Isolation and in silico characterization of novel esterase gene with β-lactamase fold isolated from metagenome of north western Himalayas

Avneet Kour Sudan · Jyoti Vakhlu

Received: 6 August 2014 / Accepted: 17 September 2014 / Published online: 19 October 2014 © The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract An esterase-producing clone Aph2 was isolated from the Apharwat soil metagenomic library, a mountain peak in NW Himalayas. ORF 2 (Est Ac) of clone Aph2 corresponds to 271 aa protein and showed 26 % sequence similarity to carboxylesterase gene of Synechococcus sp. JA-2-3B. Est Ac contains nucleophilic Ser in S\(^{68}\)-X-X-K\(^{71}\) motif of β-lactamases with Tyr Y\(^{103}\). The conserved sequences are common with family VIII carboxylesterase and class C β-lactamase sequences. Phylogenetic analysis revealed that Est Ac sequence is closely related to esterase than to β-lactamases. In silico 3D protein structure of Est Ac was generated using MODELLER software (9.10 version). Model was generated on the basis of carboxylesterase template (PDB:1C18) of Est B (Burkholderia gladioli) and the stereochemical parameters of the model generated were satisfactory. Docking with diisopropyl-fluorophosphate confirmed catalytic activity of Ser\(^{68}\) present in S-X-X-K motif.

Keywords β-Lactamases · Docking · DFP · Esterases · Homology modelling · Multiple displacement amplification

Introduction

Genes with wide industrial applications have been isolated from animals, plants and microbes with the largest source being the microbes (Wiseman 1995). Microbial genes are isolated and engineered by cultivation-dependent or cultivation-independent approaches. Cultivation-independent, metagenomic approach is favourable for the isolation of novel genes, due to absence of bias which is introduced by selective cultivation, cloning and PCR (Fuhrman 2012; Daniel 2005). The list of novel genes isolated using metagenomics is long and to name a few antibiotics, oxidoreductases, nitrile converting enzymes, glycerol hydratases, proteinases, lipases, esterases, amylases and antibiotic resistance genes (Gillespie et al. 2002; Knietsch et al. 2003; Liebeton and Eck 2004; Schmeisser et al. 2007; Morrooshi et al. 2011; Berlemont et al. 2013; Vidya et al. 2011; Mullany 2014). The global enzyme market is expected to reach USD 7,652 million by 2020, growing at a CAGR of 8.3 % from 2014 to 2020 (http://www.grandviewresearch.com). Lipolytic enzymes being the most important class of enzymes, isolated and exploited for various industrial purposes, covering ~5 % of the global enzyme market (Vakhlu and Kour 2006). Lipolytic enzymes are categorised into eight families (I–VIII) on the basis of catalytic triad with residues Ser-His-Asp (Glu) in which the Ser-O\(_\gamma\) acts as a nucleophile present in pentapeptide G-X-S-X-G or GDSL motif (Arpigny and Jaeger 1999). Carboxylesterases (EC 3.1.1) are a subtype of lipolytic enzymes which hydrolyse carboxyl ester molecules to release alcohol and carboxylic acid. It is family VIII of carboxylesterases that shows major similarity with class C β-lactamases, peptidases and penicillin-binding protein and their primary sequence contains highly conserved S-X-X-K motif (where X is any amino acid) located at N-terminus of primary structure rather than the conventional G-X-S-X-G (Mokoena et al. 2013). The Ser residue which is known to act as catalytic nucleophile for family VIII carboxylesterase activity is present in the S-X-X-K motif and most of the members of this family are
reported to lack activity against standard β-lactam substrates (Wagner et al. 2002). Prokaryotic genes are reported to show convergent evolution, wherein homologous catalytic motif having similar functions which may be present in different classes of hydrolases. Both classes, similarly family VIII carboxylesterases and β-lactamases possess catalytic residues in S-X-X-K motif of β-lactamase which is as a result of convergent evolution (Gherardini et al. 2007). In the present study, the isolation of a novel esterase gene from soil metagenome of Apharwat mountain peak in NW Himalayas having catalytic motif of β-lactamases is being reported. The niche is anthropogenically isolated and temperature variation is from 18°C in summers to -20°C in winters. The gene pool of the microbes in the Himalayas could be rich source for the novel enzymes and its potential has not been explored so far.

Materials and methods

Collection of soil sample

The soil sample was collected from Apharwat mountain (4267.2 m) with latitude 34.209° and longitude 74.368° during May 2007. The soil from a 5-cm-deep hole was collected in aseptic plastic bags that were then placed in containers to ensure that the microbial load was not disturbed and retained its natural form. Hands, trowels, and ice axes were treated with 70% ethanol immediately before use. The samples were transported to the laboratory in dry ice and finally stored at -20°C (Foght et al. 2004).

Construction of metagenomic library

In the present study, DNA isolated using Wechter protocol (2002) contained impurities that hampered the manipulation of DNA. To reduce the interfering impurities and to increase the quantity of DNA, whole metagenome was diluted 50-fold and then multiple displacement amplification (MDA) was performed using phi (φ)29 DNA polymerase as per the method standardised previously (Sudan and Vakhlu 2012). Subsequent cloning and transformation generated 10,000 clones (Sudan and Vakhlu 2012).

Sequencing and sequence analysis using bioinformatic tools and prediction of signal peptide

The sequencing was performed following Sanger’s dideoxy termination method at CIF South campus, Delhi University, India. Bioinformatic tools ORF Finder, BLASTP (http://www.ncbi.nlm.nih.gov.Blast) were used for functional ORF detection and sequence alignment, respectively. Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment was done using CLUSTALX (Larkin et al. 2007). The sequences selected for multiple alignment were retrieved from NCBI. Class C β-lactamase included that from Candidatus kori (Yp_589716), Caulobacter sp.K31 (Yp_001682441) and family VIII carboxylesterase of that of Pseudomonas sp. (gi AAA25813.1), esterase A Streptomyces anulatus (gi CAA78842.1) and Est B Burkholderia gladioli (PDB:1CI8) sequences. Evolutionary relationship between Est Ac, family VIII carboxylesterase and class C β-lactamase was inferred using neighbour joining method conducted with CLC software. Functional ORF was submitted to GenBank under accession number JX068525.1.

Detection of signal peptide

Potential signal peptide in present case was retrieved using SIGNALP 3.0 server (Bendtsen et al. 2004).

Comparative modelling of putative protein

Comparative modelling for protein 3D model generation was carried out by MODELLER 9.10 version software (Eswar et al. 2006). Est B was selected as template. Sequence of Est B was retrieved from PDB database. The modelling was based on alignment between sequences of Est B and Est Ac. The generated model was subjected to validation.

Validation of model

Validation of the model generated was carried out using SAVs server including PROCHECK (Laskowski et al. 1993) and ProSA (Wiederstein and Sippl 2007). These softwares were used to calculate the stereochemical evaluations.

Ligand preparation and molecular docking

For in silico docking first the catalytic pocket was detected by CASTp software (Binkowski et al. 2003). The 3D descriptor (in SDF format) of diisopropyl-fluorophosphate (DFP) was retrieved from pubchem database (http://pubchem.ncbi.nlm.nih.gov). For in silico docking with diisopropyl-fluorophosphate (Wagner et al. 2002), the best orienting binding pose of ligand into the active site of the protein was estimated by using GOLD software in terms of its fitness score (Verdonk et al. 2003). The best fit pose generated a dock score which should be more than 30 (default). And ligand protein interaction, was indicated as hydrogen bonding and lipophilic interaction with other...
residues which was detected by LIGPLOT software (Wallace et al. 1995).

**Results and discussion**

Metagenomic library of soil metagenome isolated from Apharwat mountain peak in the north western Himalayas was constructed and 10,000 clones were screened for esterase activity on the tributyrin plates. Three clones showing esterase activity namely Aph2, Aph4 and Aph7 were isolated and preserved. An esterase-producing clone Aph4 has already been characterised (Sudan and Vakhlu 2012). In the present study, characterization of another esterase-producing clone Aph2 is being reported (Fig. 1a).

The insert in clone Aph2 harbours two non-overlapping functional ORFs as revealed by ORF Finder tool provided by NCBI (National Centre for Biotechnology Information) and the size of ORF 1 and ORF 2 of Aph2 was 356 and 811 bp, respectively. The nucleotide sequence of ORF 1 was 71% similar to Endopeptidase lactopectin gene (Bacillus sp.), whereas ORF 2 (Est Ac) showed 26% similarity with carboxylesterase gene (Synechococcus sp.) on BLASTP analysis (Table 1). Est Ac on analysis with NCBI tool (ORF Finder) suggested that it to encode for 271 aa enzyme that possess S-X-X-K motif of \( \beta \)-lactamases, wherein active catalytic Ser was conserved. Multiple sequence alignment of sequences of family VIII esterases, class C \( \beta \)-lactamases with Est Ac was carried out using CLUSTALX software. The sequences for multiple analysis were retrieved from NCBI (mentioned in “Materials and methods”, “Multiple sequence alignment” section). On multiple sequence alignment, the signature motif S-X-X-K and proton donor (Tyr) were found to be conserved in ORF 2 of clone Aph2. Ser present was found to be at position S\(^{68}\)-X-X-K\(^{71}\) that acts as catalytic nucleophile characteristic of family VIII carboxylesterase. The nonconserved residues, however, in conserved motif of Est Ac were Thr and Phe making the signature sequence of this esterase, S\(^{68}\)-T-F-K\(^{71}\) (Fig. 1b). Similar residues were found to be present in other cases like that of esterase Est C from Burkholderia gladioli (Wagner et al. 2002), lipase (LipBL) from halophilic bacterium Marinobacter lipolyticus SM19 (Perez et al. 2011), esterase Est M-N1 and Est M-N2 isolated from metagenomic DNA of Arctic soil sample (Yu et al. 2011), esterase Est C isolated from leachate fosmid shotgun library (Mokoena et al. 2013). However, the G-X-S-X-G motif of carboxylesterases and lipases that is known to be present along with S-X-X-K is absent in the present sequence. The esterase activity is dependent on S-X-X-K motif and G-X-S-X-G motif of lipases is reported to be non-functional (Petersen et al. 2001; Wagner et al. 2002; Perez et al. 2011; Yu et al. 2011; Mokoena et al. 2013).

A phylogenetic tree was constructed using CLc Sequence software, wherein Est Ac is grouped into family

![Fig. 1](link-to-image)
VIII carboxylesterase instead of β-lactamases (Fig. 1c) as Est Ac clustered with reported sequences belonging to this family. The distance of Est Ac with selected family VIII carboxylesterase sequences were found to be remarkable, indicating Est Ac to be a novel protein.

In silico modelling and signal peptide & catalytic centre detection

The nucleotide sequence was translated into amino acid sequence in silico using ORF Finder software. Est Ac was predicted to have 23 aa long N-terminus signal peptide using SIGNALP 3.0 server (Bendtsen et al. 2004). This could be cleaved to form mature protein (with a maximum cleavage site probability of 0.7 between Ala23 and His24). This could be a transmembrane protein as reported in similar study by Rashamuse and co-worker wherein 398 aa esterase Est C was obtained after the cleavage of 29 aa signal peptide from a protein of 423 aa (Rashamuse et al. 2009, 2011).

Homology modelling of Est Ac using MODELLER software

In addition to primary sequence analysis, structural insights were also generated by homology modelling using MODELLER (9.10 version) software (Eswar et al. 2006). Template IC18 (Est B isolated from Burkholderia gladioli that was confirmed by X-ray crystallography) was retrieved from PDB database and was used for model generation. Energies of models generated were calculated by the software in terms of their dope scores. In the present study, a total of five models were generated, model with least dope score, i.e. (−28911.082031) (Fig. 2a) was selected for future analysis. The model was visualised by using PyMOL software. Stereoechemical parameters of the selected model was validated using Ramachandran plot that showed no residues in the “disallowed regions” of ϕ/ψ space. The energy capacity was found to be accurate for the model which was calculated in terms of z score that determined overall quality which was −6.67. The local energy, however, determined by ProSA server that was also found to be negative. These results confirmed the model of good quality. The CPH model generated by ProSA software showed to contain all the residues in same place as that of the template thus validating the model (Fig. 2c). Wagner et al. (2002) constructed model for their protein Est B. The stereoechemical parameters were found to be satisfactory. Ramachandran plot showed no residues in disallowed region of ϕ/ψ space with Ala74 and Asp294 being present in disallowed region.

For carrying molecular docking a positive pocket possessing functional residues were determined using catalytic site atlas (CSA) software and CASTp software (Binkowski et al. 2003). In Est Ac catalytic residues Ser68 Lys71 Tyr103 were found in pocket number 37. It was found to be potent pocket for carrying docking with diisopropyl-fluorophosphate (DFP), a Ser inhibitor using GOLD software (Verdonk et al. 2003). It was observed that Ser68 found in S-X-X-K motif formed covalent interaction with DFP molecule through Oγ moiety (Fig. 2c). The GOLD score for the interaction determined was 48.0805. To study hydrogen bonding and lipophilic interactions between isolated protein and inhibitor (DFP), LIGPLOT software was used (Wallace et al. 1995). LIGPLOT generated hydrogen bonds of Ser68 along with other residues that may influence the catalytic power of the protein. The residues forming hydrogen bonds were Ala235 and Glu102 as is indicated with green dotted lines. However, the hydrophobic interaction were found to be seen in case of Lys232, Val234, Ser128, Asn168, Met67, Asn130, Glu164, Tyr103 (Fig. 2d). RMSD (root mean square deviation) difference of Ca atoms between the native structure and DFP derivative was calculated to be 0.20 Å. This interaction between Ser68 and ligand proved Ser68 to be potent binding catalytic residue for this inhibitor. Wagner and his group also reported the potent catalytic residue in their protein (Est B) which was found to be Ser75 present within β-lactamase Ser-X-X-Lys motif that acted as catalytic nucleophile and further suggested that covalent attachment of the inhibitor (diisopropyl-fluorophosphate) with their protein was made through Oγ of

### Table 1 ORF description of Aph2 gene

| S. no. | ORF name | Length | Frame | Accession no. | Similarity with | E value | Percentage of similarity |
|--------|----------|--------|-------|---------------|----------------|--------|-------------------------|
| 1      | ORF 1 (118 a.a) | 563–919 bp | −2 | ZP_07104368.1 | Endopeptidase lactopectin (Bacillus sp.) | 2e−22 | 71 |
| 2      | ORF 2 (271 a.a) | 130–2,116 | +3 | YP_478112.1 | Carboxylesterase with S-X-X-K motif of esterase (Synechococcus sp.) | 4e−07 | 26 |
Ser\textsuperscript{75}. RMSD difference of C\textalpha{} atoms between the native structure and the DFP derivative was 0.27 Å (Wagner et al. 2002) similar to the case reported in the present study. However, while conducting molecular docking it seemed that the acquisition of the inhibitor was from the top rather than from the front as reported by Wagner and his colleagues which further satisfied the claim of Est Ac to be an esterase and not a \(\beta\)-lactamase. The model generated falls well within the permitted limits suggesting it to be a valid model. The docking studies found Ser present in S-X-X-K motif as a potential nucleophile aiding in catalysis. Similar residues and similarities found in both Est Ac and Est B are given in Tables 3 and 4, respectively.

**Conclusion**

The present study was conducted to explore the untreated Aphanwat soil metagenome for its potential to isolate novel gene/s. A novel esterase gene was isolated which showed nucleotide sequence similarity largely with \(\beta\)-lactamases, but activity was seen on tributyrin plates. Sequence alignment studies and structural analysis proved that the
isolated protein was a novel esterase with active catalytic residues Ser\(^{68}\) Lys\(^{71}\) Tyr\(^{103}\), also found in class C \(\beta\)-lactamase. This may be because of convergent evolution wherein the catalytic Ser is present in S-X-X-K motif (\(\beta\)-lactamase motif), common in case of \(\beta\)-lactamases and family VIII esterases.

**Acknowledgments**  The project was sponsored by the Department of Science and Technology (DST), Government of India. The authors acknowledge resources sanctioned by DST, FIST and PURSE programmes. We are also grateful to the Department of Biotechnology (DBT) as some instruments funded by DBT were also used in present case. The bioinformatics analysis was performed at Jawaharlal Nehru University (INU), New Delhi, and at Centre for Bioinformatics, School of Biotechnology, University of Jammu.

**Conflict of interest**  The authors declare that they have no conflict of interest in publication.

**Open Access**  This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

**References**

Arpigny JL, Jaeger KE (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177–183

Bendtsen JD, Nielsen H, Heijne VG, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Berlemont R, Spice O, Delsaute M, Lara Y, Simon C, Power P, Daniel R, Galleni M (2013) Novel organic solvent tolerant esterase isolated by metagenomics: insight into lipase/esterase classification. Rev Argent Microbiol 45:3–12

Binkowski TA, Naghibzadeh S, Liang J (2003) CASTp: computed Atlas of surface topography of proteins. Nucleic Acids Res 31:3352–3355

Biver S, Portetelle D, Vandenbol M (2013) Characterization of new oxidant stable serine protease isolated by functional metagenomes. Springerplus 2:410

Daniel R (2005) The metagenomes of soil. Nat Rev 3:471–478

Eswar N, Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A (2006) Comparative protein structure modeling with MODELLER. Curr Protoc Bioinform 15:1–30

Foght J, Aislalibe J, Turner S, Brown C, Ryburn J, Saul D, Lawson W (2004) Culturable bacteriain subglacial sediments and ice from two southern hemisphere glaciers. Microb Ecol 47:329–340

Fuhrman JA (2012) Metagenomic and its connection to microbial community organization. P1000 Biol Rep 4:15

Gherardini PF, Wass MN, Citterich MH, Sternberg MJ (2007) Convergent evolution of enzyme active sites is not a rare phenomenon. J Mol Biol 372:817–845

Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, Handelsman J (2002) Isolation of antibiotic turbomycin A and B from a metagenomic library of soil microbial DNA. Appl Environ Microbiol 68:4301–4306

Knietsch A, Waschkwitz T, Bowien S, Henne A, Daniel R (2003) Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductases activity on *Escherichia coli*. Appl Environ Microbiol 69:1408–1416

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Lasowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Liebeton K, Eck J (2004) Identification and expression in *E. coli* of novel nitrite hydrolases from metagenome. Eng Life Sci 4:57–62

Mokoena N, Mathiba K, Tsikoe T, Steenkamp P, Rashamuse K (2013) Functional characterization of a metagenome derived family VIII esterase with a deacetylation activity on \(\beta\)-lactam antibiotics. Biochem Biophys Res Commun 437:342–348

### Table 2

| Structural refinement and model | Est Ac (native) |
|-------------------------------|-----------------|
| Residue no.                   | 271 aa          |
| Bond length (%)               | 99.4            |
| Bond angle (%)                | 94.1            |
| Ramachandran plot (with procheck) |                |
| Core region (%)               | 93.6            |
| Generously allowed regions (%)| 0.9             |
| Not allowed regions (%)       | 0.0             |
| G factors dihedral            | 0.07            |
| Covalent                      | −0.10           |
| Overall                       | 0.00            |
| M/C bond angle (%)            | 94.1            |
| Overall quality factor (ERRAT2)| 86.957         |
| ProSA overall model quality (Z score) | −6.62          |

### Table 3

| Equivalent residues of Est Ac and esterase (Est B) |
|---------------------------------------------------|
| Ser75 | Ser68 |
| Lys78 | Lys71 |
| Tyr133| Tyr103|
| Tyr181| Asn168|
| Val351| Val214|

### Table 4

| Similarity in enzyme properties | Est Ac (present enzyme) | Est B (Wagner’s enzyme) |
|---------------------------------|--------------------------|-------------------------|
| 1 Catalytic motif               | S-X-X-K                  | S-X-X-K                 |
| 2 Catalytic residues            | S\(^{68}\)-X-X-K\(^{71}\), Y\(^{103}\) | S\(^{75}\)-X-X-K\(^{78}\), Y\(^{133}\) |
| 3 KTG box                       | Absent                   | Absent                  |
| 4 Docking molecule              | DFP                      | DFP                     |
| 5 Blocking residue              | Ser\(^{68}\) through O\(^{γ}\) | Ser\(^{73}\) through O\(^{γ}\) |
Morrohoshi T, Oikawa M, Sata S, Kikuchi N, Kato N, Ikeda T (2011) Isolation and characterization of novel lipases from a metagenomic library of microbial community in the pitcher fluid of carnivorous plant Nepenthes hybrida. J Biosci Bioeng 112:315–320

Mullany P (2014) Functional metagenomics for investigation of antibiotic resistance. Virulence 5:443–447

Petersen EI, Valinger G, Solkner B, Stubenrauch G, Schwab H (2001) A novel esterase from Burkholderia gladioli which shows high deacetylation activity on cephalosporins is related to beta-lactamases and DD-peptidases. J Biotechnol 89:11–25

Perez D, Martin S, Lorente GF, Filice M, Guisan JM, Ventosa A, Garcia MT, Mellado E (2011) A novel halophilic lipase, LipBL, showing high efficiency in the production of eicosapentaenoic acid (EPA). PLoS One 6:e23325

Rashamuse K, Magomani V, Ronneburg T, Brady D (2009) A novel family VIII carboxylesterase derived from leachate metagenome library exhibit promiscuous beta lactamase activity on nitrocefin. Appl Microbiol Biotechnol 83:491–500

Rashamuse K, Sanyika W, Ronneburg T, Brady D (2011) A feruloyl esterase derived from a leachate metagenome library. BMB Rep 45:14–19

Schmeisser C, Steele H, Streit WR (2007) Metagenomic, biotechnology with non culturable microbes. Appl Microbiol Biotechnol 75:955–962

Sudan AK, Vakhlu J (2012) Isolation of a thioesterase gene from the metagenome of a mountain peak, Apharwat, in north western Himalayas. 3 Biotech 65:13205–13212

Vakhlu J, Kour A (2006) Yeast lipases: enzyme purification, biochemical properties and gene cloning. Electron J Biotechn 9:22–25

Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein-ligand docking using GOLD. Proteins 52:609–623

Vidy J, Swaroop S, Singh SK, Sukumaran RK, Pandey A (2011) Isolation and characterization of a novel a amylase from metagenomic library of Western Ghats of Kerala. India Biologia 66:939–944

Wagner UG, Peterson EI, Schwab H, Kratky C (2002) Est B from Burkholderia gladioli: a novel esterase with a β-lactamase fold reveals steric factors to discriminate between esterolytic and β-lactam cleaving activity. Protein Sci 11:467–478

Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three dimensional structures of proteins. Nucleic Acids Res 35:407–410

Wiseman A (1995) Introduction to principle. In: Wiseman A (ed) Handbook of enzyme biotechnology, 3rd edn. Ellis Horwood Ltd, Padstow, pp 3–8

Yu EY, Kwon MA, Lee M, Oh JY, Choi JE, Lee JY, Song BK, Hahn DH, Song JK (2011) Isolation and characterization of cold-active family VIII esterase from arctic soil metagenome. Appl Microbiol Biotechnol 90:573–581