Journal of Biotechnology 194 (2015) 67–80

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Two-stage gene assembly/cloning of a member of the TspDTI subfamily of bifunctional restriction endonucleases, TthHB27I

Daria Krefft[a], Agnieszka Zylicz-Stachula[a], Ewa Mulkiewicz[b], Aliaksei Papkov[a], Joanna Jezewska-Frackowiak[a], Piotr M. Skowron[a,*]

[a] Department of Molecular Biotechnology, Institute for Environmental and Human Health Protection, Division of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland
[b] Department of Environment Analysis, Institute for Environmental and Human Health Protection, Division of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland

ARTICLE INFO

Article history:
Received 11 September 2014
Received in revised form
22 November 2014
Accepted 27 November 2014
Available online 6 December 2014

Keywords:
Restriction endonuclease
Methyltransferase
Type IIS
Type IIG
Thermus

ABSTRACT

The *Thermus* sp. family of bifunctional type IIS/IIG/IIC restriction endonucleases (REase)–methyltransferases (MTase) comprises thermo-stable TaqII, TspGWI, TspDTI, TsoI, Tth111II/TthHB27I enzymes as well as a number of putative enzymes/open reading frames (ORFs). All of the family members share properties including a large protein size (*ca.* 120 kDa), amino acid (aa) sequence homologies, enzymatic activity modulation by S-adenosylmethionine (SAM), recognition of similar asymmetric cognate DNA sites and cleavage at a distance of 11/9 nt. Analysis of the enzyme aa sequences and domain/motif organisation led to further *Thermus* sp. family division into the TspDTI and TspGWI subfamilies. The latter exhibits an unprecedented phenomenon of DNA recognition change upon substitution of SAM by its analogue, sinefungin (SIN), towards a very frequent DNA cleavage. We report cloning in *Escherichia coli* (*E. coli*), using a two-stage procedure and a putative *tthHB27IRM* gene, detected by bioinformatics analysis of the *Thermus thermophilus* HB27 (*T. thermophilus*) genome. The functionality of a 3366 base pair (bp)-/1121 aa-long, high GC content ORF was validated experimentally through the expression in *E. coli*. Protein features corroborated with the reclassification of TthHB27I into the TspDTI subfamily, which manifested in terms of aa-sequence/motif homologies and insensitivity to SIN-induced specificity shift. However, both SAM and SIN stimulated the REase DNA cleavage activity by at least 16–32 times; the highest was observed for the *Thermus* sp. family. The availability of TthHB27I and the need to include SAM or SIN in the reaction in order to convert the enzyme from "hibernation" status to efficient DNA cleavage is of practical significance in molecular biotechnology, extending the palette of available REase specificities.

1. Introduction

REases have become interesting model proteins for function–structure studies, as they are typically not related to each other at the aa sequence levels, and they exhibit very short and variable epitopic similarities. Nevertheless, despite the very high aa sequence diversities, they are represented by multiple variants (isoschizomers) that are capable of conducting the same function of specific DNA sequence recognition and pre-programmed cleavage. The enzymes have been classified into four Types (I, II, III and IV) on the basis of their gene and protein organisation, cofactor requirements, modes of cognate DNA sequence recognition and cleavage; although, wide internal diversification and atypical REases within the types exist (Roberts et al., 2003). One such case involves the family of *Thermus* sp. enzymes, belonging to subtype IIS/IIC/IIG REases, which have been defined previously (Skowron et al., 2003). We cloned and expressed all of the coding genes for thus-far characterised enzymes of this group (Zylicz-Stachula et al., 2009, 2012, 2014; Skowron et al., 2013; unpublished results). The enzymes can also be considered as "half" of the functionally streamlined type I enzyme, with all restriction-modification (RM) system components located within a single polypeptide (Zylicz-Stachula et al., 2009). Moreover, the *Thermus* sp. family appears to be a transitional REase group between Type I, II and III REases.

* Corresponding author.
E-mail addresses: daria.krefft@phdstud.ug.edu.pl (D. Krefft), a.zylicz-stachula@ug.edu.pl (A. Zylicz-Stachula), ewa.mulkiewicz@ug.edu.pl (E. Mulkiewicz), aliaksei.papkov@ug.edu.pl (A. Papkov), j.jezewska-frackowiak@ug.edu.pl (J. Jezewska-Frackowiak), piotr.skowron@ug.edu.pl, pmars44@gmail.com (P.M. Skowron).

http://dx.doi.org/10.1016/j.jbiotec.2014.11.030
0168-1656/

The *Thermus* sp. family shares a consecutive domain and active motif arrangement of the REase–MTase structure, an unusually large protein size for prokaryotes of approximately 120 kDa, homologies and similarities in polypeptide aa sequences and in the 5–6 bp-long asymmetric recognition sites in DNA, the same cleavage distance of 11/9 nt from cognate sites and, finally, REase activity modulated by SAM/SIN. Originally, the family members (Zylicz-Stachula et al., 2009) were found in various thermophilic *Thermus* sp. strains; thus, it is suspected that they have evolved from a common ancestor (Skowron et al., 2003). Characterised thus-far biochemically and genetically, the family members include six thermostable enzymes: TaqII, TspGWI, TspDTI, TsoI and non-identical isoschizomers Tth111II and TthHB27I (Zylicz-Stachula et al., 2002, 2009, 2011a,b, 2012; Skowron et al., 2003, 2013; Barker et al., 1984; Shinomiya et al., 1980; this work). A detailed bioinformatics and biochemical comparison has shown that the *Thermus* sp. group is further diversified into two subfamilies of more closely related enzymes: the TspGWI subfamily and the TspDTI subfamily (Zylicz-Stachula et al., 2009, 2012; Skowron et al., 2013).

In this work, we present the cloning, expression and analysis of the atypical subtype IIS/IIG/IIC REase (TthHB27I), which belongs to the TspDTI subfamily of bifunctional enzymes. Furthermore, we present biotechnological aspects of the SAM or SIN induced activation of "hibernated" thermophile-derived enzymes for useful DNA analysis/manipulation.

2. Materials and methods

2.1. Bacterial strains, plasmids, media and reagents

The TthHB27I-producing bacteria, *T. thermophiles*, was obtained from the Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH collection (Braunschweig, Germany). The bacteria were cultivated at 65 °C in microelement-enriched Luria broth, with a composition of (per 1 L) peptone (3 g), yeast extract (2 g), nitrilotriacetic acid (0.1 g), $CaSO_4 \times 2H_2O$ (60 mg), $MgSO_4 \times 7H_2O$ (0.1 g), NaCl (8 mg), KNO_3 (0.103 g), $NaNO_3$ (0.689 g), Na_2HPO_4 (0.111 g), $FeCl_3$ (2.8 mg), H_2SO_4 (0.5 μL) 95% solution, $MnSO_4 \times H_2O$ (2.2 mg), $ZnSO_4 \times 7H_2O$ (0.5 mg), H_3BO_3 (0.5 mg), $CuSO_4$ (0.016 mg), $Na_2MoO_4 \times 2H_2O$ (0.025 mg) and $CoCl_2 \times 6H_2O$ (0.046 mg). *E. coli* TOP10 {F-*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *nup*G *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*E15 *gal*K16 *rps*L(Str^R) *end*A1 λ$^-$} (Invitrogen, Carlsbad, CA, USA) was used for clone selection. For protein expression, pET21d(+) vector and *E. coli* BL21(DE3) {F–*omp*T *hsd*SB(r_{B-}, m_{B-}) *gal dcm* (DE3)} were employed (Novagen, Madison, WI, USA). Components of the LB and TB media (Green and Sambrook, 2012) were from Becton–Dickinson (Franklin Lakes, NJ, USA). Agarose was obtained from FMC (Rockland, NY, USA), Heparin–Agarose was purchased from GE Healthcare (Uppsala, Sweden), whereas Phosphocellulose P11 and DEAE–cellulose were obtained from Whatman (Springfield Mill, UK). Other reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA), POCh S.A. (Gliwice, Poland), AppliChem Inc. (St. Louis Missouri, MO, USA) or Fluka Chemie GmbH (Buchs, Switzerland). Plasmid pBR322, REases, T4 DNA polymerase and T4 DNA ligase were obtained from Thermo Fisher Scientific Baltics UAB (Vilnus, Lithuania). Plasmid pACYC184 was obtained from Piotr Skowron's collection. *E. coli* BL21(DE3) and *T. thermophilus* total DNAs were isolated using Genomic Mini AX Bacteria kit. DNA purification kits and Marathon DNA polymerase were purchased from A&A Biotechnology (Gdynia, Poland). DNA ladders were obtained from Vivantis (Subang Jaya, Malaysia). Oligodeoxyribonucleotide (oligo) synthesis and DNA sequencing were conducted at Genomed S.A. (Warsaw, Poland). An Amicon ultrafiltration unit (180 and 10 mL) was used with pressurised nitrogen for protein sample concentrations. 100 kDa cut-off RC filters for protein concentration were obtained from Millipore Corporation (Billerica, MA, USA).

2.2. Determination of TthHB27I recognition and cleavage sites

Native TthHB27I REase preparation (Fig. 1), isolated from *T. thermophilus*, was used for the recognition site and cleavage position determinations, with the use of: (i) digestion pattern analysis of plasmid DNA, (ii) digestion of a 1789 bp PCR substrate, (iii) run-off sequencing of PCR digestion products. In method (i), 0.5 μg of plasmid pBR322 DNA was digested for 1 h with 1.5 units of TthHB27I in 25 μL of the TthHB27I reaction buffer, containing 10 mM Tris–HCl (pH 7.0 at 65 °C), 6 mM $MgCl_2$, 40 mM NaCl, 6 mM β-mercaptoethanol (ßMe) and 1 mg/mL BSA, supplemented prior to the reactions with 100 μM SAM. The "base" buffer did not contain $MgCl_2$ or SAM (Fig. 2A). As all *Thermus* sp. enzymes do not digest DNA completely, 1 unit of TthHB27I REase is defined as the minimal amount of enzyme needed to hydrolyse 1 μg of λ DNA in 1 h at 65 °C in 50 μL of TthHB27I buffer, supplemented with 100 μM SAM, resulting in a stable partial DNA digestion pattern. Plasmids, genomic DNA and PCR were analysed after the digestions using agarose; this was conducted in 1× TBE buffer and stained with ethidium bromide (EtBr). The DNA band patterns were subjected to analysis using the following software: DNASIS MAX (Hitachi Software, San Bruno, CA, USA) and REBASE Tools (http://rebase.neb.com), REBsites, NEBcutter (Vincze et al., 2003) and REBpredictor (Gingeras et al., 1978). In method (ii), a PCR product, amplified from pACYC184 DNA (576–2364 bp from pACYC184), was digested with TthHB27I (Fig. 2B; Supplementary Fig. S1) and the digestion products were subjected to agarose electrophoresis. In method (iii), DNA fragments from TthHB27I digestion of the PCR substrate were subjected to agarose electrophoresis and isolated using electro-elution, phenol–chloroform/2× chloroform extractions and ethanol precipitation, followed by DNA sequencing using the ABI Prism 310 automated sequencer with ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer Applied Biosystems, Foster City, CA, USA). The run-off sequencing reactions were performed with divergently oriented primers, heading towards DNA fragment ends of the TthHB27I-digested PCR substrate. The fragments were PAGE separated and individually electro-eluted and precipitated with ethanol. Purified fragments were subjected to sequencing using the following oligo primers (in parentheses positions on pACYC184 map, bp): F-pACYCminiSEQ (1364-1382)-5′-GAGATTACGCGCAGACCAA-3′, R-pACYC (1693-1672)-5′-GAATGGACGATATCCCGCAAGA-3′.

2.3. DNA cleavage and methylation assays

The putative TthHB27I recognition sites (http://rebase.neb.com) were verified using a custom PCR substrate, amplified from pACYC184, containing two variants of the TthHB27I cognate sequence: 5′-CAAACA-3′ and 5′-CAAGCA-3′ (see Supplementary Fig. S1). This substrate, if cleaved to completion, would result in three well-resolved on-gel bands of 872, 602 and 311 bp (Fig. 2B). The PCR was amplified with a pair of primers (forward: 5′-CATCAGCGCTAGCGGAGTGTA-3′ and reverse: 5′-CGAGGGCGTGCAAGATTCC-3′). The reactions were performed in 100 μL volumes in a 2720 Thermal Cycler (Perkin Elmer Applied Biosystems) and contained 1× Marathon PCR buffer, 0.4 mM dNTPs, 0.5 μM of each primer, 35 ng of template DNA and 1.5 units of Marathon DNA polymerase. The cycling profile included: 97 °C for 4 min, 89 °C for 20 s (addition of the polymerase), then 30 cycles of 94 °C for 30 s, 56 °C for 30 s and 68 °C for 2 min, and a final step of 68 °C for 1.5 min. PCR products were analysed and purified using