On chip quadruplex priming amplification for quantitative isothermal diagnostics

C. Pease 1 · G. E. Plum 2 · B. Kankia 3 · J. J. Kwiek 4 · R. Sooryakumar 1

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
Nucleic acid testing is a common technique for medical diagnostics. For example, it is used to detect HIV treatment failure by monitoring viral load levels. Quadruplex Priming Amplification (QPA) is an isothermal nucleic acid amplification technique that requires little power and few chemical reagents per assay, all features that make QPA well suited for point-of-care (POC) diagnostics. The QPA assay can be further optimized by integrating it with microfluidic devices that can automate and combine multiple reaction steps and reduce the quantity and cost of reagents per test. In this study, a real-time, exponential QPA reaction is demonstrated for the first time in a microfluidic chip, where the reaction was not inhibited and supported performance levels comparable to a commercially-available, non-microfluidics setup.

Keywords QPA · Quadruplex priming amplification · Isothermal amplification · Microfluidics · Lab on a Chip

1 Introduction
The World Health Organization (WHO) estimated in 2017 that 37 million people are already living with HIV (UNAIDS 2017), of whom 91% live in low and middle income countries (UNAIDS 2010). Increased access to anti-retroviral therapies (UNAIDS 2016) has redefined HIV infection from being an acute, fatal disease into a chronic condition that must be managed. The WHO recommends that people living with HIV periodically check for treatment failure by monitoring the amount of HIV virions in their blood (World Health Organization 2013) through nucleic acid amplification tests (NAATs) of the genomic RNA in HIV virions. These tests are often carried out in large, centralized laboratories making access difficult in rural and resource-limited settings (Mazzola and Perez-Casas 2015).

To increase access, there is a general agreement that NAATs need to be made available close to the point-of-care (POC) of the patient, which requires the assay to be affordable when performed in small batches. POC NAATs, however, remain expensive with costs averaging $28–29 per test (Mazzola and Perez-Casas 2015).

Quadruplex priming amplification (Adams et al. 2014; Gogichaishvili et al. 2014; Kankia 2011; Loh et al. 2014; Partskhaladze et al. 2015; Taylor et al. 2013) (QPA) is a simple nucleic acid amplification assay for real-time quantification (Craw and Balachandran 2012) that can exponentially amplify DNA from attomolar initial concentrations (Partskhaladze et al. 2015). QPA is cheaper per reaction than other NAATS like Polymerase Chain Reaction (PCR) and thus may be used instead to reduce the cost for HIV viral load monitoring. Unlike PCR-based reactions, QPA reacts isothermally (~65 °C) which reduces its power demand and it uses few reagents allowing for a cheap cost-per-assay. The reaction can occur isothermally because double-stranded DNA separates spontaneously to form stable quadruplex structures and does not need to be melted through thermal cycling. Few reagents are required since QPA does not need additional components to amplify RNA or separate fluorescent probes for detection. Either DNA or RNA can be used as the amplification target without additional steps and the reaction is accompanied by an intrinsic increase in fluorescence from the quadruplex formation itself (Gogichaishvili et al. 2014).

To increase the affordability of POC NAATs, an active research area is the integration of NAATs with microfluidic
chips (Andresen et al. 2009; Basha et al. 2017; Fang et al. 2010; Gulliksen et al. 2004; Lutz et al. 2010; Mauk et al. 2017; Nalluru et al. 2001; Ramalingam et al. 2009; Westin et al. 2000) in order to decrease the volume and cost of reagents, reduce cross-contamination between subsequent reactions with cheap, disposable chips, and create less variation between measurements with automated protocols. For POC HIV viral load monitoring, the WHO recommends a lower limit of detection of 1000 copies/mL (1 copy/μL) (World Health Organization 2013) constraining the reaction volume to be at least tens of microliters (μL). Yet, to our knowledge, the QPA reaction has not been integrated with a microfluidic chips. Integration would be especially advantageous for QPA, because its current version has disadvantages that can be overcome using microfluidic technologies including a low excitation wavelength (348 nm) at which many polymers and plastics autofluoresce (Galla et al. 2012; Piruska et al. 2005) potentially creating additional noise and the requirement that multiple reagents be mixed in sequentially.

This report describes the integration and detection of the real-time, exponential QPA assay and a parallel control reaction in a microfluidic chip. The results from the reaction on-chip are found to be comparable to findings from a commercial tube-based system demonstrating that the chip does not inhibit QPA. These results demonstrate that QPA has the potential to be used in microfluidic devices for POC applications and lays the foundation for QPA to be used to monitor HIV viral load.

2 Methods and materials

2.1 QPA

In order to run the QPA assay (Fig. 1a), the target nucleotide sequence of the pathogen is sequentially bound to the QPA linker and then mixed with the reaction mixture. In the presence of the linker, the reaction mixture exponentially creates fluorescent quadruplex structures. From the change in fluorescence, the threshold time $T_r$ required for the fluorescent signal to increase above the background noise is measured. By comparing $T_r$ with calibrated values, the initial concentration of the pathogen can be determined. $T_r$ also determines the duration necessary to finish a single test and the rate at which a single detector can process multiple samples.

In the presence of certain ions, the G3T sequence of DNA (GGG(3MI)GGTGTTGGTGGG) folds into a stable quadruplex structure (Fig. 1b). QPA is isothermal because it uses the formation of G3T quadruplexes as its driving mechanism rather than thermal cycling.

The exponential QPA reaction (Partskhaladze et al. 2015) uses two primers: the QPA primer and the 2nd primer. The QPA primer forms quadruplexes and drives the reaction, whereas the 2nd primer is necessary for the reaction to be exponential. The QPA linker (Fig. 1c) contains the complement to the target nucleotide sequence of the pathogen as well as a copy of the 2nd primer and the binding site of the QPA primer. When the linker is mixed with the QPA reaction reagents, there are analogous reaction cycles for the QPA and 2nd primer (Fig. 1d). The QPA primer, which is the G3T sequence minus a few bases, binds to and extends on the linker creating a 2nd primer binding site. Once fully extended, the primer spontaneously breaks away to form a quadruplex that has a free 2nd primer binding site attached to it. The 2nd primer can then undergo a similar annealing/extension/separation cycle to create a new QPA primer binding site. Once either cycle is complete, the number of binding sites has doubled causing the reaction to be exponential. For each QPA primer cycle, a free QPA primer turns into a quadruplex yielding a 50-fold increase (Gogichaishvili et al. 2014) in fluorescence emitted by the fluorescent base analog 3-Methyl Isoxanthopterin (3MI) embedded within the QPA primer. The progress of the assay can thus be tracked from the change in fluorescence.

The QPA reaction reagents (Partskhaladze et al. 2015) include the linker (ATTATTTCATTATCTTTATATATTCCCCACCCGCCCC, 100 pM for 100 μL reaction), QPA primer (GGG(3MI)GGGCGGGTGG, 300 nM), 2nd primer (ATTATTTCATTATCTTTATATATCTCCC, 400 nM), buffer (10 mM KCl, 40 mM CsCl, 2 mM MgCl₂, 10 mM Tris), deoxynucleoside triphosphates (dNTPs, 800 μM total), and the enzyme Vent(exo-) (0.06 U/μL). Because the part of the linker that reacts with the QPA reaction mixture is pathogen independent, for our tests, the linker was the target and was not attached to any pathogen nucleotide sequences before the reaction. The control only included the fluorescent QPA primer and salt buffer which creates the same background fluorescence as the reaction reagents but without the possibility of amplification.

2.2 Reaction well

During the reaction, the QPA reagents and the control are placed in the chip (Fig. 2) which is sealed and pressurized. The chip is placed on a heater for one minute, set at 86.9 °C, to preheat the solution and melt any free quadruplexes in solution. The temperature of the heater is decreased for three additional minutes to the reaction temperature (~65 °C) where it subsequently remains. The sample and control are excited by an LED and their emission is measured in real-time by a camera above the chip.

Reaction wells (Fig. 2) are made from polydimethylsiloxane (PDMS). Two adjacent wells (5.5 mm diameter) separately contain the QPA reaction mixture and the control. PCR tape (Lightcycler 480 Sealing Foil) seals the top and bottom of the wells to prevent leakage. Two glass slides held together with
rubber bands sandwich the reaction well and provided a source of pressure to reduce bubble formation and keep the reaction well cover flat to ensure incident light reflects correctly.

### 2.3 Temperature controller

The temperature of the metal-ceramic resistive plate heater is monitored by a thermistor attached underneath the heater in 0.8 °C increments and a proportional-integral-derivative (PID) controller maintains the temperature of the microfluidic chip within an increment.

### 2.4 Optical detection system

The optical system designed to monitor the progress of QPA is depicted in Fig. 2. Light from the LED (centered at 365 nm) reflects off the dichroic mirror to the sample which excites the 3MI fluorophore (excitation centered at 348 nm) (Gogichaishvili et al. 2014). Orienting the dichroic mirror at 35° (rather than 45°) reflects the incident light off the flat reaction cover away from the camera. The excitation of the 3MI fluorophore (center at 431 nm) (Gogichaishvili et al. 2014) passes up through the...
dichroic mirror and the highpass filter, which further suppress any scattered incident light, to reach the camera.

Custom software saves images from a 16-bit monochrome camera (QSI 504) of the microfluidic chip every 20 s and measures the fluorescent signal from the reaction and control wells as a function of time. The LED is turned off between measurements to reduce photobleaching.

3 Results and discussion

The measured fluorescence (Fig. 3a) commences with a flat baseline signal before increasing exponentially when the signal from the reaction product is greater than the background. The threshold time $T_T$ is the time required for the fluorescence to substantially increase above the baseline and is set at when the signal is $10^x$ greater than the standard deviation of the baseline (Life Technologies 2012). For quantitative diagnostics, the initial concentration of the target can be determined by comparing $T_T$ to a standard reference curve (Fig. 3b) created by measuring $T_T$ for known initial concentrations (Fig. 3c).

3.1 Temperature calibration

The reaction rate, as measured by the threshold time $T_T$, is sensitive to the reaction temperature (Partskhaladze et al. 2015). Fig. 4 shows the value of $T_T$, as determined from the average of three to four measurements with a 100 μL reaction volume, at each temperature. $T_T$ was optimal at 14.7 min in the temperature range between 65.8 °C and 67.3 °C.

The standard deviation of $T_T$ increased with temperature, resulting perhaps from the formation of a greater number of bubbles caused by the decreased solubility of gases in water at higher temperatures. At 68.9 °C, the reaction was especially erratic indicating that it is likely near the highest viable reaction temperature for QPA. The optimal reaction temperature was determined to be 65.8 °C and was used in subsequent reactions because it minimized $T_T$ and associated error.

3.2 Accuracy

To determine the accuracy of the on-chip setup, results were compared with those from a commercial instrument (ESEQuant Tube Scanner), which measures the fluorescence from a reaction while the reagents are held, at a stabilized temperature (68 °C), within 100 μL tubes. Unlike the on-chip setup, the commercial device cannot be extended to integrate multiple steps or reduce fluid volumes. To better compare the setups, QPA on-chip was also measured in a 100 μL reaction well.

Fig. 5 presents outputs from the commercial and on-chip setups. The fluorescent signal is shown once the setup reached its reaction temperature after the initial pre-heat (3 min for the commercial device and 4 min on-chip). On-chip, $T_T$ was 14.7 ± 2.0 min over three measurements while for the commercial instrument, $T_T$ was 13.7 ± 3.3 min over four reactions. The additional minute in $T_T$ for the on-chip setup corresponds with...
the extra time required to reach its reaction temperature (65.8 °C). The increased error in the commercial setup results from the differences in temperature stability. The reaction temperature was stable within ±0.5 °C on-chip and ±1.5 °C in the commercial unit. The reduced error on-chip is especially important at low initial target concentrations where small changes in $TT$ yield large changes in the expected initial concentration (Fig. 3c). Altogether, this comparison confirms that the QPA reaction is uninhibited while on-chip.

3.3 Reaction well depth

Reducing the reaction volume decreases the cost per reaction because smaller reagent quantities are consumed. For POC HIV viral load monitoring, the WHO recommends a lower limit of detection of 1000 copies/mL (1 copy/μL) (World Health Organization 2013). Therefore, while the reaction volume may be reduced from 100 μL, it is still likely constrained to the tens of microliters (μL). The effect of reducing the well height on the reaction rate is unclear because it changes the amount of fluorescent material measured by the camera within a vertical slice, heat transfer times, and the target concentration if the number of copies is fixed as the well depth changes.

In order to determine the effect of the well depth, the original 100 μL well (3.5 mm well depth), was compared with one that holds 60 μL (2 mm well depth). For the latter, the reagent concentrations were all kept the same except for the target, i.e. the QPA linker. The same number of copies of target was used in both wells yielding 166 pM of target for the 60 μL well rather than the 100 pM in the 100 μL well. An example of a run at each height is shown in Fig. 6.

Reducing the well volume slightly increased the threshold time from 15 ± 2 min. to 20 ± 3 min. and significantly decreased the total change of fluorescence from 1100 ± 400 A.U. to 400 ± 20 A.U., as determined by the average of three runs in the 100 μL well and four runs in the 60 μL well. The fluorescence did not decrease proportionally with height, but rather was reduced nearly 60% for a 40% decrease in height (3.5 mm to 2 mm). The slowdown in $TT$ (from 15 to 20 min) likely resulted from this large decrease in fluorescence, as more amplification cycles are needed before the signal can be detected above the baseline.

4 Conclusion

QPA is a simple reaction that uses few reagents for an isothermal nucleic acid amplification technique. A system for controlling the temperature and measuring the change in fluorescence was constructed allowing QPA and a control to be integrated with a microfluidic chip in parallel for real-time detection. This approach allowed the concentration of a 100 pM target sample to be determined from the measured $TT$. This setup yielded comparable reaction times as a tube-based commercial instrument indicating that the reaction was not
inhibited on-chip. The on-chip setup also had less error in $T_T$, due to smaller temperature fluctuations, which is especially important at low initial concentrations where small changes in $T_T$ result in large changes in the expected initial concentrations. Moreover, reducing the reaction volume (from 100 μL to 60 μL) by decreasing the reaction well depth was found to increase the threshold time (from 15 to 20 min) necessary to determine initial target count.

This integration of QPA and microfluidics demonstrates the feasibility of bringing the advantages of the QPA reaction to POC diagnostic devices and lays the foundation for the future extension of the microchip platform to allow for the parallel measurement of HIV viral load with QPA.

Acknowledgements The authors gratefully acknowledge the financial support provided by the Ohio State University Public Health Preparedness for Infectious Diseases Program (OSU-PHPID).

References

P. Craw and W. Balachandran, Lab Chip 12, 2469–2486 (2012)
UNAIDS, AIDS by the Numbers (2016)
UNAIDS, UNAIDS Data 2017 (2017)
World Health Organization, Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection (2013)