Gene 660 (2018) 8–12

Contents lists available at ScienceDirect

Gene

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

Research paper

New bifunctional restriction-modification enzyme AloI isoschizomer (PcoI): Bioinformatics analysis, purification and activity confirmation

Marina L. Mokrishcheva[a], Attila Kertesz-Farkas[b], Dmitri V. Nikitin[a,c,*]

[a] Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Prospekt Nauki, 5, Pushchino, Moscow region 142290, Russia
[b] National Research University Higher School of Economics, Kochnovskiy Proezd 3, Moscow 125319, Russia
[c] M.V. Lomonosov's Moscow State University, Biology Faculty, Leninskie Gory, 1, b. 12, Moscow 119234, Russia

ARTICLE INFO

Keywords:
Restriction-modification system evolution
Bifunctional restriction-modification enzymes
Restriction endonuclease
DNA-methyltransferase

ABSTRACT

Type II restriction endonucleases and modification DNA-methyltransferases are key instruments of genetic engineering. Recently the number of proteins assigned to this group exceeds 8500. Subtype IIC organizes bifunctional endonuclease-methyltransferase enzymes and currently consists of 16 described members. Here we present phylogenetic tree of 22 new potential bifunctional endonucleases. The majority of them are thought to be fusions of a restriction nuclease with a DNA-methyltransferase and a target recognition subunit of type I restriction-modification systems (R-M-S structure). A RM.AloI isoschizomer from *Prevotella copri* DSM-18205, PcoI, has been cloned, purified and its REase activity demonstrated. It cuts DNA in magnesium-dependent manner and demonstrates high affinity to DNA, which probably reflects its mechanism of action. This work provides additional proves that gene fusion might play an important role in evolution of restriction-modification systems and other DNA-modifying proteins.

1. Introduction

Restriction-modification systems (RMS) serve as prokaryotic defense tools against external DNAs invasion (Williams, 2003). They stimulate microbial biodiversity, while reducing horizontal gene transfer. Usually, they consist of a restriction endonuclease (REase) and a modification DNA methyltransferase (MTase) enzyme recognizing the same short 4–8 nucleotide sequence. MTase transfers methyl groups to adenines or cytosines within the target sequence, thus preventing modified DNA from hydrolysis by a cognate REase. In a corresponding database, REBASE, over 20,000 different characterized and putative RMS could be found (Vincze et al., 2010). Some of them have head-to-tail R and M gene orientation, thus providing, by our hypothesis, the possibility of gene fusion through point mutations or genome rearrangements (Wilson, 1991). Similar events could be involved into origin of bifunctional restriction enzymes of type IIC (Roberts et al., 2003), bifunctional MTases and regulatory SsoII-related MTases (Karyagina et al., 1997). Previously, we have proven the possibility of bifunctional enzyme origin through gene fusion in the case of Eco29kI RMS with intact REase and MTase activities (Mokrishcheva et al., 2011; Zakharova et al., 1998; Nikitin et al., 2003; Nikitin et al., 2007).

Bifunctional proteins of type IIC have interesting practical applications for the creation of enzymes with altered recognition sites. Recently, three approaches have been proposed: methylation activity-based selection (MABS, (Rimseliene et al., 2003)), target recognition domain (TRD) reassortment (Jurenaite-Urbanaviciene et al., 2007) and rational specificity engineering of type IIL enzymes (Morgan and Luyten, 2009). Following the first two, type IIC enzymes such as Eco57I, AloI, PpiI, and TstI have been successfully manipulated to change their recognition sequences. The critical point is that both of their activities can be operated on the same target sequence, allowing using the DNA-modification activity of these enzymes for the selection of mutants with new recognition site. With our recent work we have increased the list of bifunctional REases available for these manipulations.

In our previous work (Mokrishcheva et al., 2012) we have searched protein databases and found 22 new potential bifunctional REases. Their phylogenetic analysis is presented in this report. A potential RM enzyme from *Prevotella copri* DSM-18205 has been cloned, purified and its REase activity confirmed. These data provide additional evidences that gene fusion plays an important role in evolution of restriction-modification systems and other DNA-modifying proteins.

Abbreviations: ORF, open reading frame; REase, restriction endonuclease; MTase, methyltransferase; RMS, restriction-modification system; TRD, target recognition domain; aa, amino acids; SAM, S-adenosylmethionine

* Corresponding author at: Leninskie Gory, 1, b. 12, Moscow 119234, Russia.
E-mail address: dvnikitin@rambler.ru (D.V. Nikitin).

https://doi.org/10.1016/j.gene.2018.03.069
Received 15 November 2017; Received in revised form 19 March 2018; Accepted 20 March 2018
Available online 21 March 2018
0378-1119/

2. Materials and methods

2.1. Phylogenetic tree construction

Previously, 325 restriction enzyme sequences have been downloaded from the REBASE database (Vincze et al., 2010; Mokrishcheva et al., 2012) and searched against the non-redundant protein database. The local pairwise alignments hits have then been filtered using the following criteria: a match to a restriction enzyme was kept if its E-value was less than 1e−140, the sequence identity within the aligned region was greater than 80% and the length of the subject sequence was at least 1.5 times longer than the query. With this algorithm we have selected 22 candidate matches (Mokrishcheva et al., 2012; Altschul et al., 1990). The phylogenetic tree has been built using the neighbour joining method with the Jukes-Cantor distance algorithm (Felsenstein, 2003) from the Bioinformatics toolbox V3.4 of Matlab 2009b.

2.2. New potential RM enzyme from Prevotella copri DSM-18205 gene cloning, protein purification and testing

Potential gene ZP_06253288.1, coding for bifunctional RM protein, has been amplified from *Prevotella copri* DSM-18205 genomic DNA by the following primers:

ZP_06253288.1
PC_fw: TTTGTCGACATGGTAAAAGAACAAATAGGAG;
PC_rev: CCTGGATCCTTACTTTAAATATTTGTCCAAGAT

BamHI and SalI sites, used for subsequent cloning, are underlined. The amplified DNA fragment has been inserted into pET19mod (Ap^r) vector (Alexandrov et al., 1999) and its correctness has been verified by sequencing.

This construct has been introduced into *E. coli* BL21 Star (DE3) cells via calcium transformation method (Sambrook et al., 1989). Fresh cells have been inoculated into LB liquid medium containing 100 μg/μl of Ampicillin and grown till OD_{600} = 0.6. Then 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) has been added and cells have been grown for 16 h at 18 °C, collected by centrifugation and stored at −70 °C.

The protein has been purified via Ni-NTA and heparin column chromatography, dialysed against storage buffer (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 7 mM β-mercaptoethanol, 0.1 mM EDTA, 0.1% Triton X-100, 50% glycerol (v/v)) and stored at −20 °C. Quality of the purified protein has been assessed by SDS-PAGE according to the method of Laemmli with Coomassie Blue R-250 conventional staining (Laemmli, 1970). For endonuclease activity testing 0.3–0.5 μg of phage λ DNA has been treated by 1–3 μl of the corresponding enzyme in Buffer R (10 mM Tris-HCl, pH 8.5, 10 mM $MgCl_2$, 100 mM KCl, 0.1 mg/ml BSA; Thermo Scientific) for RM.AloI and RM.PcoI proteins. RM.AloI and RM.PcoI reactions have been conducted at 30 °C during 0.5–1 h. Then hydrolysis patterns have been analyzed in 0.8% agarose gels. To prevent band retardation due to RM.AloI and RM.PcoI enzymes binding to DNA molecules, 0.1% SDS has been added to the loading buffer. To conduct RM.PcoI reactions in absence of magnesium ions the reaction mixtures have been reproduced missing $MgCl_2$ component.

3. Results and discussion

3.1. Phylogenetic analysis of new bifunctional restriction endonucleases

Fig. 1 shows a phylogenetic tree of new and previously described type IIC/IIS/IIG enzymes. As can be seen from the figure, there are three main clusters on the tree: AloI/HaeIV-related, CjeI-related and TspDTI-related enzymes. The first cluster includes 7 members from different genera, which implies horizontal gene transfer between distantly related and environmentally isolated species. The second consists of 10 representatives mainly of *Campylobacter* genus. The third one organizes 7 enzymes from *Thermus* genus, divided on two subfamilies, TspGWI-related and TspDTI-related. For them it could be suggested that fusion occurred in some common ancestor and then genes, coding the resulting bifunctional proteins, diverged and spread throughout the different microbial taxons. As can be seen from the Fig. 1, MmeI-related type IIS enzymes, both bifunctional (RM.BseRI and RM.BseMII) and nonbifunctional (R.FokI (Sanders et al., 2009; Kim et al., 1996; Catto et al., 2008) and R.BsmI), have separated from AloI/HaeIV-related and CjeI-related branches earlier in the evolution history. Based on the phylogenetic analysis and similarity with well characterized type IIC/IIS/IIG REases, biochemical properties of the new bifunctional enzymes can be predicted. It should be noted that none of these enzymes was annotated as a bifunctional protein and our phylogenetic analysis could help to establish cellular functions for these putative GenBank entries.

The tree represents 22 new bifunctional REases that we have previously succeeded to filter from NCBI database (Mokrishcheva et al., 2012). According to their domain organization, 20 of them represent the fusion of a REase with a MTase and a target recognition subunits of the type I restriction-modification systems (R-M-S structure). This pattern is common for the known type IIC bifunctional enzymes such as AloI (Cesnaviciene et al., 2001), CjeI (Vitor and Morgan, 1995), MmeI (Morgan et al., 2008), PpiI (Jurenaite-Urbanaviciene et al., 2007), TstI (Jurenaite-Urbanaviciene et al., 2007), TspGWI (Zylicz-Stachula et al., 2009), TspDTI (Żylicz-Stachula et al., 2012), TsoI (Żylicz-Stachula et al., 2012), TthHB27I (Żylicz-Stachula et al., 2012) and Tth111II (Żylicz-Stachula et al., 2012). Type I RMS enzymes function as a single protein complex, quaternary structure of which consists of R, M and S subunits (Roberts et al., 2003). The S subunit defines their DNA sequence specificity. The R subunit catalyses cleavage (restriction) and the M subunit - the methylation reaction. Their organization into one polypeptide via covalent bonds is not thought to interfere with their catalytic activities, thus providing a base for successful fusion. Currently, the REBASE contains over 8000 entries of type I RMS. Hypothetically, any of them could be fused, naturally or artificially, giving a new bifunctional RMS.

Only two REases from the tree have an alternative domain organization. An RMS from *Bacteroides* sp. D22, belongs to type III enzymes, having conserved motifs similar to Eco57I protein (Rimseliene et al., 2003). In type III systems Mod and Res subunits interact together to get an active enzyme (Roberts et al., 2003). That is why potential in-frame fusion of their genes would not probably interfere with their activities, being also favorable for new bifunctional protein origin. An RMS from *Arthrospira maxima* CS-328 has probably originated from a fusion of head-to-tail oriented type II REase and MTase genes. Successful possibility of this joining via point mutations has been previously shown in our recent work on Eco29kI RMS (Mokrishcheva et al., 2011). In this study it was clearly demonstrated that an artificial fusion of separate Eco29kI REase and MTase genes can yield an active bifunctional enzyme, capable both to cut and to modify its target sequence with efficiencies just slightly different from native enzymes (Nikitin et al., 2003; Nikitin et al., 2007). By a similar mechanism, a new bifunctional RMS could appear from other head-to-tail oriented type II RMS such as AccI, BanI, Bsp6I, BsuBI, Cfr9I, DdeI, EagI, EcoPI, EcoP15, EcoRI, FnuDI, HaeIII, HgiBI, HgiCI, HgiCII, HgiDI, HgiEI, HgiGI, HhaII, HincII, HindIII, HinfI, HpaI, MboII, MwoI, NcoI, NdeI, NgoMI, NgoPII, NlaIII, PaeR7I, RsrI, SalI, Sau3A, Sau96I, TaqI, TthHB8I, XbaI, XmaI etc. (Wilson, 1991).

3.2. Potential bifunctional enzyme from Prevotella copri DSM-18205 purification and testing

To prove functional activity of some new RM enzymes, we have constructed overproducing strain for ZP_06253288.1 protein from *Prevotella copri* DSM-18205. Its schematic organization is shown on Fig. 2. This enzyme has been successfully expressed and purified in