Journal of Applied Microbiology 2001, **91**, 110–117

A repeatable laboratory method for testing the efficacy of biocides against toilet bowl biofilms

B. Pitts, A. Willse[3], G.A. McFeters[2], M.A. Hamilton[3], N. Zelver and P.S. Stewart[1]

Center for Biofilm Engineering and Departments of [1]Chemical Engineering, [2]Microbiology and [3]Mathematical Sciences, Montana State University, Bozeman, MT, USA

517/08/00: received 21 August 2000, revised 31 January 2001 and accepted 22 February 2001

B. PITTS, A. WILLSE, G.A. MCFETERS, M.A. HAMILTON, N. ZELVER AND P.S. STEWART. 2001.

Aims: The purpose of this study was to develop a laboratory biofilm growth reactor system that simulated the toilet bowl environment and which could be used for biocide efficacy testing.

Methods and Results: A microbial biofilm reactor system incorporating intermittent flow and nutrient provision was designed. The reactor system was open to the air and was inoculated with organisms collected from toilet bowl biofilms. Once per hour, reactors were supplied with a nutrient solution for a period of 5 min, then flushed and refilled with tap water or tap water amended with chlorine. Quantitative measures of the rate and extent of biofilm accumulation were defined. Biofilm accumulated in untreated reactors to cell densities of 10^8 cfu cm^{-2} after approximately 1 week. Biofilm accumulation was also observed in reactors in the continuous presence of several milligrams per litre of free chlorine. Repeatability standard deviations for the selected efficacy measures were low, indicating high repeatability between experiments. Log reduction values of viable cell numbers were within ranges observed with standard suspension and hard surface disinfection tests. Biofilm accumulated in laboratory reactors approximately seven times faster than it did in actual toilet bowls. The same ranking was achieved in tests between laboratory biofilms and field-grown biofilms with three of the four measures, using three different concentrations of chlorine.

Conclusions: This reactor system has been shown to simulate, in a repeatable way, the accumulation of bacterial biofilm that occurs in toilet bowls. The results demonstrate that this system can provide repeatable assays of the efficacy of chlorine against those biofilms.

Significance and Impact of the Study: The laboratory biofilm reactor system described herein can be used to evaluate potential antimicrobial and antifouling treatments for control of biofilm formation in toilet bowls.

INTRODUCTION

The control of bacterial biofilms in environments ranging from household drains to cooling water towers depends heavily on the use of biocides. There are two significant frustrations to this strategy. The first is the inherent resistance of micro-organisms in the biofilm mode of growth to antimicrobial agents of all kinds (Brown and Gilbert 1993; Gilbert *et al.* 1997; Costerton *et al.* 1999). The second difficulty is the shortage of accepted methods for measuring biocide efficacy against biofilms. Because micro-organisms in a biofilm and cells in free aqueous suspension respond so differently to disinfectants, conventional planktonic culture methods cannot be used to select appropriate control agents or to design biofilm dosing protocols. Consequently, a biofilm testing method that adequately simulates the field setting of interest is required.

Laboratory biofilm systems have been designed to evaluate biocides in different environments, including the human mouth (Bradshaw *et al.* 1996), oilfield water injection systems (Green *et al.* 1987) and cooling water or distribution systems (LeChevallier *et al.* 1990; Green and Pirrie 1993). In this article, we describe a biofilm reactor system designed to simulate the toilet bowl environment. The reactors were

Correspondence to: P.S. Stewart, Department of Chemical Engineering, Montana State University, Bozeman, MT 59717, USA (e-mail: phil_s@erc.montana.edu).

initially inoculated with a consortium of four bacteria isolated from naturally occurring toilet bowl biofilms. The design incorporated intermittent nutrient provision and periodic flow and draining. The laboratory system was designed to quantitatively assess the efficacy of toilet cleaning products that are dropped into the tank and thereby delivered to the bowl.

Replicate experiments were conducted to determine the repeatability standard deviation and the sources of variability. Field experiments were conducted to assess the relevance of the laboratory method.

MATERIALS AND METHODS

Biofilm growth reactors and conditions

The reactor vessels consisted of glass 1-l beakers fitted with drain spouts. A magnetically-driven rotor was placed at the bottom of each vessel. The rotor was constructed from a star-head magnetic stir bar to which a Teflon and silicone rubber disc was attached (Fig. 1). Six removable porcelain ceramic discs were inserted into each base. These growth surfaces (Tyler Research Instruments, Edmonton, Alberta, Canada) were 1·27 cm in diameter, 0·4 cm thick and mildly bead blasted to promote bacterial adhesion. Reactor vessels were placed on magnetic stirrers to drive the rotors, providing complete mixing of the liquid contents. The rotors were calibrated with a tachometer to spin at 500 rev min^{-1}. Nutrient solutions were pumped into the vessels while tap water was gravity fed using solenoid valves to control flow. Reactors were operated open to the air.

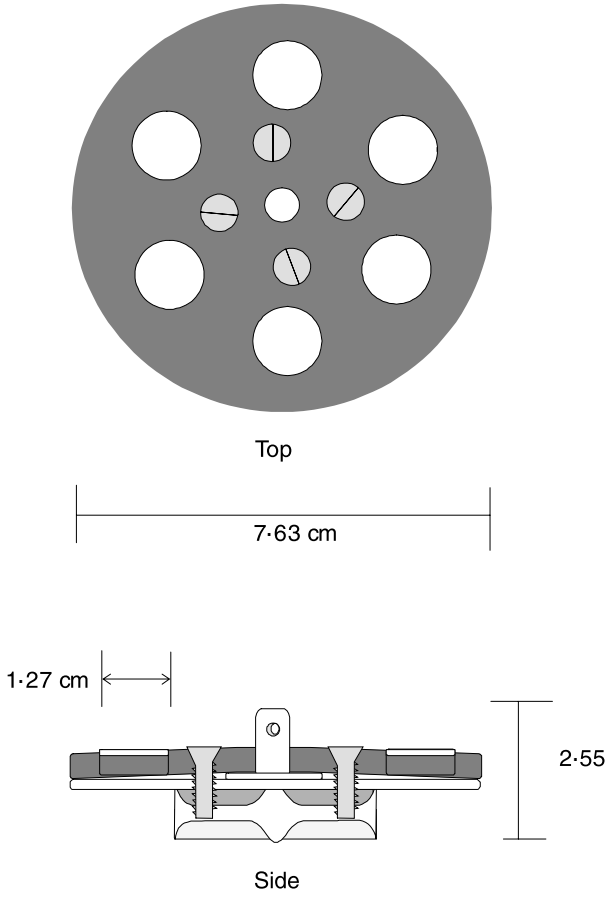

Fig. 1 Rotating disc reactor rotor

Bacterial inoculum and procedures

Overnight cultures of four toilet bowl isolates (Pitts *et al.* 1998) were combined with sterile glycerol and peptone such that the final concentrations were 20% and 2%, respectively. This mixture was frozen in 2-ml volumes. For reactor inoculation, a 2-ml vial was thawed and 20 μl added to a reactor vessel filled with 400 ml 10 g l^{-1} tryptic soy broth (TSB; Difco, Detroit, MI, USA). The reactor stir plate was turned on at 500 rev min^{-1} for a 4-h attachment period, after which reactors were operated under cyclic fill-and-draw conditions.

Once per hour, a concentrated sterile TSB nutrient solution was added to the 400 ml working volume to produce a final concentration of 1 g l^{-1}. The TSB was prepared at 120 g l^{-1} with 80 g l^{-1} NaCl added to inhibit the growth of contaminating organisms in the supply bottle and plumbing. After 5 min, the nutrient-containing solution was stirred and drained. When the vessel was empty, a fresh supply of refill tap water, with or without added sodium hypochlorite, was delivered to a final volume of 400 ml. This cycle, controlled by a programmable timing module (Chrontrol, San Diego, CA, USA), was repeated once per hour for the duration of the experiment.

Chlorine solutions were prepared by diluting household bleach with tap water. Fresh solutions were prepared daily during an experimental run and solution concentrations were assayed for total chlorine immediately after preparation and again 24 h later, prior to being discarded. Total chlorine was determined using the N,N-diethyl-*p*-phenylenediamine colourimetric method and absorbance measurements were taken with a spectrophotometer. Each experiment included at least two untreated control reactors, in which chlorine was neutralized in the incoming tap water by addition of sodium thiosulphate. Residual chlorine was assayed by the above method.

Discs were sampled for biofilm accumulation at regular intervals over the experimental period. The disc positions in each reactor were numbered and the numbers randomly permuted to determine the order in which discs were removed from that reactor. Removed discs were not replaced during the experiment. Once removed from reactors, disc surfaces were scraped until dry with sterile wooden

applicator sticks which were then stirred vigorously in sterile phosphate-buffered saline. The resulting bacterial suspensions were prepared for viable plate counting as described in Pitts *et al.* (1998). Appropriate dilutions were spread on R2A agar (Difco) plates in triplicate and incubated at room temperature for 6 d. After 6 d colonies were counted, converted to colony-forming units per square centimetre (cfu cm^{-2}) and transformed to a $\log_{10}$ scale. The average of the log-transformed densities over the three plates provided a measure of log areal viable cell density.

Four separate experiments were run. Table 1 gives the number of reactors used for each chlorine concentration in each experiment as well as the durations of the experiments.

The experiments were designed to provide estimates of the three main components of variability: the variability among discs within reactors, the variability among reactors within experimental runs and the variability among experimental runs. Thus, it was necessary to sample (at times) at least two discs from the same reactor and, for each chlorine concentration, to use at least two reactors in an experimental run. In each experiment, duplicate reactors were run at each chlorine concentration (except at the 27 mg l^{-1} concentration in experiments 2 and 3). Duplicate discs were occasionally sampled within reactors. Each chlorine concentration was tested in at least two different experiments, allowing assessment of the variability between experiments.

Statistical methods

The goals of the statistical analysis were to (i) provide summaries of the effects of time and hypochlorite concentrations on viable cell densities, (ii) determine the extent to which such results can be repeated from experiment to experiment and (iii) compare the data collected from the reactor system with data collected from real toilet bowls.

For a specified reactor, let $y(t)$ denote the density at time t (d) and $L(t)$ denote the mean of log densities at time t.

Descriptive statistics. For each concentration of chlorine in each experiment, a plot of $L(t)$ against t was constructed. A smooth growth curve is described by the empirical model of eqn 1

$$L(t) = \beta_0 - \beta_1 e^{-t/\theta} \quad (1)$$

where $\beta_0 \geq \beta_1 \geq 0$, $\theta \geq 0$.

Equation 1 describes a concave curve that reaches a plateau as time increases. The parameter β_0 represents a horizontal asymptote (or upper bound for $L(t)$) as t gets large. The quantity $\beta_0 - \beta_1$ is the log areal cell density at time 0. The parameter θ determines the shape of the curve and its inverse is a measure of the relative (to steady state density) log growth rate.

Equation 1 was fitted to the plotted points by calculating the non-linear least squares regression estimates of the coefficients β_0, β_1 and θ. The curves provide a reasonably accurate visual summary of the observations.

It was expected that, during the first few days of the experiment, biofilm cells would accumulate according to the exponential model of eqn 2

$$y(t) = y(0)e^{bt} \quad (2)$$

where b is positive.

Equation 2 implies that $L(t)$ is a linear function of t, where the line has a slope of $0{\cdot}4343b$. For each chlorine concentration in each experiment, b was estimated by least squares regression of the observed response variable $L(t)$ on the explanatory variable t, for t not large (range of t specified below).

Four quantitative response variables were submitted to statistical analysis. The four variables, BA2, BAE, SFR and LR, are defined as follows:

- **i** BA2, biofilm accumulation by day 2 = $L(2)$;
- **ii** BAE, biofilm accumulation at endpoint, which is the average of $L(t)$ for t from 5 to 11 d;
- **iii** SFR, specific fouling rate, which is the coefficient b of eqn 2 calculated by least squares regression for t between days 2 and 6 and
- **iv** LR, log reduction for each chlorine concentration, which is the BAE for chlorine-treated reactors subtracted from the BAE for the associated untreated control reactors.

The quantity BA2 was chosen as a measure of biofouling because the growth curve plots showed that the second day of the experiment was critical. Any difference in biofilm accumulation between untreated control reactors and chlorine reactors was manifest by day 2; afterwards the growth curves were essentially parallel. Growth curve plots showed that the log density of viable cells remained relatively constant after 5 d. Therefore, BAE was chosen as the endpoint measure of biofouling. For eqn 2, the derivative of $y(t)$ at $t = 0$ is $by(0)$. Thus SFR × 100 is the percentage increase in the biofilm bacterial population per day at the start of the growth curve. The LR is commonly used as a measure of the efficacy of antimicrobial agents. It was necessary that this biofilm growth system repeatedly produce the same LR for each chlorine concentration to demonstrate its utility in evaluating antimicrobial agents.

Table 1 Experimental design for laboratory reactors

Chlorine (mg l^{-1})	Expt 1 (9 d)	Expt 2 (5 d)	Expt 3 (11 d)	Expt 4 (10 d)
0	6	3	3	4
3	0	2	2	0
9	0	2	2	2
27	0	1	1	2

Results are shown as the number of reactors operated at each chlorine concentration in the four experiments.

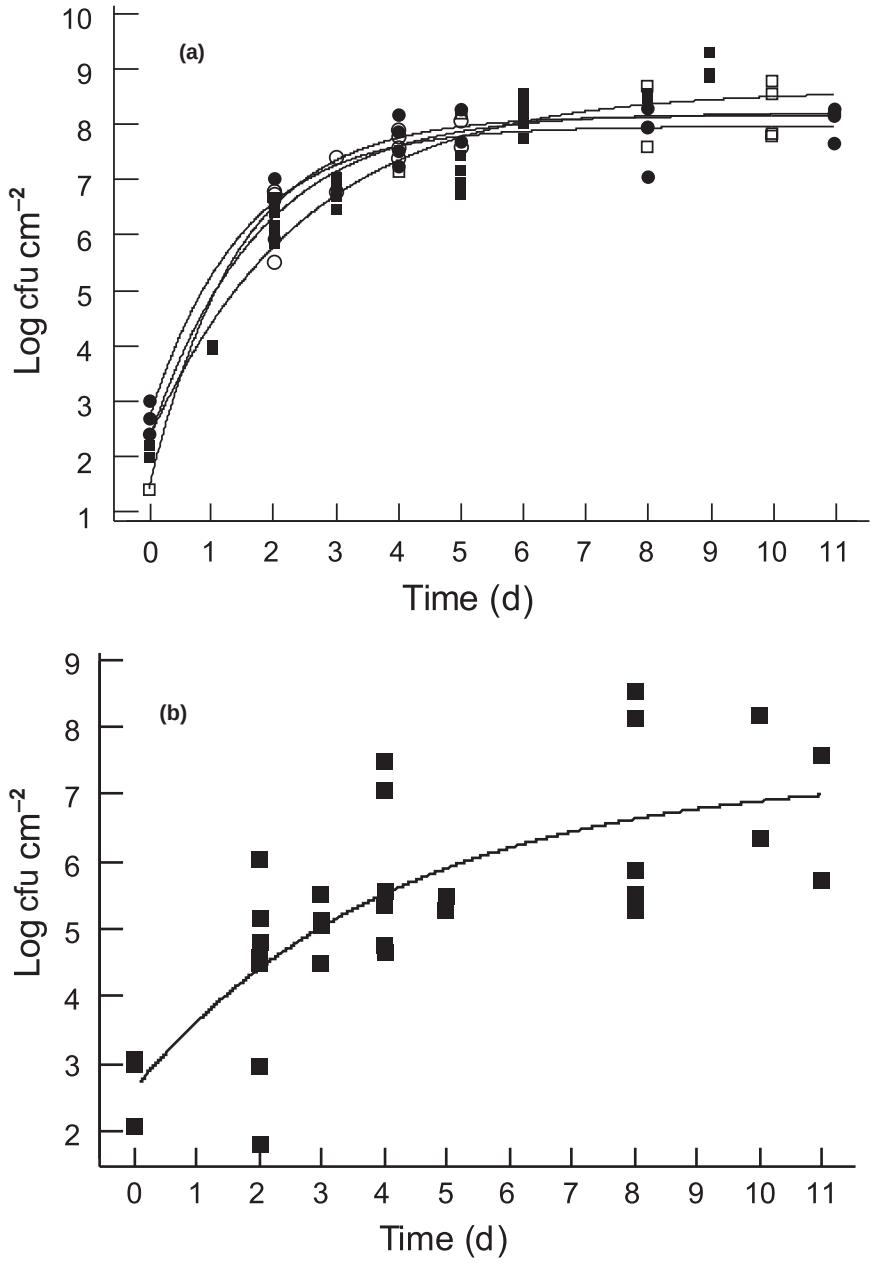

Fig. 2 (a) Biofilm accumulation curves for untreated reactors. Data points from expts 1–4 are represented by ●, ○, □ and ■, respectively. (b) Biofilm accumulation curves for 9 mg l^{-1} chlorine reactors. Each data point is a single coupon

For each experiment and chlorine concentration (3, 9 and 27 mg l^{-1}, where available), a measure of LR was computed as the difference in average log areal cell density between the treated reactors and the untreated control reactors. A concentration–response curve was constructed by plotting the computed log reduction values against the chlorine concentrations.

Repeatability assessment. It is of considerable interest to know the extent to which the measurements will be similar in repetitions of an experiment. Following Helrich (1990), we used the standard deviation of a measurement across repetitions of an experiment to indicate repeatability. A repeatability standard deviation of zero would indicate that the result is perfectly repeatable and a large repeatability standard deviation would indicate that the result is not repeatable. The standard deviation has the same units as the measure of interest (e.g. $\log_{10}$ cfu cm^{-2} for BA2). The repeatability standard deviation can be divided by the mean measurement to obtain the coefficient of variation, a dimensionless, standard version of repeatability. The coefficient of variation is useful for comparing systems that do not use the same units of measurement. For example, it is sometimes useful to

compare planktonic systems, where measurement units may be cfu ml^{-1}, with biofilm systems, where measurement units are cfu cm^{-2}.

The repeatability standard deviation includes variability within experiments as well as variability between experiments. The variability within experiments can be further decomposed into variability between reactors (within experiments) and variability between discs (within reactors). Data from repetitions of the experiment were submitted to an analysis of variance to estimate each of the three components of variance and the repeatability standard deviation was then defined as the square root of the sum of the component variances, i.e. the square root of the total variance.

For BA2 and BAE, estimates of the three variance components were obtained by the method of restricted maximum likelihood using the varcomp function in S-Plus (Insightful Corp., Seattle, WA, USA) (Venables and Ripley 1994). For SFR, the variance components were estimated using a method of moments approach described in Hyde (1980). Because LR is the difference of two independent BAE measurements, it was possible to calculate the repeatability standard deviation for LR directly from the variance components for BAE. The LR repeatability calculations are for an assay in which two discs are sampled from each of two reactors at each time point.

Comparison to toilets. For the toilet data, a plot of $L(t)$ *vs* t was constructed, where t is time in weeks. This plot was overlaid on the plot of $L(t)$ *vs* t for the reactor untreated control data, where t is time in days. The empirical growth curve (1) was fitted separately for the toilet and reactor data.

For both of the treated toilets, log reduction values were computed by subtracting the average log areal cell density at the endpoint for the treated toilet from the average log areal cell density at the endpoint for the same toilet untreated. These two data points were added to the dose–response curve constructed from the reactor system data. Finally, for additional comparison (and calibration) with the reactor data, the four chlorine concentrations were ranked on the measures BA2, BAE, SFR and LR for both the toilet and reactor data.

RESULTS

Descriptive statistics

Biofilm accumulated in untreated laboratory reactors over a period of days to endpoint densities of approximately 10^8 cfu cm^{-2} (Fig. 2a). The accumulation curves (eqn 1) for the four experiments were nearly identical. Surprisingly, biofilm also accumulated in laboratory reactors receiving 3 or 9 mg l^{-1} (Fig. 2b) chlorine continuously. Even treatment with 27 mg l^{-1} chlorine failed to completely suppress biofilm formation. The data for chlorine-treated reactors exhibited greater spread about the fitted growth curves than was observed for the untreated control (Fig. 2a), but the curves were similar between experiments.

Repeatability assessment

The amount of biofilm accumulation, as measured by BA2 and BAE, decreased steadily with increasing chlorine concentration (Table 2). For untreated control experiments, the repeatability standard deviations were 0·47 and 0·66 (log_{10} cfu cm^{-2}) for BA2 and BAE and the among experiments variance components were generally negligible. For BA2 and BAE, the repeatability standard deviations were mostly higher for chlorine-treated than for untreated control reactors. The variability among experiments was generally a minor component of total variability, although BA2 at 3 mg l^{-1} chlorine was a notable exception. The maximum repeatability standard deviations occurred at 9 mg l^{-1} chlorine.

The same mean SFR of 1·1 d^{-1} was measured for both the untreated control and 9 mg l^{-1} chlorine. In other words, the biofilm grew with a 110% per day increase in bacterial densities, which corresponds to a population doubling time of 22 h. The SFR repeatability standard deviation for the 9 mg l^{-1} chlorine experiment was three times larger than for the untreated control experiment. Overall, SFR variability

Table 2 Summary of results for biofilm accumulation at day 2 (BA2), at endpoint (BAE) and log reduction (LR) in laboratory reactors

	Chlorine (mg l^{-1})			
Summary statistics	0	3	9	27
Mean BA2 (log cfu cm^{-2})	6·4	5·4	4·3	3·2
Repeatability standard deviation	0·47	0·88	1·34	0·67
Coefficient of variation (%)	7	16	31	21
Among-experiment variability (%)	0	89	6	0
Among-reactor variability (%)	40	0	33	0
Within-reactor variability (%)	60	11	61	
Mean BAE (log cfu cm^{-2})	8·0	7·3	6·5	3·9
Repeatability standard deviation	0·66	0·61	0·98	0·88
Coefficient of variation (%)	8	8	15	23
Among-experiment variability (%)	0	0	0	36
Among-reactor variability (%)	27	26	89	48
Within-reactor variability (%)	73	74	11	16
Mean LR		0·7	1·5	4·1
Repeatability standard deviation		0·51	0·88	0·80
Coefficient of variation (%)		73	59	20
Among-experiment variability (%)		0	0	44
Among-reactor variability (%)		42	82	39
Within-reactor variability (%)		58	18	17

The variance components are shown as percentages of the total variability.

among experiments was about equal to variability within experiments.

Table 2 also shows the summary statistics for LR. The LR values increased consistently with increasing chlorine concentration.

Comparison to toilets

Biofilm accumulated much faster in laboratory reactors than it did in toilet bowls. This does not necessarily represent a deficiency in the laboratory model, as one of the desired features of laboratory methods is that they be rapid. The field and laboratory data differed on the time scale by about a factor of 7. Thus, when the biofilm reactor data are plotted in days and the corresponding toilet bowl data in weeks, the accumulation curves have similar slopes (Fig. 3). The same acceleration factor of 7 was evident when the laboratory reactors and toilet bowls were treated with 9 mg l^{-1} chlorine (Fig. 4).

The concentration–response curve of log reduction *vs* chlorine concentration was approximately linear (Fig. 5). The two data points collected from treated toilets fell close

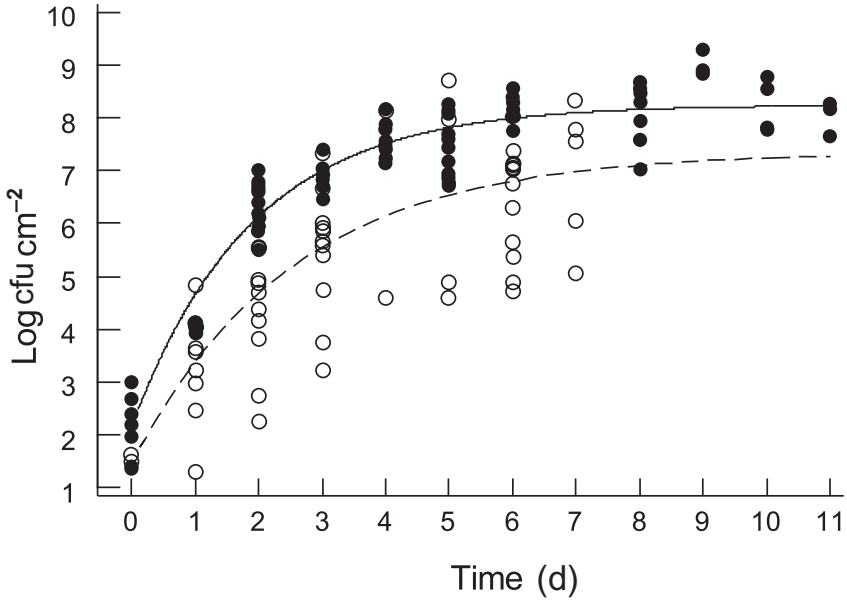

Fig. 3 Comparison of biofilm accumulation curve for all untreated control reactors (●) and for all untreated toilets (○). The time scale is in days for the reactors and weeks for the toilets

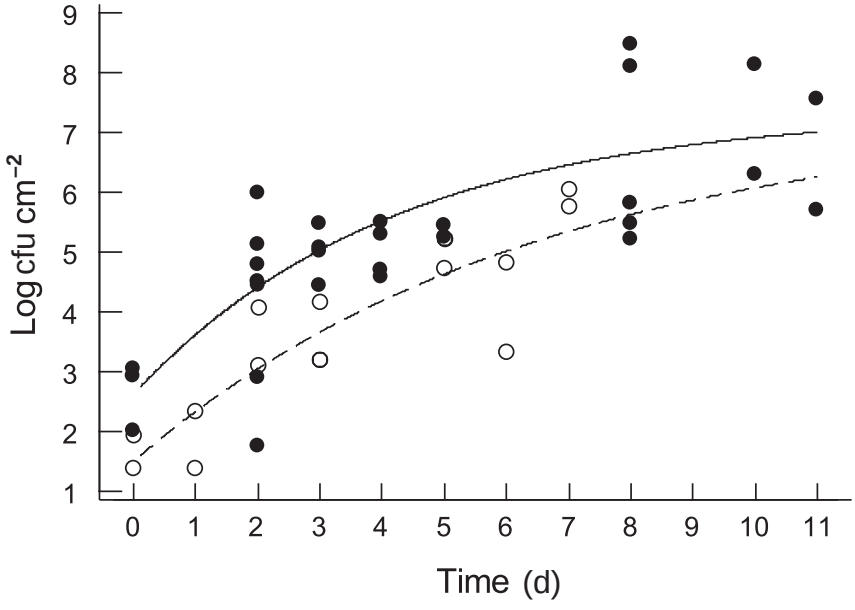

Fig. 4 Comparison of biofilm accumulation curve for 9 mg l^{-1} chlorine-treated reactors (●) and for 9 mg l^{-1} chlorine-treated toilet (○). The time scale is in days for the reactors and weeks for the toilet

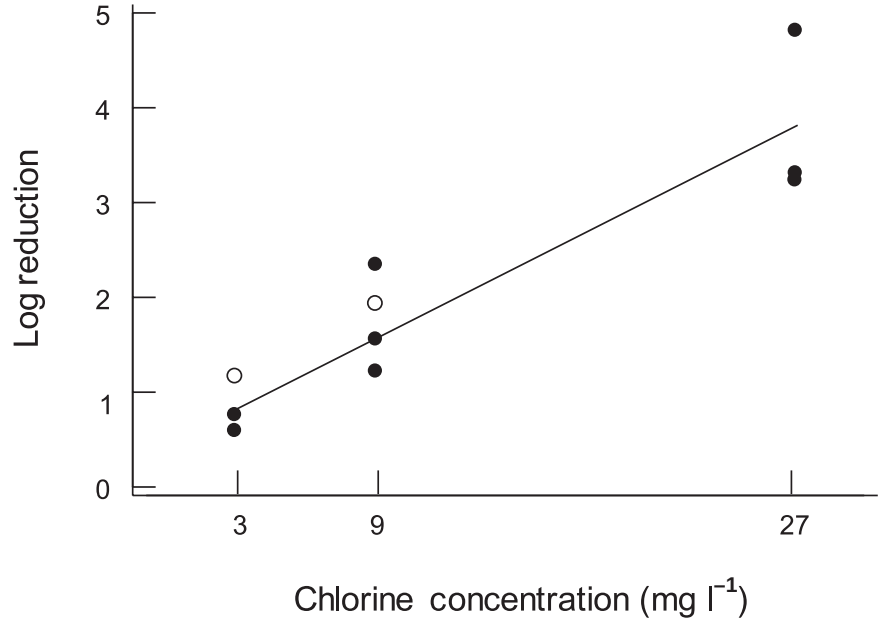

Fig. 5 Concentration–response curve (log reduction *vs* chlorine concentration) for reactors (●) and toilets (○)

Table 3 Comparison of treatments in reactors and toilets on biofilm accumulation at endpoint (BAE), at day 2 (BA2), specific fouling rate (SFR) and log reduction (LR)

	Chlorine (mg l^{-1})					
	0		3		9	
Summary statistics	Toilet	Reactor	Toilet	Reactor	Toilet	Reactor
BAE (log cfu cm^{-2})	6·8	8·0	6·4	7·3	5·0	6·5
BA2 (log cfu cm^{-2})	4·1	6·4	4·8	5·4	3·6	4·3
SFR (d^{-1})	0·26	1·1	0·29	1·4	0·25	1·1
LR	N/A	N/A	1·6	0·7	2·3	1·5

to the least squares regression line calculated from the laboratory reactor data.

Rankings of the three treatments, 0, 3 and 9 mg l^{-1} chlorine, on BAE, BA2, SFR and LR are given in Table 3. The toilet and reactor data agree on most of these rankings.

DISCUSSION

These results demonstrate that the biofilm growth system described herein can provide repeatable assays of the efficacy of antimicrobial agents against biofilms grown in intermittent flow conditions. A relatively small repeatability standard deviation was achieved for LR measurements. The standard deviations for LR are within the range of standard deviations, 0·2–1·2, observed for standard suspension and hard surface disinfectant assays (Tilt and Hamilton 1999). We expect that this biofilm growth system would produce even better repeatability of LR values if more than two reactors were simultaneously operated at each treatment level, including the untreated control.

Although the methods described in this article were tailored to simulate the toilet bowl environment, the biofilm reactor system is very flexible and could be readily modified to simulate many different situations. It can be run with no flow, intermittent flow, continuous flow or drain and fill. The relevant reactor environment can be selected by choosing the growth medium, gaseous environment (which requires sealing the top of the reactor), disc rotation speed, nutrient delivery rate and microbial inoculum. The biocide can be applied before, during or after biofilm formation and either in pulses or continuously. The growth surfaces can be made of biocide-treated material.

The persistence of microbial biofilms in chlorinated water is well known in the drinking water and cooling water industries. Drinking water distribution systems are not sterile even when continuously chlorinated. It is suspected that biofilms are responsible for the persistence of bacterial growth in the presence of the free chlorine residuals maintained in these systems, from a fraction of a mg l^{-1} to perhaps 2 mg l^{-1}. van der Wende *et al.* (1989) showed that continuously delivered concentrations of free chlorine, in the range 0·2–0·8 mg l^{-1}, effectively reduced biofilm accumulation in a simulated drinking water system. LeChevallier *et al.* (1990) reported that, while a 1-mg l^{-1} chlorine residual was sufficient to inactivate biofilm bacteria on polyvinyl chloride surfaces, a 3 mg l^{-1} residual was unable to reduce viable counts on iron pipes. Corrosion products were thought to contribute to the protection of bacteria on iron surfaces. Since there were no iron corrosion products in our biofilm reactor system

some other explanation is required for the ability of biofilms to accumulate readily in water containing 3 mg l^{-1} chlorine and to persist in water containing up to 27 mg l^{-1} chlorine.

REFERENCES

Bradshaw, D.J., Marsh, P.D., Schilling, K.M. and Cummins, D. (1996) A modified chemostat system to study the ecology of oral biofilms. *Journal of Applied Bacteriology* **80**, 124–130.

Brown, M.R.W. and Gilbert, P. (1993) Sensitivity of biofilms to antimicrobial agents. *Journal of Applied Bacteriology* **74**, 87S–97S.

Costerton, J.W., Stewart, P.S. and Greenberg, E.P. (1999) Bacterial biofilms: a common cause of persistent infections. *Science* **284**, 1318–1322.

Gilbert, P., Das, J. and Foley, I. (1997) Biofilm susceptibility to antibiotics. *Advances in Dentistry Research* **11**, 160–167.

Green, P.N., Bousfield, I.J. and Stones, A. (1987) The laboratory generation of biofilms and their use in biocide evaluation. In *Industrial Microbiological Testing* eds Hopton, J.W. and Hill, E.C. pp. 99–108. Oxford: Blackwell Scientific Publications.

Green, P.N. and Pirrie, R.S. (1993) A laboratory apparatus for the generation and biocide efficacy testing of *Legionella* biofilms. *Journal of Applied Bacteriology* **74**, 388–393.

Helrich, K. (ed.) (1990) Appendix: Guidelines for Collaborative Study Procedure to Validate Characteristics of a Method of Analysis, Section 6.5. In *Official Methods of Analysis of the Association of Official Analytical Chemists,* 15th edn. p. 681. Arlington, VA: AOAC.

Hyde, J. (1980) Determining an average slope. In *Biostatistics Casebook* ed. Miller, R.G., Efron, B., Brown, B.W. and Moses, L.E. 171–189. New York: Wiley.

LeChevallier, M.W., Lowry, C.D. and Lee, R.G. (1990) Disinfecting biofilms in a model distribution system. *JAWWA* **82**, July 87–99.

Pitts, B., Stewart, P.S., McFeters, G.A., Hamilton, M.A., Willse, A. and Zelver, N. (1998) Bacterial characterization of toilet bowl biofilms. *Biofouling* **13**(1), 19–30.

Tilt, N. and Hamilton, M.A. (1999) Repeatability and reproducibility of germicide tests: a literature review. *Journal of AOAC International* **82**(2), 384–389.

Venables, W.N. and Ripley, B.D. (1994) *Modern Applied Statistics with S-Plus* New York: Springer-Verlag.

van der Wende, E., Characklis, W.G. and Smith, M.A. (1989) Biofilms and bacterial drinking water quality. *Water Research* **23**, 1313–1322.