Elsevier Editorial System(tm) for FEBS Letters
Manuscript Draft

Manuscript Number: FEBSLETTERS-D-14-00272R1

Title: Biochemical and structural characterisation of a haloalkane dehalogenase from a marine Rhodobacteraceae

Article Type: Research Letter

Keywords: Haloalkane dehalogenase; marine Rhodobacteraceae; three-dimensional structure; catalytic activity.

Corresponding Author: Prof Jennifer Littlechild,

Corresponding Author's Institution: University of Exeter

First Author: Halina R Novak, PhD

Order of Authors: Halina R Novak, PhD; Christopher Sayer, PhD; Michail N Isupov, PhD; Dorothee Gotz, PhD; Andrew Mearns Spragg, PhD; Jennifer Littlechild

Manuscript Region of Origin: UNITED KINGDOM

Abstract: A putative haloalkane dehalogenase has been identified in a marine Rhodobacteraceae and subsequently cloned and over-expressed in Escherichia coli. The enzyme has highest activity towards the substrates 1,6-dichlorohexane, 1-bromooctane, 1,3-dibromopropane and 1-bromohexane. The crystal structures of the enzyme in the native and product bound forms reveal a large hydrophobic active site cavity. A deeper substrate binding pocket defines the enzyme preference towards substrates with longer carbon chains. Arg136 at the bottom of the substrate pocket is positioned to bind the distal halogen group of extended di-halogenated substrates.

Suggested Reviewers:

Opposed Reviewers:

Cover Letter

Biosciences,
College of Life and Environmental Sciences,
Henry Wellcome Building for Biocatalysis,
Stocker Road,
Exeter,
EX4 4QD
Telephone +44 (0)1392 263468
Fax +44 (0)1392 263434
Email J.A.Littlechild@ex.ac.uk
Web www.exeter.ac.uk

19th Feb 2014

Ref- FEBSLETTERS-D-14-00272

Dear Dr Ferguson,

We would like to resubmit the manuscript entitled 'Biochemical and structural characterisation of a haloalkane dehalogenase from a marine *Rhodobacteraceae'* Halina R. Novak, Christopher Sayer, Michail Isupov, Dorothee Gotz, Andrew Mearns-Spragg and Jennifer A. Littlechild, for publication in FEBS Letters.

We have changed the format of the references as requested.

We hope the manuscript will now be acceptable for publication in FEBS Letters.

Yours sincerely

Prof. Jenny Littlechild
Director, Exeter Biocatalysis Centre,
Prof Biological Chemistry

***Detailed Response to Reviewers**

Response to Reviewer

We are happy that the reviewer now regards that this manuscript is suitable for publication in FEBS Letters and we have addressed all of the comments previously made. We have now reformatted the references as requested by the editor.

Biochemical and structural characterisation of a haloalkane dehalogenase from a marine *Rhodobacteraceae*

Halina R. Novak[1], Christopher Sayer[1], Michail N. Isupov[1], Dorothee Gotz[2], Andrew Mearns Spragg[2] and Jennifer A. Littlechild[1]

[1] The Henry Wellcome Building for Biocatalysis, Biosciences, College of Life and Environmental Sciences, University of Exeter
Stocker Road
Exeter
EX4 4QD
UK

[2] Aquapharm Biodiscovery Ltd
European Centre for Marine Biotechnology
Dunstaffnage
Oban
Argyll
PA37 1QA

Correspondence
J.A.Littlechild
Tel: 44-1392-2634698
e-mail: J.A.Littlechild@exeter.ac.uk

Abstract

A putative haloalkane dehalogenase has been identified in a marine Rhodobacteraceae and subsequently cloned and over-expressed in *Escherichia coli*. The enzyme has highest activity towards the substrates 1,6-dichlorohexane, 1-bromooctane, 1,3-dibromopropane and 1-bromohexane. The crystal structures of the enzyme in the native and product bound forms reveal a large hydrophobic active site cavity. A deeper substrate binding pocket defines the enzyme preference towards substrates with longer carbon chains. Arg136 at the bottom of the substrate pocket is positioned to bind the distal halogen group of extended di-halogenated substrates.

Highlights

1. Isolation of a marine Rhodobacteraceae sp. from a tube worm.
2. Identification of a putative haloalkane dehalogenase enzyme.
3. Characterisation and crystal structure of the native haloalkane dehalogenase enzyme and product bound complex.
4. Relationship between the substrate specificity and the size and charge distribution within the active site cavity.

Biochemical and structural characterisation of a haloalkane dehalogenase from a marine *Rhodobacteraceae*

Halina R. Novak[1], Christopher Sayer[1], Michail N. Isupov[1], Dorothee Gotz[2], Andrew Mearns Spragg[2] and Jennifer A. Littlechild[1]

[1] The Henry Wellcome Building for Biocatalysis, Biosciences, College of Life and Environmental Sciences, University of Exeter
Stocker Road
Exeter
EX4 4QD
UK

[2] Aquapharm Biodiscovery Ltd
European Centre for Marine Biotechnology
Dunstaffnage
Oban
Argyll
PA37 1QA

Correspondence
J.A.Littlechild
Tel: 44-1392-2634698
e-mail: J.A.Littlechild@exeter.ac.uk

Abstract

A putative haloalkane dehalogenase has been identified in a marine Rhodobacteraceae and subsequently cloned and over-expressed in *Escherichia coli*. The enzyme has highest activity towards the substrates 1,6-dichlorohexane, 1-bromooctane, 1,3-dibromopropane and 1-bromohexane. The crystal structures of the enzyme in the native and product bound forms reveal a large hydrophobic active site cavity. A deeper substrate binding pocket defines the enzyme preference towards substrates with longer carbon chains. Arg136 at the bottom of the substrate pocket is positioned to bind the distal halogen group of extended di-halogenated substrates.

Keywords

Haloalkane dehalogenase; marine Rhodobacteraceae; three-dimensional structure; catalytic activity.

Abbreviations

Haloalkane dehalogenase, (HLD); L-haloacid dehalogenase (L-HAD); *Xanthobacter autotrophicus* haloalkane dehalogenase, (DhlA); *Rhodococcus rhodochrous* haloalkane dehalogenase, (DhaA)*; Mycobacterium tuberculosis* haloalkane dehalogenase, (DmbC); *Sphingomonas paucimobilis* haloalkane dehalogenase, (LinB)*; Bradyrhizobium japonicum* USDA110 haloalkane dehalogenase, (DbjA)*; Plesiocystis pacifica* SIR-1 haloalkane dehalogenase (DppA)*; Stronglocentrotus purouratus* haloalkane dehalogenase (DspA); Rhodobacteraceae family haloalkane dehalogenase, (HanR); 1-hexanol, (1HO), Benzamidine, (BAM), Phenylmethysulfonyl fluoride (PMSF).

Highlights

1. Isolation of a marine Rhodobacteraceae sp. from a tube worm.
2. Identification of a putative haloalkane dehalogenase enzyme.
3. Characterisation and crystal structure of the native haloalkane dehalogenase enzyme and product bound complex.
4. Relationship between the substrate specificity and the size and charge distribution within the active site cavity.

Introduction

Haloalkane dehalogenases (HLDs) (EC 3.8.1.5) catalyse the conversion of halogenated alkanes into their corresponding alcohol products and hydrogen halides. They have potential applications in biocatalysis, biosensors and cell imaging, as well as in the bioremediation of recalcitrant and carcinogenic halogenated by-products from organic synthetic reactions and halogenated pesticides and insecticides [1,2].

The HLDs from the bacteria *Xanthobacter autotrophicus* (DhlA) [3,4], *Rhodococcus rhodochrous* (DhaA) [5], *Sphingomonas paucimobilis* (LinB) [6,7], *Bradyrhizobium japonicum* USDA110 (DbjA) [8,9], *Mycobacterium tuberculosis* (DmbC) [10] and *Plesiocystis pacifica* SIR-1 (DppA) [11] have been biochemically and structurally characterised and are two domain proteins. The core domain belongs to the α/β hydrolase protein fold family with the active site located on the interface of the core and the cap domain, the latter determining substrate specificity. More recently, the first eukaryotic HLD from a purple sea urchin has been biochemically characterised [12].

The HLD enzymatic reaction proceeds via a S_N2 type nucleophilic substitution mechanism [1]. The first step is the nucleophilic attack by the carboxyl group of the conserved aspartate on the carbon atom attached to the halogen of the substrate. This produces an ester intermediate complex where the carboxyl oxygen of the catalytic aspartic acid replaces the halogen. The second step of the reaction is the nucleophilic attack on the carboxyl group of the ester intermediate by a water molecule to release the alcohol product. For the reaction to proceed, the catalytic aspartate oxygen which is not involved in the ester intermediate adduct has to be positioned in the oxyanion hole formed by the main chain nitrogens. The catalytic water is activated by a histidine/carboxylic acid dyad.

Crystal structures of different HLD enzymes have allowed the identification of the catalytically important residues, which have been confirmed by site-directed mutagenesis [4,6,13]. A catalytic pentad has been identified consisting of a catalytic triad of Asp, His and Asp/Glu and a Trp-Trp or Trp-Asn pair, which stabilises the halide leaving group [1]. The Trp-Trp or Trp-Asn pair of residues are essential for

activity and are responsible for binding the substrate in the correct orientation by stabilizing both the transition state and the halide ion released during the carbon-carbon bond break. The division of HLDs into three phylogenetic subfamilies was defined by Chovancova *et al*. in 2007 [14], depending on the amino acid content of the catalytic pentad. Classification of these HLDs into four groups based on substrate specificity revealed there was no clear correlation between substrate specificity and phylogenetic subfamilies [15]. More recently, the characterisation of a HLD from *Agrobacterium tumefaciens* that has a unique halide-stabilizing tyrosine residue (Tyr109) in place of the conserved Trp residue [16] has required an extension of the existing phylogenetic families.

Over the last decade the marine environment has been recognised as a potential source of novel enzymes [17]. The oceans are known to contain abundant organohalogens, which are thought to be produced mainly by marine algae. Some species of polychaete tube worms also produce a wide range of structurally diverse halogenated compounds [18]. These compounds are often cytotoxic and are thought to be part of the organism's defence mechanisms. Microorganisms living in symbiosis with algae and tube worms may have evolved detoxification enzymes such as dehalogenases.

In the search of novel dehalogenase enzymes, a marine *Rhodobacteraceae* species was isolated from the surface of a tube worm. The wild-type bacterium tested positive for L-haloacid dehalogenase (L-HAD) activity. In order to obtain sufficient quantities of this enzyme for biochemical characterisation, the *Rhodobacteraceae* genome was partially sequenced so that the gene could be cloned and the protein over-expressed [19]. In addition to the presence of a novel L-HAD which has been biochemically and structurally characterised [19], a putative HLD was also identified called HanR.

This paper describes the biochemical and structural characterisation of a HLD enzyme from a marine source which has been isolated from the surface of a tube worm.

Material and methods

Gene identification, cloning and overexpression

A *Rhodobacteraceae* family bacterium (Rhb) isolated from a Polychaeta worm was collected from Tralee beach, Argyll, UK. The Rhb genomic DNA was extracted and the genome was sequenced using an Illumina GA2 sequencer. As there was low genome identity to previously sequenced genomes, *De novo* assembly was used to arrange the 72 bp paired-end reads into 1082 contigs using Velvet, version 0.7.63 [20]. A preliminary genome of 3.77 Mbp was assembled into 795 contigs. A putative HLD gene called *HanR* was identified using the NCBI BLAST tool [21] on a Galaxy bioinformatics pipeline [22]. The *HanR* gene was cloned from genomic DNA into the pET-28a plasmid with a N-terminal His-tag. The HanR protein was over-expressed in BL21-CodonPlus (DE3) *Rosetta2 E. coli.*

The cell pellet was re-suspended in buffer A (0.1 M Tris-HCl, 0.1 M NaCl, 0.5 mM EDTA, 1 mM BAM, 1 mM PMSF, pH 8.2) containing 0.01 M imidazole, lysed by sonication and centrifuged to remove cell debris. A nickel affinity chromatography column (GE Healthcare) was equilibrated with buffer A, the cell extract loaded and the unbound protein washed off with buffer A containing 0.01 M imidazole. The bound protein was eluted in buffer A with 0.5 M imidazole.

The fractions corresponding to the protein were concentrated using a 10 kDa membrane (Vivaspin 20, Vivascience) at 3,000 x *g*, at 4ºC until the final volume reached 1 ml. The concentrated protein sample was further purified on a 120 ml Superdex 75 GF chromatography column which was eluted over 1 column volume of buffer A.

Enzyme activity

The HLD activity was measured using a modified colorimetric assay, based on the method of Holloway *et al*. [23]. The assay solution had a final concentration of 1 mM HEPES, 1 mM EDTA, 20 mM sodium sulfate, 10 mM substrate, 20 µg/ml phenol red, pH 8.2. To initiate the assay, 180 µl of dehalogenase assay solution was mixed with 20 µl of purified protein (2 mg/ml). The reaction was followed by measuring a decrease in absorbance at 540 nm over 1 h. To produce a standard curve, the dehalogenase assay solution was mixed with HCl to a final concentration between 0 to 2 mM in a total volume of 200 µl.

Crystallization

The purified HanR enzyme was concentrated until a final concentration of 10 mg/ml was reached. The concentrated protein was subjected to microbatch crystallization screens at 18°C using an Oryx Robot (Douglas Instruments). The best crystals were grown from a mixture of equal volumes of protein and precipitant solution containing 0.15 M $MgCl_2$, 0.1 M Tris-HCl and 15% PEG 4000, pH 8.0. The native crystals were frozen using a cryoprotectant containing 0.1 M Tris-HCl pH 8.0, 0.1 M NaCl, 25% PEG 4000 and 30% PEG 400. The substrate soak was conducted by soaking the HanR crystal for 1 min in 0.1 M Tris-HCl, 20% PEG 3350, 30% PEG 400, 25 mM 1-bromohexane, pH 4.0.

Data collection and structure determination

Diffraction data were collected at the Diamond Light Source synchrotron, UK. Data were processed and scaled with the programs XDS [24] and Aimless [25] using the Xia2 [26] pipeline. Further data analysis and refinement were performed using the CCP4 suite of programs [27].

The HanR native structure was solved by molecular replacement (MR) with the program MOLREP [28] using the LinB model (PDB: 1CV2; [6]). The MR solution was submitted to an automated refinement procedure using ARP/wARP version 7.0.1 [29]. The resulting model was manually rebuilt in COOT [30] and refined with REFMAC5 [31].

The structure of the product bound 1-hexanol (1HO) complex was solved by MR using the refined structure of the native HanR. The dictionary definition for the ligand was built using JLIGAND [32]. The quality of the structures was checked using the program PROCHECK [33]. The program PyMOL [34] was used to produce Figures 2a, 3-4.

Results and Discussion

Biochemical characterisation

The specific activity of the HanR was measured with a variety of substrates (Fig. 1). The HanR is active towards both chloroalkanes and bromoalkanes with preference towards longer chain linear substrates.

Activity of HanR towards brominated compounds appears to be higher than towards their chlorinated equivalents, which is not surprising, as the energy of C-Br bond is lower than that of C-Cl. This is also in line with the abundance of brominated organic compounds in the marine environment. The DppA HLD from the marine bacteria *Plesiocystis pacifica* [11] is not active towards any chlorinated compounds. Substrate variations between the HLD enzymes are thought to arise from differences in the size, shape and hydrophobicity of the cap domain [14]. The differences in substrate specificity are difficult to infer from protein sequence only. Therefore, in order to elucidate the structural features responsible for the preference of HanR towards long chain substrates, the crystal structure was determined.

HanR Structure

Overall structure

The native and product complex structures were refined to resolutions of 1.7 Å and 1.6 Å respectively. The resulting R_{free} values for the native and complex structures were 24.6% and 18.4% (Table 1). The G-factors calculated for each model confirmed these structures have acceptable stereochemical properties [33]. Proline residues 37, 216 and 244 are in the *cis* conformation in both the native and complex structures.

The approximate dimensions of the HanR monomer are 40 Å x 40 Å x 50 Å. The HanR adopts the typical α/β hydrolase fold consisting of a core and cap domain (Fig. 2a). The core domain is conserved in all of the α/β hydrolase fold superfamily proteins and is composed of 8 β-strands surrounded by 6 α-helices, two on one side of the β-sheet and 4 on the other. The β-sheet is of mixed type with direction + - + + + + + + and topology +1, +2, -1x, +2x, +1x, +1x, +1x [35]. The cap domain (residues 132-233) is located between strands 6 and 7 of the core domain β-sheet. It is built up of 6 α-helices (α3-α8) and is the most variable region of the HLD enzymes. The overall structure of the HanR is very similar to LinB, DmbC, DhaA and DbjA HLD enzymes with Cα RMSDs of less than 1Å.

The asymmetric unit of the native structure contains two enzyme monomers which appear to form a dimer related by a molecular dyad, however only a small portion of the surface area (8.7%) is buried upon its formation. This dimer differs from the

crystallographic dimer in the product complex structure by a relative rotation of several degrees between the monomers. In the dimer observed in the crystal structure the adjacent subunit hinders the access to the enzyme active site. The estimated size of the protein in solution corresponds to a monomer, as determined by size exclusion chromatography. As most related HLDs are reported to be monomers, the dimer is most likely to be an artefact of crystallisation.

Catalytic pentad

The active sites of Han enzymes are characterised by a catalytic pentad [14]. This consists of a catalytic triad and a stabilizing pair of residues. Currently four different pentads [14,16] have been identified from the various HLD enzymes characterised to date. HanR and LinB (64% amino acid sequence identical) both belong to the HLD-II subfamily. The catalytic triad (Fig. 2b) consists of Asp106 (located between β-strand 5 and α-helix 2), His271 (located between β-strand 8 and 3_{10} helix 9) and Glu130 (located within β-strand 6). The side chains of Trp107 and Asn36 bind the substrate halogen, while their main chain nitrogen atoms form the oxyanion hole.

As the crystallisation media contained chloride ions, the halide binding site of HanR contains a bound chloride ion at full occupancy. The catalytic pentad residues in HanR and the position of the chloride ion match those of LinB [36] very closely (Fig. 3).

Product binding

To map the HanR substrate pocket, crystals were soaked with 1-bromohexane. The electron density of the HanR 1-bromohexane substrate soaked crystal revealed a fully dehalogenated 1HO product in the active site. This was modelled in two conformations, each having a partial occupancy of 0.5 (Fig. 4). The 1HO molecule sits in a hydrophobic pocket formed by residues Leu247, Val210, Phe150 and Phe142.

There is little structural difference between the native and product complex structures (RMSD of 0.35 Å). The presence of the 1HO product shifts Arg145 2.7 Å away from the hydrophobic active site resulting in a larger substrate binding pocket. This observation is confirmed by the well-defined electron density and low B-factors of the Arg145 residue in both native and complex structures.

The product 1-butanol (1BO) in the LinB complex structure binds in a similar conformation to the 1HO in the active site of the HanR product complex but in HanR it is located deeper into the active site. The disorder of the 1HO product in the HanR active site is thought to be due to the wider substrate channel which can also bind cyclic haloalkanes.

Structural features responsible for substrate specificity

For HLD catalysis, the enzyme halide site should bind the substrate halogen, while the carbon attached to the halogen and the adjacent carbon should have a fixed position in the active site. To achieve this, HLDs need to bind the substrate tightly in an extended conformation. DhlA, a HLD with a small active site displays high, activity towards 1,2-dichloroethane while the enzymes with larger active site pockets (HanR, LinB) are inert towards this substrate.

The HLD active site is built up of an entrance tunnel and a hydrophobic pocket for substrate binding [37]. The residues lining the substrate pocket mainly belong to the cap domain and define substrate specificity. Minor differences in the substrate pocket can dramatically affect the substrate specificity of very similar HLDs [15].

We observe a difference in substrate specificity between the highly similar HLD-II family enzymes, HanR and LinB [15]. LinB displays a significant drop of activity as the carbon chain becomes longer from the substrates 1-bromobutane to 1-bromoxehane. This would infer a lower activity towards the longer 1-bromooctane substrate, which HanR prefers to 1-bromoxehane. There are significant differences between the entrance tunnels of these two enzymes (Fig. 3), where in HanR, Ser176 replaces Leu177 of LinB, and on the opposite side of the entrance tunnel Arg145-Ala146 replaces the Gln146-Asp147 pair of residues in LinB. Ile211 which is located close to the halide binding site in LinB is replaced by the smaller Val210 in HanR. Additionally, some conserved residues between the two enzymes such as Phe150, Leu247, and Phe142 in HanR have different conformations or are shifted from the positions of their equivalents in LinB, thereby increasing the volume of the active site cavity in HanR. Although these changes may contribute to the different substrate specificity between these two enzymes, we propose that the differences in the substrate binding pocket are more likely to account for the substrate variations observed.

The importance of the extended or flexible substrate channel for activity towards long chain substrates was demonstrated for the DhlA enzyme [38]. Specific activity of the wild type DhlA towards 1,6 dibromohexane was only 0.073 $\mu mol \cdot min^{-1} \cdot mg^{-1}$. This was increased three-fold in the F164A mutant DhlA enzyme by freeing the space occupied by the phenyl ring. Interestingly, the D170A surface mutation of DhlA, which disrupted the salt bridge holding a loop in place, showed a two-fold higher specific activity towards 1,6 dibromohexane, probably due to the increased flexibility of the substrate channel.

The most significant differences between HanR and LinB are at the bottom of the substrate pocket where Ile138, Met253 and Leu248 of LinB form a hydrophobic cluster, limiting the size of the bound substrate. Substitutions of Ile138 (LinB) by Arg136 (HanR) and Met253 by the smaller Val252, in addition to a different conformation of the conserved Leu247, significantly expand the space in the substrate channel of HanR to accommodate longer chain substrates in comparison to LinB. Positively charged Arg136 in HanR at the bottom of the active site is not observed in other HLDs. This will facilitate the binding of the distal halogen group, as demonstrated by the high activity of HanR towards 1,6 dichlorohexane (0.44 $\mu mol \cdot s^{-1} \cdot mg^{-1}$).

Summary

Substrate specificity experiments showed that the activity profile of HanR towards a range of substrates is different from other HLD enzymes. The crystal structure of this novel marine HLD has revealed a wide active site cavity and a deep substrate binding channel with a positively charged residue at the bottom. This is the second novel dehalogenase [19] enzyme from this Rhb family bacterium that we have characterised that shows interesting biochemical and structural properties. These results demonstrate the potential of the increased diversity and evolutionary divergence of organisms in the marine environment as a source for the discovery of novel enzymes.

Accession number

The *HanR* gene and its protein sequence have been deposited in Genbank: KF032932. Structure factors and coordinates of the final models have been deposited in the PDB: 4BRZ and 4C6H.

Acknowledgements

We thank Diamond Light Source for access to beamlines I03 and I24 (proposal numbers mx2032 and mx6851) and the beamline scientists. We thank Konrad Paszkiewicz, University of Exeter, for help with the genome sequencing of the *Rhodobacteraceae* organism. The Biotechnology and Biological Science Research Council, UK and Aquapharm Biodiscovery, Oban are acknowledged for PhD studentship funding (HRN). CS acknowledges a PhD GTA bursary from the University of Exeter. MI is grateful to the University of Exeter and a BBSRC funded ERA-IB grant BB/L002035/1. Funding in JALs laboratory has been supported by the Wellcome Trust, EU, BBSRC and EPSRC.

References

[1] Janssen, D. B. (2004). Evolving haloalkane dehalogenases. Curr. Opin. Chem. Biol. **8**, 150-159.

[2] Koudelakova, T., Bidmanova, S., Dvorak, P., Pavelka, A., Chaloupkova, R., Prokop, Z. and Damborsky, J. (2013). Haloalkane dehalogenases: Biotechnology applications. Biotechnol. J. **8**, 32-45.

[3] Van der Ploeg, J., Van Hall, G. and Janssen, D.B. (1991). Characterization of the haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 and sequencing of the dhlB gene. J. Bacteriol**. 173**, 7925-7933.

[4] Verschueren, K.H.G., Franken, S.M., Rozeboom H.J., Kalk, K.H. and Dijkstra, B.W. (1993). Refined X-ray Structures of Haloalkane Dehalogenase at pH 6 2 and pH 8 2 and Implications for the Reaction Mechanism. J. Mol. Biol. **232**, 856-872.

[5] Newman, J., Peat, T.S., Richard, R., Kan, L., Swanson, P.E., Affholter, J.A., Holmes, I.H., Schindler, J.F., Unkefer, C.J. and Terwillinger, T.C. (1999). Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. Biochemistry, **38**, 16105-16114.

[6] Marek, J., Vévodová, J., Smatanová, I.K., Svensson, L.A., Newman, J., Takagi, M. and Damborsky, J. (2000). Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. Biochemistry, **39**, 14082-14086.

[7] Prokop, Z., M. Monincová, Chaloupková, R., Klvana. M., Nagata, Y., Janssen, D.B. and Damborsky, J. (2003). Catalytic mechanism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. J. Biol. Chem. **278**, 45094-45100.

[8] Sato, Y., Natsume, R., Tsuda, M., Damborsky, J., Nagata, Y. and Senda, T. (2007). Crystallization and preliminary crystallographic analysis of a haloalkane dehalogenase, DbjA, from *Bradyrhizobium japonicum* USDA110. Acta Crystallogr. F **63**, 294-296.

[9] Sfetsas, C.C, Milios, L., Skopelitou, K., Venieraki, A., Todou, R., Flemetakis, E., Katinakis, P. and Labrou, N.E. (2009). Characterization of 1,2-dibromoethane-degrading haloalkane dehalogenase from *Bradyrhizobium japonicum* USDA110. Enzyme Microb. Tech. **45**, 397-404.

[10] Mazumdar, P.A., Hulecki, J.C., Cherney, M.M., Garen, C.R. and James, M.N. (2008). X-ray crystal structure of *Mycobacterium tuberculosis* haloalkane dehalogenase Rv2579. BBA-Proteins Proteome. **1784**, 351-362.

[11] Hesseler, M., Bogdanovic, X., Hidalgo, A., Berenguer, J., Palm, G.J., Hinrichs, W. and Bomscheuer, U.T. (2011). Cloning, functional expression, biochemical

characterization and structural analysis of a haloalkane dehalogenase from *Plesiocystis pacifica* SIR-1. Appl. Microbiol. Biotechnol. **91**, 1049-1060.

[12] Fortova, A., Sebestova, E., Stepankova, V., Koudelakova, T., Palkova, L., Damborsky J. and Chaloupkova, R. (2013). DspA from *Stronglocentrotus purouratus:* The first biochemically characterized haloalkane dehalogenase of non-microbial origin. Biochimie, **95**, 2091-2096.

[13] Pries, F., Kingma, J., Pentenga, M., Van Pouderoyen, G., Jeronimum-Stratingh C.M., Bruins, A.P. and Janssen, D.B. (1994). Site-directed mutagenesis and oxygen isotope incorporation studies of the nucleophilic aspartate of haloalkane dehalogenase. Biochemistry, **33**, 1242-1247.

[14] Chovancová, E., Kosinski, J., Bujnicki, J.M. and Damborsky, J. (2007). Phylogenetic analysis of haloalkane dehalogenases. Proteins, **67**, 305-316.

[15] Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J. and Damborsky, J. (2011). Substrate specificity of haloalkane dehalogenases. Biochem J. **435**, 345-354.

[16] Hasan, K., Gora, A., Brezovsky, J., Chaloupkova, R., Moskalikova, H., Fortova A., Nagata, Y., Damborsky, J. and Prokop, Z. (2013). The effect of a unique halide-stabilizing residue on the catalytic properties of a halolalkane dehalogenase DatA from *Agrobacterium tumefaciens* C58. FEBS J. **280**, 3149-3159.

[17] Trincone, A. (2011). Marine biocatalysts: enzymatic features and applications. Mar. Drugs, **9**, 478-499.

[18] Fielman, K.T., Woodin, S.A. and Lincoln, D.E. (2001). Polychaete indicator species as a source of natural halogenated organic compounds in marine sediments. Environ. Toxicol. Chem. **20**, 738-747.

[19] Novak, H.R., Sayer, C., Isupov, M.N., Paszkiewicz, K., Gotz, D., Spragg, A.M. and Littlechild, J.A. (2013). Marine *Rhodobacteraceae* L-haloacid dehalogenase contains a novel His/Glu dyad that could activate the catalytic water. FEBS J. **280**, 1664-1680.

[20] Zerbino, D.R. and Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research **18**, 821-829.

[21] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**, 403-410.

[22] Goecks, J., Nekrutenko, A., Taylor, J. Galaxy Team (2010). Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life science. Genome Biol. **11**, R86.

[23] Holloway, P., Trevors, J.T. and Lee, H. (1998). A colorimetric assay for detecting haloalkane dehalogenase activity. J. Microbiol. Meth. **32**, 31-36.

[24] Kabsch W (2010). XDS. Acta Crystallogr. D **66**, 125-132.

[25] Evans, P.R. and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. D **69**, 1204-1214.

[26] Winter, G, Lobley, C.M.C. and Prince, S.M. (2013). Decision making in *xia*2, Acta Crystallogr. D **69**, 1260–1273.

[27] Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. and Wilson, K.S. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D **67**, 235-242.

[28] Vagin, A. and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallogr. D **66**, 22-25.

[29] Langer, G.G, Hazledine, S., Wiegels, T., Carolan, C. and Lamzin, V.S. (2013). Visual automated macromolecular model building. Acta Crystallogr. D **69**, 635-641**.**

[30] Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D **66**, 486-501.

[31] Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D **67**, 355-367.

[32] Lebedev, A.A., Young, P., Isupov, M.N., Moroz, O.V., Vagin, A.A. and Murshudov, G.N. (2012). JLigand: a graphical tool for the CCP4 template-restraint library. Acta Crystallogr. D **68**, 431-440.

[33] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283-291.

[34] DeLano, W.L. (2002). The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA.

[35] Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. Adv. Prot. Chem. **34,** 167-339.

[36] Oakley, A.J., Prokop, Z., Bohac, M., Kmunicek, J., Jedlicka, T., Monincova, M., Kuta-Smatanova, I., Nagata, Y., Damborsky, J. and Wilce, M.C.J. (2002). Exploring the structure and activity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26: evidence for product-and water-mediated inhibition. Biochemistry, **41**, 4847-4855.

[37] Chaloupková, R., Sykorová, J., Prokop, Z., Jensenská, A., Monincová M., Pavlová, M., Tsuda, M., Nagata, N. and Damborsky. J. (2003). Modification of activity and specificity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26 by engineering of its entrance tunnel. J. Biol. Chem. **278**, 52622-52628.

[38] Holloway, P., Knoke, K.L., Trevors, J.T. and Lee, H. (1998). Alterations of the substrate range of the haloalkane dehalogenase by site-directed mutagenesis. Biotechnol. Bioeng. **59**, 520-523.

[39] Vaguine, A.A., Richelle, J. and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr. D **55**, 191-205.

[40] Gouet, P., Robert, X. and Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucl. Acids Res. **31**, 3320-3323.

Table 1

Crystal	Native HanR	Complex HanR
Beamline (Diamond)	I24	I03
Resolution (Å)	44.4-1.67 (1.71-1.67)[a]	47.84-1.61 (1.66-1.61) [a]
Wavelength (Å)	0.9778	0.97625
Space group	$P2_1$	$P2_12_12$
Cell dimensions (Å),	a, b, c= 55.0 Å, 75.5 Å, 64.9 Å; α, γ=90º β=92.9º	a, b, c= 64.5 Å, 71.41 Å, 59.75 Å; α, β, γ=90º
No. of protomers in A.U.	2	1
Solvent content (%) V_M (Å^3 Da^{-1})	40 (2.1) [a]	41.1(2.1) [a]
Unique reflections	61186	36192
Redundancy	3.4 (3.3) [a]	12.8 (12.6) [a]
Completeness	99.6 (99.6) [a]	99.9 (100) [a]
<(I)/ σ (I)>	14.2 (2.0) [a]	26.0 (3.6) [a]
R_{sym} (%)	4.3 (51.0) [a]	6.1 (74.7) [a]
Overall R-factor (%)	20.4	15.3
R_{free} (5 % total data) %	24.6	18.7
Residues modelled	A (T3-L292) B (T3-L292)	A (T2-L292)
No. of waters modelled	466	352
No of chlorines modelled	2	1
No of ligands modelled	-	1
RMSD bond length (Å)	0.007 (0.019) [b]	0.013 (0.022) [b]
RMSD bond angle (º)	1.24 (1.95) [b]	1.35 (1.95) [b]
Wilson B factor (Å^2)	30.0	27.4
Average B factor		
Protein (Å^2)	25.8	22.8
Water (Å^2)	34.2	38.8
Chlorine (Å^2)	33.2	16.23
Ligand (Å^2)	-	33.71
REFMAC RMS error estimate (Å^2)	0.126	0.09
Ramachandran analysis (% of residues)		
Most favoured	88.9	91.2
Additionally allowed	10.5	8.4
Generously allowed	0.6	0.4
Disallowed	0	0
G-factor	0.1	0

Table 1. The data processing and refinement statistics for the HanR structures.

[a] Values for the outer resolution shell are given in brackets.

[b] Target values are given in parentheses. $R_{sym}=\sum_h\sum_J|<I_h>-I_J(h)|/\sum_h\sum_J I(h)$, where I(h) is the intensity of the reflections h, $\sum_h$ is the sum over all the reflections and $\sum_J$ is the sum over J measurements of the reflections. $R_{cryst}=\sum||Fo|-|Fc||/\sum|Fo|$. Wilson B-factor was estimated by SFCHECK [39]. The Ramachandran plot analysis and G-factor calculation were performed by PROCHECK [33].

Figure Legends

Figure 1. The specific activity of HanR towards halogenated compounds with varying structures, carbon chain length and different positions and degree of chlorination or bromination displayed in $\mu mol \cdot s^{-1} \cdot mg^{-1}$.

Figure 2. (a) Folding of the HanR monomer presented as a cartoon diagram. The catalytic Asp106 residue is displayed as a grey CPK model and α-helices (H), β-strands (S) and loops are coloured in turquoise, magenta and pink respectively. (b) The sequence alignment of HanR and LinB. The secondary structure of HanR is shown above the sequence with a spiral and an arrow representing α-helices and β-strands respectively using ESPript [40]. The numbering of α-helices and β-strands is consistent with Figure 2a. Conserved residues are shown in red boxes and residues with similar properties are in blue boxes. The HLD catalytic pentad is highlighted with stars in cyan beneath the sequence alignment.

Figure 3. A superimposition of the active site residues in the HanR (pink) and the LinB (grey), with the chlorine ion in the active site cavity shown as a green sphere. The significantly different residues between the active site of these HLDs are highlighted in magenta in the HanR structure.

Figure 4. A stereo diagram showing the active site cavity of the HanR with the $2F_o$-F_c electron density map contoured at 1 σ. The 1HO product modelled in two conformations is displayed as black sticks. The chloride ion is shown as a green sphere. Active site residues are shown as sticks.

Figure1

Click here to download high resolution image

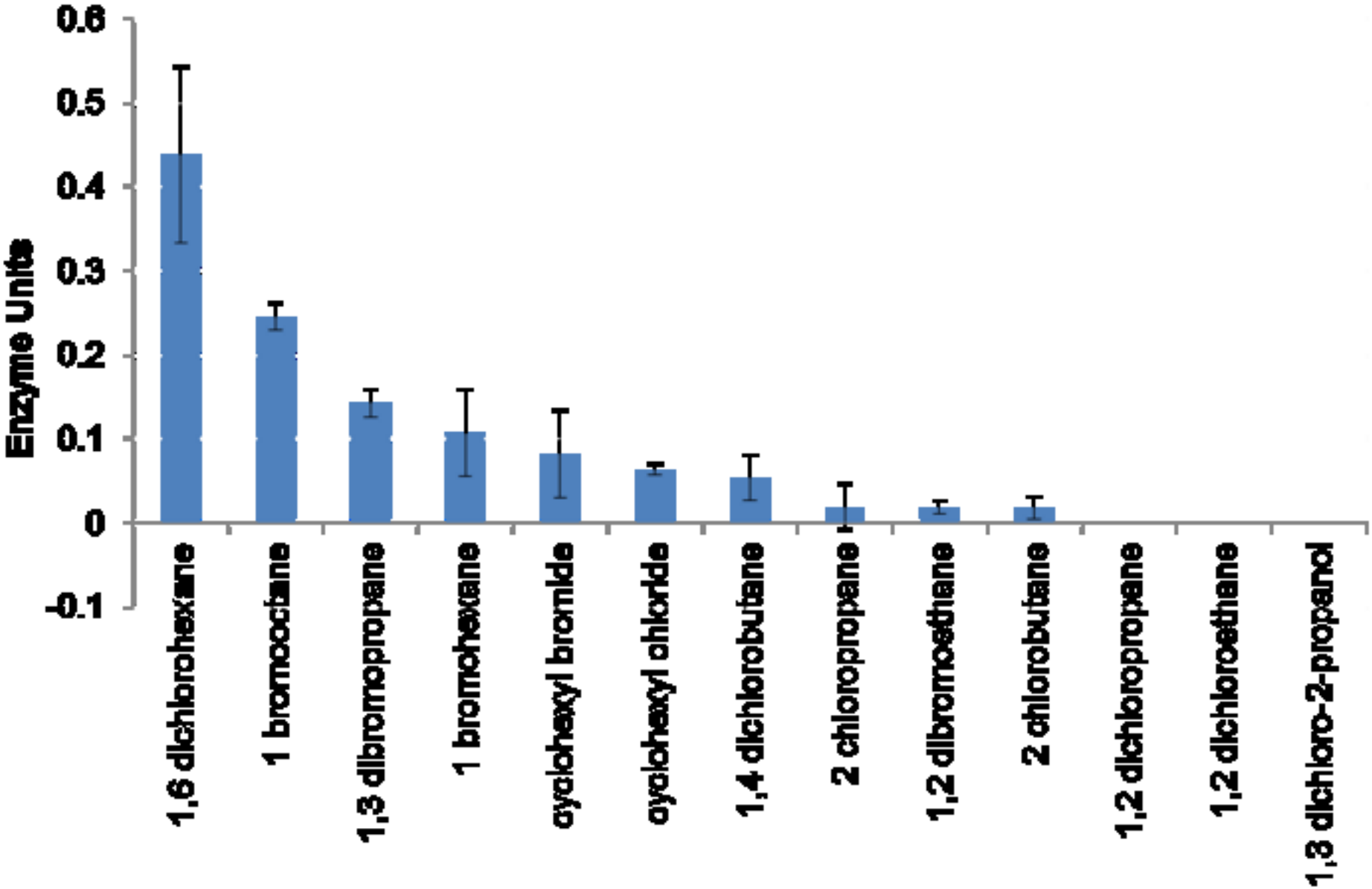

Figure2a

Click here to download high resolution image

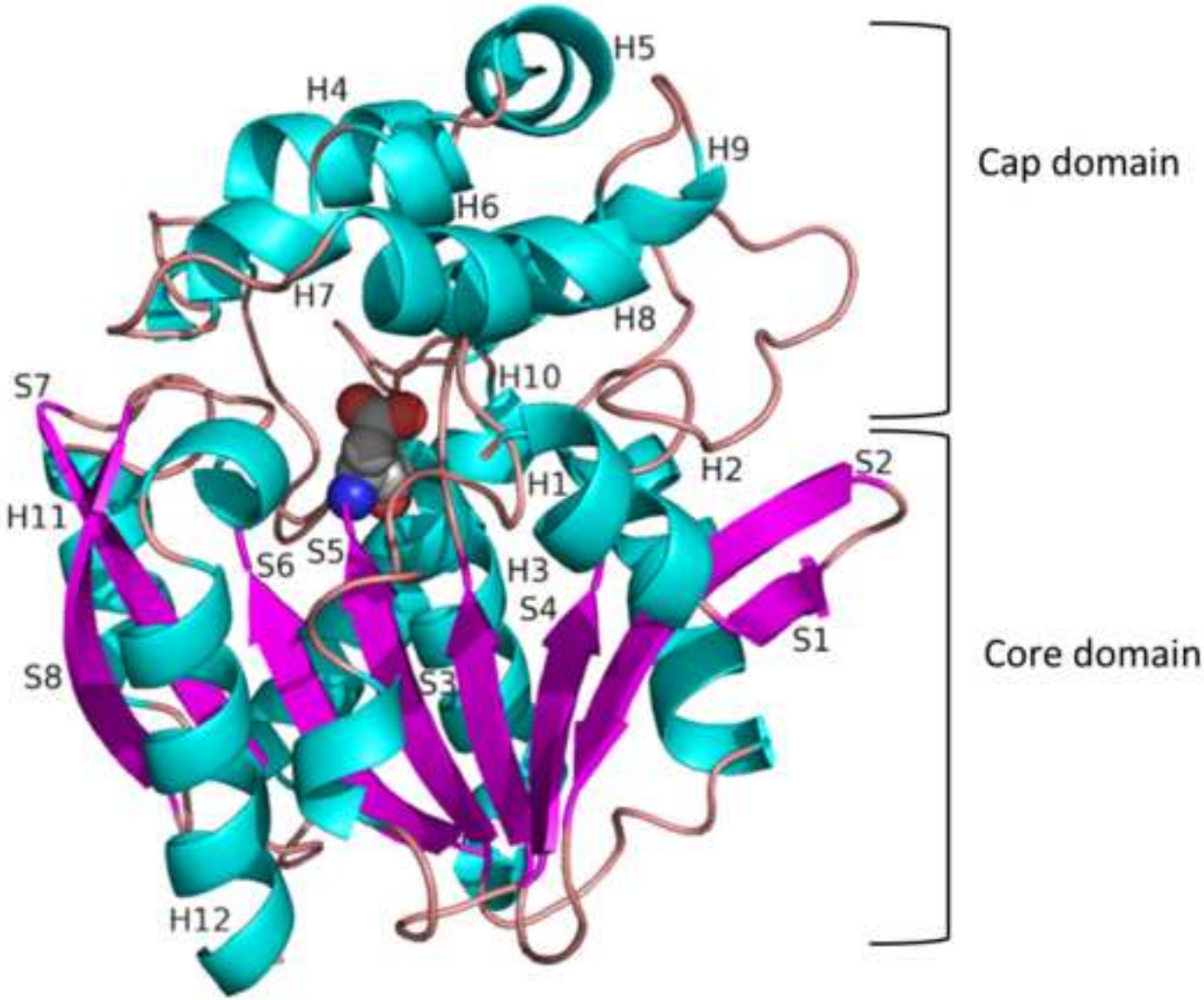

Figure2b

Click here to download high resolution image

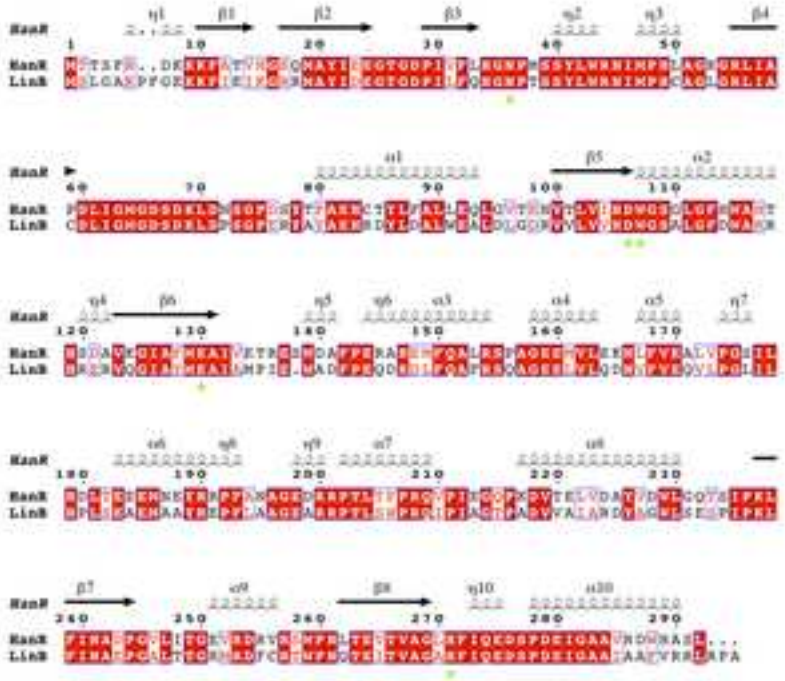

Figure3

Click here to download high resolution image

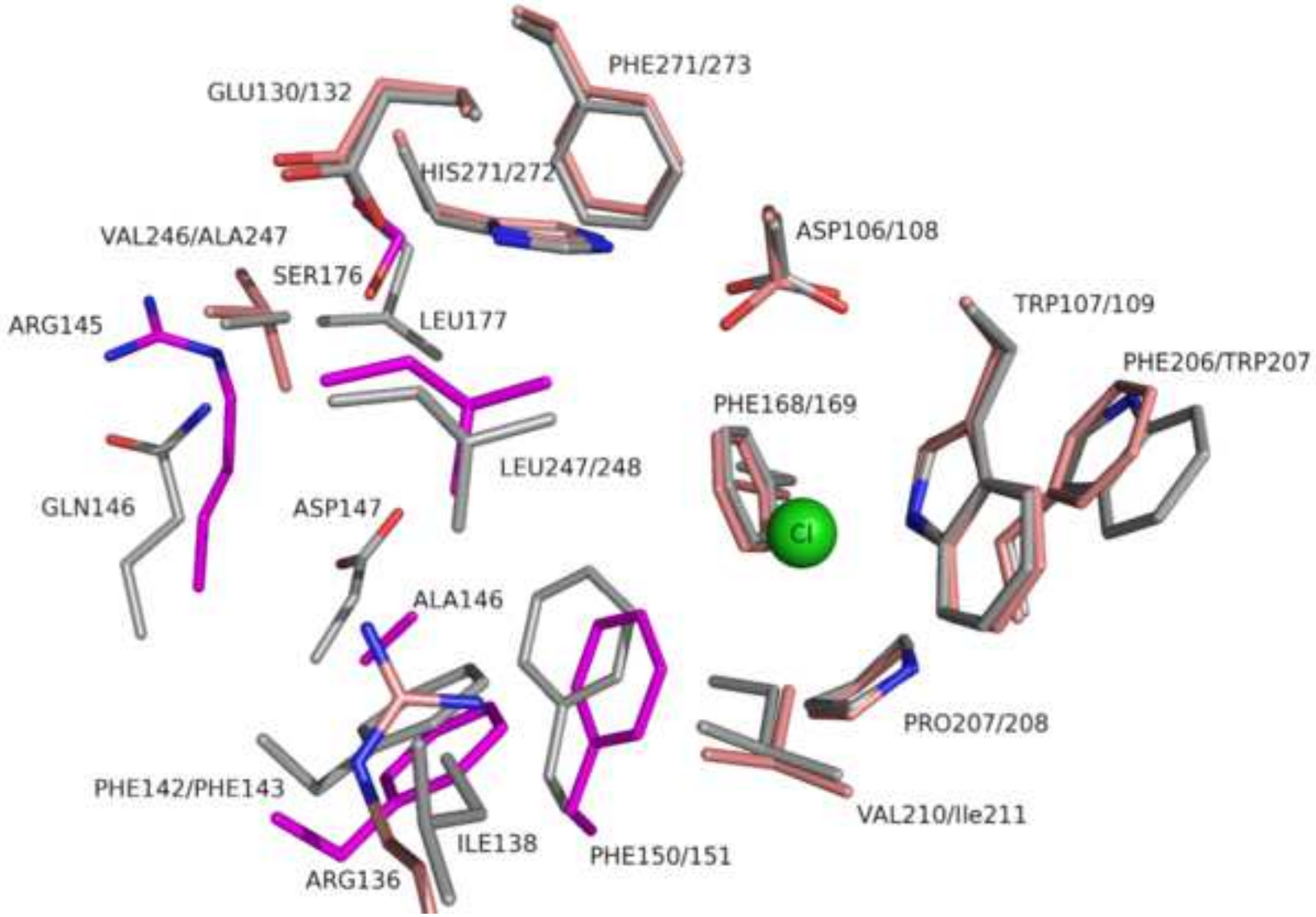

Figure4

Click here to download high resolution image