Characterization of a novel ω-transaminase from a Triassic salt mine metagenome

Stephen A. Kelly, Timofey Skvortsov, Damian Magill, Derek J. Quinn, John W. McGrath, Christopher C.R. Allen, Thomas S. Moody, Brendan F. Gilmore

School of Pharmacy, Queen’s University Belfast, Belfast, BT9 7BL, N. Ireland, UK
School of Biological Sciences, Queen’s University Belfast, Belfast, BT9 7BL, N. Ireland, UK
Arran Chemical Company Limited, Unit 1 Monksland Industrial Estate, Athlone, Co. Roscommon, Ireland

ABSTRACT

Chiral amines are valuable building blocks for the pharmaceutical industry, and are increasingly synthesized by transaminase-mediated (TAm) synthesis. Currently available TAsms, primarily isolated from the genomes of cultured mesophilic bacteria, often suffer from a number of drawbacks, including poor substrate range and an inability to tolerate the harsh conditions often demanded by industrial processes. These characteristics have, in part, driven the search for novel biocatalysts from both metagenomic sources and extreme environments. Herein, we report the isolation and characterization of an ω-TAm from a metagenome of a Triassic salt mine in Kilroot, N. Ireland, an extremely hypersaline environment formed circa 220–250 mya. The gene sequence was identified based on homology with existing bacterial TAsms, synthesized within a pET28a(+) plasmid and expressed in E. coli BL21 DE3 cells. The resultant 49 kDa protein accepted (S)-methylbenzylamine (MBA) as amino donor and had a specific activity of 0.54 U/mg using ɑ-ketoglutarate (AKG) as substrate. Molecular modeling and substrate docking indicated the presence of key residues, conserved in a number of other TAsms. Despite the hypersaline environment from which it was isolated, the enzyme displayed low halotolerance, highlighting that not all biocatalysts will demonstrate the extreme characteristics associated with their source environment. This study does however reinforce the viability of mining metagenomic datasets as a means of discovering novel and functional biocatalysts, and adds to a currently scant list of such examples in the field of TAsms.

1. Introduction

A climate of increasing environmental awareness and the pursuit of optimized process efficiency has driven the rise of biocatalysis as an alternative for the production of chiral amines, versatile building blocks for the pharmaceutical industry [1]. Transaminases (TAsms) have emerged as an efficient and effective candidates for this purpose, and have been subject to a number of recent reviews [2–4]. TAsms are pyridoxal 5'-phosphate (PLP) dependent enzymes which catalyze the conversion of an amine group to a prochiral ketone, resulting in the formation of an optically active amine. The potential of TAsms for chiral amine synthesis was arguably best demonstrated with the production of the antidiabetic drug sitagliptin from its prochiral ketone precursor [5]. In this example the use of an engineered TAm resulted in increased yield, reduced waste and removed the need for the toxic rhodium-catalyzed hydrogenation and associated expensive high-pressure equipment.

Despite high-profile successes such as these, TAsms can suffer from narrow substrate specificity and range, and an inability to perform under the challenging conditions encountered in some industrial processes. Such drawbacks emphasize the need for new enzymes or improvements within the existing catalogue of TAsms via protein engineering. A number of groups have demonstrated the benefits of rational design of known TAsms for improvements in activity and substrate acceptance [6–8].

Approaches have also included mining microorganisms which are tolerant to extreme conditions [9–11], with the view that enzymes which can survive such extreme environments can also tolerate harsh conditions often demanded of industrial processes.
Mining of metagenomic datasets has also emerged as an exciting approach in the search for novel enzymes, providing access to sequences beyond those which can be obtained using a culture-based approach alone. Although still in its relative infancy, metagenomic approaches have recently been successfully employed in the search for TAmS [12], including thermostable TAmS from hot spring metagenomes [13].

In this study we characterize a previously unreported ω-TAm, whose gene was identified in metagenomic DNA obtained from a saturated brine pool from a Triassic salt mine (formed circa 220–250 mya). We have previously reported an ω-TAm cloned and expressed from a salt mine isolate following a culture-based approach [9]. Obtaining environmental DNA from this hypersaline environment has allowed us to expand the search for novel enzymes beyond those accessible using standard culture-based techniques, and access biocatalyst-coding genes in the DNA of uncultivable microorganisms.

2. Materials and methods

2.1. DNA extraction and metagenomic sequencing

Metagenomic DNA was extracted from a brine sample taken from Kilroot salt mine (54.7331° N, 5.7509° W). Brine was sterile filtered using a 0.2 μm nitrocellulose filter membrane. This membrane was cut into several pieces and DNA was extracted using GenElute™ Bacterial Genomic DNA Extraction kit (Sigma-Aldrich, UK) as per manufacturer’s instructions. A sample of extracted DNA was sent to MR DNA Lab (Shallowater, TX, USA) for sequencing library preparation (Nextera XT) and metagenomic sequencing on the Illumina MiSeq platform.

2.2. Gene identification and preparation of pET28a(+)KMG-TAm4 plasmid

Initial quality control was performed on the metagenomic dataset with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and reads were processed with BBMap v33.54 (http://sourceforge.net/projects/bbmap/). All reads with an average Q-score <10 or containing Ns were discarded, with reads then trimmed of adaptors and quality-trimmed. The assembly of the processed reads into contigs was performed with SPAdes genome assembler v 3.10.1 [14] in—meta mode with default parameters. Prodigal v2.6.3 [15] was used to predict genes in the assembled contigs. BLAST [16,17] search with default parameters was carried out to identify full-length coding sequences homologous to the genes of amine TAmS from Vibrio fluvialis (UniProt: F2XBU9) and Halomonas sp. CSM-2 (Refseq: WP_085918404.1). The pET28a(+) KMG-TAm4 plasmid was synthesized via GenPlus cloning by GenScript, comprising the KMG-TAm4 gene sequence in pET28a(+) vector, with flanking NdeI and XhoI restriction sites. Plasmids were reconstituted in molecular grade water to a concentration of 0.2 μg/μL.

2.3. Transformation, expression and phylogenetic analysis

Following reconstitution, 2 μL of pET28a(+)KMG-TAm4 was used to transform 20 μL One Shot® BL21 DE3 Chemically Competent Escherichia coli cells as per manufacturer’s instructions. Cells were also transformed with empty pET28a(+) vector for use as a negative control. Transformants were grown and induced for expression, cell-free extract produced for enzymatic assays, and analysed by SDS-PAGE and Western Blotting using standard protocols as described previously (Supplementary information from Ref. [9]).

ClustalW [18] was used to perform multiple sequence alignment (MSA) of the amino acid sequences of KMG-TAm4 and a number of other TAmS. Boxshade (https://embnet.vital-it.ch/software/BOX_form.html) was used to visualise conserved residues based on this alignment and MEGA v7.0.26 [19] was used to generate a neighbour-joining tree, with bootstrap values calculated from 1000 replicates.

2.4. Molecular modeling and docking

Molecular models of the KMG-TAm4 were constructed using the I-Tasser server followed by energy minimization using Yasara [20,21]. Molecular models of pyridoxamine 5’-phosphate (PMP) and α-ketoglutarate (AKG) were constructed in Ascalaph Designer and subject to 1000 steps of steepest descent energy minimization. The fidelity of the KMG-TAm4 model was assessed using the tools of the Whatif server along with Ramachandran analysis. Molecular models and docked complexes were viewed and manipulated in PyMol [22]. Structural comparisons and corresponding quantitative similarity was determined utilizing the TM-Align algorithm [23]. Protein comparisons to the PDB database were carried out using 3D-BLAST [24].

Autodock Vina was utilised for the docking of both PMP and AKG in two separate docking events. Yasara energy minimization was conducted following each docking step resulting in the production of the final docked complex [25]. Construction of residue interaction maps was carried out using Ligula’s followed by visualization in PyMol.

2.5. HPLC-based screening assays

(S)-Methylbenzylamine (MBA) (25 mM) and PLP (1 mM in KPi buffer, 100 mM, pH 7.0), and AKG (5 mM) were added to a 96-well deep well plate to give a total volume of 180 μL. The reaction was initiated by the addition of 20 μL rehydrated KMG-TAm4 cell-free extract to a final working concentration of 20 μg/mL. After incubation for 16 h at 30 °C and 1200 rpm, the reaction was quenched with 800 μL 62.5% acetonitrile and denatured protein was removed by centrifugation. 200 μL of supernatant was removed for analysis by HPLC using an XSelect™ CSH™ C18 5 μm column (4.6 × 250 mm) at 25 °C and an Agilent 1260 Infinity model, with detection of acetoephonene coproduct at 240 nm. The concentration of acetoephonene coproduct formation was determined over 11 min using a gradient of 5–100% organic solvent (where A = dH2O with 0.1% H2PO4 and B = acetonitrile with 0.1% H2PO4). Acetoephonene eluted at a retention time of 8.6 min. Assays using both KMG-TAm4 and vector only controls were carried out in triplicate, with an average taken for each and values for negative control subtracted from KMG-TAm4 values.

To assess the effect of temperature on enzyme performance, assays were set up as described above but incubated at different temperatures, ranging from 30 to 60 °C. To assess the effect of changing pH, assays were set up using universal buffer [26] in place of KPi. This contained 25 mM citric acid, 25 mM KH2PO4, 25 mM Tris, 12.5 mM Na2B4O7 and 25 mM KCl, adjusted using NaOH, with pH values from pH 6.0 to 11.0 investigated. To assess the effect of organic solvents, reactions were supplemented with organic solvents whilst maintaining overall concentrations of KPi buffer (pH 7.0) and PLP in solution at 100 mM and 1 mM respectively. Concentrations of organic solvents investigated ranged from 0 to 30%. To assess the effect of increasing salt concentration on enzyme performance, reactions were supplemented with NaCl or KCl, assessing a range of 0–4 M for each.

Please cite this article in press as: S.A. Kelly et al., Characterization of a novel ω-transaminase from a Triassic salt mine metagenome, Biochemical and Biophysical Research Communications (2018), https://doi.org/10.1016/j.bbrc.2018.08.073
2.6. Determination of specific activity

Specific activity was determined by adapting an assay described previously [27]. (S)-MBA (5 mM) and PLP (16 μM in 0.1 M KPi buffer, pH 7.0), and AKG (1 mM) were added to a 96-well plate, along with NAD⁺ (5 mM) and glutamate dehydrogenase (1 U), to give a total volume of 90 μL. The reaction was initiated by the addition of 10 μL rehydrated KMG-TAm4 cell-free extract to a final working concentration of 25 mg/mL. The reaction proceeded at 30 °C, with NADH formation measured at 340 nm over 20 min. Assays were carried out in triplicate, with an average taken for each. KMG-TAm4 concentration in cell-free extract was determined using a combination of a Qubit® 2.0 Fluorometer (Invitrogen, UK) and SDS-PAGE densitometry.

3. Results and discussion

3.1. Expression and phylogenetic analysis

BLASTp search identified an adenosylmethionine-8-amino-7-oxononanoate TAm from Halomonas utahensis as the closest neighbour to KMG-TAm4 based on the amino acid sequence (97% identity). SDS-PAGE and Western blot analysis showed the protein to be ~49 kDa, approximately commensurate with its amino acid sequence (Fig. 1A and B). Evolutionary relationships with other previously reported TAsms were investigated and these are shown in Fig. 1C. Comparisons with previously reported TAsms show sequence similarities of 29, 33, 34, 35 and 36% to TAsms from Pseudomonas putida (Pp-TAm) [28], Vibrio fluvialis (Vf-TAm) [29], Chromobacterium violaceum (Cv-TAm) [30], Hs-TAm [31] and Ad2-TAm [8] respectively. Despite the relatively low homology observed to these TAsms, KMG-TAm4 contains key conserved equivalent residues required for TAm activity and substrate specificity, namely Trp60, Tyr153 and Lys288 [32] (Fig. 2).

3.2. Molecular modeling and docking

Whatif and Ramachandran analysis of the most favorable KMG-TAm4 model revealed no evidence suggesting structural defects. The most energetically favorable representation of the docked complex is given in Fig. 3a. 3D-BLAST analysis of the KMG-TAm4 model revealed a best hit to a 7,8-Diaminopelargonic acid synthase from E. coli (similarity score of 0.8907 with TM-Align analysis). LigPlus was utilised in the identification of proposed interacting residues with both AKG and PMP. A 3D representation of the resulting interaction network is presented in Fig. 3b. Structural alignments were utilised in an attempt to identify catalytically important residues. This identified a number of residues whose equivalents have already been established to possess catalytically important roles in the literature. The residues Trp60 and Tyr153 are of particular significance, and are conserved in the active sites of a number of previously reported TAsms [13,33,34]. Here it is observed that contributions to substrate specificity are provided by Trp56, Tyr149, and Arg414.

The catalytic lysine was subsequently shown to be Lys284, with detailed 2D maps provided in Fig. S1. This presence of this residue is characteristic of TAsms, and is critical in Schiff base formation as it binds with PLP, subsequently releasing the coenzyme prior to PMP formation [7,32]. With respect to PMP, Ser117, Asn115, and Glu152 were predicted as being engaged in hydrogen bonding with the oxygen atoms of the phosphate moiety. This appears to orientate it within its binding pocket such that it lies in close proximity to Lys284. There are a number of active site residues in KMG-TAm4 interacting with the coenzyme in its PMP form, which are present in the active site of other TAsms. Noted previously are equivalents of Tyr149, which allows for hydrogen bonding of a hydroxyl group to the phosphate group of PLP [32] and Ile257 [7]. Taken together, our predicted complex provides a feasible representation of the binding mode between KMG-TAm4, PMP, and AKG.

3.3. Evaluation of enzyme performance under varying reaction parameters

Enzyme performance was characterized using (S)-MBA as amino donor and AKG as substrate under a number of reaction conditions. Similar trends for relative conversion were observed when the reaction was supplemented with varying concentrations of both NaCl and KCl (Fig. 4a). KMG-TAm4 performed best in both cases when no salt was added to the reaction mixture, with sizeable reductions in relative conversion noted as salinity increased, decreasing to 32.1% and 29.5% relative conversion when 1 M NaCl and KCl were added respectively.
Fig. 2. Multiple alignment of amino acid sequences of KMG-TAm4 with a number of previously reported TAm sequences. Key residues for activity and substrate specificity are conserved across all sequences, including equivalents of Trp60, Tyr153 and Lys288 (highlighted in red).

Fig. 3. Molecular model of KMG-TAm4 and Predicted Complex Between α-KG and PMP. a. Predicted final model of the complete complex with α-KG shown in blue and PMP shown in yellow. b. Enlarged area of the active site shown in a with 3D representation of residues interacting with the cofactor and substrate as predicted from the docked complex. Relevant distances are shown (in units of Å). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
Despite being derived from an extremely hypersaline environment, KMG-TAm4 itself appears to possess little salt tolerance, with relative conversions decreasing substantially with increased salinity. This can be explained by the relative adaptive mechanisms employed by haloarchaea and halotolerant bacteria. Based on the BLASTp results, where the closest neighbour was a protein from the halotolerant bacterium, *Halomonas utahensis*, coupled with characterization of enzyme performance under varying reaction parameters, it can be assumed that KMG-TAm4 is derived from the DNA of a bacterial species. Unlike their haloarchaeal counterparts, the enzymes and intracellular machinery of halotolerant bacteria aren't necessarily adapted to high salt concentrations, largely excluding salt from the cytoplasm and replacing it with other ‘compatible solutes’ [39–41]. This perhaps provides a cautionary tale that genomic mining from extreme environments will not necessarily deliver enzymes which exhibit extreme adaptations.

KMG-TAm4 had a pH optimum of 7.0, although the enzyme retained reasonable activity over the range of pH 6.0–8.0, with relative conversions of 88.9% and 76.7% seen for pH 6.0 and 8.0 respectively, relative to optimum (Fig. 4b).

KMG-TAm4 showed an ability to function in the presence of organic co-solvent. Better conversions were achieved with 10% of each co-solvent compared to no co-solvent in the reaction, with the exception of THF. The greatest conversions were observed with 10% DMF in the reaction mixture. Enzyme performance deteriorated rapidly with increasing co-solvent concentrations (20 and 30%), with the exception of DMSO, where the decrease was less pronounced (Fig. 4c). The additions of DMF and DMSO have previously been applied in API synthesis and other TAM-containing processes [42,43]. The greatest conversion was seen at 30 °C, which decreased with increasing temperature (Fig. 4d).

KMG-TAm4 was assayed using a variety of structurally diverse ketone and aldehyde substrates (Fig. S2.), but displayed minimal activity to substrates other than σ-KG. This is in contrast to our previous work, in which a TAM from a halotolerant bacterium exhibited a much broader substrate scope [9]. The acceptance of (S)-MBA as amino donor is a positive characteristic, given its established and widespread use as an amino donor molecule in a number of TAM-catalyzed processes [44–47]. Only the (S)-isomer of MBA was accepted by KMG-TAm4, showing the enzyme to be an (S)-selective TAM.
3.4. Determination of specific activity and overall conversion

Specific activity for KMG-TAm4 using (S)-MBA as amino donor and AKG as amino acceptor was calculated to be 0.54 U/mg, with one unit defined as the amount of enzyme producing 1 μmol of NADH per minute. The highest conversion observed was 23.9% conversion of AKG with DMF as co-solvent at 10% concentration.

Metagenomic mining provides access to a vast resource of TAm-coding DNA sequences, previously unobtainable using culture-based search methods. When applied to extreme environments, metagenomic mining has the potential to unlock highly adapted, tolerant enzymes with robust properties. Caution must also be urged with metagenomic samples from hypersaline environments, as the respective adaptive mechanisms of archaea and bacteria mean that although source organisms may be halophilic, not all enzymes derived from these will have extreme characteristics. This work provides a further example, in the relatively scant literature to date, of an active and functional ω-TAm isolated and expressed using a metagenomic mining approach, and highlights the feasibility of using this approach in the quest for novel biocatalytic enzymes.

Acknowledgements

This work was supported by Biotechnology and Biological Sciences Research Council (BBSRC) through an ICASE Training Grant with Almac Group Ltd (Grant reference: BB/L017083/1).

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.bbrc.2018.08.073.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.bbrc.2018.08.073.

References


