Analytical Biochemistry 384 (2009) 136–144

Contents lists available at ScienceDirect

Analytical Biochemistry

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

Site-specific labeling of T7 DNA polymerase with a conformationally sensitive fluorophore and its use in detecting single-nucleotide polymorphisms

Yu-Chih Tsai, Zhinan Jin, Kenneth A. Johnson *

Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, University of Texas, Austin, TX 78712, USA

ARTICLE INFO

Article history:
Received 14 July 2008
Available online 14 September 2008

Keywords:
DNA polymerase
Fluorescence labeling
Fidelity
Conformational changes
Induced fit
Single-nucleotide polymorphisms

ABSTRACT

Like most enzymes, DNA polymerases undergo a large conformational change on the binding of a correct nucleotide. To determine how the conformational change contributes to substrate specificity, we labeled the T7 DNA polymerase with a conformationally sensitive fluorophore at a position that provides a signal coincident with structural changes following nucleotide binding and distinguishes correct base pairs from incorrect ones by the sign of the fluorescence change. Here we describe methods to document that only one site on the polymerase was labeled with the fluorophore based on mass spectral analysis of tryptic peptides. In addition, we show by equilibrium titrations of opposing signals that mismatches and correct bases compete for the same site. This analysis forms an essential basis for characterization of a fluorescently labeled enzyme intended for mechanistic studies. Finally, we show that the labeled enzyme can be used to identify single-nucleotide mutations in a procedure that could be automated.

DNA polymerases successively incorporate nucleotides according to the sequence of the template and can discriminate against mismatched nucleotides with high accuracy [1,2]. Early kinetic studies using T7 DNA polymerase, *Pol*I Klenow fragment, and HIV-RT each suggested a two-step nucleotide binding and recognition sequence with a conformational change step preceding chemistry [3–8]. Structural studies of these enzymes later revealed a large nucleotide-induced conformational change observed in a ternary enzyme–DNA–dNTP complex [9–14]. However, structural data do not define the role that the conformational change plays in nucleotide selectivity, which is a purely kinetic phenomenon dependent on k_{cat}/K_m. To understand how polymerases achieve such extraordinary specificity and efficiency of catalysis, we and others have sought a fluorescence signal to monitor the kinetics of the nucleotide-induced conformational change [15–20].

Different methods have been developed using fluorescent signals to resolve the kinetics of the polymerase conformational changes. One method relies on labeling the DNA template with a 2-aminopurine (2-AP)[1] at the T + 1 position adjacent (5′) to the templating base. Nucleotide-induced fluorescence changes of the *Escherichia coli Pol*I Klenow fragment, DNA polymerase β, and T4 DNA polymerase were studied using this approach [16–18], leading to the conclusion that the observed change was faster than chemistry. The fluorescence signal is attributed to the unstacking of the 2-AP at the unique T + 1 position where the template undergoes an unusual bend as it enters the polymerase active site. Accordingly, the signal reports changes in DNA structure that could arise during translocation of the DNA or could be correlated with nucleotide-induced changes in protein structure. Another method used for studying the conformational change of the Klenow fragments of Taq polymerase and *Pol*I was based on Förster resonance energy transfer (FRET) between two fluorophores: one on the fingers domain and another on the DNA substrate [19,20]. Again, placing one fluorophore on the DNA introduces ambiguity in the interpretation of results in that the signal could be due to changes in protein structure or to movement in the DNA, a problem that is compounded by the possible partitioning of the DNA between polymerase and exonuclease sites.

Because of ambiguities in the interpretation of signals arising from labels positioned on the DNA, we sought a different method that relied on the placement of a single, environmentally sensitive fluorophore at a position that could sense a change in structure without perturbing the reaction. We previously reported the construction of an E514C mutant of T7 DNA polymerase with eight native cysteines replaced with either alanine or serine (the E514C-8C mutant). This enzyme was then labeled on the nucleotide recognition domain with the fluorophore MDCC (7-diethylamino-3-[([(2-maleimidyl)ethyl]amino)carbonyl]coumarin), and the effects of

* Corresponding author. Fax: +1 512 471 0435.
E-mail address: kajohnson@mail.utexas.edu (K.A. Johnson).

[1] *Abbreviations used:* 2-AP, 2-aminopurine; FRET, Förster resonance energy transfer; MDCC, 7-diethylamino-3-[([(2-maleimidyl)ethyl]amino)carbonyl]coumarin; MS, mass spectrometry; SNP, single-nucleotide polymorphism; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; TFA, trifluoroacetic acid; ddH_2O, doubly distilled H_2O; HPLC, high-performance liquid chromatography; UV, ultraviolet; MALDI, matrix-assisted laser desorption/ionization; TOF/TOF, tandem time-of-flight; CID, collision-induced dissociation; MS/MS, tandem MS; ssDNA, single-stranded DNA.

0003-2697/$ - see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.ab.2008.09.006

mutagenesis and labeling on the enzymatic activity were quantified [15]. We showed that the binding of a correct nucleotide (which forms canonical base pairs with the templating base) induced a decrease in fluorescence, whereas the binding of a mismatched nucleotide caused an increase in fluorescence, and we used that signal to monitor the kinetics of the nucleotide-induced conformational change. Here we describe the rationale for choice of fluorophore and selection of the site for labeling, and we report mass spectrometry (MS) analysis of tryptic peptides to show that only a single site was labeled. In addition, we describe a quantitative equilibrium competition assay to demonstrate that both correct and incorrect bases compete for the same site while inducing opposite changes in fluorescence. These methods constitute a minimal set of assays necessary to document the fluorescently labeled protein, and we propose that these assays should be applied as part of the analysis of any fluorescently labeled enzyme. Finally, we explore different DNA sequence contexts and demonstrate that the fluorescence signal can be the basis for a simple assay to detect single-nucleotide polymorphisms (SNPs).

Materials and methods

Construction, purification, and labeling of MDCC–E514C-8C T7 DNA polymerase

The plasmid encoding an exonuclease-deficient (exo$^-$) mutant of T7 DNA polymerase, pG5X, was used for constructing a Cys-light enzyme [21] by removing 8 of the 10 native cysteine residues and adding the mutation E514C to provide a surface-exposed cysteine for site-specific labeling with MDCC as described previously [15]. Thioredoxin was purified and used to reconstitute the native polymerase holoenzyme as described previously [6]. Purified T7 gp5 and *E. coli* thioredoxin were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and showed a 95% or higher purity by Coomassie blue staining. The enzyme concentration was determined at 280 nm with a molar extinction coefficient of 134,420 M^{-1} cm^{-1} calculated using the amino acid sequence [22]. The concentration of MDCC on the protein was measured at 419 nm in methanol with the molar extinction coefficient of 50,000 M^{-1} cm^{-1} provided by the manufacturer (Invitrogen). The MDCC labeling efficiency was estimated to be approximately 90 to 95% based on these absorbance measurements.

Tryptic digestion of MDCC–E514C-8C exo$^-$T7 DNA polymerase

The MDCC labeled protein, stored in the final dialysis buffer (40 mM Tris–HCl [pH 7.5], 0.1 mM ethylenediaminetetraacetic acid [EDTA], 50 mM NaCl, 50% glycerol, and 1 mM dithiothreitol [DTT]), was mixed with trypsin (sequencing-grade modified trypsin, Promega) at a ratio of 20:1 (w/w) and then was incubated overnight at 37 °C. The tryptic peptides were separated by reverse phase HPLC.

Reverse phase HPLC

A POROS R2 perfusion column (PerSeptive Biosystems) was used to separate the tryptic peptides. The column was equilibrated with buffer A (0.1% trifluoroacetic acid [TFA], 2% acetonitrile, and doubly distilled H_2O [ddH_2O]). Then 100 μl of sample was loaded onto the column using an ÄKTA high-performance liquid chromatography (HPLC) instrument (Amersham Pharmacia Biotech). The column was washed with buffer A, and peptides were eluted with an acetonitrile gradient (buffer A to buffer B [0.08% TFA, 80% acetonitrile, and ddH_2O] at a 1.2-ml/min flow rate). The eluted peptide peaks were monitored by 220 nm ultraviolet (UV) absorption, and the presence of MDCC was monitored by 425 nm absorption. The fractions corresponding to the absorbance peaks at 425 nm were collected for MS analysis.

MS and tandem MS analysis

The collected samples from HPLC were frozen in liquid nitrogen and dried with a Savant SpeedVac concentrator (Forma Scientific) and then dissolved in 10 μl of a solution containing 50% acetonitrile, 50% H_2O, and 1% TFA. The matrix solution was made of α-cyano-4-hydroxycinnamic acid supersaturated in a solution of 70% acetonitrile, 30% H_2O, 0.1% TFA, and 5 mM $(NH_4)_2HPO_4$. The dissolved samples were mixed with the matrix solution at a 1:1 ratio, and 0.5 μl of the mixture was spotted onto a matrix-assisted laser desorption/ionization (MALDI) stainless-steel target. The mass spectra were obtained by an ABI 4700 Proteomics analyzer MALDI tandem time-of-flight (TOF/TOF) instrument (Applied Biosystems). To verify the identities of the ions in the mass spectra, the high-energy collision-induced dissociation (CID) was used to fragment selected ions, producing tandem MS (MS/MS) spectra for the derivation of peptide sequences.

Fluorescence emission profile of MDCC–E514C-8C T7 DNA polymerase at different substrate-bound states

DNA duplexes formed with a 27mer primer (5′-GCC TCG CAG CCG TCC AAC CAA CTC AACdd-3′) and 45mer templates (5′-GGA CGG CAT TGG ATC GA**N** GTT GAG TTG GTT GGA CGG CTG CGA GGC-3′) with varying bases at position 18 (**N**) were custom synthesized by IDT and used in the nucleotide binding assays. The enzyme–DNA complex was formed using 200 nM enzyme, 300 nM DNA, 4 μM thioredoxin, and 12.5 mM $MgCl_2$ in the T7 reaction buffer [15]. The fluorescence emission intensity was recorded by exciting the enzyme–DNA complex at 425 nm and monitoring the fluorescence intensities at 460 nm before and after the addition of 1 mM dNTP using a fluorometer from Photon Technology International. No correction for inner filter effects was necessary at these wavelengths.

Equilibrium titration experiments

A solution containing 200 nM MDCC–E514C-8C T7 DNA polymerase in the T7 reaction buffer and 12.5 mM $MgCl_2$ was preincubated in the presence of 300 nM 27ddC/45-18G DNA duplex [15]. Solutions containing nucleotides and equal concentration of $MgCl_2$ were used to titrate the enzyme–DNA complex using a KinTek TMX titration module (http://www.kintek-corp.com). Fluorescence intensities at equilibrium were monitored continuously, while a solution of nucleotide was added at a rate of 4 μl/min, and were corrected for the small dilution. The wavelength of excitation was set at 425 nm, and a 450-nm bandpass filter was used for emission detection. The overall dissociation constant at equilibrium state for nucleotide binding was determined by nonlinear regression to either a quadratic equation or a hyperbolic equation shown below, where A is the amplitude of fluorescence change and $[E_0]$ is the enzyme concentration:

Quadratic equation:

$$Y = A \times \frac{(K_d + [E_0] + [dNTP]) - \sqrt{(K_d + [E_0] + [dNTP])^2 - 4[E_0][dNTP]}}{2[E_0]} + C$$