A Direct Droplet Digital PCR Method for *E. coli* Host Residual DNA Quantification

Jeremy Anderson, Musaddeq Hussain

Merck & Co., Inc., Kenilworth, New Jersey, USA
Email: musaddeq.hussain@merck.com

Abstract

Injectable drugs manufactured in *E. coli* must be tested for host residual DNA (hr DNA) impurity in ensuring drug purity and safety. Because of low allowable hr DNA as impurity, highly sensitive methods are needed. Droplet digital PCR (ddPCR) is a new method where the reaction is partitioned into about 20,000 nanoliter-sized droplets and each droplet acts as individual PCR reaction. After completion of end-point PCR, droplets are analyzed for fluorescence and categorized as positive or negative and DNA quantified using Poisson statistics. Here we describe development of a direct *E. coli* hr DNA dd PCR method where the drug is directly added to the ddPCR reaction. We show that the ddPCR method has acceptable precision and high accuracy, works with different biologic drugs, and compared to qPCR shows higher tolerance of drug matrices. The method does not require DNA extraction or standard curves for quantification of hr DNA in unknown samples.

Keywords

E. coli Host, Host Residual DNA, Droplet Digital PCR, Direct Method, Biologic Drugs, Injectable Drug

1. Introduction

Many recombinant therapeutic proteins are produced in *E. coli*, e.g., insulin, human growth hormones, insulin-like growth factors, interferons, and interleukins [1]. *E. coli* is preferred as host because of well characterized genome, simplicity associated with cell cultures, rapid growth rates, high levels of expression, and low cost [2] [3] [4]. However, biopharmaceuticals manufactured in host cells contain host residual protein and DNA (hr DNA) as impurity and must be removed and quantified in the purified drug before it can be used in patients. The WHO guidelines recommend the hr DNA levels to be <10 ng per daily dose

[5] [6]. Typically, the biopharmaceutical industry uses qPCR, which is a sensitive and accurate method for quantification of hr DNA [7] [8]. Since qPCR requires generation of a standard curve for each experiment and often DNA extraction, both of which can be time consuming, we evaluated direct droplet digital PCR (ddPCR) as an alternative method [9]. In the ddPCR method the PCR reaction is partitioned into about 20,000 individual nanoliter-sized droplets using micro-fluidics. After the PCR is run to endpoint, the droplet fluorescence is read and analyzed, the number of positive and negative droplets is categorized, and the DNA amount is quantified based on Poisson statistics. After an initial standard curve is generated, future experiments can be run without the need for a DNA standard curve.

2. Materials and Methods

2.1. Materials

We tested two human insulin derived drugs in development at Merck & Co., Inc., Kenilworth, NJ USA (RP-IG and RP-IR) and three protein molecules, human interferon-α A (IFN-α, Cat# I4276), interferon-γ (IFN-γ, Cat# I3265) and insulin like growth factor (IGF, Cat# I3769) purchased from Sigma; all were manufactured in *E. coli* host. Total DNA from *E. coli* strain B purchased from Sigma-Aldrich (Cat# D4889) was used as standard. TaqMan Universal PCR Master Mix II with UNG (Cat# 4440038) was from Applied Biosystems (Foster City, CA). The ddPCR Supermix RDQ (Cat# 186-4037) was from Bio-Rad (Hercules, CA).

2.2. QPCR Method

The 3'-end of rrsA gene of *E. coli* (16s rRNA gene, Gen Bank Acquisition # CP001969.1; region 228440 to 229981) was the amplification target for quantification of hr DNA by PCR. The forward and reverse primers, designed in-house, were 5'-gttcccgggccttgtacacac-3' and 5'-aagtggtaagcgccctcccg-3', respectively. The probe, also designed in-house, was FAM-ccgtcacaccatgggagtgggt-TAMRA. A 10× Assay Mix of the primers and probe was prepared containing 10 μM of each primer and 2.5 μM probe. The mix was stored at –20˚C. Prior to each testing, a qPCR reaction mix was prepared with TaqMan Universal Master Mix II with UNG and 10× Assay Mix so that each 20 μL mix per PCR reaction contained 15 μL of TaqMan Universal Master Mix II, 3.0 μL of 10x Assay Mix and 2.0 μL water. Each PCR reaction contained 20 μL mix and 10 μL of sample with a total volume of 30 μL per well in 96-well optical reaction plate. Generally, 5.0 μL of drug or water and 5.0 μL of standard DNA or water constituted the 10 μL of sample. The drug and the DNA standard were prepared such that the desired amount was present in the 5.0 μL added. PCR cycling conditions were: 2 min at 50˚C, then 10 min at 95˚C followed by 40 cycles each consisting of 15 sec at 95˚C and 1 min at 60˚C. The PCR plates were covered with optical adhesive sheets and placed for amplification in 7500 Fast Real-Time PCR System (Applied

Biosystems) using the Accu SEQ Real-Time PCR Detection Software v2.1.

2.3. DDPCR Method

We developed the ddPCR method for *E. coli* hr DNA based on the qPCR method; the same 10× Assay Mix mentioned above was used with Supermix RDQ. Prior to each testing, a ddPCR reaction mix was prepared so that each 25 μL mix per PCR reaction contained 12.5 μL of Supermix RDQ, 2.5 μL 10x Assay Mix, and 10 μL of sample in wells of a 96-well PCR plate. The amount of drug and standard DNA was added such that the desired amounts were present in the 20 μL picked by Bio-Rad Automate Droplet Generator (ADG) to make droplets. The plate was sealed with Pierceable Foil Heat Seal using a Bio Rad PX1 PCR Plate Sealer for 5 seconds. The plate was briefly spun and put in the ADG for making droplets according to the manufacturer's protocol. The ADG made droplets by mixing with oil, and delivered in identical wells of a fresh PCR plate. The new PCR plate with the droplets was carefully removed from the ADG, sealed, and put in a thermocycler for PCR. The PCR cycling conditions were: 10 min at 95°C, one cycle followed by 40 cycles consisting of 30 s at 94°C with 2°C/sec ramp rate and 1 min at 60°C with 2°C/sec ramp rate; then 10 min at 98°C one cycle and hold at 4°C indefinitely. After PCR, the plate was transferred to the Bio Rad QX200 Droplet Reader and the fluorescence of individual droplets was read following the manufacturer's protocol. The ddPCR data was analyzed by QuantaSoft ver. 1.7.4.0917 software with threshold manually set at 1000 after looking at the 1D scatter of the droplets.

2.4. Data Analysis

Precision and determination of DNA quantity was measured from at least three replicate PCR wells and expressed as %RSD. The accuracy was determined by measuring the DNA spike recovery and expressed as %Recovery by the following calculation:

((DNA quantity in spiked sample) – (DNA quantity in unspiked sample))*100/Spike amount

3. Results and Discussion

The above-mentioned drugs were directly added to the PCR wells for qPCR or ddPCR. The ddPCR results for RP-IG and RP-IR are shown in Figure 1. The positive and negative droplets were well separated with low fluorescence in the negative droplets and no discernable effect of drug on the ddPCR. The ddPCR data analysis software has auto select function for the threshold, either for individual wells or combined wells. However, because of issues with auto settings [10] [11], we decided to set the threshold manually at 1000, which we used for all experiments reported here. We tested the linear range of the ddPCR method by serially diluting the *E. coli* DNA standard from 1e5 fg to 1.0 fg and performing ddPCR in triplicate over several days. Results exhibit a linear range from 1.0 fg to 1e5 fg of DNA per PCR reaction (Figure 2) and the LOQ was set at
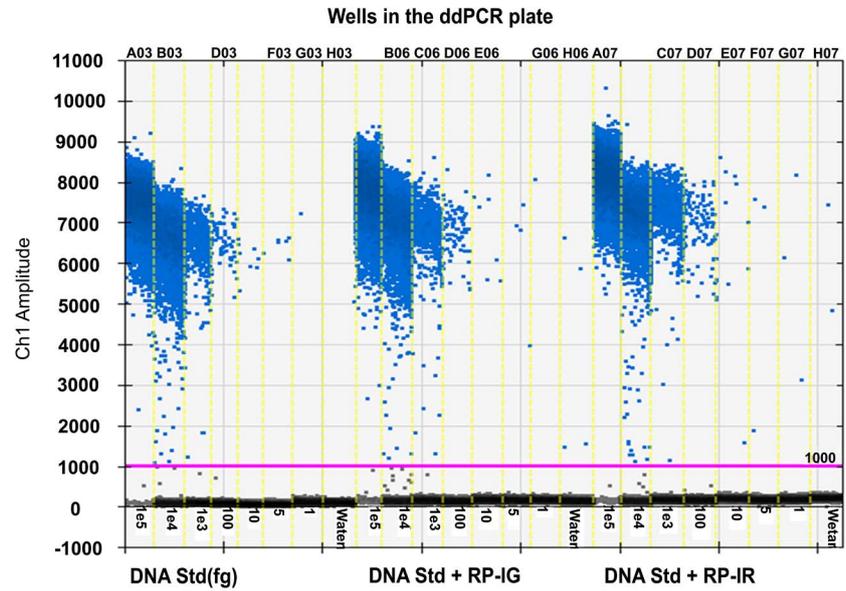
Figure 1. Droplet fluorescence of *E. coli* ddPCR. The amount of serially diluted *E. coli* DNA standard alone (Std) and spiked to drugs in ddPCR is shown on the x-axis. Two purified drugs, RP-IG and RP-IR at 5.0 μg each, were used. The fluorescence amplitude (of Ch 1 set for FAM) for each droplet after PCR in this 1D concentration plot, analyzed with Bio-Rad QuantaSoft, is shown on the y-axis. The vertical dashed yellow lines separate the individual wells (noted in some cases at the top) of the ddPCR plate. The horizontal red line is the threshold separating positive (blue dots) and negative droplets (black dots) was created manually and set at 1000.
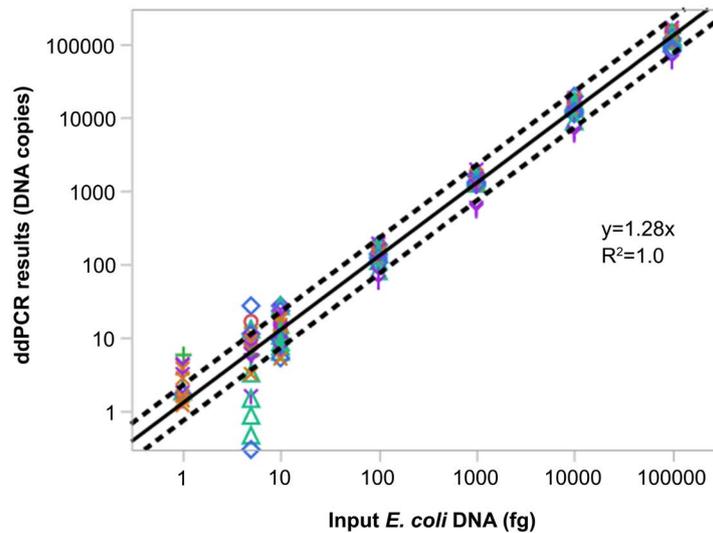

Figure 2. Linearity of *E. coli* ddPCR. The linear range was determined with DNA standards only (Y), and DNA standards with 5.0 μg of drugs RP-IG (x), RP-IR (Δ), IFN-α (+), IGF (o) or 0.5 μg IFN-γ (◊). The *E. coli* DNA added is shown in the x-axis. The ddPCR results are shown in the y-axis as copies of DNA detected. Based on the DNA standards only, the solid line shows the linear trend line for the mean. The dotted lines show the 95% confidence interval for the individual measurements, calculated by 2.0 RMSE of a log/log fit with slope = 1. The conversion factor of DNA copies to weight in fg was calculated from the inverse of the slope of the standard curve: $1/1.28 \approx 0.8$.

10 fg based on precision of 17.6% RSD. The linearity was maintained with drug added to the DNA standards (Figure 2) with varying precision at 10 fg level, e.g., RSD < 30% for RP-IG, IFN-α and IGF; 33.9% for RP-IR; and 53.7% for IFN-γ. The spike recovery, as a measure of accuracy of the method from the different drugs in the linear range, was about 100% as seen in Figure 2. The ddPCR determined copies of DNA can be converted to weight as shown in Figure 2. Since the size of the *E. coli* genome is approximately 4.7 Mbps, or about 5.18 fg, the data showed that about 7 copies of the 16s rRNA gene target were present in the *E. coli* genome. A literature search showed that 1 to 15 copies of the 16 s rRNA gene were present in different bacterial genomes, with an average of seven copies found in *E. coli* [12] [13]. In order to assess sample matrix effect, RP-IR process-intermediate sample and purified drug substance were serially diluted 2-fold and 5.00 to 0.31 μg were tested in qPCR and ddPCR (Table 1). The results showed that qPCR worked only when the samples were diluted to 0.63 μg or lower but ddPCR worked at all levels starting from 5.0 μg. An advantage of ddPCR over qPCR is a relative lack of matrix effect of crude, process intermediate and purified drug on the ddPCR method, as a result allowing testing larger amount of drug (Table 1). However, the hrDNA values determined from the samples by either qPCR or ddPCR were reasonably close to each other (Table 2).

We successfully developed a ddPCR method for *E. coli* hrDNA without the

Table 1. Sample matrix effect on qPCR and ddPCR.

Drug Samples	Drug Amount Tested (μg)	qPCR results		ddPCR results	
		DNA Amount Quantified (fg)	Recovery of Spiked DNA (%)	Observed DNA Copies	Recovery of Spiked DNA (%)
Process intermediate	5.00	0.0	0	3680.0	96
	2.50	0.0	0	3220.0	84
	1.25	0.0	0	3660.0	95
	0.63	1872.6	97	3480.0	90
	0.31	1888.5	98	3880.0	101
Purified drug substance	5.00	0.0	0	3540.0	92
	2.50	0.0	0	3680.0	96
	1.25	0.0	0	3680.0	96
	0.63	2007.7	104	3740.0	97
	0.31	2073.3	107	3800.0	99

Table 2. Quantification of *E. coli* hr DNA.

Sample	qPCR result (pg DNA/ mg drug)	ddPCR result (pg DNA/ mg drug)
RP-IR (Purified)	0.47	1.90
RP-IR (Crude)	43.54	45.82
RP-IR (Process Intermediate)	2.10	6.14

need for DNA extraction or a standard curve. This method is accurate, precise, and sensitive; tolerant of different sample matrices allowing testing higher drug amounts, and can be used for a wide variety of drug substances produced in *E. coli*. This method is currently being used for routine testing of samples in our lab.

Acknowledgements

References

[1] Sanchez-Garcia, L., Martin, L., Mangues, R., Ferrer-Miralles, N., Vazquez, E. and Villaverde, A. (2016) Recombinant Pharmaceuticals from Microbial Cells: A 2015 Update. *Microbial Cell Factories*, **15**, 33. https://doi.org/10.1186/s12934-016-0437-3

[2] Baeshen, N., Baeshen, M., Sheikh, A., Bora, R., Ahmed, M., Ramadan, H., Saini, K. and Redwan, E. (2014) Cell Factories for Insulin Production. *Microbial Cell Factories*, **13**, 141. https://doi.org/10.1186/s12934-014-0141-0

[3] Baeshen, M., Al-Hejin, A., Bora, R., Ahmed, M., Ramadan, H., Saini, K., Baeshen, N. and Redwan, E. (2015) Production of Biopharmaceuticals in *E. coli*: Current Scenario and Future Perspectives. *Journal of Microbiology and Biotechnology*, **25**, 953-962. https://doi.org/10.4014/jmb.1412.12079

[4] Kamionka, M. (2011) Engineering of Therapeutic Proteins Production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, **12**, 268-274. https://doi.org/10.2174/138920111794295693

[5] World Health Organization (WHO) (1997) WHO Expert Committee on Biological Standardization. Highlights of the 46th Meeting, October 1996. *The Weekly Epidemiological Record* (*WER*), **72**, 141-145.

[6] Wang, X., Moran, D., Wang, G. and Mozier, N. (2012) Residual DNA Analysis in Biologics Development: Review of Measurement and Quantitation Technologies and Future Directions. *Biotechnology and Bioengineering*, **109**, 307-317. https://doi.org/10.1002/bit.23343

[7] Mamnoon, B., Farivar, T., Kamyab, A., Ilghari, D., Khamesipour, A. and Arzenani, M. (2016) Quality Control of Widely Used Therapeutic Recombinant Proteins by Novel Real-Time PCR Approach. *Iranian Biomedical Journal*, **20**, 56-62.

[8] Hussain, M. (2015) A Direct qPCR Method for Residual DNA Quantification in Monoclonal Antibody Drugs Produced in CHO Cells. *Journal of Pharmaceutical and Biomedical Analysis*, **115**, 603-606. https://doi.org/10.1016/j.jpba.2015.03.005

[9] Hussain, M., Fantuzzo, R., Mercorelli, S. and Cullen, C. (2016) A Direct Droplet Digital PCR Method for Quantification of Residual DNA in Protein Drugs Produced in Yeast Cells. *Journal of Pharmaceutical and Biomedical Analysis*, **123**, 128-131. https://doi.org/10.1016/j.jpba.2016.01.050

[10] Trypsteen, W., Vynck, M., De Neve, J., Bonczkowski, P., Kiselinova, M., Malatinkova, E., Vervisch, K., Thas, O., Vandekerckhove, L. and De Spiegelaere, W. (2015) ddpcRquant: Threshold Determination for Single Channel Droplet Digital PCR Experiments. *Analytical and Bioanalytical Chemistry*, **407**, 5827-5834. https://doi.org/10.1007/s00216-015-8773-4

[11] Hussain, M. and Bowers, J. (2017) A Droplet Digital PCR Method for CHO Host

Residual DNA Quantification in Biologic Drugs. *Journal of Analytical & Pharmaceutical Research*, **4**, 00107. https://doi.org/10.15406/japlr.2017.04.00107

[12] Pei, A., Oberdorf, W., Nossa, C., Agarwal, A., Chokshi, P., Gerz, E., Jin, Z., Lee, P., Yang, L., Poles, M., Brown, S., Sotero, S., DeSantis, T., Brodie, E., Nelson, K. and Pei, Z. (2010) Diversity of 16s rRNA Genes within Individual Prokaryotic Genomes. *Applied and Environmental Microbiology*, **76**, 3886-3897. https://doi.org/10.1128/AEM.02953-09

[13] Vetrovsky, T. and Baldrian, P. (2013) The Variability of the 16s rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLOS One*, **8**, e57923. https://doi.org/10.1371/journal.pone.0057923