucose (DGlc), and monobromoacetate (MBA), monochloroacetate (MCA) shown as rectangular gray shading. (B) Expanded view of the binding site residues (Gln^{137}, Gln^{250}, Gln^{251}, Asn^{256}, and Trp^{367}) complexed with DGlc (purple), MBA (yellow), and MCA (blue) in the ball and stick representation. Blue lines represent hydrogen bonds as computed by the UCSF Chimera 1.11.2rc (Pettersen *et al.*, 2004) software.

Ligand	Affinity (kcal/mol.)		Distance from best mode (RMSD)	
	GlcPse	DehrP	GlcPse Lower/ upper bounds	DehrP Lower/ upper bounds
DGlc	-4.6	-4.2	37.064/ 38.093	12.051/13.486
MBA	-2.9	-2.9	53.11/ 53.426	17.618/17.822
MCA	-3.2	-2.9	10.986/11.084	12.488/12.804
DBA	-	-3.6	-	17.534/ 18.321
DCA	-	-3.5	-	18.615/ 19.141
TCA	-	-3.9	-	19.169/ 19.810
2, 2-DCP	-	-4.0	-	19.556/ 20.219

Supplementary Table 1: The affinities of DehrP and GlcPse complexes generated by AutoDock vina 1.1.2(Trott and Olson, 2010).

The AutoDock vina (Trott and Olson, 2010) binding free energies (see Supplementary Table 1) for the seven DehrP-ligands complexes formed shows some variations in values. The best docked conformation and affinity of DehrP-DGlc complex is comparable with that of GlcPse-DGlc complex. DehrP-MBA and DehrP-MCA complexes had the lowest

binding affinity. This could be because the two ligands are less substituted with halogen. GlcPse-MBA and GlcPse-MCA complexes showed similar affinity. Increase in the number of halogens corresponded with an increase in the affinity of DehrP for the ligands. Perhaps, the halogens contributed some weak form of interaction between the binding site residues as illustrated in Figure 7. Halogen bonds (C–X···D–Z) between halogenated ligands and binding site residues containing nitrogen and oxygen significantly alters the ligand binding affinity to the target protein receptor (Olsen *et al.*, 2004), depending on the lipophilicity substituent constant (π) of the halogen (Bois *et al.*, 1998). The higher affinity of DehrP for DBA compared to DCA might be due to the higher lipophilicity substituent constant (π) of bromine atom. Docking of DCA, DBA, TCA and 2,2-DCP with GlcPse shows that the ligand could not interact with the glucose binding site residues. This is consistent with the fact that GlcPse is not a transporter of haloacids and this can be attributed to the difference in the binding site residues in DehrP and GlcPse. The presence of electron-rich halogens in haloacids and the halogen bond-acceptor side chains in DehrP (Trp^{34} Phe^{37}, Phe^{38}, Glu^{33} and Glu^{370}) may have contributed to the greater affinity of DehrP for haloacids. This is consistent with the affinity of halogens for O and aromatic rings (Voth Regier and Ho Shing, 2007, Matter *et al.*, 2009, Hardegger *et al.*, 2011). Among all the haloacids docked with DehrP, 2,2-DCP had the highest binding affinity, may be because it is the original substrate for *Rhizobium* sp. RC1. Further analysis of DehrP binding site conformations using Discovery Studio v.16.1.0 (BIOVIA, 2016) shows that the haloacid ligands interacts with DehrP binding site residues majorly through van der Waals interactions (Figure 8). In addition to the six putative binding site residues (Glu^{33}, Trp^{34}, Phe^{37}, Phe^{38}, Gln^{165}, and Glu^{370}). Ala^{165} and Leu^{373} appears to be involved in the stabilization of DehrP-haloacid complexes. Leu and Ala residues are reported to have a high frequency of invlovement in halogen bond formation (Sirimulla *et al.*, 2013), and leu is considered to be an important halogen interaction initiator (Wilcken *et al.*, 2012).

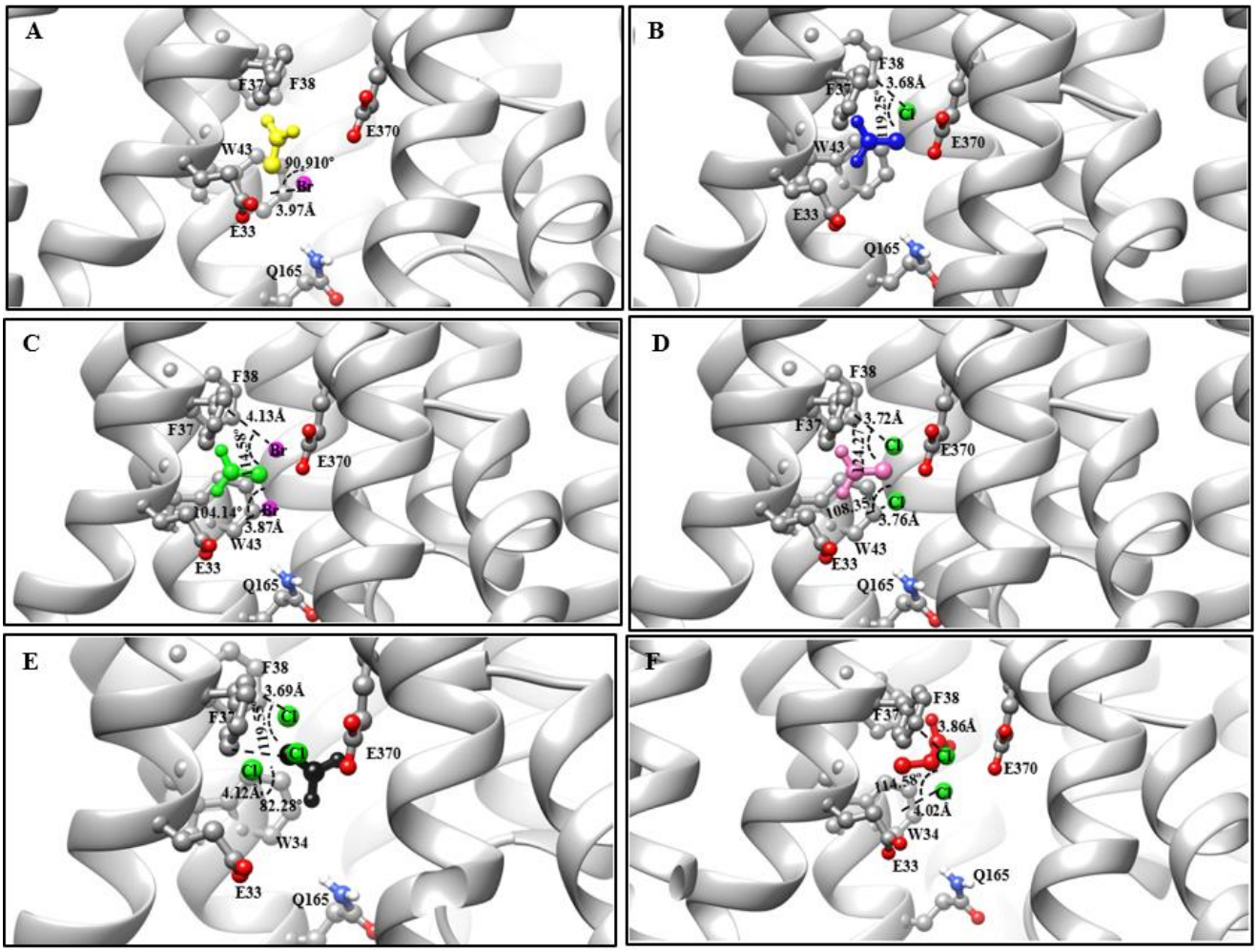


Figure 7: Illustration of halogen bonds and aromatic interactions in DehrP-haloacid complexes. The ligands and binding site residues are shown in the ball and stick presentation. The protein is presented as a cartoon model and coloured by atom type with carbon atom in grey, hydrogen in white, oxygen in red, nitrogen in blue, chloride in green and bromide in purple. (A) Monobromoacetate (MBA in yellow) ligand forming Br···aromatic ring interaction with Trp^{34}. (B) Monochloroacetate (MCA in blue) ligand forming Cl···aromatic ring interaction with Phe^{38}. (C)

Dibromoacetate (DBA in green) ligand forming Br···aromatic ring interactions with Trp[34] and Phe[38]. (D) Dichloroacetate (DCA in pink) ligand forming Cl···aromatic ring interactions with Trp[34] and Phe[38]. (E) Trichloroacetate (TCA in black) ligand forming Cl···aromatic ring interactions with Trp34, Phe[37], and Phe[38]. (F) 2,2-dichloropropionate (2,2-DCP in red) ligand forming Cl···aromatic ring interactions with Trp[34] and Phe[38]. The figure was prepared using Chimera 1.11.2rc (Pettersen *et al.*, 2004) software.

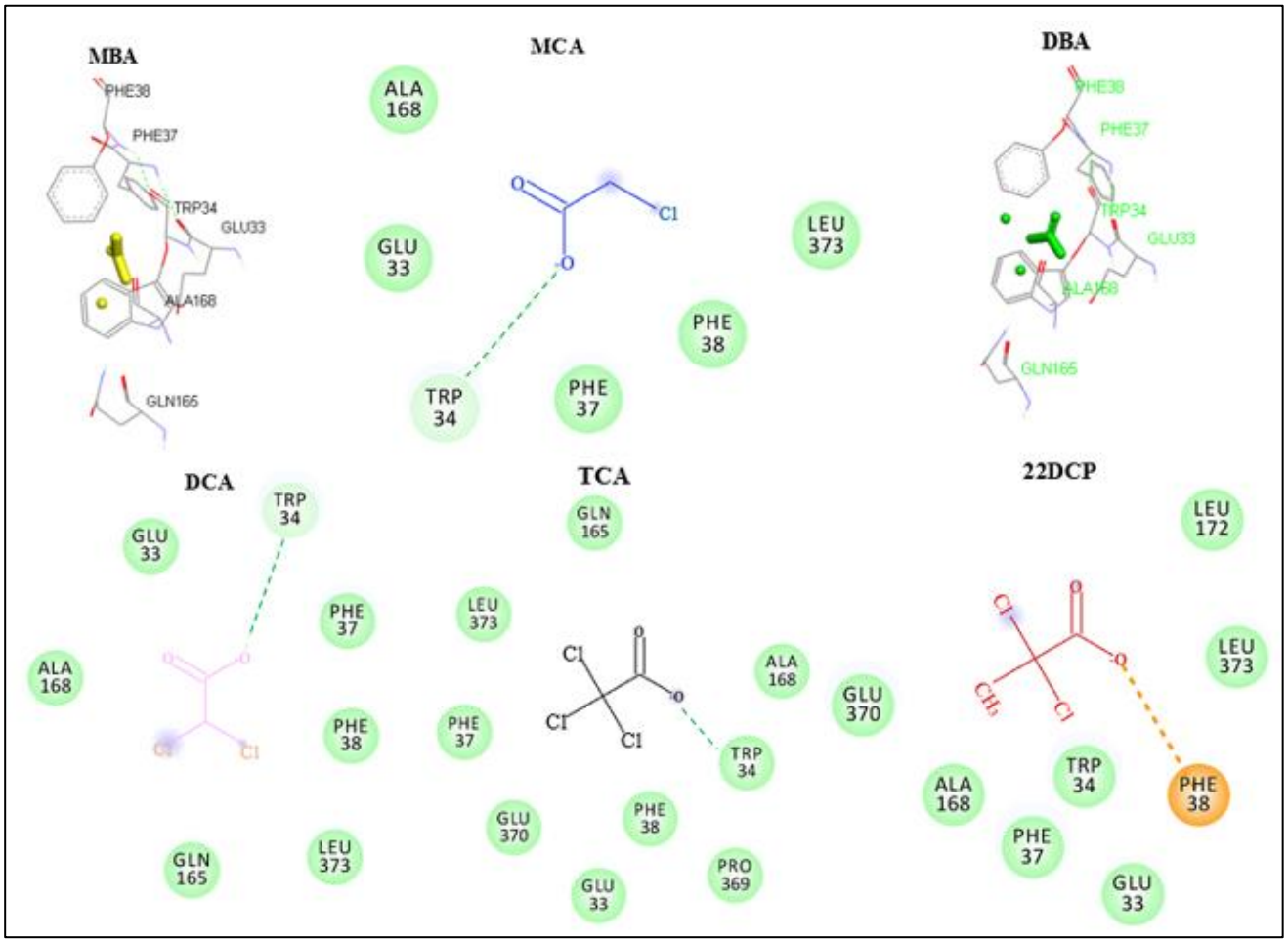


Figure 8: Overall DehrP interacting residues with haloacid ligands. 2-D Schematics generated by Discovery studio v.16.1.0 (BIOVIA, 2016) shows DehrP residues that contacts monobromoacetate (MBA), monochloroacetate (MCA), dichloroacetate (DCA) trichloroacetae (TCA) and 2,2-dichloro-propionate (2,2-DCP), respectively. Green dotted lines indicate van der Waals interactions, and the pi-anion interaction encountered is coloured orange.

CONCLUSION

Rhizobium sp. RC1 haloacid permease (DehrP) has a Major Facilitator Superfamily (MFS) that can be folded into twelve transmembrane helices, represented as two-fold pseudosymmetrical domains connected by a relatively large cytoplasmic loop between helices transmembrane helices 6 and 7. To model the structure with the large periplasmic loop linking transmembrane helices 9 and 10, experimental study is needed. The proposed DehrP haloacid binding site residues that might be important in haloacids transport are Glu[33], Trp[34], Phe[37], Phe[38], Gln[165], and Glu[370] and may interact with haloacids majorly through Van der Waals and halogen bond interactions. Haloacid transport might be coupled with H^+ transport via the H^+ binding site residues; Asp[36], Arg[130] and Gln[133]. Docking analysis around the putative haloacid binding residues shows that DehrP has more affinity for 2,2-DCP than it has for MBA, MCA, DBA, DCA, and TCA.

ACKNOWLEDGMENT

This work was supported by the Research University Grant (RUG) Vot Q.J130000.2545.12H91, awarded to Universiti Teknologi Malaysia by the Ministry of Education Malaysia. The funding body had no role in the design of the study, analysis, and interpretation of data and in writing of the manuscript. MAM thanks the Kaduna State Government of Nigeria Scholarship for studentship award.

REFERENCE

Abramson, J., Kaback, H.R. and Iwata, S. (2004), "Structural comparison of lactose permease and the glycerol-3-phosphate antiporter: members of the major facilitator superfamily", *Current Opinion in Structural Biology,* Vol. 14, No. 4, pp. 413-419.

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R. and Iwata, S. (2003), "Structure and mechanism of the lactose permease of *Escherichia coli*", *Science,* Vol. 301, No. 5633, pp. 610-615.

Allison N., S., A. and Cooper, R. (1983), "The dehalogenases of a 2, 2-dichloropropionate degrading bacterium.", *Journal of General Microbiology,* Vol. 129, No. 5, pp. 1283–1293.

Altschup, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990), "Basic local alignment search tool", *Journal of Molecular Biology,* Vol. 215, No. 3, pp. 403-410.

Berry, E.K.M., Allison, N., Skinner, A.J. and Cooper, R.A. (1979), "Degradation of the Selective Herbicide 2,2-Dichloropropionate (Dalapon) by a Soil Bacterium", *Microbiology,* Vol. 110, No. 1, pp. 39-45.

Berthiaume, C., Gilbert, Y., Fournier-Larente, J., Pluchon, C., Filion, G., Jubinville, E., *et al.* (2014), "Identification of dichloroacetic acid degrading Cupriavidus bacteria in a drinking water distribution network model", *Journal of Applied Microbiology,* Vol. 116, No. 1, pp. 208-221.

Bhattacharya, D., Nowotny, J., Cao, R. and Cheng, J. (2016), "3Drefine: an interactive web server for efficient protein structure refinement", *Nucleic Acids Research,* Vol. 44, No. W1, pp. W406-409.

BIOVIA, D.S. (2016), "Discovery Studio Modeling Environment, Release 2017", *Dassault Systèmes, San Diego, CA.,* Vol., No.

Bois, F., Beney, C., Boumendjel, A., Mariotte, A.-M., Conseil, G. and Di Pietro, A. (1998), "Halogenated Chalcones with High-Affinity Binding to P-Glycoprotein: Potential Modulators of Multidrug Resistance", *Journal of Medicinal Chemistry,* Vol. 41, No. 21, pp. 4161-4164.

Chaudhry, G.R. and Chapalamadugu, S. (1991), "Biodegradation of halogenated organic compounds", *Microbiological Reviews,* Vol. 55, No. 1, pp. 59-79.

Colovos, C. and Yeates, T.O. (1993), "Verification of protein structures: patterns of nonbonded atomic interactions", *Protein Science : A Publication of the Protein Society,* Vol. 2, No. 9, pp. 1511-1519.

Consortium, U. (2015), "UniProt: a hub for protein information", *Nucleic Acids Research,* Vol. 43, No. Database issue, pp. D204-212.

Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., *et al.* (2010), "Structure of a fucose transporter in an outward-open conformation", *Nature,* Vol. 467, No. 7316, pp. 734-738.

Effendi, A.J., Greenaway, S.D. and Dancer, B.N. (2000), "Isolation and Characterization of 2,3-Dichloro-1-Propanol-Degrading Rhizobia", *Applied and Environmental Microbiology,* Vol. 66, No. 7, pp. 2882-2887.

Hardegger, L.A., Kuhn, B., Spinnler, B., Anselm, L., Ecabert, R., Stihle, M., *et al.* (2011), "Systematic investigation of halogen bonding in protein–ligand interactions", *Angewandte Chemie International Edition,* Vol. 50, No. 1, pp. 314-318.

Hashiramoto, M., Kadowaki, T., Clark, A. E., Muraoka, A., Momomura, K., Sakura, H., Tobe, K., Akanuma, Y., Yazaki, Y., Holman, G. D. and Kasuga, M. (1992), "Site-directed mutagenesis of GLUT1 in helix 7 residues 282 results in perturbation of exofacial ligand binding", *The Journal of Biological Chemistry,* Vol. 267, No. 25, pp. 17502–17507.

Higgins, T.P., Hope, S.J., Effendi, A.J., Dawson, S. and Dancer, B.N. (2005), "Biochemical and molecular characterisation of the 2,3-dichloro-1-propanol dehalogenase and stereospecific haloalkanoic dehalogenases from a versatile *Agrobacterium* sp", *Biodegradation,* Vol. 16, No. 5, pp. 485-492.

Huang, Y., Lemieux, M.J., Song, J., Auer, M. and Wang, D.N. (2003), "Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*", *Science,* Vol. 301, No. 5633, pp. 616-620.

Huyop, F.Z., Cooper, A.R.A. (2003), "A Potential Use of Dehalogenase D (DehD) from *Rhizobium* sp. for Industrial Process", *Jurnal.Teknologi,* Vol. 39, No. (C), pp. 1-8.

Iancu, C.V., Zamoon, J., Woo, S.B., Aleshin, A. and Choe, J.Y. (2013), "Crystal structure of a glucose/H^+ symporter and its mechanism of action", *Proceedings of the National Academy of Sciences U.S.A,* Vol. 110, No. 44, pp. 17862-17867.

Janssen, D.B., Scheper, A., Dijkhuizen, L. and Witholt, B. (1985), "Degradation of halogenated aliphatic compounds by *Xanthobacter autotrophicus* GJ10", *Applied and Environmental Microbiology,* Vol. 49, No. 3, pp. 673-677.

Jing, N., Wahab, R.A., Hamdan, S. and Huyop, F. (2010), "Cloning and DNA Sequence Analysis of the Haloalkanoic Permease Uptake Gene from *Rhizobium* sp. RC 1", *Biotechnology,* Vol. 9, No. 3, pp. 319-325.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015), "The Phyre2 web portal for protein modeling, prediction and analysis", *Nature Protocols,* Vol. 10, No. 6, pp. 845-858.

Leigh, J., Skinner, A. and Cooper, R. (1986), "Isolation and partial characterisation of dehalogenase-deficient mutants of a *Rhizobium* sp", *FEMS microbiology letters,* Vol. 36, No. 2-3, pp. 163-166.

Leigh, J.A., Skinner, A.J. and Cooper, R.A. (1988), "Partial purification, stereospecificity and stoichiometry of three dehalogenases from a *Rhizobium* species", *FEMS microbiology letters,* Vol. 49, No. 3, pp. 353-356.

Liò, P. and Bishop, M. (2008), "Modeling sequence evolution", *Bioinformatics: Data, Sequence Analysis and Evolution,* Vol., No., pp. 255-285.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., *et al.* (2003), "Structure validation by Calpha geometry: phi, psi and Cbeta deviation", *Proteins,* Vol. 50, No. 3, pp. 437-450.

Matter, H., Nazare, M., Gussregen, S., Will, D.W., Schreuder, H., Bauer, A., *et al.* (2009), "Evidence for C-Cl/C-Br...pi interactions as an important contribution to protein-ligand binding affinity", *Angewandte Chemie International Edition in English,* Vol. 48, No. 16, pp. 2911-2916.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., *et al.* (1998), "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function", *Journal of computational chemistry,* Vol. 19, No. 14, pp. 1639-1662.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., *et al.* (2009), "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility", *Journal of computational chemistry,* Vol. 30, No. 16, pp. 2785-2791.

Mueckler, M. and Makepeace, C. (2009), "Model of the Exofacial Substrate-Binding Site and Helical Folding of the Human Glut1 Glucose Transporter Based on Scanning Mutagenesis", *Biochemistry,* Vol. 48, No. 25, pp. 5934-5942.

Mueckler, M., Weng, W. and Kruse, M. (1994), "Glutamine 161 of Glut1 glucose transporter is critical for transport activity and exofacial ligand binding", *The Journal of Biology Chemistry,* Vol. 269, No. 32, pp. 20533-20538.

Musa, M.A. (2017), "Computational Analysis of Putative Haloalkanoic Permease (DehrP) from *Rhizobium* sp. RC1", Masters dissertation, Universiti Teknologi Malaysia, Malaysia..

Newstead, S., Drew, D., Cameron, A.D., Postis, V.L., Xia, X., Fowler, P.W., *et al.* (2011), "Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2", *The EMBO Journal,* Vol. 30, No. 2, pp. 417-426.

Olsen, J., Seiler, P., Wagner, B., Fischer, H., Tschopp, T., Obst-Sander, U., *et al.* (2004), "A fluorine scan of the phenylamidinium needle of tricyclic thrombin inhibitors: effects of fluorine substitution on pKa and binding affinity and evidence for intermolecular C-F[three dots, centered]CN interactions", *Organic and Biomolecular Chemistry,* Vol. 2, No. 9, pp. 1339-1352.

Pence, H.E. and Williams, A. (2010), "ChemSpider: An Online Chemical Information Resource", *Journal of Chemical Education,* Vol. 87, No. 11, pp. 1123-1124.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., *et al.* (2004), "UCSF Chimera—a visualization system for exploratory research and analysis", *Journal of computational chemistry,* Vol. 25, No. 13, pp. 1605-1612.

Plewa, M.J., Simmons, J.E., Richardson, S.D. and Wagner, E.D. (2010), "Mammalian cell cytotoxicity and genotoxicity of the haloacetic acids, a major class of drinking water disinfection by-products", *Environmental and Molecular Mutagenesis,* Vol. 51, No. 8-9, pp. 871-878.

Saier, M.H., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C. and Moreno-Hagelsieb, G. (2016), "The Transporter Classification Database (TCDB): recent advances", *Nucleic Acids Research,* Vol. 44, No. Database issue, pp. D372-D379.

Sanderson, N.M., Qi, D., Steel, A. and Henderson, P.J.F. (1998), "Effect of the D³²N and N³⁰⁰F mutations on the activity of the bacterial sugar transport protein, GalP", *Biochemical Society Transactions,* Vol. 26, No. 3, pp. S306.

Schaffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., *et al.* (2001), "Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements", *Nucleic Acids Research,* Vol. 29, No. 14, pp. 2994-3005.

Shenoy, S.R. and Jayaram, B. (2010), "Proteins: sequence to structure and function--current status", *Current Protein and Peptide Science,* Vol. 11, No. 7, pp. 498-514.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011), "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega", *Molecular systems biology,* Vol. 7, No. 1, pp. 539.

Sirimulla, S., Bailey, J.B., Vegesna, R. and Narayan, M. (2013), "Halogen interactions in protein–ligand complexes: implications of halogen bonding for rational drug design", *Journal of chemical information and modeling,* Vol. 53, No. 11, pp. 2781-2791.

Slater, J.H., Lovatt, D., Weightman, A. J., Senior, E. and Bull, A. T. (1979), "The growth of *Pseudomonas putida* on chlorinated aliphatic acids and its dehalogenase activity", *Journal of General Microbiology,* Vol. 114, No., pp. 125-136.

Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S., *et al.* (2012), "Alternating access mechanism in the POT family of oligopeptide transporters", *The EMBO Journal,* Vol. 31, No. 16, pp. 3411-3421.

Stringfellow, J.M., Cairns, S.S., Cornish, A. and Cooper, R.A. (1997), "Haloalkanoate Dehalogenase II (DehE) of a *Rhizobium* sp. — Molecular Analysis of the Gene and Formation of Carbon Monoxide from Trihaloacetate by the Enzyme", *European Journal of Biochemistry,* Vol. 250, No. 3, pp. 789-793.

Strotmann, U.J., Pentenga, M. and Janssen, D.B. (1990), "Degradation of 2-chloroethanol by wild type and mutants of *Pseudomonas putida* US2", *Archives of Microbiology,* Vol. 154, No. 3, pp. 294-300.

Su, X., Deng, L., Kong, K.F. and Tsang, J.S. (2013), "Enhanced degradation of haloacid by heterologous expression in related *Burkholderia* species", *Biotechnology and Bioengineering,* Vol. 110, No. 10, pp. 2687-2696.

Su, X., Kong, K.-F. and Tsang, J.S.H. (2012), "Transports of acetate and haloacetate in *Burkholderia* species MBA4 are operated by distinct systems", *BMC Microbiology,* Vol. 12, No. 1, pp. 267.

Su, X. and Tsang, J.S. (2013), "Existence of a robust haloacid transport system in a *Burkholderia* species bacterium", *Biochimica et Biophysica Acta,* Vol. 1828, No. 2, pp. 187-192.

Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., *et al.* (2012), "Crystal structure of a bacterial homologue of glucose transporters GLUT1-4", *Nature,* Vol. 490, No. 7420, pp. 361-366.

Trott, O. and Olson, A.J. (2010), "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading", *Journal of computational chemistry,* Vol. 31, No. 2, pp. 455-461.

Tsang, J.S.H., Sallis, P.J., Bull, A.T. and Hardman, D.J. (1988), "A monobromoacetate dehalogenase from *Pseudomonas cepacia* MBA4", *Archives of Microbiology,* Vol. 150, No. 5, pp. 441-446.

Tse, Y.M., Yu, M. and Tsang, J.S. (2009), "Topological analysis of a haloacid permease of a *Burkholderia* sp. bacterium with a PhoA-LacZ reporter", *BMC Microbiology,* Vol. 9, No., pp. 233.

van der Ploeg, J. and Janssen, D.B. (1995), "Sequence analysis of the upstream region of dhlB, the gene encoding haloalkanoic acid dehalogenase of *Xanthobacter autotrophicus* GJ10", *Biodegradation,* Vol. 6, No. 3, pp. 257-263.

Voth Regier, A. and Ho Shing, P. (2007), "The Role of Halogen Bonding in Inhibitor Recognition and Binding by Protein Kinases", *Current Topics in Medicinal Chemistry,* Vol. 7, No. 14, pp. 1336-1348.

Wiederstein, M. and Sippl, M.J. (2007), "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins", *Nucleic Acids Research,* Vol. 35, No. Web Server issue, pp. W407-W410.

Wilcken, R., Liu, X., Zimmermann, M.O., Rutherford, T.J., Fersht, A.R., Joerger, A.C., *et al.* (2012), "Halogen-Enriched Fragment Libraries as Leads for Drug Rescue of Mutant p53", *Journal of the American Chemical Society,* Vol. 134, No. 15, pp. 6810-6818.

Yan, N. (2013), "Structural advances for the major facilitator superfamily (MFS) transporters", *Trends in biochemical sciences,* Vol. 38, No. 3, pp. 151-159.

Yin, Y., He, X., Szewczyk, P. and Chang, G. (2006), "Structure of the multidrug transporter EmrD from *Escherichia coli*", *Journal of Science,* Vol. 312, No. 5774, pp. 741-744.

Yu, M., Faan, Y.W., Chung, W.Y. and Tsang, J.S. (2007), "Isolation and characterization of a novel haloacid permease from *Burkholderia cepacia* MBA4", *Applied Environmental Microbiology,* Vol. 73, No. 15, pp. 4874-4880.

-- 0 --