Biotechnology and Industrial Microbiology

Cloning and high-level expression of Thermus thermophilus RecA in E. coli: purification and novel use in HBV diagnostics

Sudarson Sundarajan, Sneha Rao, Sriram Padmanabhan*
Cancyte Technologies Pvt. Ltd., Rangadore Memorial Hospital, Sri Shankara Research Center, Shankarapuram, Bangalore, India

A R T I C L E   I N F O
Article history:
Received 4 October 2017
Accepted 13 March 2018
Available online 12 April 2018
Associate Editor: Flávio Da Fonseca

Keywords:
RecA
Hepatitis B virus
DNA virus
Thermus thermophilus

A B S T R A C T
We studied the role of Thermus thermophilus RecA (RecA) in enhancing the PCR signals of DNA viruses such as Hepatitis B virus (HBV). The RecA gene of a thermophilic eubacterial strain, T. thermophilus, was cloned and hyperexpressed in Escherichia coli. The recombinant RecA protein was purified using a single heat treatment step without the use of any chromatography steps, and the purified protein (>95%) was found to be active. The purified RecA could enhance the polymerase chain reaction (PCR) signals of HBV and improve the detection limit of the HBV diagnosis by real time PCR. The yield of recombinant RecA was ~35 mg/L, the highest yield reported for a recombinant RecA to date. RecA can be successfully employed to enhance detection sensitivity for the diagnosis of DNA viruses such as HBV, and this methodology could be particularly useful for clinical samples with HBV viral loads of less than 10 IU/mL, which is interesting and novel.© 2018 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

The recombinase A (RecA) protein (accession no. AAK15321.1) from the thermophilic bacterium Thermus thermophilus is a thermostable enzyme that plays important roles in homologous recombination and DNA repair.1 DNA recombination in prokaryotes is achieved by catalyzing the pairing of homologous DNA sequences,2,3 and DNA repair is achieved by protection of the damaged DNA ends.4,5 This protein has the activities of single-stranded DNA dependent ATPase, DNA annealing, and exchanging of strands between two recombining DNA double helices, similar to Escherichia coli RecA (accession no. AFU91764.1) protein, but the optimal temperature for the strand exchange of RecA is between 50 and 60 °C.6

RecA-like proteins have been sequenced from over 60 different prokaryotic species7,9 and it hydrolyzes ATP when bound to DNA, and its ATPase activity rate is proportional to
the amount of RecA bound to the DNA with the fact that the ATPase activity is used to measure its binding efficiency to DNA. It has been shown that the RecA protein catalyzes in vitro DNA strand exchange reactions with peak efficiency at approximately 60 °C, a temperature that is fairly close to the Tm required for thermal denaturation of dsDNA. Numerous applications of RecA are reported. These include identification of different Burkholderia cepacia genomovars, sequence-specific ligation of DNA using the E. coli RecA protein, maintaining the integrity of chloroplast DNA molecules in Arabidopsis, auto-cleavage of other serine proteases and repressors of the bacteriophage lytic cycle, and its role in swarming motility in E. coli.

Shigemori et al. have used RecA in stimulating pairing between completely matched primers and target sequences and effectively eliminating non-specific PCR products. This specificity enabled them to amplify the sequences between any sites of interest without needing to consider the optimal priming sites for each PCR. Similarly, Fukui and Kuramitsu have recently utilized MutS, a thermostable mismatch-recognizing protein from T. thermophilus, along with RecA to reduce non-specific amplifications in PCRs.

Hence, we realized the immense potential of RecA in PCR-based diagnostics and attempted to generate RecA protein by a simple purification method that would in turn make the PCR based detection of viruses more sensitive and cost-effective.

Material and methods

Reagents

All chemicals were of analytical grade. Oligonucleotides, used in this study, were from BioServe Technologies Pvt. Ltd, Hyderabad, India. Bradford protein assay reagents and IPTG were obtained from Amresco, USA. Standard RecA pure protein was procured from NEB, USA, while ATP was purchased from Thermo Fisher Scientific, USA. Protein molecular weight markers 11–180 kDa and 11–97 kDa were obtained from G-Biosciences, USA and Aristogene Life Sciences, Bangalore, India respectively.

Cloning and construction of RecA expression construct

The synthetic gene of T. thermophilus RecA was custom synthesized from Integrated DNA Technologies, Inc. (IDT), USA, and cloned into the pUC57 vector. The nucleotide sequence of the T. thermophilus RecA gene was adjusted for codon bias without altering any of the amino acids. The RecA gene was excised from the pUC57-RecA plasmid as a NdeI/HindIII fragment and cloned into the pET26b vector into similar restriction sites. The recombinants were screened by PCR using T7 promoter specific primers, forward 5′ TAATACGACTCACTATAGGG 3′ and reverse 5′ GCTAGTTATGCTCAGGG 3′, and the positive clones were further confirmed by restriction analysis and DNA sequencing. The pET26b-RecA plasmid was designated as pCAN008.

The plasmid DNA of pCAN008 was introduced into ER2566 competent cells (NEB, USA) for expression of recombinant RecA. The protocol was essentially as described by Paul et al. Briefly, the induction of RecA was effected by the addition of 1 mM IPTG when cells in LB medium (containing 20 μg/mL kanamycin) reached an OD600 value between 0.5 and 0.8. The induction proceeded for 4 h by incubating the culture at 37 °C, and then the induced cells were pelleted at 7670 × g for 10 min. The harvested cells were resuspended in 25 mM Tris, Cl, pH 7.5 and sonicated in a sonicator (Omni Sonic Ruptor 400, GA, USA) at 50% amplitude for one min on/off for 10 min in cold. The sonicated lysate was spun at 9000 × g for 20 min at 4 °C and the supernatant obtained was used for purification of the RecA expressed protein.

RecA purification by heat treatment

RecA is known to be thermostable, and we utilized this property to purify the recombinant RecA protein. Different volumes of crude lysate of RecA were heated at 75 °C for 30 min, after which the samples were centrifuged at 30,670 × g for 10 min, and the supernatants were analyzed on 12% SDS–PAGE for RecA protein followed by silver staining as described by Nesterenko et al. Protein concentrations were determined by Bradford’s method to quantify the total yield of RecA per liter of induced culture under the described experimental conditions.

ATP hydrolysis assay

ATPase activity was determined using a QuantiChrom ATPase/GTPase Assay Kit DATG-200 (EnzyChrom, USA). Phosphate standards were made by taking varying concentrations of phosphate (0–50 μM) in nuclease free water as per manufacturer’s instructions. Equal amounts of proteins of in-house purified RecA and commercial RecA (New England Biolabs, UK) were used for detecting the ATPase activity. After incubating the RecA proteins with ATP, 200 μL of the supplied reagent was added, and contents were incubated for 30 min at room temperature. The OD of the samples, reflecting free phosphate release due to ATP hydrolysis by RecA, was read at OD630nm in an ELISA plate reader (Tecan Group Ltd, Männedorf, Switzerland).

Human leucocyte antigen (HLA) PCR

Human gDNA is the source of the template for the amplification of HLA genes. The different protocols for HLA typing reported require use of gDNA between 5 and 30 ng/μL and DNA concentrations lower than 2 ng/μL have been found to compromise genotyping results severely. Since RecA is known to prevent false priming, we hypothesized that samples with lower concentrations of genomic DNA might display higher gene amplifications in presence of RecA. To validate this hypothesis, we carried out PCRs of HLA genes from human gDNA in presence and absence of RecA. The primers for amplification of HLA class I genes, namely, HLA-A-exon 2 and exon 3, HLA-B-exon 2 and exon 3 and HLA-C-exon 2 and 3, were as described by Itoh et al., while the primers for HLA class II genes such as DPB1, DQB1 and DRB1, targeting exon 2, were according to Lange et al. The PCR reaction was carried out in 25 μL reaction volume and the PCR mixture comprised 1 x ammonium sulfate buffer, pH 8.3;
2.5 mM magnesium chloride; 0.2 mM dNTPs; 5% DMSO; 0.2% Tween-20; 0.04% BSA; 0.4 µM HLA primers; and 1 unit of Taq DNA polymerase (ThermoFisher Scientific, USA) with 40 ng of gDNA. PCR conditions were as follows: Initial denaturation of 95 °C for 5 min, followed by 35 cycles of 95 °C, 30 s; 51.1 °C, 30 s; and 72 °C, 45 s. The final extension step was performed at 72 °C for 10 min, and suitable aliquots were loaded on a 2% agarose gel for visualization.

Sample collection and HIV, HBV and HCV PCR

Patients who visited Kuppam Medical College, Andhra Pradesh, India for viral testing of HIV, HBV and HCV during the period of March 2016 to March 2017 served as the study material. Blood samples were collected in EDTA coated tubes and plasma was collected by centrifuging the samples at 2000 × g for 10 min. The plasma samples were stored at −20 °C until further use. In some cases, serum was collected from the patients.

Plasma or serum samples were analyzed for HIV, HBV and HCV by routine diagnostic tests (RTD) as per manufacturer's instructions. The kits employed included the HIV-1 TRI-DOT kit, HEPACARD test kit and HCV TRI-DOT test kits. Viral positive samples were handled with adequate precautions in the BSL-II facility of Canyect Technologies Pvt. Ltd, Bangalore, India. IVD certified Nucleospin Dx virus kit (Macherey-Nagel, GmbH) was used to isolate viral nucleic acids following manufacturer's instructions.

To minimize the risk of contamination of PCR amplicons, and to avoid transfer of DNA from one sample to the other, in our laboratory, the pre and post-polymerase chain reaction (PCR) products were kept physically separated, and all biological samples were processed separately in a dedicated BSL-II biosafety hood. The DNA-free consumables and negative controls were used at every stage of the PCR analysis.

The HIV primers were designed to amplify the gag gene of the HIV (accession number KY713247.1), while the HBV and the HCV primers amplified the S gene (accession number MF967563.1) of HBV and 5′ UTR (accession number MG436884.1) region of HCV respectively. All the PCR s were carried out in 25 µL reaction volumes and the PCR products were analyzed on 2% agarose gel electrophoresis stained with ethidium bromide for visualization, unless mentioned otherwise.

The PCR mixture comprised 1× Superscript III RT/Platinum® Taq Mix buffer, HIV forward primer (5′ ATACAAGGACGCAATGCAAAT 3′) and reverse primer (5′ TACTAGTATTCTCGTATGT 3′) or HCV forward (5′ GAA AGC GTC TAG CCA TGG CG 3′) and reverse primer (5′ ACG CCC AAA TGG CCG GCC ATA GA 3′), and 0.5 µL of Super Script III enzyme. PCR was performed in a Proflex PCR machine (Applied Biosystems, USA). PCR conditions were as follows: one cycle of 50 °C for 30 min to convert RNA to cDNA, followed by a single step of denaturation at 95 °C for 5 min and 40 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min followed by a final extension at 72 °C for 10 min.

PCR for HBV detection in clinical samples was carried out in 25 µL reaction volumes. The PCR mixture comprised of 1× ammonium sulfate buffer, pH 8.3; 2.5 mM magnesium chloride; 0.2 mM dNTPs; 5% DMSO; primers 0.4 µM (FP 5′ CCCCCCAGCTGGGGCCGTGT 3′ and RP 5′ AGGAGCTCCGCCCAAGGTAC 3′) and 1 unit of Taq DNA polymerase (ThermoFisher Scientific, USA). PCR steps included an initial denaturation of 95 °C, 10 min, followed by 40 cycles of 95 °C, 30 s; 60 °C, 30 s; and 72 °C, 45 s and a final extension step at 72 °C for 10 min.

For carrying out PCRs with RecA and ATP, ATP was used at a final concentration of 0.3 mM along with 0.25 µg of either in-house purified RecA or commercially available RecA.

HBV viral load

Briefly, the reaction was carried out in a final volume of 25 µL containing 1× Thermo Scientific Maxima SYBR Green master mix, and forward (FP 5′ CCCCCACTGGGCTGGGCCTTGT 3′ and RP 5′ AGGAGCTCCGCCCAAGGTAC 3′) primers were added to a final concentration of 4 pmol per reaction. A total of 0.3 mM ATP and 0.25 µg of RecA were added to one set, and suitable samples without RecA served as negative controls.

The qPCR was performed in a Qiagen qPCR machine (Qiagen, Germany), and the conditions for qPCR were 45 cycles of PCR amplification, each comprising 95 °C for 5 min, 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s. Data collection was performed at 72 °C in each cycle. Following the amplification, a melting curve analysis was performed to verify the authenticity of the amplified products by checking their specific melting temperatures (Tm) and experiments were done in duplicates.

The recombinant plasmid containing the S gene of HBV (named pTZ57R/T-HBV) was confirmed by DNA sequencing (data not shown) and used to generate a standard curve. The concentration (ng/µL) of the purified stock of plasmid DNA was measured by a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific), and plasmid copy number was calculated using the DS DNA copy calculator using the formula: pTZ57R/T copies/µL = (plasmid concentration [ng/µL] × 6.022 × 10²³)/(nucleotide length × 1 × 10⁹ × 650), as described by Staroscik.³⁴ The plasmid stock solution was diluted with DNase/RNase-free distilled water, 10-fold serial dilutions containing 2.5 × 10⁸–250 copies/µL were made, and small aliquots were stored at −20 °C for future use in generating standard curves. The Ct values were tested for statistical significance by Student’s t-test.

Results

Cloning and expression studies

Fig. 1 shows the plasmid map of pET26b-RecA. The clones of pET26b-RecA were confirmed by DNA sequencing and restriction analysis. Fig. S1 shows the silver stained gel of whole cell lysates of two clones of pET26b-RecA. The recombinant RecA protein was restricted to the soluble fraction of the cell and constituted >50% of the total soluble protein as judged by densitometry scanning and of the expected size of ~39 kDa. The studies on the expression of RecA from pET26b-RecA clone in ER2566 cells using different concentrations of IPTG indicated that the amount of RecA expression increased with increased time of induction as expected (Fig. S2) and that the expression of RecA could be induced even with 0.125 mM IPTG effectively (Fig. S3).
Purification and total yield of RecA using heat treatment

The RecA expressed from the clone pET26b-RecA (#5) after heat treatment was found to be homogenous and was >95% pure as judged by silver staining. It is evident from Fig. S4 that heat treatment was sufficient to achieve pure RecA. The total yield of pure RecA per liter of LB was found to be 35 mg/L and its purity matched the purity of the commercially available RecA protein (Fig. 2).

ATP hydrolysis assay

The in-house purified RecA protein showed ATPase activity units of $3.0 \times 10^4$ U/$\mu$L, while the commercial RecA showed activity of $2.5 \times 10^3$ U/$\mu$L. This result indicated that the purification methodology followed for the in-house RecA protein does not compromise the function of the RecA protein.

HLA-PCR

It is clear from Fig. 3 that RecA amplifies the HLA genes of interest even when the gDNA used was as low as 5 ng per reaction, while without RecA, there was negligible amplification of all the genes under similar experimental conditions.

HIV, HBV and HCV PCR

Fig. 4 shows the PCR amplification of HIV/HBV and HCV target genes using three ELISA positive clinical samples. It is evident from Fig. 4 that RecA improved the PCR signal intensities of the HBV samples (Fig. 4, lanes 13, 14 and 15 in comparison to lanes 4, 5 and 6), while the PCR signals of target genes of RNA viruses such as HIV and HCV did not show any improvements, albeit...
there was marked reduction in levels of non-specific amplifications when RecA was included in the PCR reaction mix. Since HBV is a DNA virus, it is possible to speculate that RecA enhances the annealing of the HBV specific primers to the target S gene of the HBV, and due to the reported role of RecA in minimizing false priming, the intensity of the PCR signals for the HBV samples appear higher with RecA in comparison to the PCR of the same samples without RecA.

**HBV viral load**

Since our data showed that the PCR signals of HBV was higher in presence of RecA in comparison to that achieved without RecA, we wanted to examine if RecA would also have an effect on the HBV viral load values in the HBV positive plasma samples. From Fig. 5A, it is clear that the sensitivity of the HBV detection is enhanced and the overall HBV viral load values were higher (by 50–98%) for all the samples tested in this study when RecA was included in the PCR reaction mixture in comparison to the HBV viral load values obtained for the same samples without RecA (Table 1). Fig. 5B shows the standard curve prepared using plasmid pTZ57R/T-HBV from 250 to $2.5 \times 10^5$ copies/mL.

**Discussion**

The T7 promoter is known to be the strongest promoter known for the prokaryotic expression of proteins and in experimental E. coli cultures, consistently higher expression levels of the protein have been observed with the kanamycin marker than in cultures grown with the commonly used antibiotic ampicillin with the target protein being almost 50% of total cellular protein in some cases. It is possible to speculate that for the above reasons, we achieved high levels of RecA expression described in this study.

One could observe a small fraction of RecA protein induced even under uninduced conditions (Fig. 4, lane 1). Expression of RecA under uninduced conditions could be attributed to the leaky expression of the T7 promoter that is seen commonly in E. coli cells and can be easily controlled using pLysE plasmid, which expresses T7 lysozyme, an inhibitor of T7 RNA polymerase. Also, one could see appreciable amounts of RecA expressed even with the lowest concentration of IPTG (0.125 mM) which supports the recent report of high level expression of a leptospiral protein in E. coli using such low concentrations of IPTG.

A number of protocols have been reported for purification of wild-type RecA protein from E. coli or *Thermus aquaticus*. These methods include salt precipitation followed by hydrophobic interaction chromatography and anion
exchange column, precipitation using spermidine, purification using chitin affinity column, monoclonal antibody affinity columns, and phosphocellulose matrix. Many of these protocols require multiple steps for purification and are not very cost-effective. In an effort to develop a simple and a cost-effective protocol for RecA purification, we took advantage of the heat stable property of RecA and obtained highly pure RecA preparations by a simple heat treatment of the crude lysate without the use of any chromatographic steps and such a methodology for the purification of RecA is hitherto unreported.

A significant amount of endonuclease contamination from the host endA gene has been reported in RecA protein preparations by Singleton et al., which was removed by rigorous washing of the induced cells with detergents. The protocol described in this article utilizes an endA1 negative E. coli host, ER2566, for RecA expression, and hence, our purified RecA is virtually free of any endonuclease contamination, which is a huge advantage for its use in PCR-based diagnostics.

The RecA gene of T. thermophilus encodes a 36 kDa polypeptide whose amino acid sequence shows 61% identity with that of the E. coli RecA protein and T. thermophilus RecA is reported to be thermostable in comparison to E. coli RecA protein with similar functions. The yield of recombinant RecA (from E. coli or T. thermophilus) from various literature reports is variable and not feasible for commercial scale operations. While the protocol described by Griffith and Shores yield only 10 mg of RecA protein per liter, the protocol developed by Singleton et al. showed a yield of 28 mg of RecA/liter of E. coli culture. In this study, we have achieved 35 mg of pure RecA per liter of E. coli culture, which is the highest report of RecA yield till date. Additionally, the cost of 1 mg of commercial RecA is $575 and we are able to produce almost US$20,125 worth RecA from every liter of E. coli culture, which is interesting and useful.

Lactose can also be used as an inducer during the transition from the exponential to the stationary phase for protein induction for genes cloned under T7 promoter. Since lactose also acts as carbon source, the use of lactose for induction would prove useful for increasing the RecA expression per cell.

ATP hydrolysis serves as a means of releasing RecA promoters from the 5’ end of the filament once DNA strand exchange occurs, thus helps recycle the RecA protein. Additionally, it has been shown by Yu and Egelman that the RecA filament is inactive when the compressed filament is formed in the absence of a nucleotide cofactor such as ATP, either as a self-polymer or on a single-stranded DNA molecule. Hence, we hypothesized that ATP hydrolysis is a crucial step for the activity of RecA, and hence, the ATPase activity of RecA can serve as a measure for monitoring RecA strand exchange activity. We observed that the ATPase activity of the in-house RecA and commercial RecA preparations were substantially similar. This data indicates that the purified RecA does not have the property of aggregation, since the aggregation state of RecA protein is known to affect its ATPase activity. Shigemori et al. have performed RT-PCR for their target gene using a SYBR premix Ex Taq kit with and without RecA and observed a nearly 30% increase in PCR fluorescence signal intensity with RecA compared to that without RecA. Our observations of ~50 to 96% increase in the fluorescence intensity of the HBV PCRs with RecA supports the reported observations. In addition, we show lower Ct values in HBV qPCR with RecA, which is a new observation. Also, the limit of detection (LOD) of the HBV viral load method developed using SYBR green dye appears to be 5 virus copies/reaction which equates to 40 IU/mL. This LOD value is very close to the LOD value of the Abbott HBV viral load kit (utilizing TaqMan chemistry) of 15 IU/mL for a procedure processing 0.2 mL of the starting sample.

An assay with a lower limit of detection of 10^3 IU/mL of HBV is reported to be sufficient to monitor and manage the patient, but in some instances, a more sensitive assay with a lower limit of detection on the order of 10^2 IU/mL is recommended to ensure detection of the emergence of resistance as early as possible. Hence, our observations of enhancement of the limit of HBV viral load detection by almost 1 log copies can be extrapolated to improving the qPCR HBV LOD from 200 copies/mL to 20 copies/mL or 40 IU/mL to 4 IU/mL, which is hitherto unreported.

In this paper, we have described a real-time PCR assay for HBV DNA quantification with SYBR Green chemistry. The primers used for qPCR targets the surface antigen of HBV and have been designed after aligning nucleotide sequences of all the HBV genotypes available in the NCBI database so that one is assured of positive PCR signal for all HBV serotypes using these set of primers.

Viral load set point has long been used as a prognostic marker of disease progression and monitoring of virus load is known to predict the possible emergence of drug-resistant

### Table 1 - Hepatitis B virus (HBV) viral load values of enzyme-linked immune sorbent assay (ELISA) positive HBV samples with and without RecA.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Ct value without RecA</th>
<th>Cycle threshold (Ct) value with RecA</th>
<th>% increase in VL with RecA in comparison to VL without RecA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample # 1</td>
<td>38.515</td>
<td>35.605</td>
<td>96.66</td>
</tr>
<tr>
<td>Sample # 2</td>
<td>41.155</td>
<td>37.285</td>
<td>95.65</td>
</tr>
<tr>
<td>Sample # 3</td>
<td>15.41</td>
<td>13.65</td>
<td>86.29</td>
</tr>
<tr>
<td>Sample # 4</td>
<td>29.58</td>
<td>26.75</td>
<td>52.59</td>
</tr>
<tr>
<td>Sample # 5</td>
<td>34.185</td>
<td>31.035</td>
<td>59.28</td>
</tr>
<tr>
<td>Sample # 6</td>
<td>38.605</td>
<td>34.36</td>
<td>99.61</td>
</tr>
<tr>
<td>Sample # 7</td>
<td>41.175</td>
<td>37.67</td>
<td>95.12</td>
</tr>
</tbody>
</table>

* The Ct values were found to be statistically significant (p < 0.000001) by Student’s ‘t’ test.
HBV strains. Also, an undetectable or low viral load does not necessarily guarantee that a patient will not experience liver damage; hence methods that enhance the sensitivity of the viral load method would have immense value in clinical setting and further drug treatments. In the light of this, our observation of RecA enhancing the LOD of the viral load estimation method is novel and interesting.

We are presently focusing on generating novel RecA fusions to examine whether the heat stability of RecA can be imparted to heterologous genes and whether such heterologous RecA fusion proteins could be purified by heat treatment as described in this article.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors wish to thank Dr. C. Nagaraj, Prof. of Microbiology, PESIMSR, Kuppam, India, for the supply of clinical samples of HIV, HBV and HCV for our studies. The authors also wish to thank Dr. Sriddhar, K.N. and Dr. Gautham Nadig for their constant support and encouragement. Canyce Technologies Pvt. Ltd. is fully supported by Sri Sringeri Sharada Peetam, India.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bjm.2018.03.007.

REFERENCES

35. Tegel H, Ottosson J, Hober S. Enhancing the protein production levels in Escherichia coli with a strong promoter. FEBS J. 2011;278:729–739.