



Biomass distributions, activity, growth, and carbon utilization in heterotrophic bacterial communities
by Brian David Ellis

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Microbiology

Montana State University

© Copyright by Brian David Ellis (1999)

Abstract:

Bacteria can play a deleterious or beneficial role in the distribution or treatment of safe drinking water. The main mode of growth is sessile, attached to the distribution system pipelines or filter media, respectively. This report describes efforts to understand the activity of microorganisms growing in drinking water.

Methods including microscopy, activity measurements, carbon analysis, and rigorous statistical procedures were used to investigate bacterial growth parameters in model biofilm reactors (annular reactors). Amino acids, carbohydrates, and humic material were investigated as model growth compounds at three different concentrations. Control and chlorinated reactors were operated in parallel. Biologically filtered drinking water was used as inoculum and dilution water for the reactors. Long-term experiments were run, permitting for development of steady state conditions within the reactors.

Biomass/biovolume distributions of heterotrophic bacteria resulting from colonization of the reactors were investigated using mathematical probability distribution analyses. Sixty-eight of seventy-two datasets comprising approximately 36,000 individual cell volume measurements were found to follow a 3-parameter generalized Pareto distribution. The average biomass (mass/volume) conformed to the same distribution. These results show that biomass increases with decreasing cell size. An empirical non-linear model was developed that takes into account practical and theoretical limits for bacterial size to convert biovolume to biomass.

Activity measurements by microautoradiography and a suite of fluorescent stains were in good agreement. Substrate type had a larger effect on activity than did substrate concentration and chlorination had the largest impact.

Bacterial carbon production ranged from 0.003 - 1.74 grams carbon/L•day, yields from 0.034 - 0.25 grams cell carbon/gram carbon, and doubling times from 1 - 15.4 days. Specific growth rates and yields decreased with increasing carbon concentration. Doubling times were lower in the presence of chlorine and independent of the bulk fluid substrate concentration in both chlorinated and control biofilms. Therefore, biofilm growth was zero order. Although humic material is generally considered recalcitrant we found that control and chlorinated biofilm communities could remove 78% and 58% of the influent humic concentration at steady state and a hydraulic detention time of 2.1 hours. Biological filtration was as effective as chlorine in controlling bacterial growth in the annular reactors.

**BIOMASS DISTRIBUTIONS, ACTIVITY, GROWTH, AND CARBON
UTILIZATION IN HETEROTROPHIC BACTERIAL COMMUNITIES**

by

Brian David Ellis

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

March 1999

D378
EL 589

APPROVAL

of a thesis submitted by

Brian David Ellis

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Gordon A. McFeters

G.A. McFeters 4/6/99
(Signature) (Date)

Approved for the Center for Biofilm Engineering

Dr. Anne K. Camper

Anne K. Camper 4/6/99
(Signature) (Date)

Approved for the Department of Microbiology

Dr. Seth H. Pincus

Seth H. Pincus 4/6/99
(Signature) (Date)

Approved for the College of Graduate Studies

Dr. Bruce R. McLeod

Bruce R. McLeod 4-6-99
(Signature) (Date)

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature B. D. Ellw

Date 04/05/99

ACKNOWLEDGEMENTS

I thank Gordon McFeters and Anne Camper for their support, enthusiasm, and interest in my project. Their encouragement and knowledge of the subject matter has significantly helped in the completion of the work presented here. I would also like to thank my committee members, Dr. Warren Jones, Dr. Dave Ward, Dr. Bill Costerton, and Dr. John Lisle.

Thank you to Phil Butterfield for your help in running these experiments. This work could not have been performed without your aid. As well, I would like to thank Marty Hamilton for his patience and unending help when I came to him with questions on the statistical analyses performed in this dissertation.

Thank you to all of the graduate students, both from Microbiology and the Center for Biofilm Engineering, who made my experience here more rewarding.

Lastly, I would like to thank my parents, Doug and Jeannine Ellis, who have encouraged me in studies throughout my life.

This research was supported by the Center for Biofilm Engineering at Montana State University, a National Science Foundation-sponsored engineering research center (cooperative agreement ECD-8907039) and the American Water Works Research Foundation.

TABLE OF CONTENTS

Chapter	Page
1. INTRODUCTION.....	1
2. BIOMASS DISTRIBUTIONS IN HETERTROPHIC BACTERIAL COMMUNITIES.....	5
Introduction.....	5
Materials and Methods.....	11
Experimental System.....	11
Nutrient Amendments.....	14
Reactor Operation.....	15
Bacterial Counts.....	16
Biovolume Determinations.....	16
Determination of Cell Carbon.....	17
Statistical Analyses.....	19
Results.....	24
Distribution Identification.....	24
Biovolume and GPD Parameter Analysis.....	29
Relationship Between Biovolume and Carbon per Biovolume.....	39
Discussion.....	42
Probability Distributions.....	42
Biovolume Conversion.....	44
Treatment Effects.....	51
Carbon Source and Concentration.....	51
Chlorination Effects.....	51
Conclusions.....	54
Appendix I.....	56
Moments and their Meanings.....	56
L-moments.....	58
Probability Weighted Moments (PWMs).....	59
Direct Estimators for L-Moments.....	60
Generalized Pareto Distribution.....	62
Data Tables.....	65
References Cited.....	68

3. RELATIONSHIPS BETWEEN TOTAL AND ACTIVE CELLS IN HETEROTROPHIC BACTERIAL POPULATIONS SUBJECTED TO CONTROLLED CARBON SOURCES, CARBON CONCENTRATIONS, AND CHLORINATION.....	73
Introduction.....	73
Material and Methods.....	78
Experimental System.....	78
Sampling Protocol.....	79
Fluorescent Activity Stains.....	79
Microautoradiography (MAR).....	80
Statistical Analyses.....	82
Activity Determination.....	83
Results.....	85
Discussion.....	98
Statistical Models.....	98
Activity Measurements.....	99
Conclusions.....	105
References Cited.....	107
Appendix II.....	111
Statistical Model for Active and Total Cells/mL	111
Summarized Data.....	112
4. EFFECTS OF CARBON SOURCE, CARBON CONCENTRATION, AND CHLORINATION ON GROWTH RELATED PARAMETERS OF HETEROTROPHIC BIOFILM BACTERIA.....	114
Introduction.....	114
Materials and Methods.....	118
Experimental Overview.....	118
TOC Measurements.....	120
Carbon Balance Across Reactors.....	120
Carbon Source Utilization.....	122
Influent and Effluent Cell Sampling.....	123
Specific Growth Rates, Yield, and Bacterial Carbon Production.....	123
Biofilm Sampling.....	125
Statistical Analyses.....	126
Results.....	127
GAC/BAC Filter and BDOC Column Performance.....	127

Control Stage (500 ppb Carbon Feed Level) Annular Reactors.....	128
Effluent.....	128
Biofilm.....	129
Test Stage (1000 and 2000 ppb Carbon Feed Levels) Annular Reactors.....	130
Biofilm.....	131
Effluent.....	133
Growth Rates, Yields, and Bacterial Carbon Production.....	135
Carbon Source Utilization by Biolog Plates.....	140
Comparative Analysis of Methods used for Growth Rate Estimations.....	144
Carbon Substrate Removal.....	145
Discussion.....	148
Control Stage Carbon Level.....	149
Test Stage Carbon Levels.....	151
Growth Rates, Production, and Yield.....	156
Control Communities.....	158
Chlorinated Communities.....	161
Comparison of Growth Rates Measured by Different Methods.....	164
Conclusions.....	165
References Cited.....	167
5. SUMMARY.....	174

LIST OF TABLES

Chapter 2	Page
Table	
2.1. Comparison of distribution type, parameters, and method of analysis.....	23
2.2. Mean and median volumes for each of the carbon sources at both 1000 ppb C and 2000 ppb C. Control refers to the unchlorinated reactor. # - numbers in parentheses are standard deviations. n* - is the number of experiments and the values in parenthesis are the number of individual datasets. See text for further details.....	30
2.3. Parameters of the Pareto distribution for biovolume. Numbers in parenthesis refer to \pm standard deviation. See Table 2.2 for n values and text for further details. 1, $p = 0.004$ and 2, $p = 0.003$. All other values are not significantly different between carbon levels of 1000 ppb and 2000 ppb. P-values were determined by a GLM ANOVA (unbalanced) at an α -level of 0.05 and the carbon levels were the factors.....	31
2.4. Results from an ANOVA performed on the normalized (each volume divided by the smallest volume) volume measurements. Reported values are means \pm standard deviation. Shows that the carbohydrates were the only substrates that produced a significantly different response in the control versus chlorinated reactor. The inability to induce homogeneity of variance for the parameter β (carbohydrates = *) did not allow for an ANOVA analysis of the effect of chlorination in this system. However, non-parametric analyses (Mood's median and Kruskal-Wallis) indicated that a significant difference ($p < 0.0005$) was apparent for this parameter. Both α and κ were found to be significantly different for carbon source type ($p < 0.0005$ for both). na – not applicable, see Table 2.3 and text for further details.....	34
2.5. Normalized parameters for the distribution of volume for the two different type II Pareto distributions. Greek letters refer to the generalized Pareto distribution (GPD) and English letters refer to the modified original Pareto distribution. AU^2 is the modified Anderson-Darling test statistic and SE is the fit standard error determined by non-linear regression using TableCurve 2D. r^2 -values correspond to the goodness of fit statistic on the same horizontal line. Note that normalization does not change either of the shape parameters κ or C	35

- 2.6. Dataset from Figure 2.3, Panel E and conversion to average C_V and total carbon in the sample. 1- the equation is $C_C = e(1.12 \times \ln(\text{Vol}) + 4.28)$ and division by volume gives C_V . 2 – the equation is $C_V = 3.05 \times (10^{-4.952} \times \text{Vol}) \times 1000$. 3 – the equation is $C_V = 25.73 \times (\text{Vol}^{-0.981})$. 4 – the equation is $C_C = 0.09 \times (\text{Vol}^{0.9}) \times 1000$ and division by volume gives C_V . 5 – the equation is $C_V = 1/[0.0013 + 0.0299 \times (\text{Vol}^{1.5})]$. 6 and 7 are single value conversion factors and are the same by either method of calculation. Method 1 used the average biovolume to find C_V by the appropriate conversion factor. Method 2 used the conversion factor on all biovolumes and averaged these C_V values. Both values for C_V were then multiplied by the total biovolume of the 1097 cells measured..... 49
- 2.7. Relationship between the number of observations (m) and the number of items (k) that produce the increasingly large number of possible linear combinations involved in the estimates for the mean (k = 1), variance (k = 2), skewness (k = 3), and kurtosis (k = 4)..... 62
- 2.8. Individual data values determined from the biovolume distribution with humics as the carbon source. n – number of cells in dataset, [C]- carbon concentration in ppm, rt- 1 is control (unchlorinated), 2 is test (chlorinated), α - location parameter, β - origin parameter, κ – shape parameter, α^* - parametric variance for α , β^* - parametric variance for β , X^1 – computed mean biovolume, X^2 – observed mean biovolume, med – median (observed), λ_2^3 – L-moment measure of dispersion, σ^2 – parametric variance of the mean, t^3 – L-moment sample skewness, t^4 – L-moment sample kurtosis, τ_3 – L-moment parametric skewness, τ_4 – L-moment parametric kurtosis. * - value does not exist..... 65
- 2.9. Individual data values determined from the biovolume distribution with amino acids as the carbon source. See Table 2.7 for an explanation of symbols..... 66
- 2.10. Individual data values determined from the biovolume distribution with amino acids as the carbon source. See Table 2.7 for an explanation of symbols..... 67

Chapter 3

- 3.1. Fluorescent activity stains used. Final concentration refers to the concentration that samples were incubated in for the indicated amount of time. DiBAC₄(3) was stored in ethanol at -20°C. The Rh123 working solution was TE (10 mM Tris-HCl, 1 mM disodium EDTA, pH 7.5) and the stock solution was stored in phosphate buffered saline (PBS) at pH 7.5. * - BacLight was used as per the manufacturers instructions..... 80

3.2.	Summarized data according to carbon level, reactor, and substrate. 1 - rate is the turnover rate of labeled leucine in percentage of the total amount of added radioactivity, numbers in parenthesis are the standard deviation of the rate and r^2 -values respectively. 2 - colony forming units/mL($\times 10^5$) on R2A agar, parenthesis are standard deviations. 3 - amino acids, 4- carbohydrates, 5- mixture, and 6- humics, nd - not determined.....	92
3.3.	Pearsons Correlation matrix for the various measurements and factors. Substrates were ranked in the order: (1) amino acids, (2) carbohydrates, (3) humics, and (4) mixture. Reactors were ranked in the order: (1) control, (2) chlorinated. Numbers in parenthesis are p-values. ns - not significant.....	95
3.4.	Control reactor results for MAR experiments. 1 - the numbers in parentheses refer to the standard deviation (in %) of the slope and the r^2 value for the fitline respectively. 2 - the number in parenthesis is the standard deviation in percentage of total active cells.....	112
3.5.	MAR results for the chlorinated reactor. 1 - the numbers in parentheses refer to the standard deviation (in %) of the slope and the r^2 value for the fitline respectively. 2 - the number in parenthesis is the standard deviation in percentage of total active cells.....	113

Chapter 4

4.1.	Pearson product moment correlation coefficients (r^2) between total cell counts and culturable cells and the percentage of culturable cells normalized to total cells in reactor effluents at the two stages of analysis. 500 ppb C refers to the control level of the experiments and is the combined results for both reactors. Test level is the 1000 and 2000 ppb C loading levels. $p < 0.0005$ for all correlations except 1 - $p = 0.005$, 2 - $p = 0.037$, and 3 - $p = 0.011$. Numbers in brackets for r^2 are n-values and numbers in brackets for ratio are standard deviations in percent. n was the same for both r and ratio. Influent cells had a percentage culturability of 25 (21) and $r^2 = 0.349$ (185). Control effluents were not significantly different at the mid and high carbon loading levels.....	131
4.2.	SGRs for amino acids at the mid and high carbon feed levels calculated by different methods. One is the corrected growth rate and 2 is the uncorrected growth rate for fg C/ μm^3 , μm^3 , and fg C/cell. For cells, 1 is total cells and 2 is active cells. 3 - growth rate measured by dividing BCP ($\mu\text{g C}/\text{cm}^2\cdot\text{hr}$) by total biofilm biomass ($\mu\text{g C}/\text{cm}^2$).....	145
4.3.	Linear regression data for carbon flux ($\mu\text{g Carbon used}/\text{cm}^2\cdot\text{hr}$) as a function of influent substrate loading rate ($\mu\text{g Carbon}/\text{cm}^2\cdot\text{hr}$). The slope	

represents the fraction of the influent carbon loading removed across the reactors in the 2.1 hour detention time of reactor operation. n_1 is the number of experiments for which TOC data was obtained and n_2 is the number of data points used in the regression. Model II regressions were performed by the method of geometric means. 95 % confidence interval is for the slope of the regression. Results for the mean normalized specific carbon removal rates are reported at the level of substrate and reactor only.....

147

- 4.4. Mean substrate specific removal rate normalized to biofilm biomass. Values are backtransformed from the log transformed data. Analysis was performed separately on the chlorinated and control reactors. The substrate control was analyzed separately. Data used for the chlorinated reactor was 2 levels, 1000 and 2000 ppb C, after chlorination was initiated.....

147

LIST OF FIGURES

Chapter 1	Page
Figure	
1.1. Schematic of the experimental design used to evaluate substrate type, substrate concentration, and chlorination in this dissertation. Samples can be divided into control and chlorinated which can be nested within the next higher level of substrate concentration which contained three treatment levels. This middle treatment level of substrate concentration can also be nested within the upper treatment level of substrate type which also contains three treatment levels.....	3
Chapter 2	
2.1. Schematic of experimental setup. The circles with P inside refer to pumps. See text for further details.....	12
2.2. Carbon and chlorine dosing protocol for all experiments.....	15
2.3. Frequency distribution of cell volume as a percentage of the total number of cells measured. Note the highly skewed nature of the data. Panels A-C are for the control reactor and panels D-F are for the chlorinated reactors. Vertical reference line was arbitrarily set at $0.15 \mu\text{m}^3$ and the percentage of the population below this volume is indicated above the histogram. Each panel consists of the cumulative datasets for the sampling period (3 samples/dataset). Representative data from one experiment using each of the substrates at 1000 ppb C/L. Histogram width is $0.05 \mu\text{m}^3$	26
2.4. GPD probability plots. α , β , and κ were determined by probability weighted L-moments. Solid line is the fit and symbols are the data points. Panel A dataset is from a carbohydrate experiment at a carbon concentration of 1000 ppb C, chlorinated reactor. Panel B is data from the corresponding control reactor from the same sampling period. $AU^2 = 0.479$, $r^2 = 0.985$ for panel A and $AU^2 = 1.368$, $r^2 = 0.984$ for panel B. Both the modified Anderson-Darling test statistic and r^2 values refer to the fit obtained from the linearized form for the generalized Pareto Distribution (see statistics section, equation (3)). Both the y-axis and x-axis are plotted on log scales. Note that the biovolume axis is different between the panels.....	27

- 2.5. Panel A is an L-moment diagram compiled from 72 datasets showing the overall trend of L-skewness vs. L-kurtosis for bacterial biovolume measurements. The solid line is the theoretical distribution of L-skewness vs. L-kurtosis for the GPD. Open symbols refer to individual datasets. Solid symbols refer to the means for each carbon type and the large cross refers to the grand mean for all data. Panel B – Generalized Pareto probability plot. α , β , and κ were determined by probability weighted L-moments. Dataset is compiled from all biovolume and carbon measurements. $AU^2 = 0.072$ and $r^2 = 0.9976$ for Panel B and refers to the linear form for the generalized Pareto distribution..... 28
- 2.6. Comparison of the raw and normalized data fits for the GPD1 (see equation 2) and the GPD2 (see equation 4). Refer to Table 2.5 for parameters and fits for the 2 graphs..... 36
- 2.7. Probability plots on a log-log scale from the average Pareto parameters for the control reactor (Panel A) and chlorinated reactor (Panel B). Horizontal lines at 10^{-1} , 10^{-2} , and 10^{-5} indicate the probability of finding a cell with a volume equal to or greater than the indicated volume. Volumes for the humics experiments are above and to the left of the curves, for the amino acids they are below and to the left of the curves, and for the carbohydrates they are above and to the right of the curves. Calculated mean biovolumes (observed mean biovolume) for amino acids, carbohydrates, and humics are 0.285 (0.260), 0.731 (0.753), and 0.202 (0.201) μm^3 for the control and 0.151 (0.147), 0.361 (0.278), and 0.131 (0.149) μm^3 for the chlorinated communities respectively. ESD is the equivalent spherical diameter or the diameter of a sphere with the indicated volume..... 37
- 2.8. Probability plots for biovolume and fg carbon/ μm^3 for the Pareto distribution derived from Table 2.5 and Figure 2.6. Plot is derived from all of the experimental data irrespective of treatments. Upper curve is fg carbon/ μm^3 and the lower curve is biovolume. Note that all axes are scaled in \log_{10} and no relationship between volume and C_V can be inferred from this graph. The calculated mean (observed mean) for C_V was 268 fg C/ μm^3 (268) and for biovolume the mean was 0.316 μm^3 (0.298)..... 38
- 2.9. Relationships between biovolume and C_V . Panel A shows the linear fit for the log-log transformed data, implying a power relationship. Four regression lines are shown: GM, geometric mean method, Major axis method, Bartlett's 3-group method, and ordinary least squares. Panel B shows the fitted relationship for predictive purposes when the data is grouped into biovolume classes (see text). The fit is from a non-linear regression performed on the weighted (1/variance) data, $n = 68$ 41

Chapter 3

- 3.1. Panel A, radioactivity counts per cell and Panel B, total cells per mL according to reactor, substrate, and carbon concentration. Error bars are 95% confidence intervals and plotted values are the back-transformed geometric means..... 88
- 3.2. Panel A shows the turnover rates and dpm/cell for the chlorinated and control reactors. Turnover rates and dpm/cell were significantly different ($p = 0.006$ and < 0.0005 , respectively) between control and chlorine. Panel B shows the total cells, plate counts and active cells for the chlorinated and control reactors. All values were significantly different in the two reactors ($p = 0.017$, < 0.0005 , and 0.005 respectively). Error bars are 95% confidence intervals..... 89
- 3.3. An example of the linear uptake of tritiated leucine in cells from the effluent of the carbohydrates-fed reactors at 2000 ppb carbon. Solid lines and diamonds refer to the percentage turnover (hr^{-1}) of the labeled leucine while the dashed line and squares refer to the normalized (dpm/cell) amount of label incorporated. Error bars are standard deviations. Panel A is for the chlorinated reactor while Panel B is for the control reactor..... 91
- 3.4. Panel A – active cells/mL in reactor effluents. Panel B shows the distribution of activity in the high and low activity classes. Results are back-transformed geometric means resulting from an ANOVA. Error bars are 95% confidence intervals..... 93
- 3.5. Activity distribution normalized to total cell counts and separated into high and low fractions according to treatments..... 94
- 3.6. Activity measurements by the various methods, normalized to the total cell counts as determined by DAPI staining. PI and DiBAC₄(3) counts were converted to active cells by subtraction from the total cells prior to normalization. Error bars are standard deviations..... 97
- 3.7. Activity stains as a function of the various combinations of live plus dead components of the individual methods. Error bars are standard deviations..... 98

Chapter 4

- 4.1. Cell counts in the effluent and biofilm from the control stage of the experiments separated by substrate type. Total cells/mL are significantly different for all substrates while the amino acids and carbohydrates are equivalent for culturable cells/mL (Tukey's pairwise comparisons). $n = 30$, 30, 20, and 5 for the amino acids, carbohydrates, humics, and mixture respectively for the cells/mL data. n for the biofilm data is 6 for all substrates

	except the mixture (n = 3). See text for further details.....	130
4.2	Total and culturable cells per square cm in the biofilm for the four substrates and two higher carbon feed levels. Reported values are means and standard deviations resulting from a GLM ANOVA on the separate reactors. Panel A is the control reactor and Panel B in the chlorinated reactor. Error bars are standard deviations. See text for further details.....	133
4.3.	Total and culturable cells per mL in reactor effluents. Panel A is control reactor effluent and Panel B is the chlorinated reactor effluent. Solid lines are total cells/mL and dashed lines are cfu/mL.....	135
4.4.	Line plots showing observed yield ($\mu\text{g cell C}/\mu\text{g C}$), bacterial carbon production (BCP, $\mu\text{g C}/\text{cm}^2\cdot\text{hr}$, $\times 10^3$), and specific growth rate (SGR, hr^{-1}) vs. nominal carbon feed concentration with respect to the separate substrates for the control reactors. Error bars are 95% confidence intervals. Data points are back-transformed means from the log transformed data. Values for n are 15, 15, 10, and 5 for the amino acids, carbohydrates, humics, and mixture respectively. P-values are the results from a one-way ANOVA (Tukey's pairwise comparisons) with carbon concentration as treatment levels (500, 1000, and 2000 ppb). Stars indicate where the significant difference arises between the carbon feed levels.....	137
4.5.	Line plots showing observed yield ($\mu\text{g cell C}/\mu\text{g C}$), bacterial carbon production (BCP, $\mu\text{g cell C}/\text{cm}^2\cdot\text{hr}$, $\times 10^3$), and specific growth rate (SGR, hr^{-1}) vs. nominal carbon feed concentration with respect to the separate substrates for the chlorinated reactors. Error bars are 95% confidence intervals. Data points are backtransformed means from the log transformed data. Values for n are 15, 15, 10, and 5 for the amino acids, carbohydrates, humics, and mixture respectively. P-values are the results from a one-way ANOVA with treatment levels of 1000 and 2000 ppb Carbon.....	138
4.6.	Dye reduction rates averaged over the four amino acids and four carbohydrates used for biofilm development. Y-axis is the substrate source used for biofilm development and the bars are the response of those cells to either the amino acids or carbohydrates present in the BIOLOG-GN microplates. Error bars are standard deviations.....	144

ABSTRACT

Bacteria can play a deleterious or beneficial role in the distribution or treatment of safe drinking water. The main mode of growth is sessile, attached to the distribution system pipelines or filter media, respectively. This report describes efforts to understand the activity of microorganisms growing in drinking water.

Methods including microscopy, activity measurements, carbon analysis, and rigorous statistical procedures were used to investigate bacterial growth parameters in model biofilm reactors (annular reactors). Amino acids, carbohydrates, and humic material were investigated as model growth compounds at three different concentrations. Control and chlorinated reactors were operated in parallel. Biologically filtered drinking water was used as inoculum and dilution water for the reactors. Long-term experiments were run, permitting for development of steady state conditions within the reactors.

Biomass/biovolume distributions of heterotrophic bacteria resulting from colonization of the reactors were investigated using mathematical probability distribution analyses. Sixty-eight of seventy-two datasets comprising approximately 36,000 individual cell volume measurements were found to follow a 3-parameter generalized Pareto distribution. The average biomass (mass/volume) conformed to the same distribution. These results show that biomass increases with decreasing cell size. An empirical non-linear model was developed that takes into account practical and theoretical limits for bacterial size to convert biovolume to biomass.

Activity measurements by microautoradiography and a suite of fluorescent stains were in good agreement. Substrate type had a larger effect on activity than did substrate concentration and chlorination had the largest impact.

Bacterial carbon production ranged from 0.003 - 1.74 grams carbon/L·day, yields from 0.034 - 0.25 grams cell carbon/gram carbon, and doubling times from 1 - 15.4 days. Specific growth rates and yields decreased with increasing carbon concentration. Doubling times were lower in the presence of chlorine and independent of the bulk fluid substrate concentration in both chlorinated and control biofilms. Therefore, biofilm growth was zero order. Although humic material is generally considered recalcitrant we found that control and chlorinated biofilm communities could remove 78% and 58% of the influent humic concentration at steady state and a hydraulic detention time of 2.1 hours. Biological filtration was as effective as chlorine in controlling bacterial growth in the annular reactors.

CHAPTER 1

INTRODUCTION

The overall goal of the project was to determine the key biodegradable organic matter (BOM) constituents responsible for bacterial growth in drinking water distribution systems. Three scales of observation were included in this research plan: (1) full-scale monitoring to identify treatment processes which are responsible for the removal of BOM, (2) pilot scale studies to determine the fate of these BOM compounds in conventional and advanced treatment processes, and (3) laboratory studies designed around the key BOM compounds responsible for microbial growth and determination of the kinetics of growth on these constituents. Synthetic BOM amendments were used in the laboratory scale experiments in the presence of natural organic matter (NOM). Choice of carbon substrate amendments were based on their role in bacterial metabolism and common presence in a number of aquatic environments such as lakes, rivers, and oceans.

The primary objective of the bench scale experiments at MSU was to measure the response of heterotrophic bacterial biofilm communities to different types and concentrations of BOM in the presence or absence of chlorine. In conjunction with this primary aim, the measurement, prediction, and modeling of heterotrophic biomass was a principal objective of this study and is reported in Chapter 2. Chapter 2 is titled "Biomass Distributions in Heterotrophic Bacterial Communities". The authors are Brian D. Ellis, Phillip Butterfield, Warren L. Jones, Gordon A. McFeters, and Anne K. Camper. An

abbreviated and condensed version of this chapter has been submitted to the journal *Microbial Ecology*.

The relation of total biomass to active biomass comprises Chapter 3. This Chapter is titled "Relationships Between Total and Active Cells in Heterotrophic Bacterial Populations Subjected to Controlled Carbon Sources, Carbon Concentrations, and Chlorination." This chapter will be submitted to the journal *Applied and Environmental Microbiology*. The authors are the same as those for Chapter 2.

Chapter 4 reports on the growth rates, bacterial carbon production, and observed yields of the communities subjected to the various treatment factors. This chapter is titled "Effects of Carbon Source, Carbon Concentration, and Chlorination on Growth Related Parameters of Heterotrophic Biofilm Bacteria" and the authors are the same as for Chapter 2. An abbreviated and condensed version of Chapter 4 has been submitted to the journal *Microbial Ecology*.

We suspected that a significant amount of sampling would be required to detect differences in the response of bacterial communities to the different treatments. As such, we attempted to structure the experiments so that they were amenable to more powerful statistical tests than can be ordinarily performed in studies involving environmental microbiology. Three major treatment levels were applied: (1) substrate type (amino acids, carbohydrates, and humics), (2) substrate concentration (500 ppb C, 1000 ppb C, and 2000 ppb C), and (3) a base level treatment of chlorination. All experiments were performed in parallel annular reactors with one acting as a control and the other receiving chlorine. This experimental design is presented in Figure 1.1 and is amenable to a nested

analysis of variance (ANOVA). With this experimental design, treatment evaluation could be compared both within and between the various treatments and their levels. Nesting of the sampling strategy refers to the fact that all lower level treatments can be analyzed within the level or levels of treatment above that level. Thus, sampling level is nested within substrate concentration (at 3 levels) and both sampling and substrate concentration are nested within substrate (also at 3 levels). This allowed us to analyze the differential effects of 3 levels of treatment (sampling level, substrate concentration, and substrate type) simultaneously or separately.

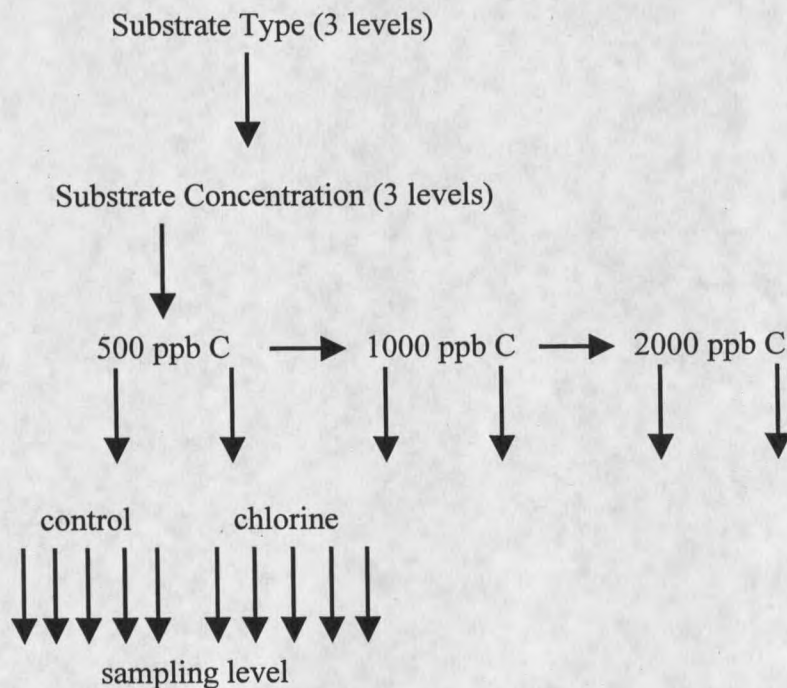


Figure 1.1 – Schematic of the experimental design used to evaluate substrate type, substrate concentration, and chlorination in this dissertation. Samples can be divided into control and chlorinated which can be nested within the next higher level of substrate concentration which contained three treatment levels. This middle treatment level of substrate concentration can also be nested within the upper treatment level of substrate type which also contains three treatment levels.

The experiments performed for this dissertation were funded by an American Water Works Association Research Foundation (AWWARF) grant to Dr. Anne Camper (Montana State University, Bozeman, MT), Dr. Peter Huck (University of Waterloo, Ontario, Canada), and Dr. Mark LeChevallier (AWWA Service Co., Voorhess, NJ). The project was titled "Investigation of Biological Stability of Drinking Water in Treatment Plants and Distribution Systems." The experiments performed at the Center for Biofilm Engineering, MSU, were the laboratory scale component of a larger study involving full scale treatment facilities in Canada and the United States.

CHAPTER 2

BIOMASS DISTRIBUTIONS IN HETEROTROPHIC BACTERIAL COMMUNITIES

Introduction

Bacteria play an important role in the turnover and cycling of elements, in addition to providing secondary carbon production in aquatic ecosystems. Microbial biomass is most commonly estimated from cell numbers and quite often from biovolume measurements (Norland, 1993; Bjornsen, 1986; Lee and Fuhrman, 1987; Nagata, 1986; Troussellier, et al., 1997; Fagerbakke, et al., 1996). Direct conversion of cell numbers to biomass through use of mass per cell or conversion of biovolume based on mass per volume are the standard methods of choice. In microbial communities comprising highly heterogeneous cell sizes, the number of volume measurements needed to adequately describe the distribution can become quite large. An ability to predict bacterial biomass that does not require an excessive number of observations of individual cells would be advantageous in elucidating the significance of heterotrophic bacterial activity in aquatic environments. Conversion factors used for adjusting biovolume or cell numbers to biomass influence the calculation of turnover rates and nutrient fluxes in almost all environmental microbiological studies.

Variability in the conversion factors for carbon per cell (C_c) or carbon per volume (C_v) range over more than an order of magnitude both between and within different studies (Fagerbakke, et al., 1996; Bratbak, 1985). Due to this large variability, most investigators have pointed out the advisability of determining the conversion factor for the particular environment under study. An excellent compilation of biovolume conversion factors through 1996 is given in Fagerbakke, et al. (1996). This paper provides values corresponding with environment (freshwater, marine, cultured, etc.) or organism (native communities, supplemented cells, cultured, etc.). Since 1996 a few more studies have added to the literature on conversion factors (Robertson, et al., 1998; Fukuda, et al., 1998; Loferer-Kroßbacher, et al., 1998; Troussellier, et al., 1997). Of particular interest is the study by Troussellier, et al. (1997), where previous speculation on the non-linear relationship between bacterial carbon content and cell volume was experimentally validated for 10 different Proteobacteria (marine and non-marine) under starvation conditions. This correlation, where smaller cells have a higher content of C_v , than larger cells had been tentatively confirmed in previous studies (Norland, 1993; Norland, et al., 1987; Simon and Azam, 1989; Lee and Fuhrman, 1987; Kroer, 1994; Psenner, 1990). Troussellier, et al. (1997) and Kroer (1994) expressed this relationship on a C_v basis, whereas Simon and Azam (1989), Norland, et al. (1987), and Lee and Fuhrman (1987) expressed the relationship on a C_c basis. Both methods produce different results and give differential weights (i.e. scaling factor) to organisms of different size.

While most of the investigations tend to infer that the relationship between carbon content and cell size is allometric, this occurs due to the common practice of

plotting the data on log-log scales. Allometry refers to the relationship between relative growth of a part in relation to the whole organism. Inherently, the distribution underlying the log-log transformation is non-linear and this should not be forgotten when attempting to use conversion factors. Given a large variability in conversion factors as well as the non-linear relationship it becomes important to determine whether a single conversion factor is appropriate or if a non-linear function that applies differential values to cells of different volumes is more appropriate. A simple allometric relationship would imply that the use of a single conversion factor for individual bacteria in communities and populations of variable biovolume is appropriate; however, doing so can result in either under or overestimation of the total biomass (Troussellier, et al., 1997; Lee and Fuhrman, 1987; Psenner, 1990; Verity, et al., 1992; Kroer, 1994; Vidondo, et al., 1997). Modeling of the biomass/biovolume interdependence is generally achieved with a power function due to the non-linear form of the relationship. The function normally used to describe this type of data is C_c or $C_v = bVol^a$, where a is termed the location parameter and b is a constant that relates the amount of carbon present at unit volume. Underlying this power relationship is a probability distribution of the Pareto type (Johnson, et al., 1987; Vidondo, et al., 1997). Log-log plots of biovolume vs. C_v have generally been used to ascertain the correlation between these two variables. Troussellier, et al. (1997) found such a relationship between biovolume and C_v , implying that the use of a single value for the conversion of biovolume to biomass is appropriate for an accurate determination of biomass distributions in bacterial communities.

Similar to investigations of bacterial biomass, a variety of models based on biomass-size distributions in aquatic organisms such as fish, phytoplankton, and zooplankton have been derived (Boudreau, et al., 1991; Borgmann, 1987; Ahrens and Peters, 1991; Gaedke, 1993; Sheldon, et al., 1972; Dickie, et al., 1987; Harris, 1994). These models work quite well in describing predator/prey relationships and the transfer of energy between trophic levels. These models are also of the power form and generally presented as log-log plots. Vidondo, et al. (1997) recently showed that by using the normalized biomass-size spectra (NB-SS) models in analyzing size distributions, investigators have unknowingly been assuming that the underlying probability distribution follows a Pareto relationship. Originally, the Pareto distribution was developed in order to model income distributions in populations, but it has subsequently been used to model the distribution of the largest and smallest values for natural phenomena such as waves, pollutant concentrations, demographics, and floods (Castillo and Hadi, 1997; Vogel, et al., 1993; Johnson, et al., 1987; Zelterman, 1992; Lewis and Chatwin, 1997; Hosking and Wallis, 1987; Vidondo, et al., 1997). Since bacteria are considered the smallest of living organisms (neglecting viruses) and seem to demonstrate size extremes, then perhaps, biovolume/biomass can also be modeled by a Pareto distribution. As well as noting the equivalency of the NB-SS and the Pareto distribution, Vidondo, et al. (1997) stated that this type of data should be viewed and analyzed as a probability distribution. If the data conform to a probability distribution then we can take advantage of advances in the analysis, identification, and description of this type of data. For example, methodological advances have been made for identifying and describing the

parameters underlying different types of probability distributions using L-moment diagrams (Vogel and Fennessey, 1993; Hosking, 1990; Wang, 1996; Vogel, et al., 1993). L-moments are analogous to conventional moments (i.e. mean, variance, skewness and kurtosis), but they can characterize a larger variety of distributions and are more robust against outliers when estimated from samples. The ability to discriminate between different distributional hypotheses is made easier with L-moments. They are often superior alternatives for summarizing probability distribution in environmental samples. In this study we adopted methods used in hydrology for the analysis of the probability distributions of flood frequencies and applied these to bacterial biovolumes, C_c , and C_v .

While predictions based on the biomass-size spectra models used in macroecology are effective at higher trophic levels, bacteria have been difficult to model with this method (Gaedke, 1993). The first reason is that the models assume that carbon flow proceeds from small to large organisms. Trophic status in bacterial communities can be controlled either by top-down pressures due to grazing or bottom-up pressures due to resource limitation (Gaedke, 1993; Psenner and Sommaruga, 1992; Gasol, et al., 1997; Meyer, 1994; Pernthaler, et al., 1996; Felip, et al., 1996; Emerson, et al., 1994). Secondly, the models assume an allometric relationship, which has been investigated and shown to occur for cultured cells and bacteria from natural environments. However, data in the literature supporting this assumption are scarce. Thirdly, a coupling between metabolic activity and biomass, as is also implied by the allometric biomass-size spectra models used in oceanography and limnology, has generally not been detected in bacteria. Highly variable measurements on the active fraction of bacteria in aquatic environments

have been reported by a number of investigators (Tuomi, et al., 1995; Pernthaler, et al., 1996; Grossmann, 1994; Karner and Fuhrman, 1997).

In this study we investigate the biomass-size probability distribution of mixed heterotrophic bacteria in a semi-controlled, nutrient-supplemented and carbon-limited biofilm bioreactor originally intended to model drinking water distribution systems (Van Der Wende, et al., 1989). Oligotrophic drinking water with an assimilable organic carbon (AOC) content of approximately 25 μg acetate carbon equivalents/L was supplemented with 3 different carbon sources comprising those most frequently detected in aquatic systems. Investigated substrates included amino acids, carbohydrates, and humics supplied in separate experiments at three different carbon loading rates (250, 500 and 1000 μg carbon/L \cdot hour) simulating different trophic states. Parallel reactors were run simultaneously with one acting as a control and the second receiving chlorine to investigate the effects of chlorine stress on the biomass distribution in the presence of different carbon sources. Probability distributions for biovolume and biomass were determined and found to follow a type II 3-parameter generalized Pareto distribution (GPD) in 68 of 72 datasets comprising a total of approximately 36,000 individual measurements. Carbon source, loading rate, and chlorination were found to significantly affect the mean and median biovolumes, as well as the parameters that describe the biovolume distribution. An empirical relationship between cell volume and C_v is derived that takes into account theoretical limitations on bacterial size and C_v . This relationship produces estimates that are more reasonable for C_v at the extreme small and large end of the biovolume spectrum when compared to models based on power functions.

Materials and Methods

Experimental System

The generation of mixed heterotrophic bacterial populations was accomplished with the use of a model biofilm reactor system comprised of rotating annular reactors (Van Der Wende, et al., 1989). Polycarbonate based, the reactors consist of a stationary outer cylinder with a rotating inner drum forming an annulus space between the inner drum and the outer cylinder. Twelve removable polycarbonate slides are located in machined grooves in the inside wall of the outer cylinder, forming a relatively smooth, even surface between the cylinder wall and the slides. The total submerged surface area in the reactors was 0.18 m^2 and the liquid volume was 630 mL. The reactor top has twelve holes with stoppers that are not in contact with the liquid phase and can be removed to allow access to the slides without interrupting reactor operation. The rotational speed of the inner drum can be varied to create the desired shear stress on the wall of the outside cylinder. In this study a rotational speed of 30 revolutions per minute was used in all work, simulating the shear stress at the wall of a 4-inch-diameter (102-mm-diameter) pipe with a flow velocity of 1.0 foot per second (0.3 meters per second). Draft tubes in the inner rotating drum assist in providing complete mixing within the reactor. These reactors can be modeled as continuously stirred tank reactors (CSTRs). The total flow rate to the reactor was set to provide a hydraulic detention time of 2.1 hours for all experiments. The colonized areas of the annular reactors were immersed in a

