J. Microbiol. Biotechnol. (2010), **20**(10), 1463–1470
doi: 10.4014/jmb.1004.04035
First published online 11 August 2010

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Quantitative Detection of Residual *E. coli* Host Cell DNA by Real-Time PCR

Lee, Dong Hyuck[1], Jung Eun Bae[1], Jung Hee Lee[1], Jeong Sup Shin[2,3], and In Seop Kim[1]*

[1]*Department of Biological Sciences, Hannam University, Daejeon 305-811, Korea*
[2]*Quality Control Unit, Green Cross Corp., Chungbuk 363-883, Korea*
[3]*Department of Molecular Science and Technology, Ajou University, Suwon 443-749, Korea*

Received: April 22, 2010 / Revised: July 8, 2010 / Accepted: July 9, 2010

***E. coli* has long been widely used as a host system for the manufacture of recombinant proteins intended for human therapeutic use. When considering the impurities to be eliminated during the downstream process, residual host cell DNA is a major safety concern. The presence of residual *E. coli* host cell DNA in the final products is typically determined using a conventional slot blot hybridization assay or total DNA Threshold assay. However, both the former and latter methods are time consuming, expensive, and relatively insensitive. This study thus attempted to develop a more sensitive real-time PCR assay for the specific detection of residual *E. coli* DNA. This novel method was then compared with the slot blot hybridization assay and total DNA Threshold assay in order to determine its effectiveness and overall capabilities. The novel approach involved the selection of a specific primer pair for amplification of the *E. coli* 16S rRNA gene in an effort to improve sensitivity, whereas the *E. coli* host cell DNA quantification took place through the use of SYBR Green I. The detection limit of the real-time PCR assay, under these optimized conditions, was calculated to be 0.042 pg genomic DNA, which was much higher than those of both the slot blot hybridization assay and total DNA Threshold assay, where the detection limits were 2.42 and 3.73 pg genomic DNA, respectively. Hence, the real-time PCR assay can be said to be more reproducible, more accurate, and more precise than either the slot blot hybridization assay or total DNA Threshold assay. The real-time PCR assay may thus be a promising new tool for the quantitative detection and clearance validation of residual *E. coli* host cell DNA during the manufacturing process for recombinant therapeutics.**

Keywords: *E. coli* host cell DNA, real-time PCR, slot blot hybridization assay, total DNA Threshold assay

*Corresponding author
Phone: +82-42-629-8754; Fax: +82-42-629-8751;
E-mail: inskim@hnu.kr

Among the many systems available for heterologous protein production, the Gram-negative bacterium *E. coli* remains one of the most attractive because of its ability to grow rapidly with a high density on inexpensive substrates, its well-characterized genetics, and the availability of an increasingly large number of cloning vectors and mutant host strains [1, 8, 20].

The use of *E. coli* as the host for recombinant therapeutic proteins is subject to many regulatory issues, one of which is the clearance of the host cell DNA [3, 5, 18]. Although the significance of DNA-based contaminants in biopharmaceutical products remains unclear, there is a possibility that residual host cell DNA could transmit genetic information to patients receiving the products. Theoretically, entry of contaminant DNA into the genome of recipient cells could have serious clinical implications. These associated risks include the alteration of the level of expression of cellular genes, or the expression of a foreign gene product [22]. As a consequence, the regulatory authorities state that manufacturers of biopharmaceuticals should control and quantify the amount of residual host cell DNA in final products and validate the clearance of residual host cell DNA during the downstream process. The acceptable residual amount of DNA in the U.S.A., as specified in the Food and Drug Administration (FDA) guidelines, is 100 pg per dose, utilizing testing procedures that can detect 10 pg [5]. The limit permitted by the World Health Organization (WHO) and the European Union (EU) is up to 10 ng per dose [4].

For a number of years, the species-specific DNA hybridization assay and the total DNA Threshold assay have been used to quantify the amount of residual host cell DNA in biopharmaceuticals [26]. The basic principle of a hybridization assay is based on the binding of the DNA probe to immobilized and denatured host cell DNA. The concentration of hybridized DNA in the test sample is evaluated by comparing the hybridization signal generated by the test sample with that of the control standard. The

total DNA Threshold assay is primarily based on the sequence-independent binding of two proteins specific to single-stranded DNA (ssDNA) [11]. In the latter, one binding protein acts as an antibody, and the other as a so-called single-stranded DNA binding (SSB) protein. First, a reaction complex is formed when the biotinylated SSB protein and the anti-ssDNA antibody (conjugated to urease) bind to single-stranded host cell DNA. A filtration stage follows, during which the strong affinity of streptavidin for biotin is utilized to capture and concentrate the reaction complex onto a biotinylated membrane. For detection, the membrane is placed into a reader that contains the substrate urea. Inside, the urea is hydrolyzed by urease to produce a pH change, which is relative to the amount of host cell DNA in the sample.

In the biopharmaceuticals industry, real-time PCR has been applied to amplify and simultaneously quantify a targeted DNA molecule. This enables both the detection and the quantification of a specific sequence in a DNA sample. Practically, real-time PCR has been used to characterize and detect numerous bacterial, fungal, and viral loads in protein therapeutics [9, 10, 13, 16, 24, 25]. Applications of real-time PCR for the specific detection of residual *E. coli* host cell DNA in plasmid preparations have also been reported [15, 21]. However, in the latter examples, the sensitivity of the assays was found to be at the 1 pg level, which is nearly the same as those of the slot blot hybridization assay and the total DNA Threshold assay. Recently, a real-time PCR based on SYBR chemistry was developed to specifically and quantitatively detect residual Chinese hamster ovary (CHO) host cell DNA [17]. The real-time PCR method was in these cases found to be highly sensitive and specific. The method could detect CHO genomic DNA to 300 fg. This level of sensitivity is higher than those of both the DNA hybridization method and the total DNA Threshold assay.

The objective of this study is to develop a highly sensitive and specific detection method of *E. coli* host cell DNA using real-time PCR as an alternative means to the conventional slot blot hybridization assay and total DNA Threshold assay. In order to develop a convenient, rapid, and sensitive way of measuring the residual *E. coli* host cell DNA, a real-time PCR assay based on SYBR chemistry is proposed and will then be compared with the slot blot hybridization assay and the total DNA Threshold assay so as to validate the overall capability of these varying methods.

Materials and Methods

Bacterial Strain and Culture Medium

The strain used in this study was *E. coli* KCTC 1102 harboring a plasmid pET 21b carrying the gene for the granulocyte colony-stimulating factor (Novogen Ltd., U.S.A.). The *E. coli* strain was grown in LB medium containing 50 μg/ml of ampicillin at 37°C.

Preparation of Genomic DNA

Genomic DNA was extracted from *E. coli* using an SV mini kit (General Bio System Inc., Korea) in accordance with the manufacturer's instructions. DNA integrity and concentration were determined by spectrophotometric analysis at 260 nm and 280 nm (UV-1650 PC; Schimadzu Corp., Japan).

Primer Design and PCR Specificity Test

Oligonucleotide primers against the 16S rRNA gene (GenBank Accession No. J01859.1) were designed for the detection of *E. coli* DNA by real-time PCR using Primer3 [19]. The primers were synthesized by Bioneer Corp. (Korea). To determine the efficiency of the primers, genomic DNA extracted from *E. coli* was serially diluted 10-fold from 42,000 pg to 0.042 pg. A PCR reaction was then carried out with each primer pair using the templates of serially diluted genomic DNA. The PCR was performed in a Palm-Cycler (Corbett Research Ltd., Australia) using the following conditions: initial heat denaturation at 95°C for 2 min, followed by 40 cycles each of 95°C for 30 s, 54°C for 30 s, and 72°C for 35 s. Two μl of genomic DNA was amplified in a total volume of 25 μl mixture of 10 μM forward primer (1 μl), 10 μM reverse primer (1 μl), 2× GoTaq Green Master Mix (Promega Corp., U.S.A.) (12.5 μl), and nuclease free water (8.5 μl). To ensure complete extension, the reaction mixture was further incubated for 5 min at 72°C. Amplified DNA was analyzed by gel electrophoresis using a 1.5% (w/v) agarose gel (Sigma Corp., U.S.A.).

Optimization of Quantitative Real-Time PCR Assay

Real-time PCR was performed with a Rotor-Gene 3000 (Corbett Research Ltd., Australia) using the following conditions: an initial heat denaturation at 95°C for 15 min, followed by 40 cycles each of denaturation at 95°C for 10 s, annealing at different temperatures of 52, 54, 56, 58, or 60°C for 20 s, and an extension at 72°C for 30 s. Two μl of genomic DNA was amplified in a total volume of a 20 μl mixture of a 10 μM forward primer (0.5 μl), a 10 μM reverse primer (0.5 μl), 4× AccuPower Greenstar PCR PreMix containing Hot-Start *Taq* DNA polymerase, SYBR Green I, and deoxynucleotide triphosphate mix in a 5 μl quantity (manufactured by Bioneer Corp., Korea), and nuclease free water (12 μl). To ensure complete extension, the reaction mixture was further incubated for 10 min at 72°C. Immediately following PCR, a melting curve analysis was performed by raising the incubation temperature from 72 to 95°C in 0.2°C increments with a hold of 1 s at each increment. Real-time PCR conditions in relation to the primer concentration, annealing temperature and time, and $MgCl_2$ concentrations were optimized. Negative controls were run with each experiment. All reactions were run in duplicate.

Determination of Sensitivity and Reproducibility of Real-Time PCR Assay

To obtain a standard curve and to verify the sensitivity of the real-time PCR assay, serial 10-fold dilutions from 42,000 pg to 0.042 pg of *E. coli* genomic DNA, were amplified using the optimized conditions. A standard curve for quantification was generated by plotting the log of the DNA concentration of the known standard against the threshold cycle (Ct) value.

DNA Quantification Using Slot Blot Hybridization Assay

E. coli genomic DNA standards were labeled with digoxigenin-dUTP by using the DIG DNA Labeling and Detection Kit (Roche, Basel, Switzerland). The general slot blot hybridization method, using DIG-labeled probes, was performed in accordance with the instructions of the DIG DNA Labeling and Detection Kit. *E. coli* genomic DNA standards were denatured (boiling for 10 min and rapidly cooling in ice water) and slot-blotted to a nylon membrane (Hybond-N+; Amersham Ltd., U.K.) and then fixed by baking at 80°C for 2 h under vacuum. The membrane was prehybridized in an appropriate volume of DIG Easy Hyb (20 ml/100 cm^2 filter) at 50°C for 30 min. Denatured DIG-labeled DNA probes (boiling for 10 min and rapidly cooling in ice) were added in pre-heated DIG Easy Hyb (3.5 ml/100 cm^2 membrane) and hybridized at 50°C for 6−8 h. The membrane was washed after hybridization (i) in 2×SSC, 0.1% SDS at room temperature for 5 min twice, and (ii) in 0.1×SSC, 0.1% SDS for 15 min at 65°C twice. The DNA hybrid was detected using an HRP-conjugated anti-DIG antibody. The membrane were then washed in washing buffer once for 1−5 min, incubated in a 100 ml blocking solution and in a 20 ml antibody solution once by turns, washed in a washing buffer twice for 15 min, and then equilibrated in a 20 ml detection buffer. The substrate for alkaline phosphate was added, for development, for 3−12 h until an ideal dark-blue positive reaction appeared. All reactions were run in duplicate.

DNA Quantification Using Threshold Assay

The total DNA assay was conducted using the Threshold System and the Threshold Total DNA Assay Kit according to the instructions of the manufacturer (Molecular Devices Inc., U.S.A.). *E. coli* genomic DNA standards were heat denatured at 105°C for 15 min. In the reaction stage, a mixture containing the biotinylated SSB protein, streptavidin, and urease-conjugated monoclonal antibody against ssDNA was added to each standard and incubated for 1 h at 37°C. Reaction mixtures were transferred to wells in the manifold of the Threshold workstation for the separation stage of the assay. Mixtures were filtered through the biotin-coated nitrocellulose membranes under a controlled vacuum. Wells were washed with a wash solution, and filtration was allowed to continue under high vacuum. In the detection stage, the dipstick membranes were transferred to the Threshold reader. Captured urease contained in the DNA−protein complexes converted the urea substrate, which resulted in detectable pH changes in the substrate solution. Corresponding samples, spiking at 50 pg of calf thymus DNA, were also assayed in order to calculate spike recoveries according to the manufacturer's recommendations. Samples were assayed in triplicate. All the controls were within the range indicated on the certificate of analysis from the supplier.

Comparative Validation of Real-Time PCR, Slot Blot Hybridization, and Threshold Assays

The precision, accuracy, linearity, and detection limit of the three methods for quantitative detection of residual host cell DNA were validated according to FDA guidance for industry, the bioanalytical method validation [6]. *E. coli* genomic DNA (200 μg) was fragmented using the restriction enzymes *Hin*dIII and *Eco*RI. The fragmented *E. coli* genomic DNA was used as the standard DNA for method validation. The precision of an analytical method describes the closeness of individual measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single homogeneous volume of a biological matrix. Precision was measured using samples containing 62.5 and 125 pg of *E. coli* of genomic DNA standard. The concentrations of the samples were measured 6 times on different days, and then averages of standard deviations (SD), and the coefficient of variations (CV), were determined.

The accuracy of an analytical method describes the closeness of mean test results obtained by the method to the true value (concentration) of the analyte. *E. coli* genomic DNA standards were spiked to a drug substance at the concentrations of 31.25, 62.5, and 125 pg, respectively. The concentrations of the spiked samples were measured 6 times on different days, and the percentage of recovery of the spiked sample was calculated. The deviation of the mean from the true value served as the measure of accuracy.

The linearity of an analytical procedure is its ability (within a given range) to obtain test results that are directly proportional to the concentration (amount) of analyte in the sample. Standard solutions containing 1,000, 500, 250, 125, 62.5, 31.25, 15.6, and 7.8 pg of *E. coli* genomic DNA fragments, and the negative standard solutions, were prepared for the validation of slot blot hybridization and threshold assays, and standard solutions containing 4,200, 420, 42, 4.2, 0.42, and 0.042 pg of *E. coli* genomic DNA fragment and negative standard solution were prepared for the validation of the real-time PCR. The concentrations of the samples were measured 6 times on different days. Standard curves were generated for each experiment, and the correlation coefficient was evaluated by the regression of the standard curve using the method of least squares.

The detection limit is determined by the analysis of samples with known concentrations of analyte and by establishing the minimum level at which the analyte can be reliably detected. Standard solutions containing 1,000, 500, 250, 125, 62.5, 31.25, 15.6, and 7.8 pg of *E. coli* genomic DNA fragments, and negative standard solution, were used for the validation of slot blot hybridization and threshold assays, and standard solutions containing 4,200, 420, 42, 4.2, 0.42, and 0.042 pg of *E. coli* genomic DNA fragments and

Table 1. Sequences of oligonucleotide primer sets used in the detection of *E. coli* host cell DNA.

Forward primer	Reverse primer	Nucleotide position[a]	Amplicon size
ER-F1 CAAGACATCATGGCCCTTAC	ER-R1 ACTTCATGGAGTCGAGTTGC	1194−1334	141
ER-F2 AGAAGCTTGCTCTTTGCTGA	ER-R2 CTTTGGTCTTGCGACGTTAT	78−197	120
ER-F3 GCTCGTGTTGTGAAATGTTG	ER-R3 GTAAGGGCCATGATGACTTG	1067−1213	147
ER-F4 TCGAAGTCGAACGAAGCACTTTA	ER-R4 GCAGGTTACCCACGCGTTAC	61−197	137
ER-F5 GTCCAAAGCGGCGATTTG	ER-R5 CAGGCCAGAAGTTCTTTT TCCA	1148−1297	150

[a]*E. coli* 16S ribosomal RNA gene (GenBank Accession No. J01859.1).

negative standard solutions were used for the validation of the real-time PCR. The detection limit was calculated based on a visual evaluation method or the equation 3.3 σ/S, where σ is the standard deviation and S is the slope of the standard curve [7].

RESULTS

Primer Selection

The 16S rRNA gene is present in multiple copies in the genomes of all known bacteria that belong to the eubacterial kingdom. Many bacterial species contain up to seven copies of these genes [2]. A gene target that is present in multiple copies can increase the sensitivity of the assay. Therefore, oligonucleotide primers against the 16S rRNA gene were designed for the detection of *E. coli* DNA, using Primer3 (Table 1). Although all five primer pairs could specifically amplify the targeted genes, the primer pair ER-F2 and ER-R2 showed a greater sensitivity and efficiency (data not shown). In addition, there was minimal primer dimer formation.

Optimization of Real-Time PCR

The specificity and sensitivity of real-time PCR depends on the annealing temperature and time, and the concentrations of cations and primers in the reaction buffer. To improve the specificity and sensitivity of real-time PCR, such parameters were optimized. Fig. 1 shows the Ct values at different annealing temperatures. The optimal annealing temperature was found to be 54°C. The optimal magnesium concentration was chosen to be 2 mM (data not shown). The optimal annealing time and primer concentration were 20 s and 0.25 μM, respectively.

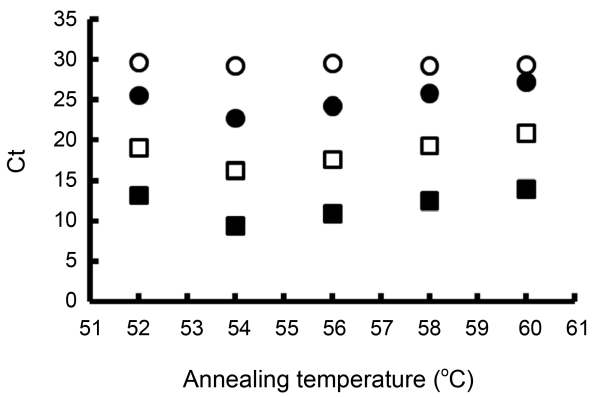

Fig. 1. Optimization of the annealing temperatures.
The Ct value refers to the cycle number at which the fluorescence of the PCR reaction rises above a set threshold and is inversely proportional to the amount of the starting target. (■) 42,000 pg of *E. coli* genomic DNA; (□) 420 pg of *E. coli* genomic DNA; (●) 4.2 pg of *E. coli* genomic DNA; (○) buffer control.

Sensitivity and Reproducibility of Real-Time PCR Assay

The sensitivity and reproducibility of the real-time PCR assay were determined. Serial 10-fold dilutions, from 42,000 pg to 0.042 pg of the *E. coli* genomic DNA were prepared and amplified using optimized conditions for the generation of the standard curve. Fig. 2 shows an example of the real-time profile of the *E. coli* genomic DNA amplification reaction, with a melting curve analysis of the

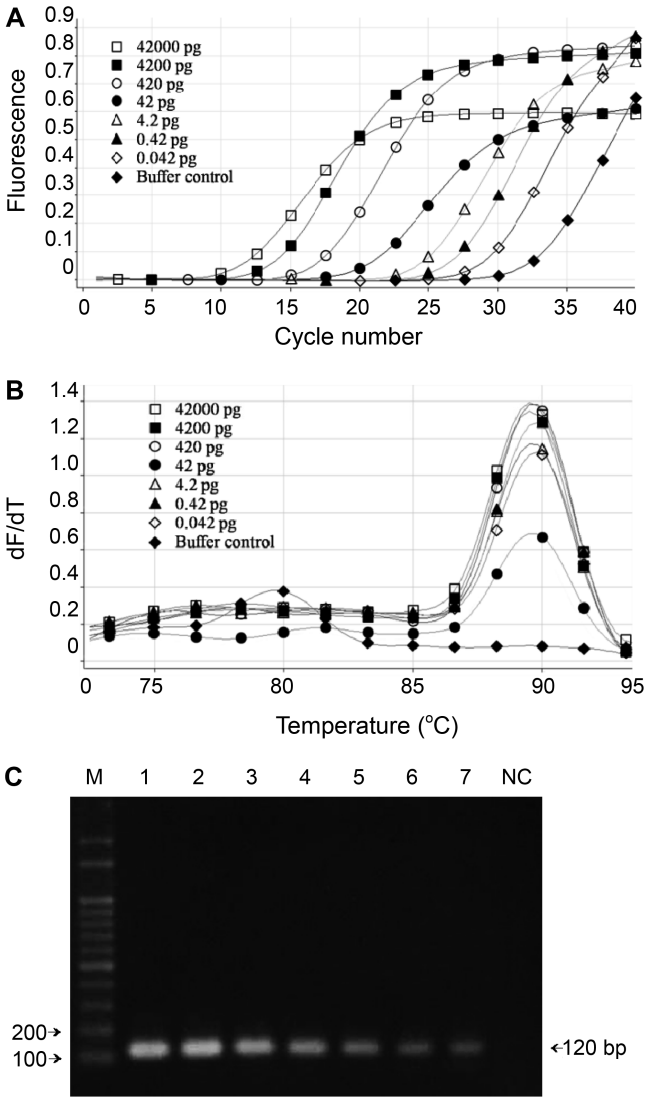

Fig. 2. Sensitivity of real-time PCR assay for the quantitative detection of *E. coli* host cell DNA.
The *E. coli* host cell DNA of 42,000 pg was serially diluted and cycle-by-cycle detection of *E. coli* host cell DNA performed with SYBR Green I. **A**. Amplification plots obtained with 10-fold serial dilutions of *E. coli* host cell DNA. **B**. Melting curve analysis of the amplification plot. **C**. Agarose gel electrophoresis of amplified products. Lanes: M, 100 bp DNA ladder; 1, 42,000 pg; 2, 4,200 pg; 3, 420 pg; 4, 42 pg; 5, 4.2 pg; 6, 0.42 pg; 7, 0.042 pg; NC, buffer control.

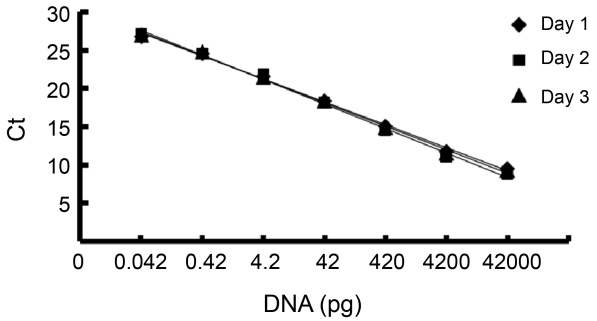

Fig. 3. Reproducibility of the real-time PCR assay for quantitative detection of *E. coli* host cell DNA.
The standard curves were obtained by the regression analysis of Ct values versus initial *E. coli* host cell DNA amounts. These results were obtained from three independent assays performed on different days.

amplification plot, and agarose gel electrophoresis of the amplified products. The sensitivity of the assay was found to be 0.042 pg. A melting curve analysis of the PCR products showed the specific identity of the PCR products. Agarose gel electrophoresis of the amplified products also showed that the real-time PCR specifically amplified the target gene. Standard curves for quantification were generated by plotting the log of the DNA concentration of the known standard against Ct values (Fig. 3). The standard curves were obtained from three independent assays performed on different days. The log concentration of *E. coli* genomic DNA on the Ct value between the standard curves was Day 1, $y=-3.0036x+30.201$ ($r^2=0.996$); Day 2, $y=-3.1976x+30.746$ ($r^2=0.995$); Day 3, $y=-3.0868x+30.492$ ($r^2=0.996$). The mean value, SD, and CV (%) of the slopes of the standard curves, were −3.096%, 0.996%, and 3.14%, respectively. In addition, the mean value, SD, and CV (%) of the *y*-intercepts of the standard curves were 30.480%, 0.272%, and 0.89%, respectively.

Comparative Validation of Real-Time PCR, Slot Blot Hybridization Assay, and Threshold Assay

In order to compare the overall capability of the real-time PCR with the slot blot hybridization and Threshold assays, the precision, accuracy, linearity, and detection limit of the three methods were validated. Precision was determined by measuring the concentration of the two spiked standard samples (62.5 and 125 pg) 6 times on different days (Table 2). The CVs determined from the real-time PCR assay were 1.86% for a 62.5 pg-spiked standard and 1.22% for a 125 pg-spiked standard, which were lower than those calculated for both the slot blot hybridization assay and Threshold assay. This result indicates that the real-time PCR assay is a highly precise method for the quantification of residual DNA when compared with either the slot blot hybridization assay or the Threshold assay.

The accuracy was assessed by analyzing the percentage of recovery of the spiked DNA standard at three different concentrations (31.25, 62.5, and 125 pg) (Table 3). The concentrations of the spiked DNA standards were determined 6 times on different days. The mean values of percentage recovery of slot blot hybridization were 102.34, 101.85, and 106.93, respectively. Those of the Threshold assay were 101.60, 100.54, and 101.88, respectively. Finally, those of the real-time assay were 101.25, 100.62, and 101.40, respectively. The mean values of percentage recovery obtained from the three different methods were all within 15% of the actual value, indicating that the three methods were all accurate. However, the standard deviation of percentage recovery obtained from the real-time PCR was lower than those obtained from both the slot blot hybridization and the Threshold assay. This result illustrates that the real-time PCR assay is more accurate than both the slot blot hybridization and Threshold assays.

The linearity of the analytical procedures was evaluated *via* the calculation of a regression line of the standard curve, using the method of least squares (Table 4). The mean values of the correlation coefficients of the slot blot hybridization assays, Threshold assays, and real-time PCRs, obtained from six independent experiments, were 0.976, 0.986, and 0.996, respectively. The higher correlation coefficient for the real-time PCR in comparison with the slot blot hybridization and the Threshold assay indicates a higher linearity of the measured concentrations from the standard curve.

The detection limits of the slot blot hybridization assay, the Threshold assay, and the real-time PCR were found to be 2.42, 3.73, and 0.042 pg of DNA, respectively (Table 4). The range of the standard curve was from 2.42 to 1,000 pg of DNA for the slot blot hybridization assay and from 3.73 to 200 pg of DNA for the Threshold assay. However, the

Table 2. Validation of the detection methods for *E. coli* host cell DNA: Precision.

Concentration of spiked DNA (pg)	Concentration of measured DNA (mean±SD)		
	Slot blot hybridization assay	Threshold assay	Real-time PCR
62.5	63.66±5.02 [7.89%][a]	62.84±4.75 [7.56%]	62.89±1.17 [1.86%]
125	133.66±9.38 [7.02%]	127.35±3.22 [2.53%]	126.75±1.54 [1.22%]

E. coli genomic DNA standards were spiked to a drug substance at the concentration of 62.5 and 125 pg, respectively. The concentrations of the spiked samples were measured 6 times on different days. The average, standard deviation, and coefficient of variation of the measured concentrations were determined. [a]Values in square brackets are the coefficient of variation (CV).

Table 3. Validation of the detection methods for *E. coli* host cell DNA: Accuracy.

Concentration of spiked DNA (pg)	% Recovery of spiked DNA (mean±SD)		
	Slot blot hybridization assay	Threshold assay	Real-time PCR
31.25	102.34±8.26	101.60±10.01	101.25±2.12
62.5	101.85±8.04	100.54±7.60	100.62±1.87
125	106.93±7.50	101.88±2.58	101.40±1.24

E. coli genomic DNA standards were spiked to a drug substance at the concentration of 31.25, 62.5, and 125 pg, respectively. The concentrations of the spiked samples were measured 6 times on different days. Percentage recovery of the spiked sample was calculated.

range of the calibration curve for the real-time PCR was from 0.042 to 42,000 pg of DNA.

DISCUSSION

Residual *E. coli* host cell DNA in biopharmaceuticals has been identified as a potential risk factor [12]. Hence, the FDA and other regulatory agencies have provided specific quality control and safety criteria, which require that quantification be carried out on all samples at intermediate points in the process, as well as on the final products [23]. Therefore, it is necessary to perform routine testing on residual host cell DNA on recombinant products or to show, in validation studies, which steps contribute to what extent in the removal of the DNA burden. Because DNA levels and matrix conditions vary during purification, residual DNA analysis in biopharmaceuticals can be very difficult and the methods have to be carefully validated before use [27]. No matter which approach is used, the assays and methods involved in the determination of residual host cell DNA must fulfill the validation requirements issued by the regulatory agencies [4-6]. Obviously, these techniques must be sensitive enough to detect very low levels of contamination. Generally the slot blot hybridization assay and total DNA Threshold assay have been used for the quantification of residual host cell DNA in the biopharmaceutical industry, although these methods are time consuming, expensive, and relatively insensitive [26].

Table 4. Validation of the detection methods for *E. coli* host cell DNA: Linearity and detection limits.

Item	Slot blot hybridization assay	Threshold assay	Real-time PCR
Linearity (r^2)	0.976	0.986	0.996
Detection limit (pg)	2.42	3.73	0.042

For the determination of linearity and detection limits, the concentrations of the standard samples were measured 6 times on different days. Standard curves were generated for each experiment and the correlation coefficient (r^2) was evaluated by regression of the standard curve using the method of least squares. Detection limits were determined according to the ICH guide [7]. The average of correlation coefficients and detection limits calculated from six independent experiments are presented.

In this study, a highly sensitive, rapid, and specific detection method for *E. coli* host cell DNA was developed using real-time PCR based on SYBR chemistry. Although there are many types of probes labeled with fluorescent molecules, such as the Molecular beacon, the Taqman probe, the FRET probe, and the Scorpion probe for real-time PCR, these probes are more expensive than SYBR Green I. A drawback in the use of the latter technique is the formation of primer dimers, as these are capable of binding the SYBR Green I dye and registering fluorescence. Therefore, primer design and optimization of real-time PCR are essential [14].

The sense and antisense primers were selected to amplify the 120-bp fragment of the highly conserved 16S rRNA gene. This primer set showed a higher sensitivity and a minimal primer dimer formation. One of the important considerations in optimizing a real-time PCR assay is cation concentration. Cations, especially Mg^{2+}, critically influence the melting behavior of DNA and therefore also affect the hybridization of the primers to the target template. The Mg^{2+} ion binds to the negatively charged phosphate groups on the backbone of the DNA. This weakens the electrorepulsive forces between the target DNA and the primer and stabilizes the primer–template complex. Excessive Mg^{2+} concentrations can lead to the amplification of nonspecific products and primer dimers, compromising PCR specificity. Although PCR efficiency was not influenced over a broad range of Mg^{2+} concentrations, from 2 mM to 6 mM under these reaction conditions when using the AccuPower Greenstar PCR PreMix, the optimal Mg^{2+} concentration was chosen to be 2 mM. The primer concentration also influences the specificity and efficiency of PCR. A high primer concentration allows for more efficient primer annealing during the annealing phase. However, concentrations that are too high will also increase the probability of nonspecific primer binding and primer dimer formation. The optimal primer concentration was therefore chosen to be 0.25 μM.

PCR amplicons were quantified by following the change in fluorescence of the DNA binding dye SYBR Green I, using a hot-start protocol. Standard curves were generated, by serial dilution of *E. coli* genomic DNA

(42,000 pg to 0.042 pg). Negative controls (no template) were run with each standard curve. A linear plot, with a correlation coefficient higher than 0.995, was typically obtained within the investigated range of genomic DNA quantities. Standard curves presented Ct numbers between 9.14±0.36 and 26.94±0.17. The sensitivity of the assay was found to be 0.042 pg of genomic DNA, which is equivalent to 62 copies of the 16S rRNA gene (equivalent to 8.8 *E. coli* genomes). The gene copy number was calculated in consideration of the seven genomic copies of the 16S rRNA gene and a molecular mass of the *E. coli* genome of 2.86×10^9 Da (4,639,221 bp). This sensitivity was 30-fold higher than previous reports for the real-time PCR detection of *E. coli* genomic DNA [15, 21]. This higher sensitivity may be due to the design of the highly sensitive primer set within highly conserved 16S rRNA genes and optimization of real-time PCR conditions; that is, concentrations of primer sets, magnesium concentrations, ionic strength, and the PCR cycle (annealing temperature and reaction time).

In order to verify the overall capabilities of the real-time PCR, the precision, accuracy, linearity, and detection limits of the real-time PCR were validated and then compared with those of the slot blot hybridization and Threshold assays. The real-time PCR was found to be more precise and more accurate than both the slot blot hybridization and Threshold assays. Moreover, the real-time PCR had a larger dynamic range and a higher linearity of the measured concentration from the standard curve. The detection limit of the real-time PCR was 0.042 pg, which was much lower than those of either the slot blot hybridization assay (2.42 pg) or the Threshold assay (3.73 pg). These validation results show that the real-time PCR is far more robust than the slot blot hybridization and Threshold assays. This study represents the first comparative validation of quantitative detection methods for residual *E. coli* host cell DNA.

In summary, the real-time PCR developed in this study shows a higher sensitivity, accuracy, and precision and yields a lower limit of detection when compared with either the slot blot hybridization or Threshold assays. This assay could be easily applied to the quantitative detection of residual *E. coli* host cell DNA both during the process, as well as on the final biopharmaceutical products. In addition, this quantitative method could be applied for the validation of the production process to establish that, at given steps of the purification scheme, *E. coli* host cell DNA is removed in a consistent and reproducible manner to an acceptable level.

Acknowledgments

This research was financially supported by the Ministry of Knowledge Economy (MKE) and Korea Industrial Technology Foundation (KOTEF) through the Human Resource Training Project for Strategic Technology, and a research fund from the Korea Food and Drug Administration (KFDA).

REFERENCES

1. Banexy, F. 1999. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol*. **10:** 411–421.
2. Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller. 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Bacteriol*. **62:** 293–300.
3. CPMP. 1997. Position Statement on DNA and host cell proteins (HCP) impurities – routine testing versus validation studies. Biotechnology Working Party, Committee for Proprietary Medicinal Product, The European Agency for the Evaluation of Medicinal Product, Human Medicines Evaluation Unit, EU.
4. CPMP. 2001. Position statement on the use of tumourigenic cells of human origin for the production of biological and biotechnological medicinal products. Biotechnology Working Party, Committee for Proprietary Medicinal Product, The European Agency for the Evaluation of Medicinal Products: Evaluation of Medicinal Products for Human Use, EU.
5. FDA. 1997. Points to consider in the manufacture and testing of monoclonal antibody products for human use. US Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research, U.S.A.
6. FDA. 2001. Guidance for industry, bioanalytical method validation. US Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research, Center for Veterinary Medicine, U.S.A.
7. ICH. 1996. Q2B Validation of analytical procedures: Methodology (guidance for industry). *International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use*, Geneva, Switzerland.
8. Jana, S. and J. K. Deb. 2005. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol*. **67:** 289–298.
9. Jeong, H. S., J. H. Shin, Y. N. Park, J. Y. Choi, Y. L. Kim, B. G. Kim, *et al*. 2003. Development of real-time RT–PCR for evaluation of JEV clearance during purification of HPV type 16 L1 virus-like particles. *Biologicals* **31:** 223–229.
10. Kubista, M., J. M. Andrade, M. Bengtsson, A. Forootan, J. Jonák, K. Lind, *et al*. 2006. The real-time polymerase chain reaction. *Mol. Aspects Med*. **27:** 95–125.
11. Kung, V. T., P. R. Panfili, E. L. Sheldon, R. S. King, P. A. Nagainis, J. B. Gomez, D. A. Ross, J. Briggs, and R. F. Zuk. 1990. Picogram quantitation of total DNA using DNA-binding proteins in a silicon sensor-based system. *Anal. Biochem*. **187:** 220–227.
12. Lahijani, M., M. Duhon, E. Lusby, H. Betita, and M. Marquet. 1988. Quantitation of host cell DNA contaminate in pharmaceutical-grade plasmid DNA using competitive PCR and enzyme linked immunosorbent assay. *Hum. Gene Ther*. **9:** 1173–1180.
13. Lee, D. H., H. S. Jeong, T. E. Kim, S. H. Oh, J. Lee, and I. S. Kim. 2008. Real-time RT–PCR for validation of reovirus type 3

safety during the manufacture of mammalian cell culture-derived biopharmaceuticals. *Kor. J. Microbiol*. **44:** 228–236.
14. Lovatt, A. 2002. Applications of quantitative PCR in the biosafety and genetic stability assessment of biotechnology products. *Rev. Mol. Biotechnol*. **82:** 279–300.
15. Martins, S. A. M., D. M. F. Prazeres, J. M. S. Cabral, and G. A. Monteiro. 2003. Comparison of real-time polymerase chain reaction and hybridization assays for the detection of *Escherichia coli* genomic DNA in process samples and pharmaceutical-grade plasmid DNA products. *Anal. Biochem*. **322:** 127–129.
16. Monis, P. T. and S. Giglio. 2006. Nucleic acid amplification-based techniques for pathogen detection and identification. *Infect. Genet. Evol.* **6:** 2–12.
17. Nissom, P. M. 2007. Specific detection of residual CHO host cell DNA by real-time PCR. *Biologicals* **35:** 211–215.
18. Rathore, A. S., S. E. Sobacke, T. J. Kocot, D. R. Morgan, R. L. Dufield, and N. M. Mozier. 2003. Analysis for residual host cell proteins and DNA in process streams of a recombinant protein product expressed in *Escherichia coli* cells. *J. Pharm. Biomed. Anal*. **32:** 1199–1211.
19. Rozen, S. and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol*. **132:** 365–386.
20. Swartz, J. R. 2001. Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol*. **12:** 195–201.
21. Vilalta, A., V. Whitlow, and T. Martin. 2002. Real-time PCR determination of *Escherichia coli* genomic DNA contamination in plasmid preparations. *Anal. Biochem.* **301:** 151–153.
22. Walsh, G. 2002. *Proteins Biochemistry and Biotechnology*, pp. 205–206. John Wiley & Sons Ltd, West Sussex, England.
23. Wang, K. U., Y. J. Guo, S. H. Sun, K. Shi, S. Zhang, K. H. Wang, Y. Zhang, and Z. H. Chen. 2006. 16S rRNA gene probe quantitates residual host cell DNA in pharmaceutical-grade plasmid DNA. *Vaccine* **24:** 2656–2661.
24. Wit, C., C. Fautz, and Y. Xu. 2000. Real-time quantitative PCR for retrovirus-like particle quantification in CHO cell culture. *Biologicals* **28:** 137–148.
25. Wolf, J. J., L. Wang, and E. Wang. 2007. Application of PCR technology in vaccine product development. *Expert Rev. Vaccines* **6:** 547–558.
26. Wolter, T. and A. Richter. 2005. Assay for controlling host-cell impurities in biopharmaceuticals. *BioProcess Int*. **3:** 40–46.
27. Workman, W. E. 1995. Sample preparation and residual DNA analysis of biopharmaceuticals. *Pharmacopeia Forum* **21:** 479–483.