Biochemical and Biophysical Research Communications 398 (2010) 254–259

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Fusion of GFP to the M.EcoKI DNA methyltransferase produces a new probe of Type I DNA restriction and modification enzymes

Kai Chen, Gareth A. Roberts, Augoustinos S. Stephanou, Laurie P. Cooper, John H. White, David T.F. Dryden *

School of Chemistry, University of Edinburgh, The King's Buildings, Edinburgh, EH9 3JJ, UK

ARTICLE INFO

Article history:
Received 12 June 2010
Available online 19 June 2010

Keywords:
DNA restriction/modification
DNA methyltransferase
Forster resonance energy transfer
Time-resolved fluorescence anisotropy
Time-resolved fluorescence
Green fluorescent protein

ABSTRACT

We describe the fusion of enhanced green fluorescent protein to the C-terminus of the HsdS DNA sequence-specificity subunit of the Type I DNA modification methyltransferase M.EcoKI. The fusion expresses well *in vivo* and assembles with the two HsdM modification subunits. The fusion protein functions as a sequence-specific DNA methyltransferase protecting DNA against digestion by the EcoKI restriction endonuclease. The purified enzyme shows Förster resonance energy transfer to fluorescently-labelled DNA duplexes containing the target sequence and to fluorescently-labelled ocr protein, a DNA mimic that binds to the M.EcoKI enzyme. Distances determined from the energy transfer experiments corroborate the structural model of M.EcoKI.

1. Introduction

Since their introduction into genetic engineering, the green fluorescent protein (GFP) and its many spectral variants have proved to be extraordinarily useful probes of protein structure and function both *in vitro* and *in vivo* [1]. In particular, Förster resonance energy transfer (FRET) to measure distances between two fluorophores, a donor and an acceptor, has been the subject of many uses of GFP despite its complex photophysics and its relatively large size compared to more traditional small molecule fluorophores such as fluorescein [2].

Sequence-specific DNA-binding enzymes such as methyltransferases (MTases) and endonucleases comprising bacterial restriction–modification (R/M) systems would seem to present excellent targets for analysis via fusion to GFP given that many of them introduce complex rearrangements of DNA structure including for example DNA looping to bring distant sites on a single DNA molecule into close proximity. However, as yet few investigations of R/M systems have utilised these versatile fluorescent probes [3].

Bacterial host restriction endonucleases (REase) attack invading foreign DNA lacking the imprinted modification pattern characteristic of the host DNA [4]. R/M systems typically comprise a REase that recognises a specific nucleotide sequence prior to cleavage, and a cognate DNA MTase able, by methylating adenine or cytosine within the same sequence, to confer protection from the REase. The REase cuts unmethylated DNA but not hemimethylated DNA, the substrate for the MTase. R/M systems are classified according to their subunit composition, recognition site, cofactor requirement and DNA cleavage position. The R/M systems display an extraordinary diversity in structure and activity leading to four distinct groupings [5]. The most common R/M systems are the Type II R/M systems, which primarily consist of separate MTase and REase enzymes that recognise 4–8 base pair (bp) palindromic sequences.

In contrast, Type I R/M enzymes [4] such as EcoKI are complex hetero-oligomers of two REase (HsdR) subunits, two MTase (HsdM) subunits and one DNA sequence-specificity (HsdS) subunit. Depending on the methylation status of the DNA substrate, this complex functions as either a REase or an MTase. These enzymes recognise an asymmetric, bipartite sequence (13–15 bp) and require ATP to affect cleavage at a distant site. Over 600 confirmed and putative Type I R/M systems are known and they appear to be as widely spread in bacteria as the Type II R/M systems [6]. The complex of 2 HsdM and 1 HsdS, M_2S_1, forms an active MTase, M.EcoKI, and is the core part of the Type I R/M enzyme. The M.EcoKI MTase recognises the sequence AACNNNNNNGTGC and the methylation status of the adenines at the underline locations. A detailed structural model of M.EcoKI in complex with DNA has recently been proposed based upon electron microscopy of the complex and crystallographic structures of the individual subunits [7].

The genes for R/M systems are found in virtually every sequenced bacterial and archaeal genome and many genomes contain multiple R/M systems [6] often with the capability to switch between different systems and DNA specificities depending upon conditions [8,9]. R/M systems are also extensively represented

* Corresponding author. Fax: +44 (0)131 650 6453.
E-mail address: david.dryden@ed.ac.uk (D.T.F. Dryden).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.06.069

within clinical strain collections such as the *Escherichia coli* ECOR collection [10]. Given that resident R/M systems limit phage propagation in a bacterial population by factors reaching 10^8 (for EcoKI), there is a huge evolutionary pressure on mobile genetic elements such as phage and conjugative plasmids and transposons to evolve 'anti-restriction' counter measures including, for example, the acquisition of proteins which inhibit DNA binding by the R/M enzymes [4]. These inhibitors are structural and electrostatic mimics of double stranded DNA with the gene 0.3 protein, ocr, from phage T7 and the ArdA protein from conjugative Tn*916* mimicking 24 base pairs and 42 base pairs, respectively [11,12]. Their tight binding to M.EcoKI physically fills the DNA binding groove on the enzyme resulting in the inactivation of the R/M system [13–17].

In this paper we demonstrate the preparation of an active M.EcoKI fused to GFP and measure via FRET the distance from the GFP to a HEX label on a duplex bound to the MTase and to a fluorescently-labelled ocr protein bound to the MTase. These distances are then compared to predictions from the structural model [7].

2. Materials and methods

2.1. Plasmid pJFMSEGFP for production of GFP-MTase

The expression construct is derived from pJFMS [18] and pEGFP-N1 (Clonetech) as detailed in supplementary information. This plasmid was named pJFMSEGFP and we call the protein GFP-MTase.

2.2. In vivo activity

pJFMSEGFP and control plasmid pBIO2 were introduced into the r^-m^- mutant, *E. coli* NM1261(DE3). This strain contains a mutation in *hsdS*. pBIO2 is a non-functional derivative of pJFMS lacking the entire *hsdM* and half of *hsdS*. In NM1261(DE3), function of the MTase was dependent upon the plasmid-encoded HsdS forming a complex with HsdM encoded on the chromosome and the plasmid. Bacteriophage lambda virulent containing unmodified EcoKI sites:$\lambda_{v.o}$) was plated on NM1261(DE3) pJFMSEGFP and plaques were picked for assay against the EcoKI tester stains *E. coli* NM1049(DE3) (r^+m^+) and NM1261(DE3). Serial dilutions of plaques resuspended in phage buffer were spotted in 10 μl aliquots on the tester strains plated on BBL top agar supplemented with carbenicillin, 100 μg/ml. Titres were scored after overnight incubation at 37 °C [19]. Note that heterologous gene expression was not induced by addition of IPTG in these experiments but instead relied upon leaky expression from the promoter.

2.3. Purification of GFP-MTase

GFP-MTase was purified to homogeneity after overexpression in *E. coli* BL21(DE3) cells using His-tag affinity, gel filtration and anion exchange chromatography as detailed in supplementary information. The protein occurred in both M_1S_1 and M_2S_1 forms as found for the native protein [18] with the M_2S_1 form being used in further experiments. All subsequent measurements were performed at 20 or 25 °C in 20 mM Tris–HCl pH 8.0, 6 mM $MgCl_2$, 7 mM 2-mercaptoethanol supplemented with NaCl when stated.

2.4. DNA binding activity in vitro

DNA binding was measured using FRET and employed 21 base pair duplexes labelled at their 5′ ends with hexachlorofluorescein (HEX). The interaction of these duplexes with M.EcoKI has been previously analysed using fluorescence anisotropy [15]. Two duplexes were used: 21TH21B has the top "21TH" strand sequence 5′-HEX-GCC TAA CCA CGT GGT GCG TAC-3′ with the complementary unlabelled bottom strand ("21B") and 21T21BH has the same sequence but the HEX label is on the 5′ end of the bottom "21BH" strand.

A range of solutions containing GFP-MTase from 0 to 200 nM and NaCl concentrations of 0, 25, 50 and 100 mM, were prepared. In addition, solutions containing different proportions of GFP-MTase and 21TH21B, where the sum of the concentration of the two components was 200 nM, were prepared. The emission spectrum of each solution was then recorded and the intensity of the emission peak plotted against the mole fraction of GFP-MTase after subtracting the intensity of the GFP-MTase alone. The binding affinity was determined using the continuous variation method [20].

2.5. Preparation of ocr mutant proteins and their interaction with GFP-MTase

Site directed mutagenesis and protein purification was performed as described previously [17] to create the single substitutions, E20C, S68C and E117C in the ocr protein. 1 ml samples of 10 μM of each mutant ocr protein (assuming an ocr dimer) were incubated overnight at 4 °C in the dark with a 20-fold molar excess of Dylight549 Maleimide (Molecular Probes) in 100 mM sodium phosphate buffer, 150 mM NaCl, 1 mM EDTA, pH 7.2. Unreacted probe was removed by extensive dialysis. The concentration of Dylight549 bound to the ocr dimer was calculated from absorption using a molar extinction coefficient of 150 000 $M^{-1}\,cm^{-1}$ at 562 nm. The concentrations of all ocr mutant proteins were calculated using a molar extinction coefficient of 31,860 $M^{-1}\,cm^{-1}$ at 280 nm for the ocr dimer [13]. The concentrations of the labelled proteins were calculated from the absorption spectra at 280 nm after subtracting the Dylight549 absorbance at this wavelength (12,150 $M^{-1}\,cm^{-1}$). A comparison of the concentration of Dylight549 with the concentration of ocr then allowed the degree of labelling to be calculated. Labeling levels of 81.5%, 77.0% and 86.3% were achieved for E20C, S68C and E117C mutant ocr proteins, respectively.

Binding of the labelled mutant ocr proteins to the GFP-MTase was assessed using size exclusion chromatography as previously described [14].

2.6. Fluorescence measurements

Steady state fluorescence intensity measurements were performed on an Edinburgh Instruments FS900 spectrofluorometer (Edinburgh Instruments) with a 5 nm bandwidth. The cuvette path lengths were 3 mm.

Time correlated single photon counting was performed with a home built time-resolved fluorimeter equipped with an Edinburgh Instruments TCC900 single photon counting card, 465 nm or 500 nm pulsed LED driven by a PDL 800-B pulsed diode laser driver (PicoQuant Gmbh) and a PMH-100-3 single photon counting photomultiplier tube (Becker & Hickl Gmbh). A 405 nm pulsed laser (Edinburgh Instruments) was also sometimes used. Emission wavelengths were selected with a monochromator. Polarisation was applied using quartz Glan–Thompson polarisers. Excitation pathlengths were 10 mm and the emission bandpass was 20 nm. Fluorescence decays were fitted using a multiexponential decay equation with the minimum number of decay components required to obtain a χ^2 value close to 1. Anisotropy decays were fitted to Eq. (1).